Available online at www.sciencedirect.com
ScienceDirect
Biochimica et Biophysica Acta 1758 (2006) 830–845
BIOCHIMICA ET BIOPHYSICA ACTA
BBA
www.elsevier.com/locate/bbamem

Review

High throughput assessment of cells and tissues: Bayesian classification of spectral metrics from infrared vibrational spectroscopic imaging data

Rohit Bhargava [a], Daniel C. Fernandez [b], Stephen M. Hewitt [c], Ira W. Levin [b,*]

[a] *Department of Bioengineering and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*
[b] *Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, MD 20892, USA*
[c] *Tissue Array Research Program, Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA*

Received 10 February 2006; accepted 9 May 2006
Available online 17 May 2006

Abstract

Vibrational spectroscopy allows a visualization of tissue constituents based on intrinsic chemical composition and provides a potential route to obtaining diagnostic markers of diseases. Characterizations utilizing infrared vibrational spectroscopy, in particular, are conventionally low throughput in data acquisition, generally lacking in spatial resolution with the resulting data requiring intensive numerical computations to extract information. These factors impair the ability of infrared spectroscopic measurements to represent accurately the spatial heterogeneity in tissue, to incorporate robustly the diversity introduced by patient cohorts or preparative artifacts and to validate developed protocols in large population studies. In this manuscript, we demonstrate a combination of Fourier transform infrared (FTIR) spectroscopic imaging, tissue microarrays (TMAs) and fast numerical analysis as a paradigm for the rapid analysis, development and validation of high throughput spectroscopic characterization protocols. We provide an extended description of the data treatment algorithm and a discussion of various factors that may influence decision-making using this approach. Finally, a number of prostate tissue biopsies, arranged in an array modality, are employed to examine the efficacy of this approach in histologic recognition of epithelial cell polarization in patients displaying a variety of normal, malignant and hyperplastic conditions. An index of epithelial cell polarization, derived from a combined spectral and morphological analysis, is determined to be a potentially useful diagnostic marker.

Keywords: Fourier transform infrared (FTIR) spectroscopy; Imaging; Biophotonics; Prostate; Tissue microarray; Bayesian statistics; Likelihood classification; Discriminant; Cancer; Histology; Pathology; ROC

Contents

* Corresponding author.
E-mail address: iwl@helix.nih.gov (I.W. Levin).

0005-2736/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbamem.2006.05.007

1. Introduction

Morphologic examination of tissue specimens using light microscopy is the method of choice for the definitive detection and grading of most human cancers. The information content of microscopy images, however, is limited to a spatial variation in optical properties and requires extensive human observations to recognize both the constitutive histologic entities and the pathologic state. Correlations of morphologic and biochemical tissue differences in conjunction with assessments based on chemically-specific spectroscopic techniques have been suggested as routes to augment studies in pathogenesis [1]. Non-destructive vibrational, infrared and Raman, spectroscopic techniques [2] are especially attractive as they afford an abundance of chemical-specific information without the requirement of extrinsic contrast agents. Significant numbers of studies have involved the measurement of average infrared (IR) spectra of intact tissues [3], exfoliated cells [4], carefully extracted cellular constituents [5] and bodily fluids [6]. Though spatially-resolved infrared micro-spectroscopic characterization of intact biological tissue has also been a topic of continual interest for over 50 years, the past decade has yielded the most systematic studies of various contributory factors in the microspectroscopic analyses of tissue [7]. Since tissue is microscopically heterogeneous, microspectroscopy is the sampling mode most capable of representing accurately tissue complexity in histopathologic determinations. The development of clinical protocols for the routine examination of tissue histology or of localized tumors using IR microspectroscopic methods, however, has not been possible due to several related issues.

First, microspectroscopic instrumentation has conventionally employed point detection and square apertures, measuring several tens of micrometers in size per edge, to restrict radiation incident upon the sample. The coarse spatial specificity of conventional point detection techniques is typically larger than characteristic dimensions of most cells. Consequently, the measured spectra consist of an unpredictable mixture of spectral contributions from neighboring ensembles of other, neighboring cell types. Increasing spatial specificity by decreasing aperture sizes to typical cell size dimensions (~5–10 μm) requires substantial increases in data acquisition times to offset the decreased throughput [8]. These long measurement times, in turn, limit both the number of spectral observations in a particular sample and the number of samples that can be analyzed in a reasonable time interval. This lack of throughput in the number of spectra and samples that can be analyzed has limited large population validations of spectroscopic studies that report promising initial results and has led to numerous different claims of biospectroscopic markers for a variety of tissue types and diseases. Careful validation of initial results in larger studies and the meticulous delineation of confounding variables [9] is critical to the establishment of robust protocols and may often reveal the failure of biomarkers based on a small number of observations. While limiting the statistical validity of obtained results, the lack of testing with large cohorts has also led to significant confusion regarding the potential of spectroscopy for biomedical decision-making. Specifically, the roles of patient to patient variation, variation due to sampling methodology [10] and the role of intra-patient heterogeneity arising from cellular turnover [11] has led to questions of whether vibrational spectroscopy can, in fact, be employed for histopathologic determinations.

The analysis of spectroscopic data for diagnostic decision-making has also been an area of considerable activity. While many approaches employed in the chemometrics of well-defined systems can be adapted, consensus is emerging that the optimal data analysis algorithm would employ spectral data as input and result in a determination that can be represented in terms of quantitative numbers or images interpretable by personnel across several disciplines. Extensive spectral processing is *de rigueur;* baseline corrections, normalization and derivative or orthogonal factor calculations for every spectrum are typically required. In this manuscript, we present a detailed description of how impediments to the vibrational biospectroscopic imaging of tissues and cells can be addressed through rigorous, quantitative tests by an integrated platform of multichannel, multiplexed spectroscopy, high throughput microarray sampling and fast numerical analysis incorporating explicit user control. The application of this approach to prostate histopathology [12] has been recently reported and we explain here previously unreported aspects of the approach. We further describe the use of a spatial index, formed using spectral data, of specific cell types in prostate tissue to examine pathologic conditions.

2. An integrated, high-throughput platform

Fourier transform infrared (FTIR) spectroscopic imaging, tissue microarrays and pattern recognition technologies form the core of the described approach to cell and tissue analysis. We briefly review the technology that we have employed and describe novel features that allow the work reported in this manuscript.