Selected Topics in Electrical Engineering: Flow Cytometry Data Analysis

Bilge Karaçalı, PhD
Department of Electrical and Electronics Engineering
Izmir Institute of Technology

Syllabus

Meeting times: Friday13:30, 14:30, 15:30

Text : Marion G. Macey, ed., "Flow Cytometry: Principles and Applications," Springer, 2007

Instructor : Bilge Karaçalı, PhD

Office : EEE Building Room 209

Phone : 6534

E-mail : bilgekaracali@iyte.edu.tr

Summary :

Principles of flow cytometry; Cell preparation; Fluorochromes and fluorescence; Experimental design and fluorescence quantitation; Cell sorting; Compensation; Statistical analysis: Probability binning; Data analysis using machine learning.

Grading :

Class Participation	50%
Project	50%

URL :

http://web.iyte.edu.tr/~bilgekaracali/EE563/

Course Outline:

Week 1: Introduction to flow cytometry

Week 2: Fluorochromes and fluorescence

Week 3: Experimental design and fluorescence quantitation

Week 4: Compensation and gating

Week 5: Normalization

Week 6: Comparing Univariate Cell Distributions

Week 7: Probability Binning

Week 8: Readings on flow cytometry data analysis

Week 9: Readings on flow cytometry data analysis

Week 10: Readings on flow cytometry data analysis

Week 11: Readings on flow cytometry data analysis

Week 12: Readings on flow cytometry data analysis

Week 13: Readings on flow cytometry data analysis

Week 14: Overview

Outline

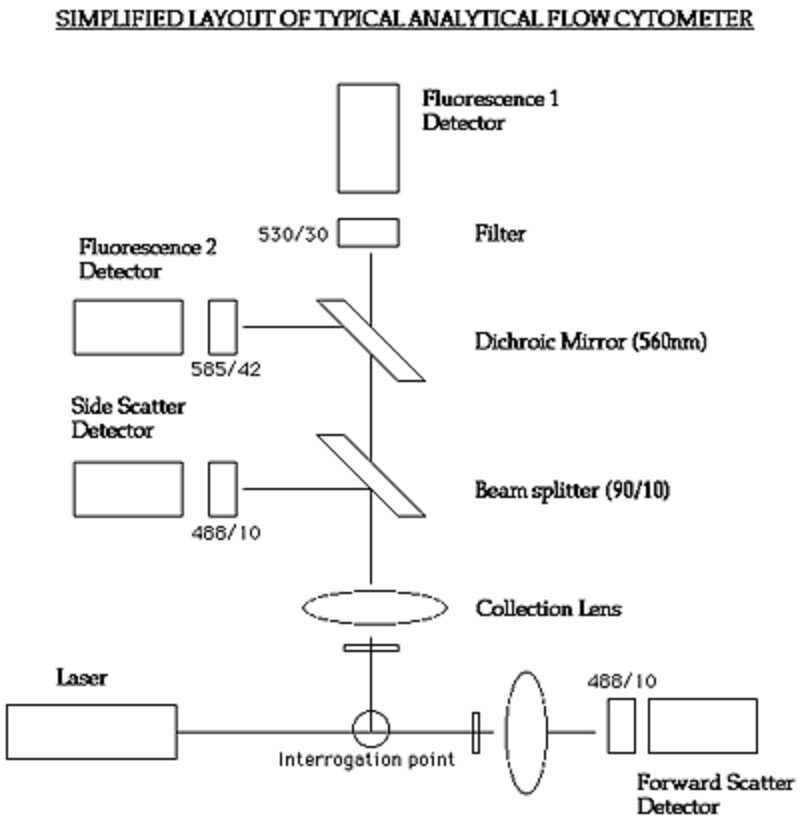
- Overview
- Elements of polychromatic flow cytometry
  - Monoclonal antibodies
  - Fluorochromes
- Flow analysis
  - Fluidics
  - Optics
  - Electronics
  - Sorting
- Basic data processing
  - Compensation
  - Visualization
  - Gating
- Advanced data processing
  - Comparing cell distributions
  - Identifying new subsets
- Examples
  - CD4+ T-cell counting in HIV
  - Cell cycle analysis

Outline

- Overview
- Elements of polychromatic flow cytometry
  - Monoclonal antibodies
  - Fluorochromes
- Flow analysis
  - Fluidics
  - Optics
  - Electronics
  - Sorting
- Basic data processing
  - Compensation
  - Visualization
  - Gating
- Advanced data processing
  - Comparing cell distributions
  - Identifying new subsets
- Examples
  - CD4+ T-cell counting in HIV
  - Cell cycle analysis

Overview

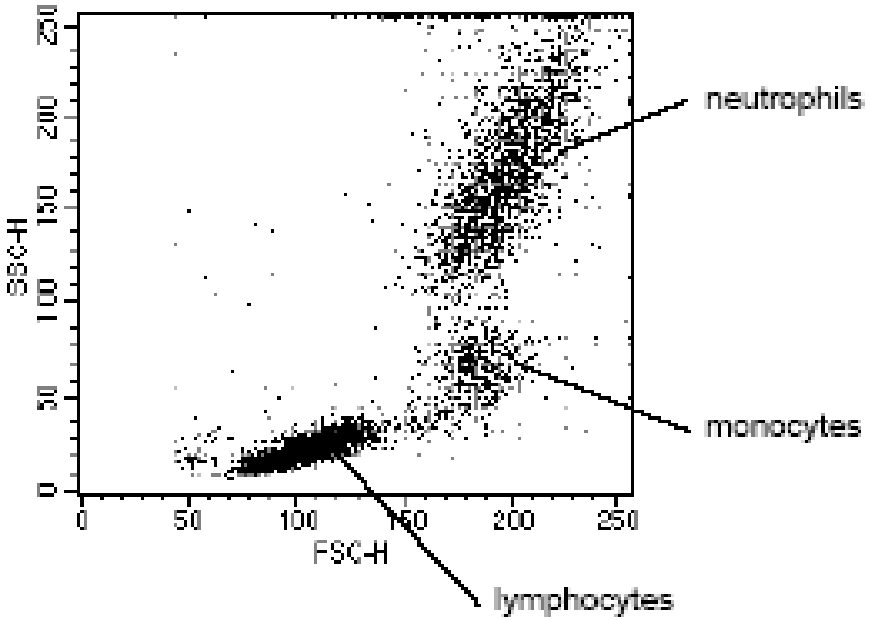
- To quantify the presence of molecular markers in cells and cell membranes among heterogeneous cell populations
- Allows cell phenotyping with
  - High throughput*
  - Precise quantification on the basis of individual cells
  - Multivariate measurements*

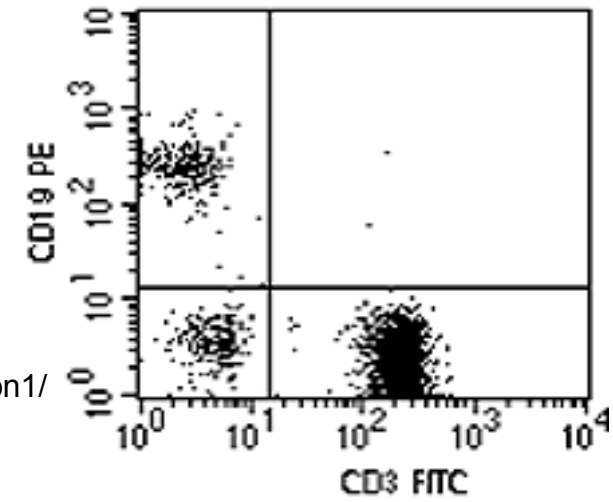


Source: http://www.pathology.wustl.edu/html/facilitiesfacsfc.html

Example: Immunophenotyping

- Objective: to assign (blood) cells in separate groups based on their phenotypic profiles
- Phenotypic parameters via flow cytometry: a multivariate profile for each cell
  - Morphological (structural) parameters:
    - Forward scatter – related to cell size
    - Side scatter – related to granularity and surface curvature
  - Fluorescence (functional) parameters:
    - Multiple intensity parameters indicating amounts of molecular markers in cell cytoplasm or membrane
- Application: CD4+ T-cell count in AIDS monitoring





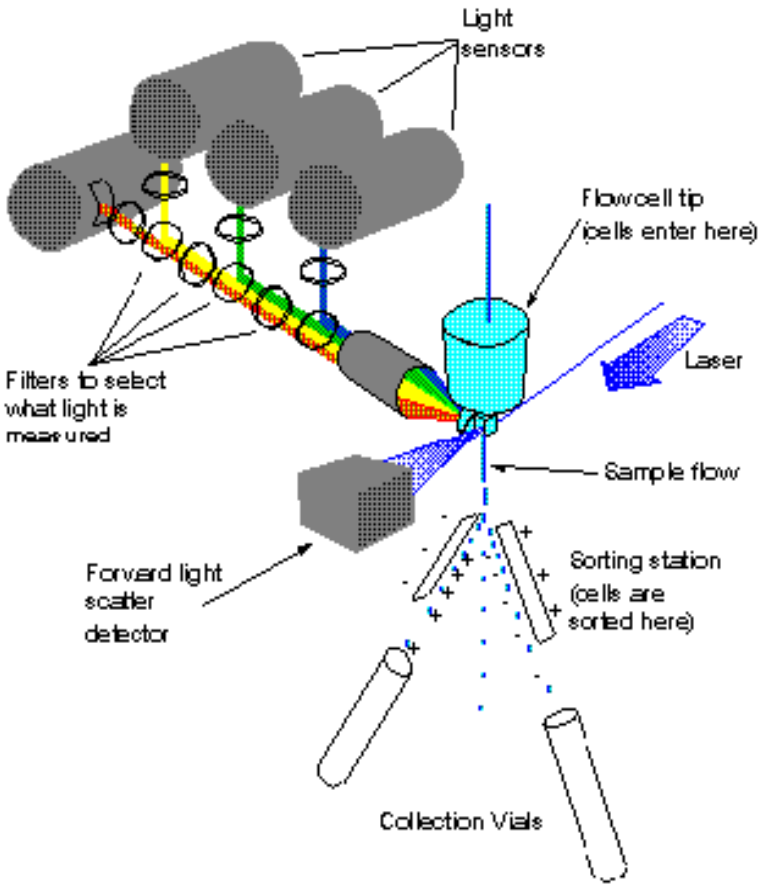
Source: http://www.med.umich.edu/flowcytometry/training/lessons/lesson1/

Outline

- Overview
- Elements of polychromatic flow cytometry
  - Monoclonal antibodies
  - Fluorochromes
- Flow analysis
  - Fluidics
  - Optics
  - Electronics
  - Sorting
- Basic data processing
  - Compensation
  - Visualization
  - Gating
- Advanced data processing
  - Comparing cell distributions
  - Identifying new subsets
- Examples
  - CD4+ T-cell counting in HIV
  - Cell cycle analysis

Elements of Flow Cytometry

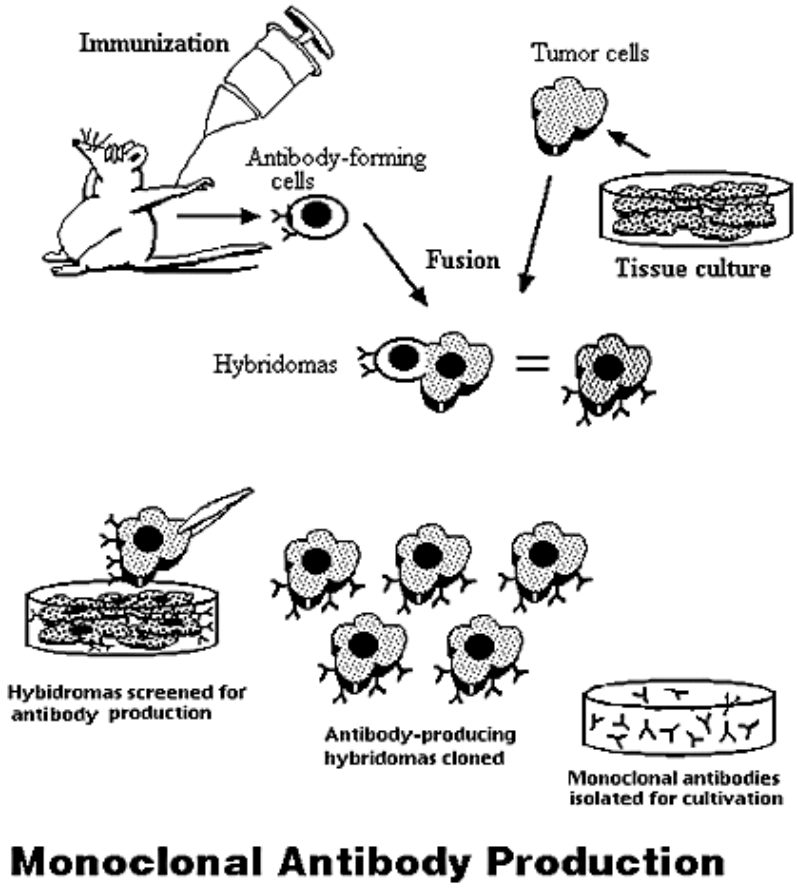
- Sample preparation
  - Specimen in single cell suspension
    - (lysed) peripheral blood
      - ammonium chloride
    - bone marrow aspirates
    - cell cultures
    - dissociated solid tissue samples
  - Fluorescence staining
    - fluorochrome-antibody conjugation
    - staining of samples with compatible fluorochrome-antibody conjugates
- Flow analysis
  - Flow cytometer operation principles
  - Cell sorting
- Data processing
  - Computation of known cell subset percentages
  - Identification of new cell subsets
  - Comparison of cell distributions across individuals and populations



Source: http://meds.queensu.ca/qcri/flow/cri-fc-getstarted.htm

Monoclonal Antibodies for Fluorescence Staining

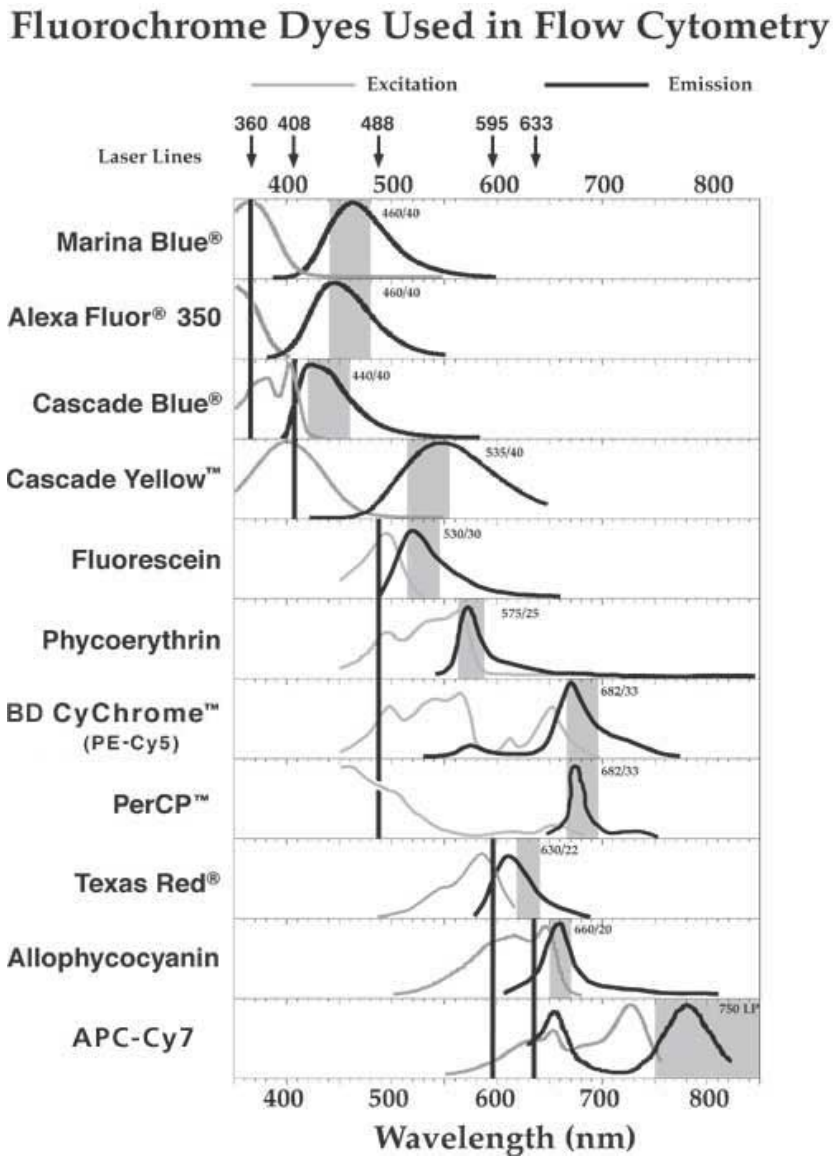
- Fluorescence staining:
  - An antibody is conjugated to a fluorescent dye
  - The conjugated antibody binds to its target molecule with high specificity
  - Concentration of fluorescent dye indicates concentration of the target molecule
- Polyclonal antibodies: yield little useable antibodies and contains undesired substances
  - live animals injected with antigen
  - serum collected
- Monoclonal antibodies: antibodies produced by continually growing cells of a single type
  - Tumor cells (e.g. myeloma) fused with cells that produce antibodies (the hybridoma cell)
  - Large quantities of antibodies in a much purer form



Source: http://www.accessexcellence.org/RC/VL/GG/monoclonal.html

Fluorochromes for Cytometry

- To mark molecules of interest with a distinct fluorescent signature
- Properties of suitable fluorochromes:
  - Biological inertness
  - High cell-associated fluorescence intensities
  - Non-overlapping spectra
  - Easy conjugation to antibodies
- Choice of cytometer (optics and lasers) affects the choice of fluorochrome combinations
  - polychromatic flow cytometry with multiple lasers
- Background fluorescence: autofluorescence (due to NADH, riboflavins, and flavin coenzymes) and “sticky” dyes (that bind nonspecifically)
  - Problematic for detecting weak fluorescent response
  - Identifying the limits of positive staining vs. the background fluorescence requires control runs



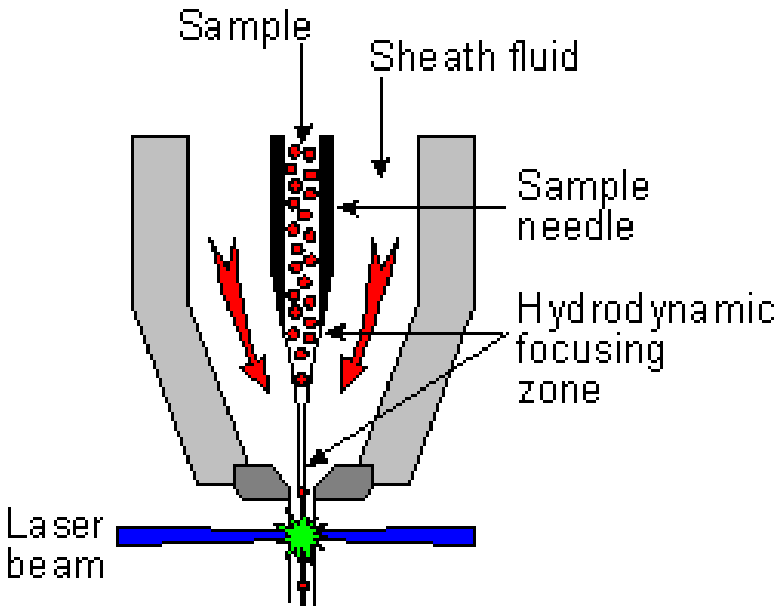
Source:http://www.bdbiosciences.com/pharmingen/protocols/Fluorochrome_Absorption.shtml

Outline

- Overview
- Elements of polychromatic flow cytometry
  - Monoclonal antibodies
  - Fluorochromes
- Flow analysis
  - Fluidics
  - Optics
  - Electronics
  - Sorting
- Basic data processing
  - Compensation
  - Visualization
  - Gating
- Advanced data processing
  - Comparing cell distributions
  - Identifying new subsets
- Examples
  - CD4+ T-cell counting in HIV
  - Cell cycle analysis

Flow Analysis: Fluidics

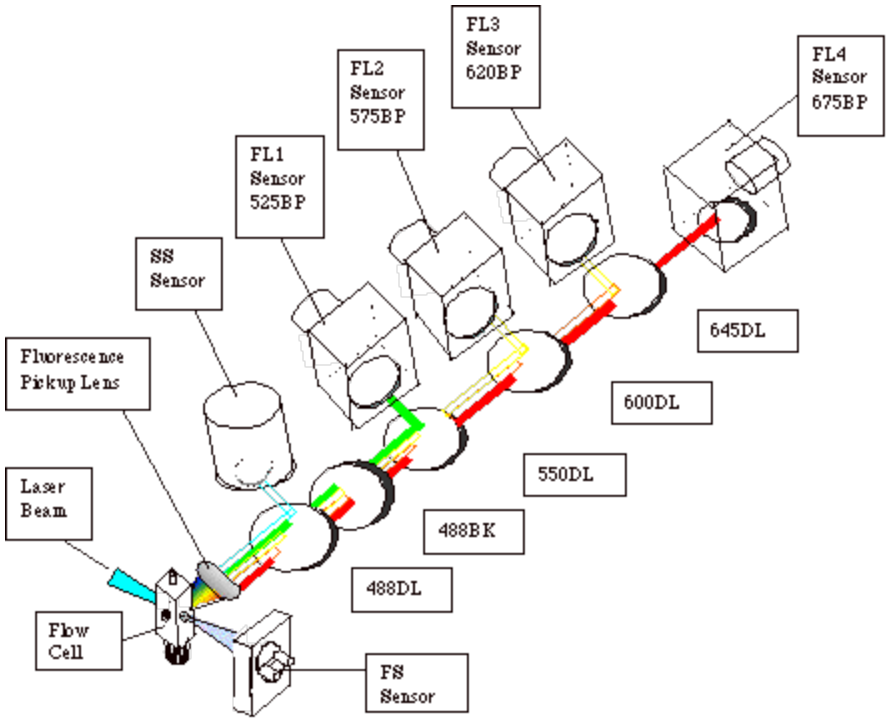
- To have stained cells in a single suspension pass through the interrogation point in a single file with a controlled speed
  - Cell suspension stream embedded in a stabilizing sheath
  - Laminar flow: flow of sample suspension doesn't mix with the sheath fluid determined by the Reynolds number (non-turbulent flow)
    - Tube diameter
    - Density, mean flow, and viscosity of the suspension fluid
  - Achieved by hydrodynamic focusing
    - Differential injection
      - Sample suspension and sheath fluid pressurized separately by gas
      - Moderate control
    - Volumetric focusing
      - Sample suspension injected by a syringe with a pump control
      - Absolute control
  - Causes artificial changes in cell shapes
    - Reorientation along the larger axis
    - Volume deformation into ellipsoidal shape



Source:http://www.med.umich.edu/flowcytometry/training/lessons/lesson1/

Flow Analysis: Optics

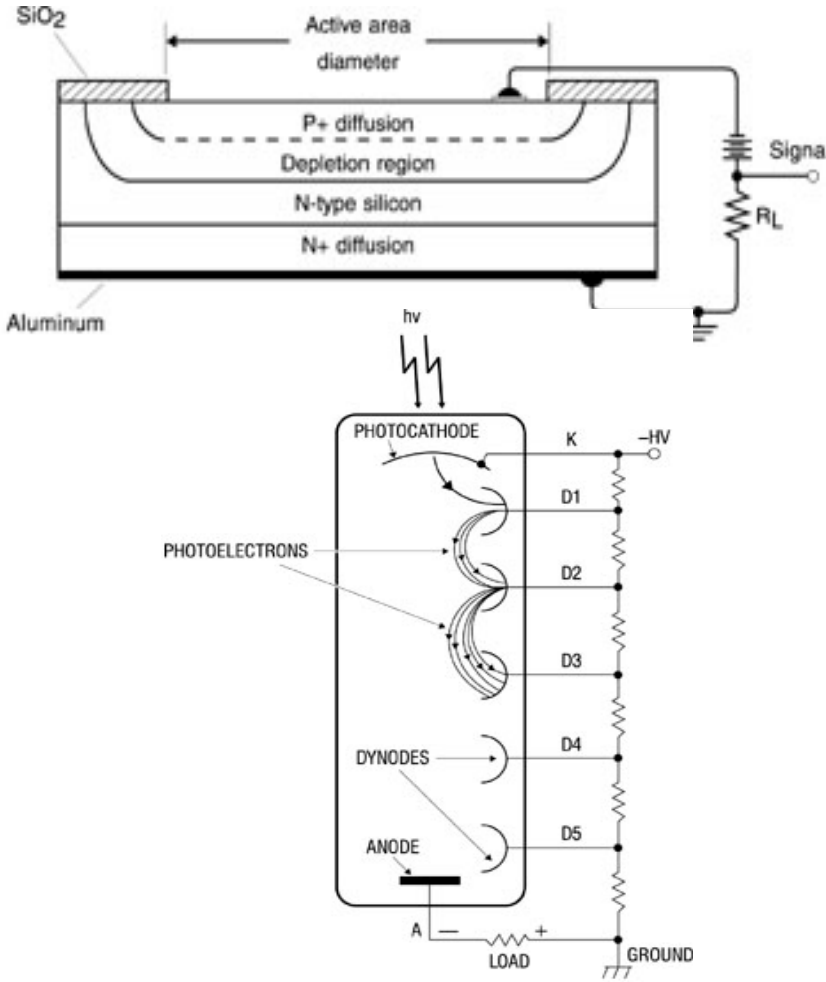
- To focus the interrogation light onto the sample suspension flow
  - Two types of light sources:
    - Lasers (single wavelength, coherent)
    - Arc-lamps (multiple wavelengths, incoherent)
  - Oblique beam for suspension flow coverage and cell separation
- To collect the scattered light into photon detectors with spectral specificity
  - Color-specific optical channels
    - Forward and side scatter channels: same frequency as the light source
    - Fluorescence channels: bandpass filtered to the peak of the fluorochrome emission spectra
  - Dichroic filters in specific order
    - Pass specific frequencies for detection
    - Reflect the rest for subsequent detection by remaining optical channels



Source:http://biology.berkeley.edu/crl/flow_cytometry_basic.html

Flow Analysis: Electronics

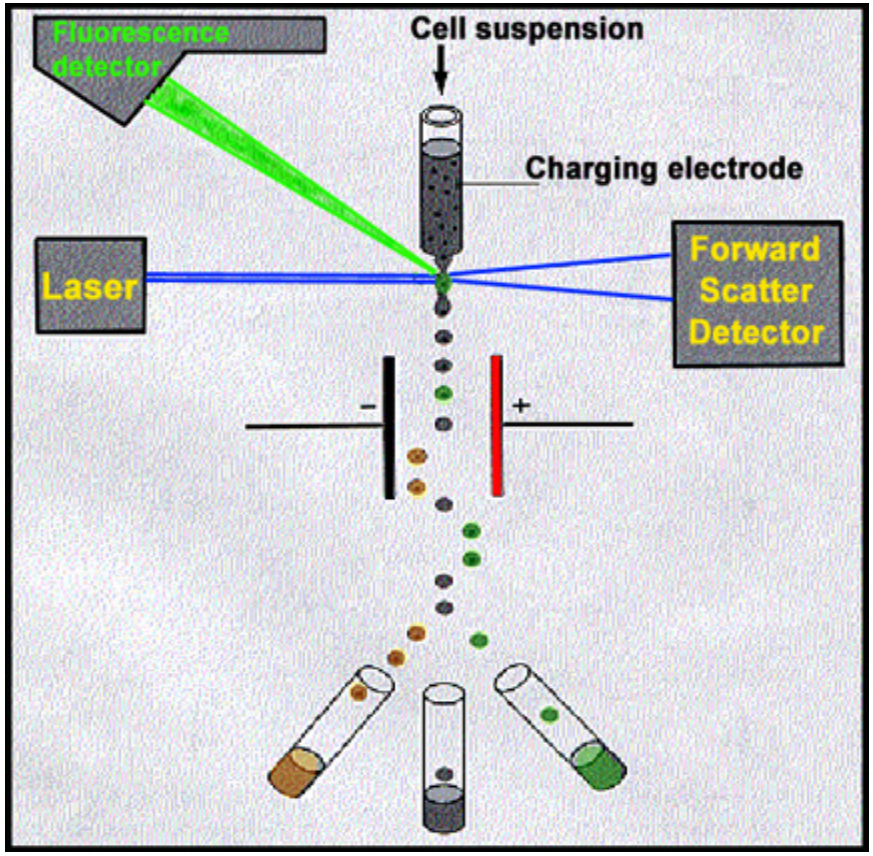
- To convert emitted colored light quantities into numbers for subsequent computer analysis
  - Light intensity measured by analog detectors as a voltage pulse
    - Photodiodes: solid state detectors operating on the principle of photovoltaic effect
      - Electron-hole pairs formed by a photon hitting the active surface
      - Electrons and holes steered to separate ends creating a current
      - Good for strong light
    - Photomultiplier tubes: photoemissive material in vacuum
      - A cathode emits electrons when hit by a photon
      - Effects amplified by a cascade of dynodes
      - Sensitive to weak light
      - Can be damaged by excessive exposure
  - Observed voltage pulse converted to integer numbers by analog-to-digital converters
    - Pulse peak identifies total fluorescence
    - Pulse width correlates with cell diameter
    - Number of digital channels determines quantization error
    - Signal can be amplified for best coverage over the available channels



Source: http://repairfaq.ece.drexel.edu/sam/CORD/leot/course04_mod08/mod04-08.html

Flow Analysis: Sorting

- To collect rare cells of interest
- Identification of cells of interest:
  - Enclosures in the space of phenotypic signatures define cells (aka gates)
    $S(x) = [x^{FSC}\ x^{BSC}\ x^{CD3}\ x^{CD4}\ x^{CD8}\ \ldots]^T$
  - Multiple gates can be combined using simple logic (and, or, nand, xor, …)
- Collection of identified cells in separate bins:
  - The flow is disrupted into separate droplets each containing a single cell
    - Sustained oscillations at a specific frequency
  - The drops containing cells that satisfy the gating logic are electrically charged
  - During their fall between electrically charged plates, the resident electrostatic force diverts the charged drops into a separate container
  - Accuracy in sorting (purity) requires exquisite timing



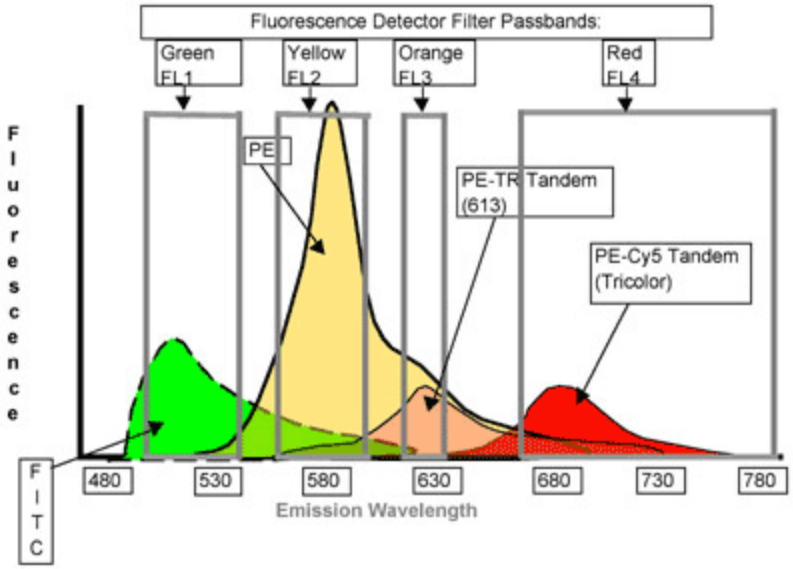
Source: http://meds.queensu.ca/qcri/flow/cellsorting.htm

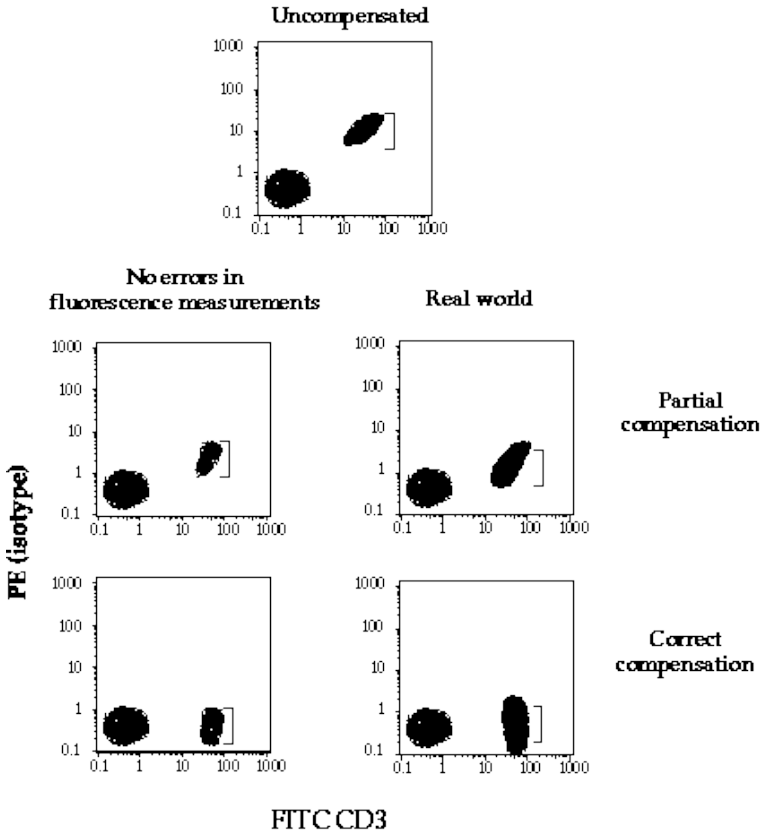
Outline

- Overview
- Elements of polychromatic flow cytometry
  - Monoclonal antibodies
  - Fluorochromes
- Flow analysis
  - Fluidics
  - Optics
  - Electronics
  - Sorting
- Basic data processing
  - Compensation
  - Visualization
  - Gating
- Advanced data processing
  - Comparing cell distributions
  - Identifying new subsets
- Examples
  - CD4+ T-cell counting in HIV
  - Cell cycle analysis

Data Processing: Compensation

- To eliminate the contribution of a fluorochrome emission into other fluorescence detectors
  - Each fluorochrome has a dedicated detector
  - Fluorochrome emission spectra overlaps with the spectral bands dedicated to other fluorochromes
  - Crucial for correct identification of antigen densities
  - Requires a dedicated tube with unstained and singly stained cells
  - Accomplished using linear algebra
    - Contribution of a fluorochrome emission into side channels computed as a fraction of its contribution in its main channel
    - True contributions obtained by solving a linear system of equations
  - Measurement error is magnified
    - Affects particularly dim channels

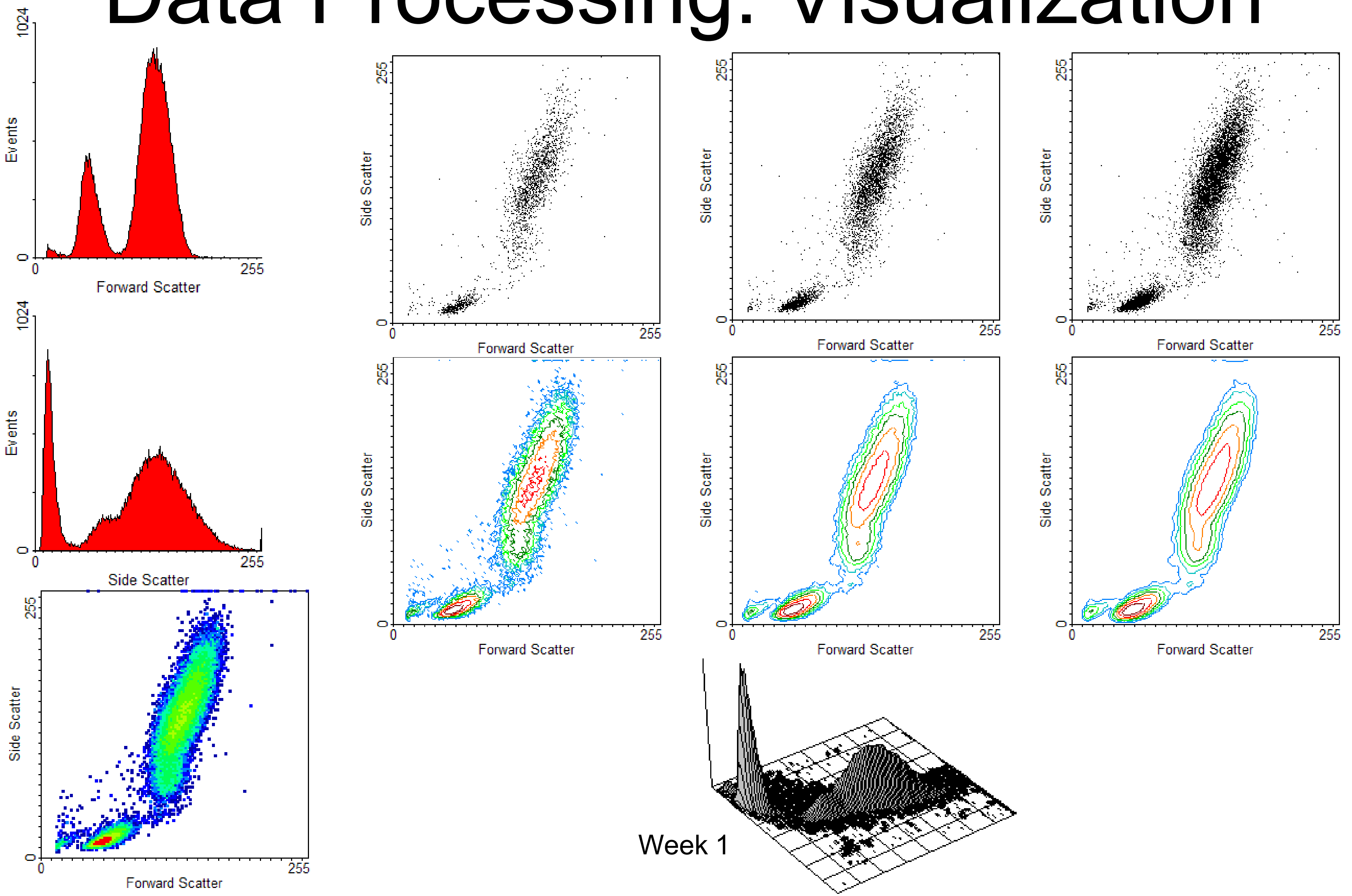




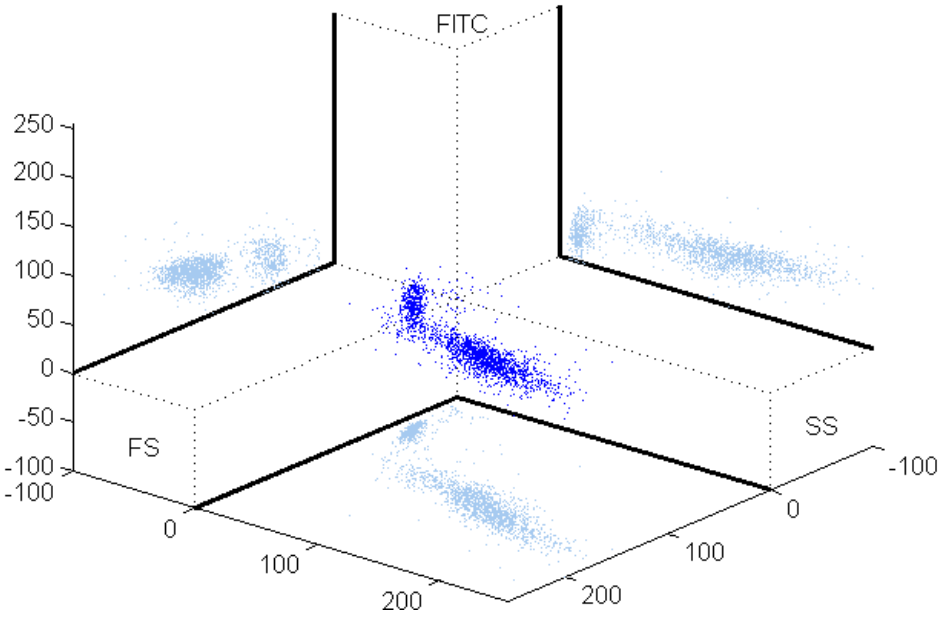
Source: http://www.drmr.com/compensation/

Source: http://biology.berkeley.edu/crl/compensation.html and Practical Flow Cytometry 3rd Ed., H.M. Shapiro, p.164

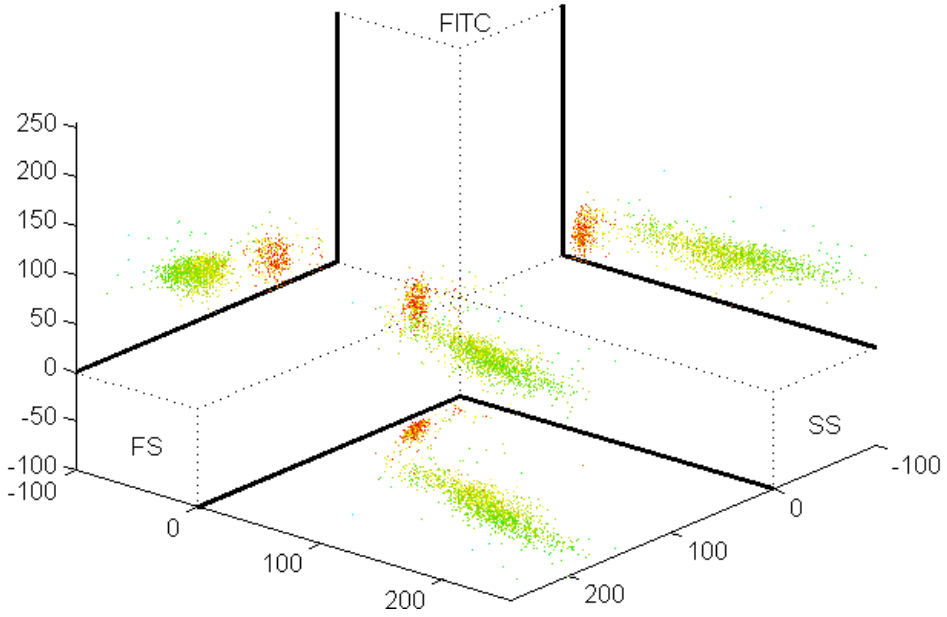
Data Processing: Visualization



Data Processing: Visualization



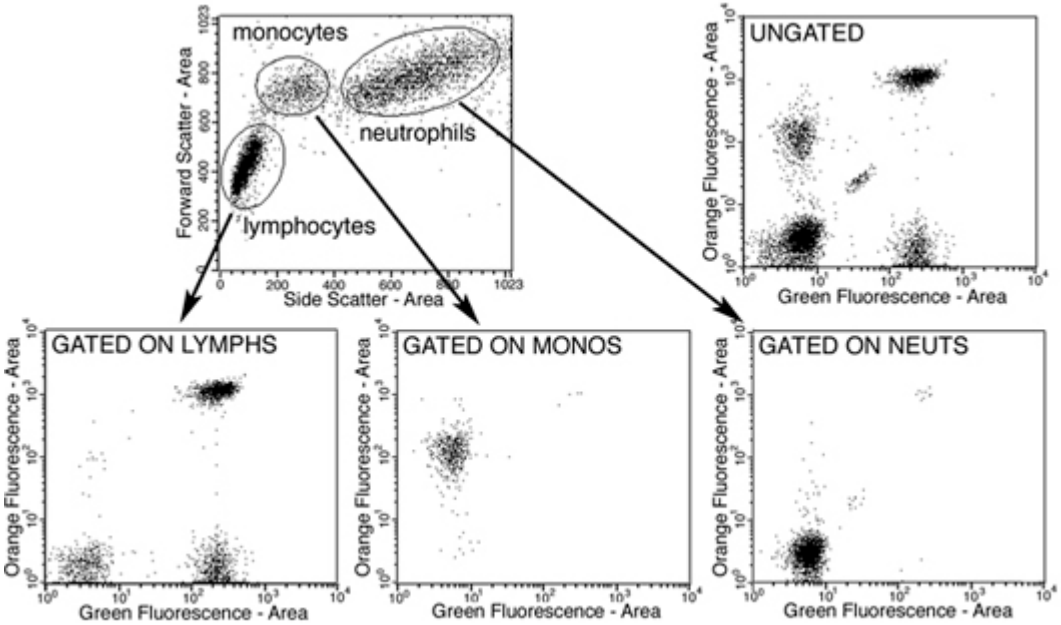
3D visualization



4D visualization

Data Processing: Gating

- To identify specific cells/particles of interest
  - Hardware gating for rare cell collection and sorting (on-the-fly – live gates)
  - Software gating for phenotyping (off-line – analysis gates)
- Simple gates
  - Quadrants
  - Rectangles and ellipsoids
  - Polygons
- Combination gates
  - Integration of multiple simple gates through logical operations such as AND, OR, NAND, etc …



Source: Givan, A.L. (2004). Flow Cytometry: An Introduction. Chapter in Flow Cytometry Protocols, 2nd edition (T.S. Hawley and R.G. Hawley, eds.). Methods in Molecular Biology, vol. 263. Humana Press Inc., Totowa, NJ, pp. 1-31.

Outline

- Overview
- Elements of polychromatic flow cytometry
  - Monoclonal antibodies
  - Fluorochromes
- Flow analysis
  - Fluidics
  - Optics
  - Electronics
  - Sorting
- Basic data processing
  - Compensation
  - Visualization
  - Gating
- Advanced data processing
  - Comparing cell distributions
  - Identifying new subsets
- Examples
  - CD4+ T-cell counting in HIV
  - Cell cycle analysis

Advanced Data Processing

- Comparing two cell distributions
  - Objective:
    - Quality control: measurements drifting due to unexpected changes in the system → instrument calibration
    - Monitoring biological response: measurements changing in parallel to a change in clinical parameters or medication
  - Univariate distributions
    - Overton cumulative histogram subtraction
    - Super Enhanced Dmax subtraction
    - K-S statistics to test the hypothesis that two histograms are different with statistical significance
    - X^2 test on each channel
    - Probability binning: cell frequencies of a test distributions at equal-probability bins of a control distribution in a X^2 sense
  - Multivariate distributions
    - Probability binning
- Comparing cell distributions across populations
  - Comparison of a test distribution against an aggregation of controls

Advanced Data Processing

- Identifying new cell subsets
  - Cells of distinct subtypes differ from each other by antigen expression
    - Dichotomous expression:
      - Some cells exhibit a given antigen while others do not
      - These cells can be differentiated using quadrant analysis
      - Each additional cluster of differentiation doubles the number of discernible subtypes
    - Continuous expression
      - Antigen expression is observed over a range of values
      - Individual expression levels do not provide clear cut separation between different cell clusters
      - Identification of different cell clusters requires joint analysis of all antigen expressions
  - Conventional (manual) techniques are insufficient to analyze polychromatic flow cytometry data
  - Unsupervised clustering techniques from the statistical learning literature are invoked
    - Cells are automatically clustered around distinct centers based on a measure of distance in the high dimensional observation space
      - Agglomerative or divisive methods
      - K-means clustering
      - Kohonen self organizing maps
    - The choice of distance measure crucial to the outcome

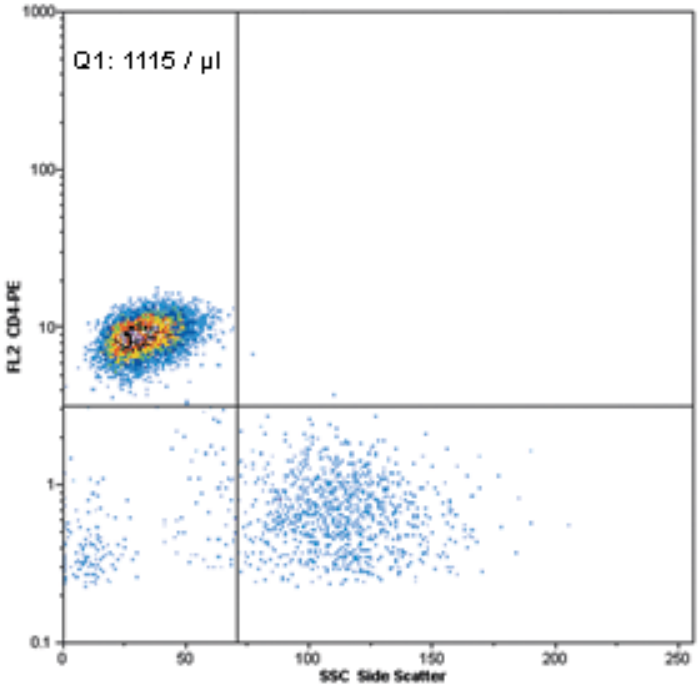
Outline

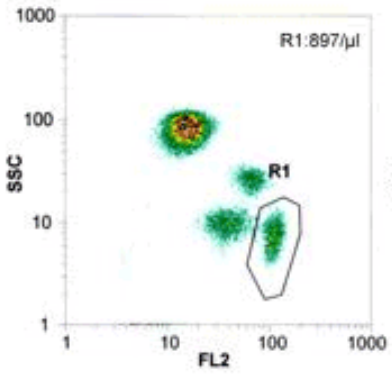
- Overview
- Elements of polychromatic flow cytometry
  - Monoclonal antibodies
  - Fluorochromes
- Flow analysis
  - Fluidics
  - Optics
  - Electronics
  - Sorting
- Basic data processing
  - Compensation
  - Visualization
  - Gating
- Advanced data processing
  - Comparing cell distributions
  - Identifying new subsets
- Examples
  - CD4+ T-cell counting in HIV
  - Cell cycle analysis

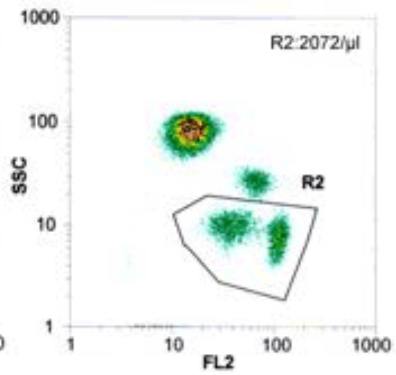
Example: CD4+ T Cell Enumeration in HIV

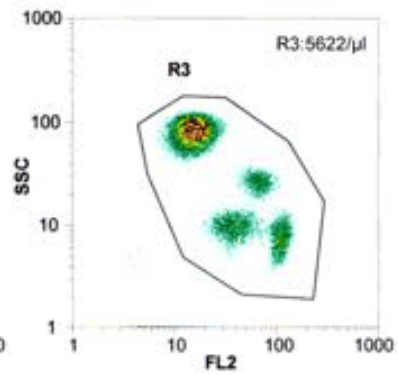
- HIV primarily infects CD4+ (T-cells) lymphocytes
- CD4+ lymphocyte (T-cell) count surrogate marker for HIV progress
  - Early symptomatic infection: CD4+ cell count > 500 cells/mm^3
  - Clinical latency: CD4+ cell count 200 - 500 cells/mm^3
  - AIDS: CD4+ cell count 50 - 200 cells/mm^3

Source: http://www.partec.de/applications/cd4cd45.html









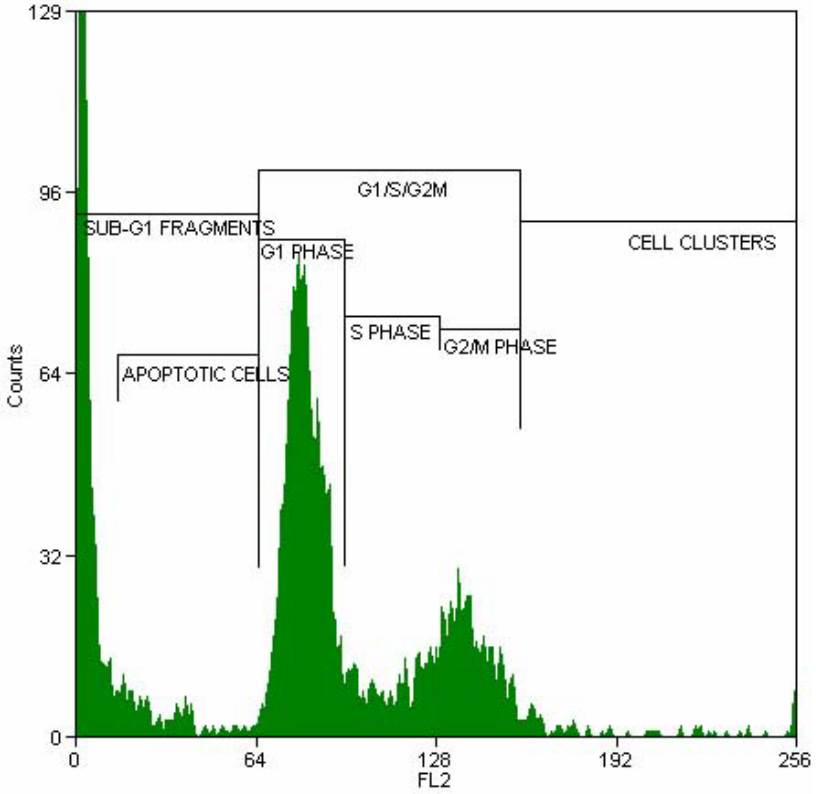
Results:

Absolute CD4 Count	897 CD4+ T-cells / µl	Cell Population in Region R1
Total Lymphocyte Count	2072 lymphocytes / µl	Cell Population in Region R2
WBC count (leukocytes)	5622 leukocytes / µl	Cell Population in Region R3
CD4% among lymphocytes	43.3%	
Lymphocyte% among WBCs	36.9%	

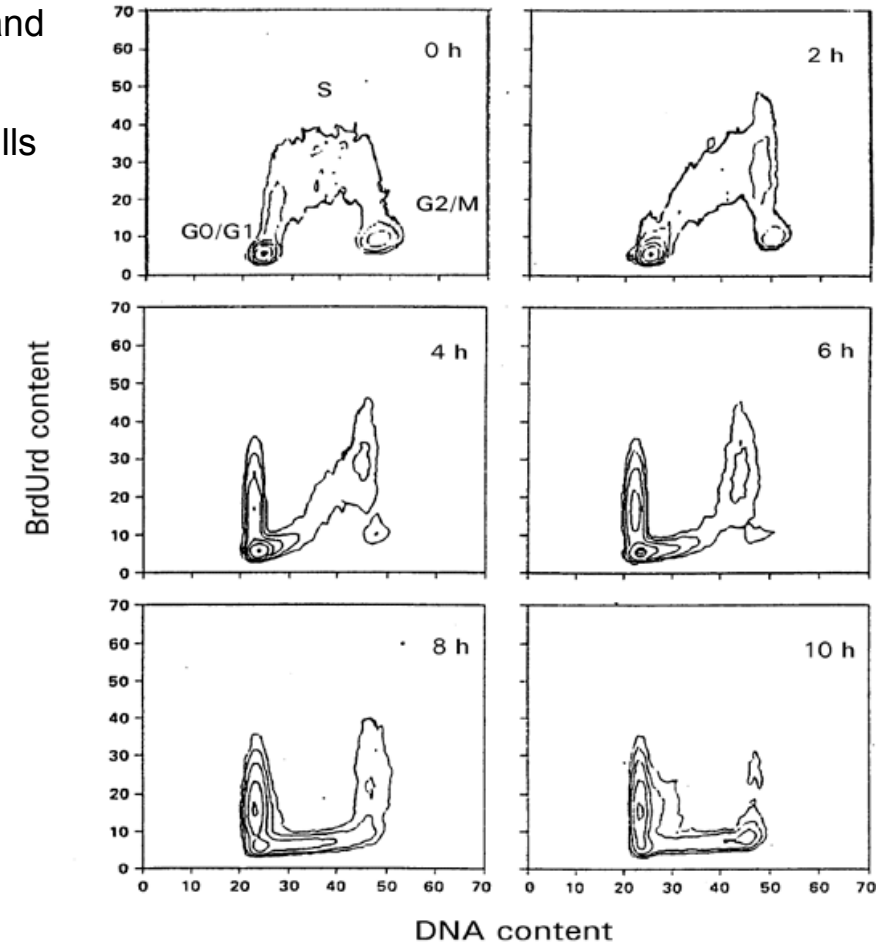
Week 1

Example: Cell Cycle Analysis

- Propidium iodide binds to double-stranded DNA (and RNA)
  - Treatment with RNase improves the DNA resolution
- Bromodeoxyuridine binds to the DNA of cycling cells
  - Kinetic information on the S-phase cells



Source: http://www.vetmed.auburn.edu/index.pl/flow_cytometry



Source:http://www.icms.qmul.ac.uk/flowcytometry/uses/cellcycleanalysis/bromodeoxyuridine/

Summary

- Flow cytometry allows a high-throughput analysis of cells suspended in an aqueous environment
  - Possessing different shape characteristics
  - Marked for different molecular receptors useful for differentiating between different cell types of interest
- High-dimensional data acquired from multi-color flow cytometers reflect various contributing effects
  - Sample preparation
  - Instrument calibration
  - Fluorescence interference
  - Digitization noise
- Data analysis must be carried out to make viable inferences regarding the clinical or biological hypothesis of interest by making the best use of the data at hand