第42卷 第3期
2023年 5月

华 中 农 业 大 学 学 报
Journal of Huazhong Agricultural University

Vol.42 No.3
May 2023,241～249

牛晓虎,冯耀泽,鲍雪,等.基于高光谱技术的细菌生物被膜分类检测[J].华中农业大学学报,2023,42(3):241-249.
DOI:10.13300/j.cnki.hnlkxb.2023.03.028

基于高光谱技术的细菌生物被膜分类检测

牛晓虎[1],冯耀泽[1,2,3,4],鲍雪[5],崔恒洁[5],王梦冉[5],岑晓旭[1],孙光全[1]

1.华中农业大学工学院,武汉 430070; 2.农业农村部水产养殖设施工程重点实验室,武汉 430070;
3.农业农村部长江中下游农业装备实验室,武汉 430070;
4.中国农业科学院深圳农业基因组研究所,深圳 518000;
5.华中农业大学动物科学技术学院、动物医学院,武汉 430070

摘要 针对现有生物被膜检测方法耗时、费力、低效的问题,以大肠杆菌、金黄色葡萄球菌、沙门氏菌为例,研究荧光高光谱技术对不同细菌生物被膜进行种类识别和成膜能力评价的可行性。采集细菌生物被膜样本荧光高光谱图像,并基于5种方法预处理后的光谱数据建立支持向量机分类(support vector classification machine,SVC)和偏最小二乘判别分析(partial least squares discriminant analysis model,PLS-DA)细菌被膜分类检测模型。利用连续投影算法(successive projections algorithm,SPA)、竞争性自适应重加权算法(competitive adaptive reweighted sampling,CARS)分别提取特征波长并建立相应简化模型。结果显示:细菌生物被膜种类识别全波长和特征波长模型中SVC性能均优于PLS-DA,最优模型为None-SPA-SVC,校正集和预测集分类准确率均为96.67%。在细菌生物被膜成膜能力的全波长模型分类判别中,SVC算法整体上分类准确率优于PLS-DA;对于简化模型,最优模型为SPA-SVC,校正集和预测集分类准确率分别为100.00%和96.67%。研究结果表明,高光谱技术可以对细菌生物被膜种类及生物被膜的成膜能力进行有效、快速、准确地分类。

关键词 高光谱技术; 生物被膜; 成膜能力; 偏最小二乘分类; 支持向量机分类

中图分类号 Q93-31 **文献标识码** A **文章编号** 1000-2421(2023)03-0241-09

生物被膜是细菌和细菌衍生物(蛋白质、多糖、肽聚糖和磷脂等物质)不可逆地附着于生命体或无生命物体表面形成的细菌群体,它将整个菌体群落包裹其中,是细菌有组织生长的聚集体,对抗生素和消毒杀菌剂等外界因素有着较强的抗性[1]。物体表面的细菌99.9%是以生物被膜的形式存在的,附着在加工食品表面的细菌通过生物被膜这一屏障直接或间接地进入食品中,对人们健康构成严重威胁[1-2]。致病性微生物引起的食源性疾病是食品安全的主要隐患[3],因此准确高效地检测并识别生物被膜对于食品安全有着重要意义。

生物被膜种类目前大多通过分子生物学方法和质谱技术进行识别。Wagner等[4]利用rRNA靶向定量PCR来检测肉类加工环境中生物被膜的存在,通过基质辅助激光解吸电离飞行时间质谱仪(matrix assisted laser desorption ionisation time of flight,MALDI-TOF)对检测到的生物被膜进行种类识别,生物被膜的检测结果和分类识别能力精确,但设备昂贵,使用复杂,检测成本较高。生物被膜成膜能力的检测主要依靠对胞外聚合物(extracellular polymeric substances,EPS)成分的分析。Wanger等[5]通过检测6种菌株3种不同水平的胞外聚合物EPS成分判断不同菌株形成生物被膜的能力,其中2种仅含碳水化合物和eDNA,3种仅含糖,所有菌株的生物膜中都含有碳水化合物。根据EPS成分的浓度可以准确地评价细菌生物膜形成能力,但该方法耗时长,且操作繁琐。为快速鉴定生物被膜种类并评估其生长状态,亟需开发一种食品加工表面细菌生物膜快速无损检测方法。

高光谱技术集光谱与图像于一体,常用于各种

收稿日期: 2022-09-13
基金项目:中央高校基本科研业务费专项(2662020GXPY003);华中农业大学深圳营养与健康研究院研发项目(SZYJY2021028)
牛晓虎,E-mail:1164556758@qq.com
通信作者:冯耀泽,E-mail:yaoze.feng@mail.hzau.edu.cn

物料的无损检测[6-8]。在细菌检测方面，Lee等[9]利用荧光高光谱图像技术在高密度聚乙烯和不锈钢表面检测识别大肠杆菌(*Escherichia coli*)和鼠伤寒沙门菌(*Salmonella typhimurium*)2种细菌的生物膜形成区域，提取了2种材料上培养的细菌生物膜光谱的像素点建立了决策树、线性判别分析、K近邻分析和偏最小二乘判别分析4种预测模型，在像素水平上对生物膜和背景分类预测准确率均超过90%，其中K-NN模型性能最优，平均准确率超过94%。余伟等[10]利用高光谱技术对培养基上大肠杆菌、李斯特菌和金黄色葡萄球菌等菌落进行快速识别和分类，建立了PLS-DA全波长模型和简化模型，全波长模型的预测集样本分类和置信预测分类的准确率分别达到100%、96%，利用竞争性自适应重加权算法(competitive adaptive reweighted sampling，CARS)提取波长的简化模型校正集和预测集的准确率分别达到100% 和 98%。上述研究中，高光谱技术一般用于细菌菌落的识别分类，还有一些研究利用高光谱技术对生物被膜进行种类识别，对于生物被膜成膜能力的分类检测大多基于生物学方法，鲜有利用高光谱技术对生物被膜成膜能力检测的研究。因此，本研究利用高光谱技术对生物被膜的种类进行识别，并对不同种类的生物被膜的成膜能力进行评价，旨在为细菌生物被膜分类识别用便携式检测设备的研制提供理论基础。

1 材料与方法

1.1 细菌活化与样本制备

试验所用细菌肠道致病性大肠杆菌(*E. coli* O157：H7，后文简写为EC)、鼠伤寒沙门氏菌(*Salmonella typhimurium* SH16SP46，后文简写为ST)、金黄色葡萄球菌(*Staphylococcus aureus* ZL0126，后文简写为SA)均采集于华中农业大学农业微生物资源发掘与利用全国重点实验室。将3种菌株在LA(Luria-Bertani agar)板上划线，37 ℃的条件下活化培养12～24 h。活化后菌株分别接种LB(Luria-Bertani)培养基，37 ℃，200 r/min条件下振荡培养8～10 h，调节LB调菌液OD_{600}=1(相当于10^8 CFU/mL)，备用。

将3种细菌菌液按照1∶100稀释，加入96孔细胞培养板中培养，每孔100 μL。由于不同细菌的成膜能力不同，设置不同培养时间和条件：大肠杆菌28 ℃静置培养24、72、120、168 h；沙门氏菌30 ℃静置培养12、24、48、80 h；金黄色葡萄球菌28 ℃静置培养6、12、24、80 h。阴性对照组为无菌培养基。

不同培养阶段完成后，取当前96孔板生物被膜待测样本，弃去其中培养液，用超纯水洗涤微孔3次并晾干，将干燥的96孔板封装在无菌盒中，用来采集光谱图像。高光谱图像采集完成之后，样本染色，每孔加入125 μL 1%的结晶紫溶液染色固定15 min，流水轻轻冲掉染液，轻甩晾干，每孔加入33%的乙酸溶液150 μL，振荡溶解10 min，使用酶标仪测定OD_{630}，得到样本OD值，除阴性对照外共得到120个样本，其中EC、ST、SA样本分别为32、64、24个。

1.2 荧光高光谱图像采集与处理

试验使用的高光谱成像系统由高光谱相机(V10E-CL，芬兰)、激发波长为365 nm的紫外灯光源、计算机、高精度电控移动平台组成，其光谱范围是300～1 100 nm，光谱分辨率为2.8 nm，光谱间隔为1.25 nm，共520个波段。荧光高光谱采集前，高光谱成像系统预热30 min，并设置相机曝光时间为100 ms，电控移动平台移动速度为1.6 mm/s，保持待测表面与高光谱相机的镜头距离为30 cm左右。对采集到的光谱图像，基于435 nm波长的图像进行感兴趣区域提取[11]，并计算感兴趣区域中全波段平均光谱，作为样本的光谱用于后续建模分析。

1.3 细菌被膜光谱数据预处理

为提升模型性能，采用卷积平滑(Savitzky-golay，SG)、矢量归一化(vector normalization，VN)、多元散射校正(multiplicative scattering correction，MSC)、标准正态变量变换(standard normalized variate，SNV)对光谱数据进行预处理以减少或消除其所包含的无关信息和噪声[12]。

1.4 细菌被膜检测模型的优化

利用SPXY[13]算法将样本数据按照3∶1的比例划分为校正集和预测集。在此基础上，分别应用偏最小二乘判别分析PLS-DA和支持向量机分类(support vector classification，SVC)建立细菌生物被膜的分类模型。其中，PLS-DA在对自变量X和因变量Y中提取主成分的同时保证了主成分之间相关性最大，通过对回归的结果进行判别分析。SVC是典型的非线性分类方法，其将低维数据映射到高维空间，通过寻找最优超平面实现样本的分类。其中，选择得当SVC的惩罚系数C和核函数g能够保证模型精

度高、泛化能力强[14]。本研究使用遗传算法优化SVC的惩罚系数C和核函数参数g对生物被膜及其成膜能力分类。

1.5　细菌被膜检测特征波长优选

本研究采用竞争性自适应重加权算法(CARS)、连续投影算法(successive projections algorithm, SPA)对波长进行选择,并建立基于优选特征波长的简化模型。竞争性自适应重加权算法通过自适应重加权采样技术选择模型中回归系数绝对值最大的波长点,再利用交叉验证选出最小的子集作为选择的波长[15]。连续投影算法(SPA)可以较好地从光谱数据中选择无共线性和冗余的特征波长,同时可以缩短建模时间。SPA在选择特征波长时不仅控制均方根误差尽可能小,而且需要总体方差没有显著性差异,本研究通过$\alpha=0.25$的F检验确定总体方差显著性[16]。

1.6　细菌被膜检测模型的评价

根据OD_{630}值对细菌生物被膜成膜能力进行分级。定义ODC(cutoff optical density)为阴性对照平均值加3倍标准差(standard deviation,SD)值。菌株OD<ODC为微弱生物被膜,Ⅰ级成膜能力;ODC<OD<2 ODC为弱生物被膜,Ⅱ级成膜能力;2 ODC<OD<4 ODC为中等生物被膜形成菌株,Ⅲ级成膜能力;OD>4 ODC为强生物被膜形成菌株,Ⅳ级成膜能力[17]。为了建模方便,对于细菌生物被膜种类识别,EC类别标签为0,ST类别标签为1,SA类别标签为2;而对于生物被膜成膜能力等级分级,微弱、弱、中等和强成膜能力的标签分别为1、2、3、4。以标签值作为Y值,光谱数据作为自变量X,在Y值附近设定1个阈值,一般为±0.5,如果预测结果落在某个Y值范围之内,则认为该结果属于对应Y值代表的样本类别。

使用正确分类率(correct classification rate, CCR)来评估模型性能[18]。CCR定义为整个样本空间中正确分类样本的百分比,CCR越高,模型性能越好。其中,校正集分类准确率(correct classification rate of correction set, CCRC)和预测集分类准确率(correct classification rate of prediction set, CCRP)分别代表校正集和预测集样本的分类准确率。混淆矩阵的每行代表真实值,每列代表预测值,通过混淆矩阵,可以更直观地表现评价指标CCR。

2　结果与分析

2.1　生物被膜光谱

图1是菌液浓度为10^8 CFU/mL的3种细菌生物被膜4个培养阶段时的平均荧光光谱图(为了清晰地区分3条曲线,将EC(*E. coli*)向上平移了100个单位,SA向下平移了100个单位)。从3种细菌生物膜荧光强度来看,它们的荧光发色基团大体相同,但基团含量、生物膜胞外物质含量以及生物膜结构特性的不同导致发射的荧光强度值有所差异[19]。

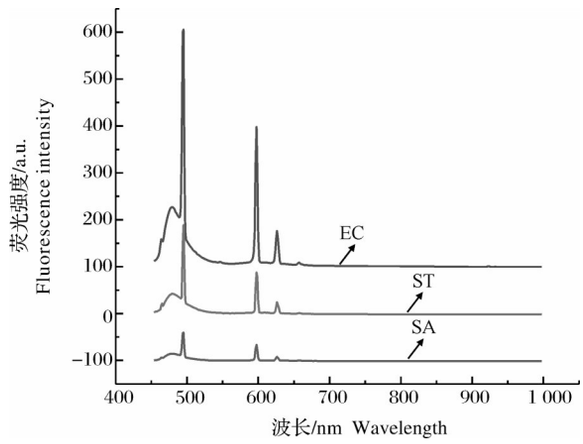

图1　EC、ST、SA生物被膜样本荧光光谱

Fig.1　EC,ST,SA biofilm sample fluorescence spectra

由图1可知,在波长为450、570和600 nm处均有明显的波峰,且在450 nm处3种细菌生物膜有着最大区分度,该波长可能对应还原酶(nicotinamide adenine dinucleotide, NADH),还原性辅酶(nicotinamide adenine dinucleotide phosphate, NADPH)的荧光响应。NADH的烟酰胺基团和NADPH的烟酰胺环基团在340 nm附近吸收光子,并在460 nm附近发射荧光,这2种酶控制细菌生物膜的代谢,参与有氧呼吸以及脂质、氨基酸、核苷酸等合成,WI-1峰和WI-2峰波长比较接近,可能对应着原卟啉基团的荧光响应[20]。WI-1峰左侧有1个明显的宽峰,可能是由于处于激发态的光量子回到基态时有不同的振动能级,能级之间间隔很小导致在400~450 nm形成1个较宽的带状峰[21]。

2.2　3种细菌生物被膜分类模型

1) 全波段PLS-DA模型结果。通过PLS-DA结合不同的光谱预处理方法,建立基于全波长信息的不同培养时间梯度下3种细菌生物被膜的分类检测模型。以细菌生物被膜分类准确率为指标,比较分析不同光谱预处理下PLS-DA分类模型的识别效

果，确定3种细菌生物被膜分类的最佳分类模型。

由表1可知，在不经过任何预处理的条件下，校正集和预测集的准确率为96.67%、100.00%。预测集的分类准确率略高于校正集，可能原因是在PLS-DA线性分类的框架下，原始数据的噪音对多样本训练集在高维空间中投影坐标及载荷确定中有一定影响从而带来误差。

表1 基于不同预处理PLS-DA的生物被膜分类识别结果

Table 1 Performance of PLS-DA models established under different preprocessing methods for classification of bacterial biofilms

预处理方法 Pre-processing method	校正集准确率/% Calibration set accuracy	预测集准确率/% Validation set accuracy	主成分数 Principal component
无 None	96.67	100.00	8
SG	98.89	96.67	10
SNV	100.00	86.67	12
VN	100.00	90.00	12
MSC	100.00	90.00	13

由表1可知，经过SG预处理的PLS-DA模型的校正集和预测集准确率最为接近，分别为98.89%、96.67%，而经过SNV、VN、MSC预处理的校正集准确率虽然都为100.00%，但是它们的预测集准确率与校正集相差较大，预测效果相对较差。因此，SG-PLS-DA为最优的PLS-DA生物被膜全波长分类识别模型。

2)全波段SVC模型结果。由表2可知，在不同的预处理条件下，SG预处理的支持向量机分类模型SG-SVC预测准确率最低，CCRC和CCRP分别为95.56%和86.67%，经过SNV、VN和不做任何预处理的支持向量机分类模型校正集和预测集的分类准确率均为100.00%。经过MSC预处理的SVC模型的校正集准确率为100.00%，但预测集准确率仅为96.67%。因此，不做任何预处理的SVC模型None-SVC为最优生物被膜分类模型，其预测集准确率为100.00%。

表2 基于不同预处理SVC的生物被膜分类识别结果

Table 2 Performance of SVC established under different preprocessing methods for classification of bacterial biofilms

预处理方法 Pre-processing method	校正集准确率/% Calibration set accuracy	预测集准确率/% Validation set accuracy	惩罚因子 Penalty actor	核函数 Kernel function
无 None	100.00	100.00	256.0	0.329 9
SG	95.56	86.67	147.03	0.189 5
SNV	100.00	100.00	27.86	0.006 8
VN	100.00	100.00	16.00	0.108 8
MSC	100.00	96.97	5.278	0.189 5

3)简化模型。为建立简化的生物被膜分类模型，分别采用CARS和SPA筛选特征波长，其分布如图2所示。图2中竖线为特征波长的分布情况，CARS和SPA分别筛选出66和10个特征波长。从图2可以看出，CARS筛选的特征波长整体上分布较为均匀，在450、550、850 nm处有着比较明显的聚集分布，SPA筛选的特征波长在400～500、500～800 nm波段之间的分布较多，在400～500 nm的分布比较均匀。

由表3可知，经过CARS筛选后的PLS-DA模型

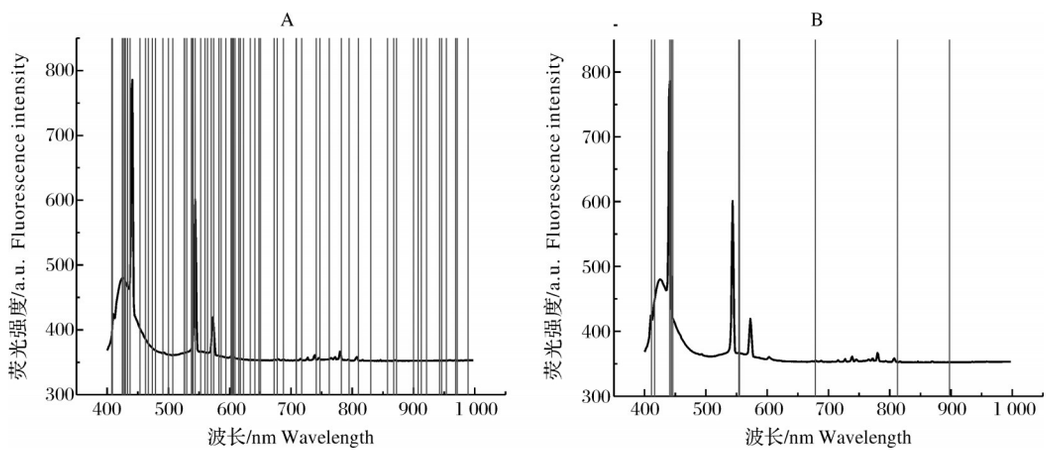

图2 CARS（A）和SPA（B）筛选出的特征波长分布图

Fig.2 Characteristic wavelength distribution map selected by CARS（A），and SPA（B）

的校正集和预测集准确率分别为82.22%和86.67%。经过SPA筛选浓度特征波长建立的PLS-DA模型的校正集和预测集准确率分别为91.11%和93.33%，略优于CARS-PLS-DA。然而，经过2种特征波长选择算法处理后的PLS-DA简化模型性能均低于全波长模型的性能。经过2种特征波长选择算法建立的SVC简化模型的校正集和预测集准确率均为96.67%，与全波长模型的性能相当。因此，对于3种细菌生物被膜分类，其最优模型为None-SPA-SVC。

图3为3种细菌生物被膜最优模型分类结果混淆矩阵，在校正集分类结果中，3个ST错分成了SA，在预测集分类结果中，1个ST错分成了SA，表明金黄色葡萄球菌生物被膜和沙门氏菌生物被膜之间的差异较小，一方面ST和SA光谱强度曲线接近（图1），另一方面，二者的成膜能力也非常相近。

表3　不同波长选择方法下建立的生物被膜分类识别简化模型性能

Table 3　Performance of a simplified model for biofilm classification and recognition with different wavelength selection methods

波长选择方法 Wavelength selection method	PLS-DA			SVC			
	校正集准确率/% Calibration set accuracy	预测集准确率/% Validation set accuracy	主成分数 Principal component	校正准确率/% Calibration set accuracy	预测准确率/% Validation set accuracy	惩罚因子 Penalty factor	核函数 Kernel function
CARS	82.22	86.67	3	96.67	96.67	84.45	5.278 0
SPA	91.11	93.33	5	96.67	96.67	256.00	3.031 4

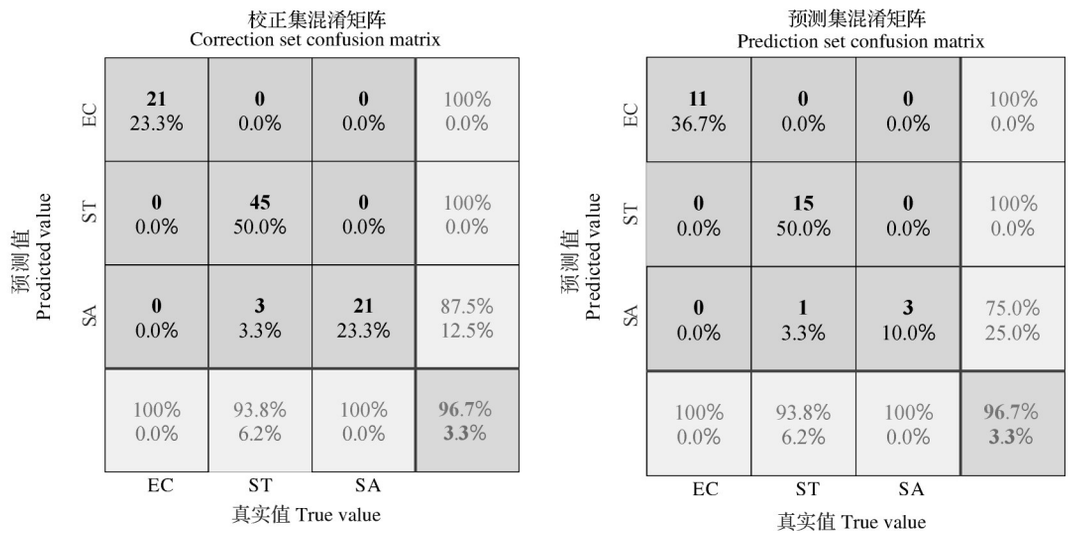

图3　3种细菌生物被膜最优模型分类结果混淆矩阵

Fig. 3　Confusion matrix of the optimal model for classification of bacterial biofilm

2.3　3种细菌生物被膜成膜能力分级模型校正集混淆矩阵

1）成膜能力PLS-DA全波长分类模型。由表4可知，对于不经过光谱预处理建立的PLS-DA模型所得到的分类校正集准确率和预测集准确率分别为83.33%和86.67%。因预测集准确率大于校正集准确率，尝试对原始数据进行预处理优化。经过SNV、VN、MSC预处理的模型校正集准确率都得到了一定程度的提高，但预测集准确率均为73.33%，与校正集的准确率有较大的差距。而经过SG预处理的SG-PLS-DA校正集和预测集准确率分别为84.44%和83.33%，具有较高的分级准确率和稳定性。因此，对于PLS-DA分类模型最优的预处理方法为SG预处理。

2）成膜能力SVC全波长分类模型。对4个培养阶段的3种细菌生物被膜光谱进行不同预处理后建立SVC分类模型（表5）。由表5可知，经过VN、MSC预处理的模型VN-SVC、MSC-SVC的校正集准确率和预测集准确率分别为100.00%、100.00%和93.33%、93.33%，二者准确率相差较大，模型稳定性欠佳。不经过任何预处理以及经过SG预处理的光

表4 基于不同预处理PLS-DA的生物被膜成膜能力分类识别结果

Table 4 Classification and recognition results of biofilm film forming ability based on different pretreated PLS-DA

预处理方法 Pre-processing method	校正准确率/% Calibration set accuracy	预测准确率/% Validation set accuracy	主成分数 Principal component
无 None	83.33	86.67	7
SG	84.44	83.33	7
SNV	87.78	73.33	9
VN	97.78	73.33	12
MSC	97.78	73.33	12

表5 基于不同预处理SVC的生物被膜成膜能力分类识别结果

Table 5 Classification and recognition results of biofilm film forming ability based on different pretreated SVC

预处理方法 Pre-processing method	校正准确率/% Calibration set accuracy	预测准确率/% Validation set accuracy	惩罚因子 Penalty factor	核函数 Kernel function
无 None	92.22	93.33	5.278	1.00
SG	92.22	93.33	5.278	1.00
SNV	96.67	93.33	3.031	0.035 9
VN	100.00	93.33	84.45	0.020 6
MSC	100.00	93.33	48.04	0.035 9

谱数据建立的模型None-SVC、SG-SVC的校正集和预测集准确率分别为92.22%、92.22%和93.33%、96.33%，而经过SNV预处理的光谱数据建立的模型SNV-SVC的校正集和预测集准确率分别为96.67%和93.33%，二者相差不大且校正集准确率高于其他预处理方法对用模型性能。因此，对于SVC分类全波长模型，SNV预处理是最优预处理方法。

3）成膜能力分级简化模型。运用CARS和SPA分别筛选出45和10个特征波长，其分布情况如图4所示。从图4可以看出，CARS和SPA筛选出的特征波段在全波段范围内分布都比较均匀，在450、550、575 nm附近分布相对集中。由表6可知，由CARS算法筛选出的特征波长建立的PLS-DA/SVC分类模型得到的模型校正集和预测集准确率均为93.33%。而由SPA波长选择算法筛选出的特征波长建立的PLS-DA分类模型校正集和预测集准确率分别为75.56%、93.33%，出现了严重的过拟合现象。对SPA-SVC分类模型，校正集准确率和预测集准确率为100.00%、96.67%，模型性能优于PLS-DA。SPA-SVC模型得到的校正集和预测集准确率均大于CARS-SVC的模型结果，虽然校正集与预测集差值更大，但SPA-SVC预测集仅预测错误1个样本，考虑到样本数量问题，最优模型为SPA-SVC。

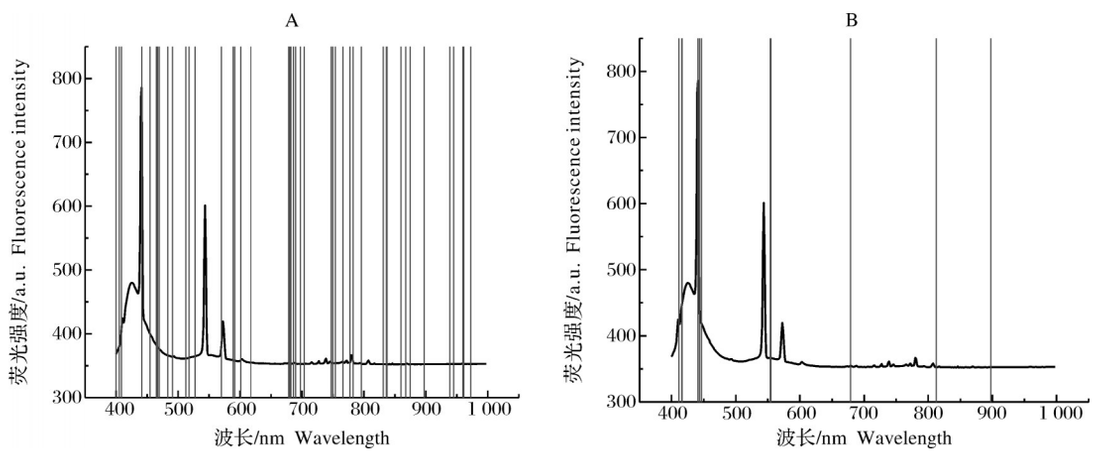

图4 CARS（A）和SPA（B）筛选出的特征波长

Fig.4 Characteristic wavelength distribution map selected by CARS（A），and SPA（B）

3 讨 论

本研究采集3种不同细菌生物被膜的荧光高光谱图像，结合不同光谱预处理方法和特征波长选择方法，建立了基于SVC和PLS-DA的不同细菌生物被膜分类模型和生物被膜成膜能力分级模型。结果显示：（1）对于3种细菌生物被膜的全波长分类模型，SG、VN、MSC、SNV等不同预处理方法对模型性

表6　不同波长选择方法下建立的生物被膜成膜能力分类简化模型性能

Table 6　Performance of a simplified model for classification of biofilm film-forming capacity under different wavelength selection methods

波长选择方法 Wavelength selection method	PLS-DA			SVC			
	校正集准确率/% Calibration set accuracy	预测集准确率/% Validation set accuracy	主成分数 Principal component	校正准确率/% Calibration set accuracy	预测准确率/% Validation set accuracy	惩罚因子 Penalty factor	核函数 Kernel function
CARS	93.33	93.33	6	93.33	93.33	147.03	1.741 1
SPA	75.56	93.33	6	100.00	96.67	256.00	5.278 0

能的影响不同，SVC最优全波长分类模型为None-SVC，PLS-DA最优全波长分类模型为SG-PLS-DA。对于3种细菌生物被膜的简化分类模型，利用CARS和SPA分别得到66、10个特征波长，基于特征变量所建立最优简化模型为None-SPA-SVC，校正和预测准确率均达到96.67%。(2)对3种细菌生物被膜的成膜能力建立了全波长分类模型和简化模型，对于SVC全波长模型，SNV是最优的预处理方法，得到的SNV-SVC模型校正集和预测集准确率分别为96.67%和93.33%，对于PLS-DA全波长分类模型，SG是最优的预处理方法，SG-PLS-DA校正集和预测集准确率分别为84.44%和83.33%，而对于简化模型，SPA波长选择算法要优于CARS，SPA-SVC为3种细菌生物被膜成膜能力的最优分类模型，校正集和预测集准确率分别为100.00%、96.67%。另外，整体上SVC建立的模型性能要优于PLS-DA。

本研究结果表明，对于细菌高光谱图像样本的分类模型，SVC在预测样本上的性能表现总体优于PLS-DA，原因是PLS-DA属于线性分类，建立在偏最小二乘法回归的基础上，在变量之间存在多重相关性、变量多、样本容量小等问题时，其具有不可替代的优势，但是对于特征数较多的三分类问题来说，只使用2个超平面将不同类别样本分开容易出错，而SVC通过将数据映射到更高维度寻找多个超平面甚至曲面将某些线性不可分的数据点分开，有着更好的拟合能力。

本研究表明，高光谱技术可以快速无损有效地对不同细菌的生物被膜进行种类识别以及评价生物被膜成膜能力，准确度高，相比生物学检测成本更低，检测速度更快。

参考文献 References

[1] 邹明明，王文骏，马晓彬，等.细菌生物被膜的研究进展[J].中国食品学报，2017，17(7)：156-164.ZOU M M，WANG W J，MA X B，et al.Research progress on bacterial biofilm[J].Journal of Chinese Institute of Food Science and Technology，2017，17(7)：156-164(in Chinese with English abstract).

[2] 马悦，吴梦洁，李卓思，等.食品接触表面生物被膜形成机制及防控方法研究进展[J].食品科学，2023，44(7)：276-285.MA Y，WU M J，LI Z S，et al.Research progress of biofilm formation mechanism and prevention and control methods on food contact surface[J].Food science，2023，44(7)：276-285(in Chinese with English abstract).

[3] 李燕杰，杨公明，朱小花，等.从生物被膜看食品机械安全性设计准则的必要性[J].农业工程学报，2008，24(11)：302-307.LI Y J，YANG G M，ZHU X H，et al.Reconsidering safety design criteria for food machinery in terms of biofilm[J].Transactions of the CSAE，2008，24(11)：302-307(in Chinese with English abstract).

[4] WAGNER E M，PRACSER N，THALGUTER S，et al.Identification of biofilm hotspots in a meat processing environment：detection of spoilage bacteria in multi-species biofilms[J/OL].International journal of food microbiology，2020，328：108668[2022-09-13].https://doi.org/10.1016/j.ijfoodmicro.2020.108668.

[5] WAGNER E M，FISCHEL K，RAMMER N，et al.Bacteria of eleven different species isolated from biofilms in a meat processing environment have diverse biofilm forming abilities[J/OL].International journal of food microbiology，2021，349：109232[2022-09-13].https://doi.org/10.1016/j.ijfoodmicro.2021.109232.

[6] 吴叶兰，管慧宁，廉小亲，等.高光谱成像的多种类柑橘病虫药害叶片检测方法研究[J].光谱学与光谱分析，2022，42(8)：2397-2402.WU Y L，GUAN H N，LIAN X Q，et al.Study on detection method of leaves with various citrus pests and diseases by hyperspectral imaging [J]. Spectroscopy and spectral analysis，2022，42(8)：2397-2402(in Chinese with English abstract).

[7] 刘友华，白亚斌，邱祝福，等.基于高光谱图像技术和波长选择方法的羊肉掺假检测方法研究[J].海南师范大学学报(自然科学版)，2015，28(3)：265-269.LIU Y H，BAI Y B，QIU Z F，et al. Hyperspectral imaging technology and wavelength selection

method for nondestructive detection of mutton adulteration[J]. Journal of Hainan Normal University (natural science), 2015, 28(3): 265-269 (in Chinese with English abstract).

[8] 王巧华,马逸霄,付丹丹.基于光谱技术的禽蛋内部品质无损检测研究进展[J].华中农业大学学报,2021,40(6):220-230. WANG Q H, MA Y X, FU D D. Progress of non-destructive detection of poultry egg internal quality based on spectroscopy[J]. Journal of Huazhong Agricultural University, 2021, 40(6): 220-230(in Chinese with English abstract).

[9] LEE A, PARK S, YOO J, et al. Detecting bacterial biofilms using fluorescence hyperspectral imaging and various discriminant analyses[J/OL]. Sensors (Basel, Switzerland), 2021, 21(6): 2213[2022-09-13].https://doi.org/10.3390/s21062213.

[10] 余伟,彭宽宽,陈伟,等.基于高光谱技术的培养基上细菌菌落分类方法研究[J].分析化学,2016,44(8):1221-1226.YU W, PENG K K, CHEN W, et al. Classification of bacterial colonies on agar plates using hyperspectral imaging technology[J]. Chinese journal of analytical chemistry, 2016, 44(8): 1221-1226(in Chinese with English abstract).

[11] LU Y, JIA B B, YOON S C, et al. Spatio-temporal patterns of *Aspergillus flavus* infection and aflatoxin B1 biosynthesis on maize kernels probed by SWIR hyperspectral imaging and synchrotron FTIR microspectroscopy [J/OL]. Food chemistry, 2022, 382: 132340 [2022-09-13]. https://doi. org/10.1016/j. foodchem.2022.132340.

[12] PANG L, WANG L M, YUAN P, et al. Rapid seed viability prediction of *Sophora japonica* by improved successive projection algorithm and hyperspectral imaging[J/OL]. Infrared physics & technology, 2022, 123: 104143[2022-09-13]. https://doi. org/10.1016/j.infrared.2022.104143.

[13] 展晓日,朱向荣,史新元,等.SPXY样本划分法及蒙特卡罗交叉验证结合近红外光谱用于橘叶中橙皮苷的含量测定[J].光谱学与光谱分析,2009,29(4):964-968.ZHAN X R, ZHU X R, SHI X Y, et al. Determination of hesperidin in tangerine leaf by near-infrared spectroscopy with SPXY algorithm for sample subset partitioning and Monte Carlo cross validation[J].Spectroscopy and spectral analysis, 2009, 29(4): 964-968(in Chinese with English abstract).

[14] BIAN X H, LI S J, FAN M R, et al. Spectral quantitative analysis of complex samples based on the extreme learning machine [J].Analytical methods, 2016, 8(23): 4674-4679.

[15] LI H D, LIANG Y Z, XU Q S, et al. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration[J]. Analytica chimica acta, 2009, 648(1): 77-84.

[16] GU P, FENG Y Z, ZHU L, et al. Unified classification of bacterial colonies on different agar media based on hyperspectral imaging and machine learning[J/OL]. Molecules (Basel, Switzerland), 2020, 25(8): 1797 [2022-09-13]. https://doi. org/10.3390/molecules25081797.

[17] 桂中玉,李琳,李冰,等.金黄色葡萄球菌耐药性与生物被膜能力的鉴定[J].现代食品科技,2014,30(3):69-75.GUI Z Y, LI L, LI B, et al. Investigation of antibiotic resistance and biofilm-forming ability in Staphylococcus aureus[J].Modern food science and technology, 2014, 30(3): 69-75(in Chinese with English abstract).

[18] MU K X, FENG Y Z, CHEN W, et al. Near infrared spectroscopy for classification of bacterial pathogen strains based on spectral transforms and machine learning[J].Chemometrics and intelligent laboratory systems, 2018, 179: 46-53.

[19] 刘宇,李增威,邓志鹏,等.激光诱导荧光光谱快速检测食源性致病菌[J].光谱学与光谱分析,2021,41(9):2817-2822.LIU Y, LI Z W, DENG Z P, et al. Fast detection of foodborne pathogenic bacteria by laser-induced fluorescence spectrometry [J]. Spectroscopy and spectral analysis, 2021, 41(9): 2817-2822(in Chinese with English abstract).

[20] KOENIG K, SCHNECKENBURGER H. Laser-induced autofluorescence for medical diagnosis[J].Journal of fluorescence, 1994, 4(1): 17-40.

[21] 姚蒙,王海水.基于线性回归分析法获得NaCl水溶液红外光谱[J].光谱学与光谱分析,2020,40(S1):27-28.YAO M, WANG H S. To obtain of infrared spectra of NaCl aqueous solutions by linear regression analysis[J].Spectroscopy and spectral analysis, 2020, 40(S1): 27-28(in Chinese with English abstract).

Classification and detection of bacterial biofilms based on hyperspectral fluorescence imaging

NIU Xiaohu[1], FENG Yaoze[1,2,3,4], BAO Xue[5], CUI Hengjie[5], WANG Mengran[5], CEN Xiaoxu[1], SUN Guangquan[1]

1. *College of Engineering*, *Huazhong Agricultural University*, *Wuhan* 430070, *China*;
2. *Key Laboratory of Aquaculture Facilities Engineering*, *Ministry of Agriculture and Rural Affairs*, *Wuhan* 430070, *China*;
3. *Agricultural Equipment Laboratory of the Middle and Lower Reaches of the Yangtze River*, *Ministry of Agriculture and Rural Affairs*, *Wuhan* 430070, *China*;
4. *Shenzhen Institute of Agricultural Genomics*, *Chinese Academy of Agricultural Sciences*, *Shenzhen* 518000, *China*;
5. *College of Animal Science and Technology*, *College of Veterinary Medicine*, *Huazhong Agricultural University*, *Wuhan* 430070, *China*

Abstract　Bacterial biofilms widely exist on the surfaces of food processing machinery, medical equipment and the environment, and bring a huge threat to human health due to strong drug resistance. *Escherichia coli*, *Staphylococcus aureus*, and *Salmonella typhimurium* were used to study the feasibility of species identification and evaluate film-forming ability of different bacterial biofilms by hyperspectral fluorescence imaging technology to solve the problems of time-consuming, laborious and inefficient detection of existing biofilms. The hyperspectral fluorescence images of bacterial biofilm samples were collected, and support vector classification machine (SVC) and partial least squares discriminant analysis (PLS-DA) models based on the spectral data preprocessed by five methods were established to classify bacterial biofilms. Characteristic wavelengths were extracted by employing successive projections algorithm (SPA) and competitive adaptive reweighted sampling (CARS) respectively, and the corresponding simplified models were established. The results showed that SVC outperformed the full-wavelength and characteristic-wavelength models of identifying bacterial biofilm species than the PLS-DA, with the optimal model being None-SPA-SVC, where the classification accuracy of calibration set (CCRC) and prediction set (CCRP) were both 96.67%. In the classification and discrimination of film-forming ability of bacterial biofilm, the full-wavelength SVC models generally outperformed PLS-DA with higher classification accuracy. For the simplified models, the optimal model was SPA-SVC, with CCRC and CCRP of 100.00% and 96.67%, respectively. It is indicated that hyperspectral fluorescence imaging technology can effectively, quickly and accurately classify the types of bacterial biofilms and the film-forming ability of biofilms.

Keywords　hyperspectral imaging; biofilm; biofilm-forming ability; partial least squares classification; support vector machine

（责任编辑：陆文昌）