SAND2007-1708C

Scanning Microarrays: Current Challenges and Future Directions

Jerilyn A. Timlin, David M. Haaland,
Michael B. Sinclair, Rachel Noek

Sandia National Laboratories*

* Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000

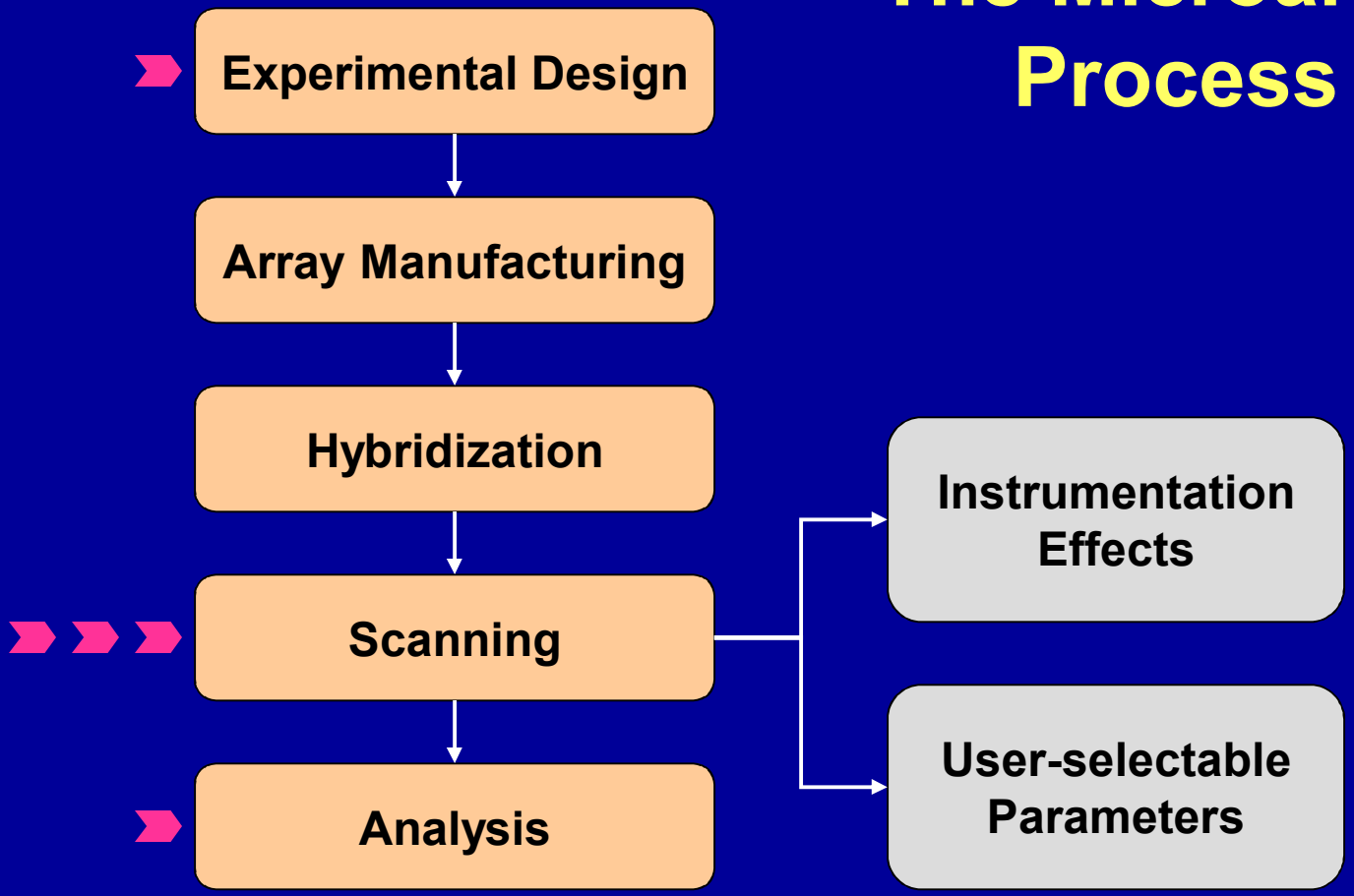
National Nuclear Security Administration

Sandia National Laboratories

Presentation Overview

- Introduction
- Current challenges
  - User
    - Slide handling
    - Display
    - Resolution, averaging
    - PMT settings
  - Instrument
    - Signal contamination, bias
    - Use of calibration slides
- Complimenting technology … spectral imaging

The Microarray Process



Sandia National Laboratories

User-Selectable Parameters – Array Handling

- Fingerprints, dust contaminate signal
- Light and ozone decrease signal
- Positioning:
  - Face-up vs. Face down
  - Scanner matters

Confocal	vs.	Wide-field
< 10 μm F.O.V.		~ 60 μm F.O.V.
↓ background		↑ background
↑ sensitivity to position		↓ sensitivity to position

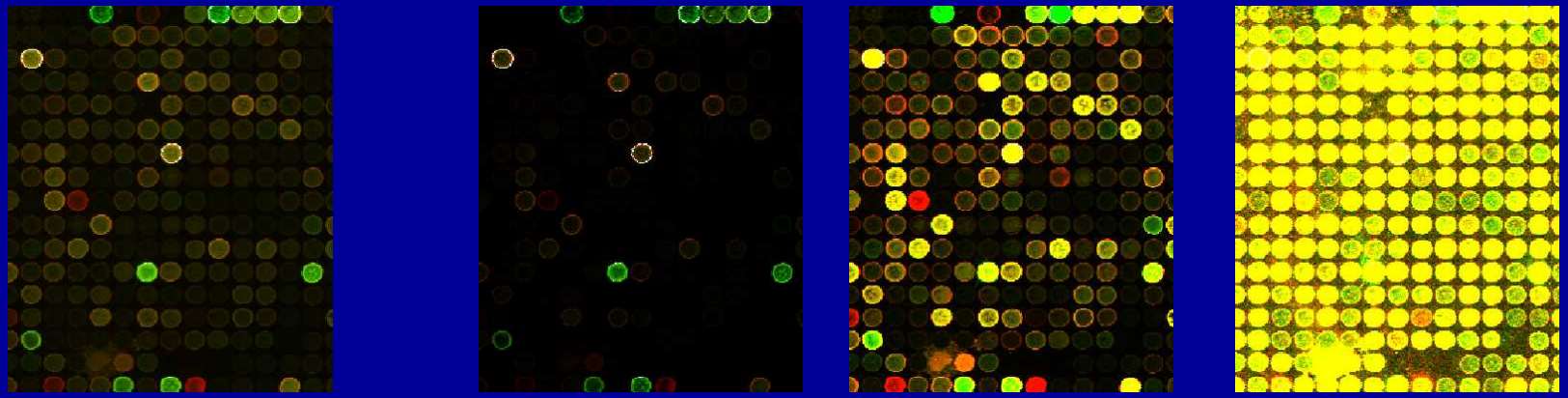
Sandia National Laboratories

User- Selectable Parameters – Image Display

16-bit data → 8-bit display

Compression

Reduction



Sq Root Transform
0 - 65535

High
256 - 65535

Middle
16 - 4096

Low
0 - 255

Sandia National Laboratories

User- Selectable Parameters – Getting Signal

- Pixel dwell time
- Spatial resolution
- Signal averaging
- Laser power
- PMT gain

User- Selectable Parameters – Getting Signal

On-the-fly PMT Gain Optimizing – not Advantageous?

Measurement Model

$$Y_{ijk} = \mu_i + \alpha_j + \beta_{k(j)} + \varepsilon_{ijk}$$

Y_{ijk} = Log2(R/G) measurement of i^{th} gene (k^{th} rep within j^{th} culture)

i = gene label

j = culture label

k = rep (withinculture) label

μ_i = overall average of i^{th} gene

α_j = global effect of j^{th} culture (across all genes)

$\beta_{k(j)}$ = global effect of k^{th} rep within j^{th} culture (across all genes)

ε_{ijk} = specific effect of i^{th} gene (k^{th} rep within j^{th} culture)

PMT Settings Adjusted

$$\hat{\sigma}_\alpha = .575\,(2\text{ dof})$$
$$\hat{\sigma}_\beta = .179\,(3\text{ dof})$$

PMT Settings Constant

$$\hat{\sigma}_\alpha = .081\,(2\text{ dof})$$
$$\hat{\sigma}_\beta = .090\,(3\text{ dof})$$

Sandia National Laboratories

Instrumentation Effects – Scanner Bias/ Stability

- Bias to one channel can arise from:
  - Dye hybridization
  - Contaminating signal *
  - PMTs *

Solutions? Transforms, Spectral imaging

- Stability
  - Lasers
  - PMTs
  - Optics
  - Scanning devices

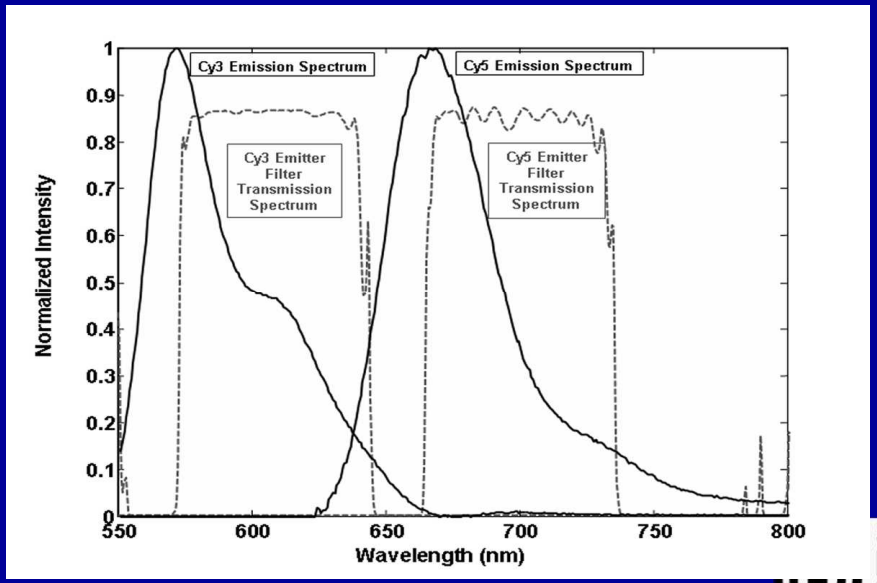
Instrument stability needs to be thoroughly checked on a regular basis → standards

Example…

Sandia National Laboratories

Instrumentation Effects – Signal Contamination

- Signal from something other than dye of interest gets detected in a channel
- Spectral crosstalk, contaminating fluorescence, stray light, etc.
- Background methods assume approximation from signal around spot – a rarity
- Even low absolute intensities lead to normalization, ratio errors



Sandia National Laboratories

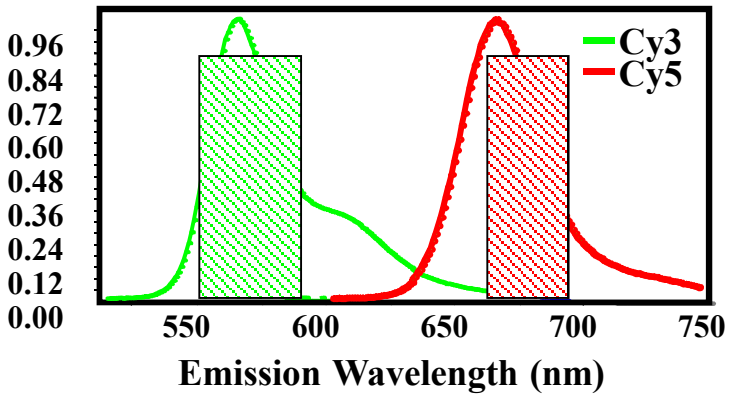
Alternative Scanning Technologies

- Surface Plasmon Resonance (SPR)
- Resonance Light Scattering (RLS)

- Hyperspectral Scanning (HSS)

Technology Comparison

Filter-based Scanner



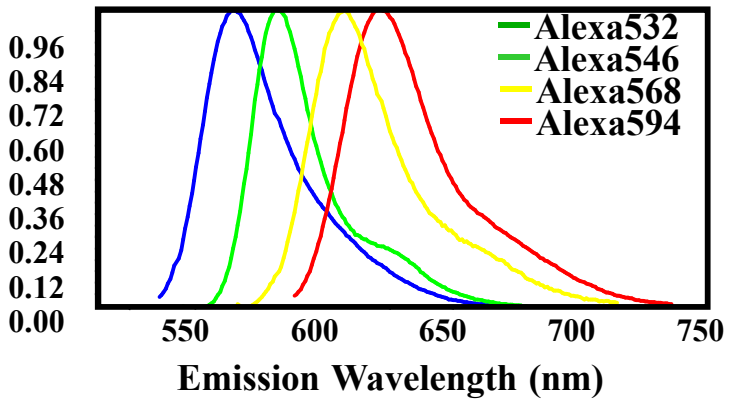
Univariate

Collect all photons within a wavelength region

Require well-separated emissions and ONE laser/label

Limited in practice to 2/3 labels

Hyperspectral Scanner



Multivariate

Collects an entire emission spectrum at each pixel

Excites multiple, overlapping dyes with ONE laser

Have shown separation of 4+

The key to HSS is Multivariate Analysis

Advantages of HSS & Multivariate Data Analysis

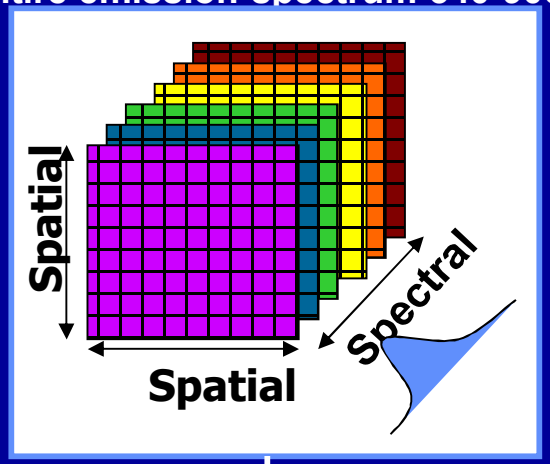
Discover & quantify all emitting species in a sample simultaneously

- Increased sensitivity
- Extended dynamic range
- Improved background/interference correction
- Improved accuracy , reliability , & quantitation
- Increase throughput – multiple overlapping dyes

Sandia National Laboratories

Hyperspectral Image Cube & Analysis

Each pixel of the image contains the entire emission spectrum 540-900 nm



Preprocess

Initial Estimates

MCR

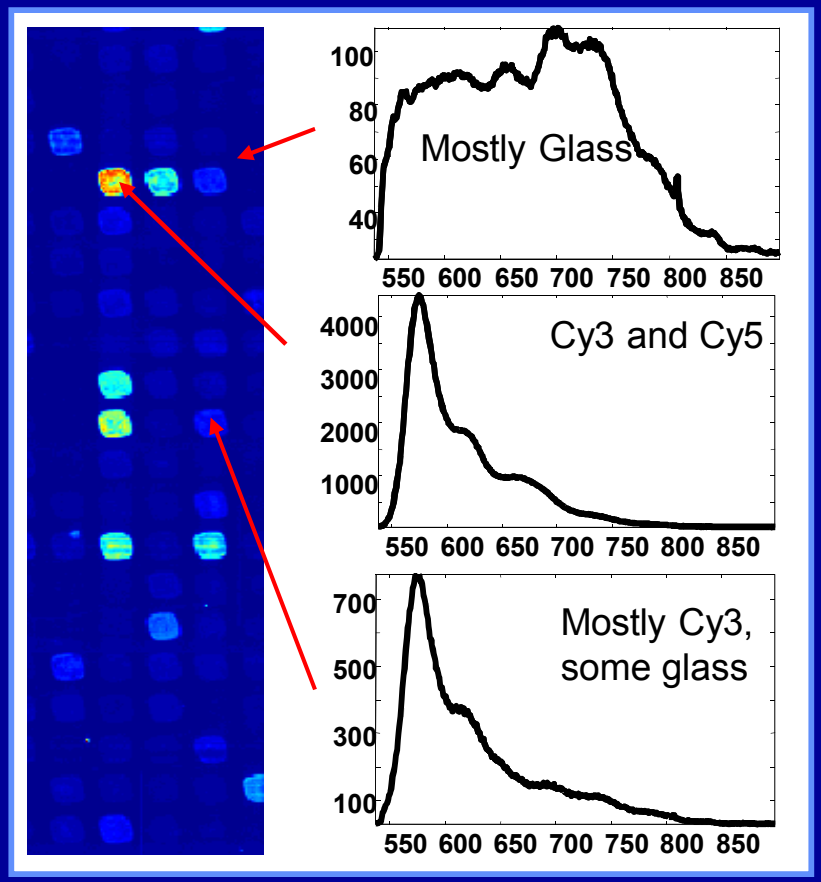
Spectral Components

Concentration Maps

WHAT?

WHERE?
HOW MUCH?

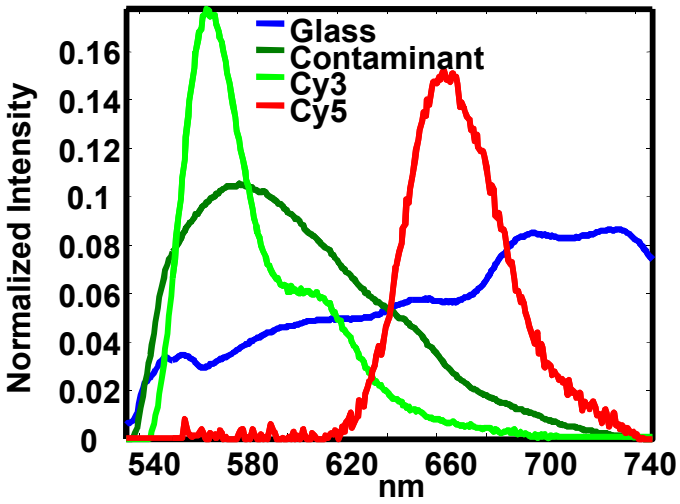
A peek at the raw data…



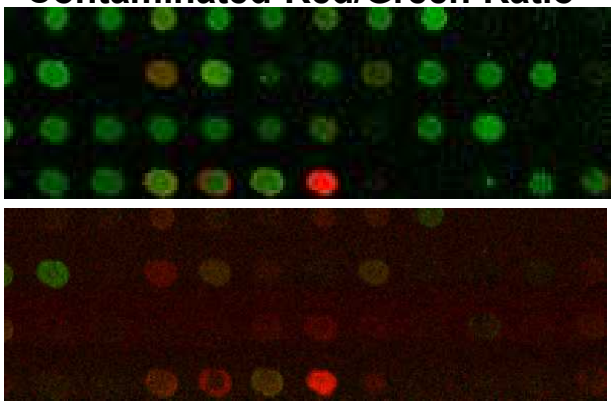
Sandia National Laboratories

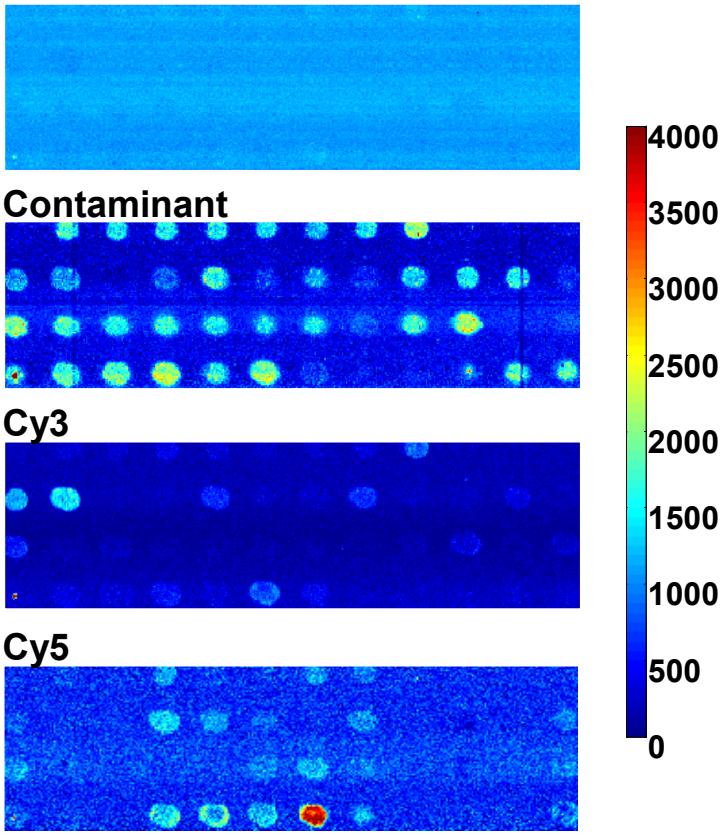
HSS Helps Understand Problems with Arrays

a. Contamination



Commercial Scanner
Contaminated Red/Green Ratio

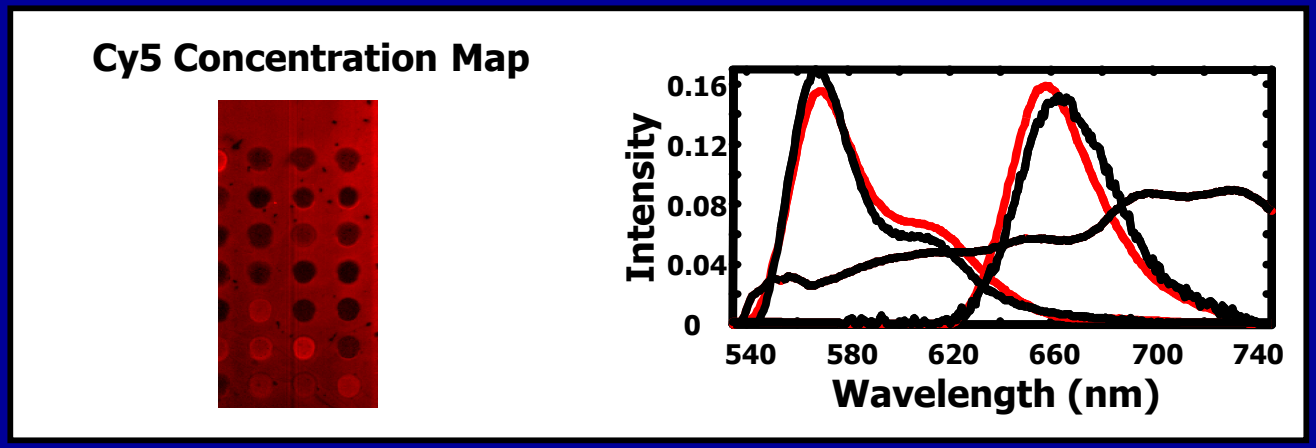




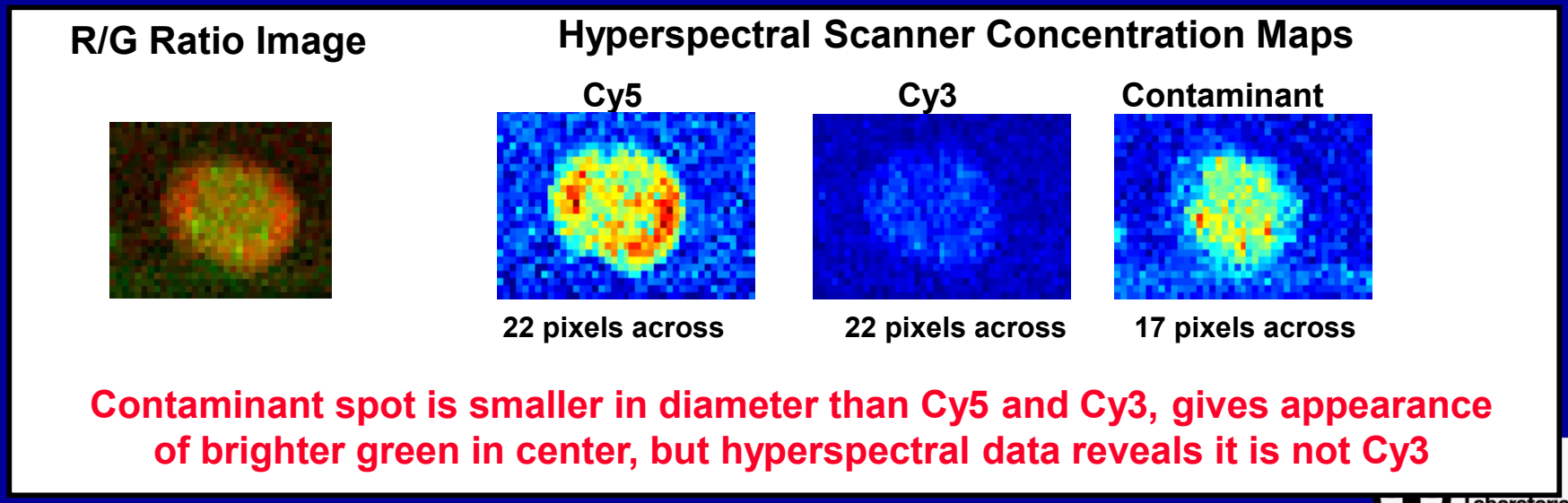
Hyperspectral Scanner
Accurate Cy5/Cy3 Ratio

HSS Helps Understand Problems with Arrays

b. Black holes – non-specific binding



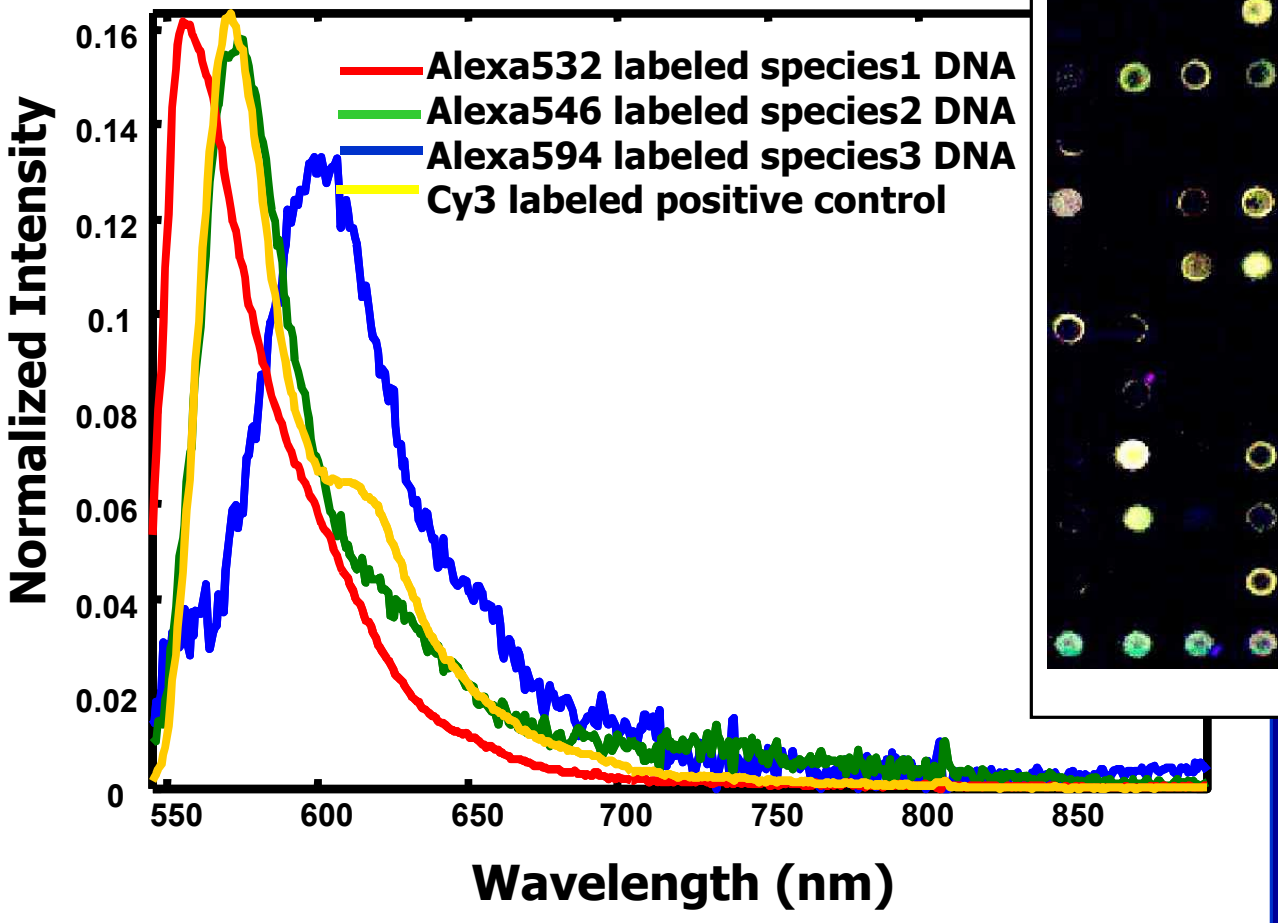
c. Dye separation?



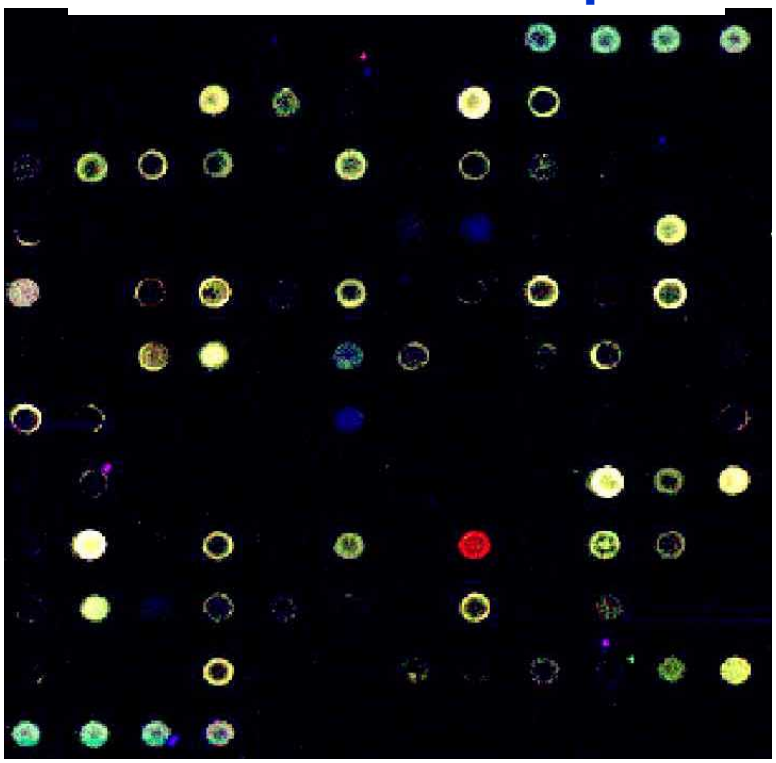
Contaminant spot is smaller in diameter than Cy5 and Cy3, gives appearance of brighter green in center, but hyperspectral data reveals it is not Cy3

Multiple Green Dyes – More Multiplexing

MCR Generated Pure Component Spectra



RGB Image of Alexa Dye Concentration Maps



In collaboration with UNM CRTC, Stephanie Ruby, Marilee Morgan

Sandia National Laboratories

In Summary

- Optimal performance realized when sources of variation are understood, controlled, minimized…
- Scanning/Scanner is a critical part of the microarray process and should not be overlooked
  - Instrument
  - User settings
  - Special considerations for large scale studies
- Emerging scanner technologies, like HSS, can compliment traditional scanning methods
  - Dye specific bias
  - Background, contaminant emissions
  - Increase multiplexing

Acknowledgments

- **Statistical analysis, exp design**
  - *Edward V. Thomas, SNL*
- **Software / algorithm development at SNL**
  - *Howland D. T. Jones*
  - *David M. Haaland*
  - *Michael R. Keenan*
  - *Paul G. Kotula*
  - *David K. Melgaard*
  - *Greg Poulter*
  - *Mark H. VanBenthem*
- **Microarray collaborations**
  - *David Bencic, Ronglin Wang (US EPA)*
  - *Brian Palenik (Scripps)*
  - *Ian Paulsen (TIGR)*
  - *Maggie Werner-Washburne (UNM)*
  - *Cheryl Willman (UNM CTC)*

$ Sandia Laboratory Directed Research and Development (LDRD) Program

$ NIH

$ US EPA

$ This work was funded in part by the US DOE's Genomes to Life program under project, "Carbon Sequestration in Synechococcus Sp.: From Molecular Machines to Hierarchical Modeling,"

$ Some experiments provided by the Keck-UNM Genomics Resource, a facility supported by the WM Keck Foundation, the State of NM, the UNM Cancer Research and Treatment Center, and the NM Center for Environmental Health Sciences."

Sandia National Laboratories

References

- Thomas, EV, et. al., "*Statistical Analysis of Microarray Data with Replicated Spots: A Case Study with Synechococcus WH8102*," 2007, submitted
- Timlin, JA, et. al., "*Hyperspectral microarray scanning: impact on the accuracy and reliability of gene expression data,*" *BMC Genomics*, 2005, 6:72.
- Sinclair, MB, et. al., "*Design, construction, characterization of a hyperspectral microarray scanner*", Applied Optics, 2004, 43, 2079-2089.
- Martinez, MJ, et. al, "*Identification and removal of contaminating fluorescence from commercial and in-house printed DNA microarrays*" Nucleic Acid Research, 2003, 31:4, e18.
- Haaland DM, et. al. "*Multivariate curve resolution for hyperspectral image analysis: applications to microarray technology*," Spectral Imaging: Instrumentation, Applications, and Analysis, 2003; 4959: 55-66.

Sandia National Laboratories