Real-time monitoring of water quality through a holistic approach to particle characterization: 'Particle-omics'

Tawnya D. Peterson, Paul G. Tratnyek, Joseph A. Needoba, Ashley N. Anhalt, Ruby Taylor

CMOP Center for Coastal Margin Observation & Prediction

Oregon Health & Science University

NSF

Monitoring for pathogens often involves time-intensive or expensive laboratory analyses

- COST: requires accurate identification through microscopy or molecular techniques
- EXPERTISE: requires extensive training in the use of advanced techniques
- RISKS: look for one thing at a time, therefore other potential risks or emerging threats not identified
- RELIABLE INDICATOR: Real-time monitoring can assist in identifying particular organisms only if a diagnostic feature or probe is available

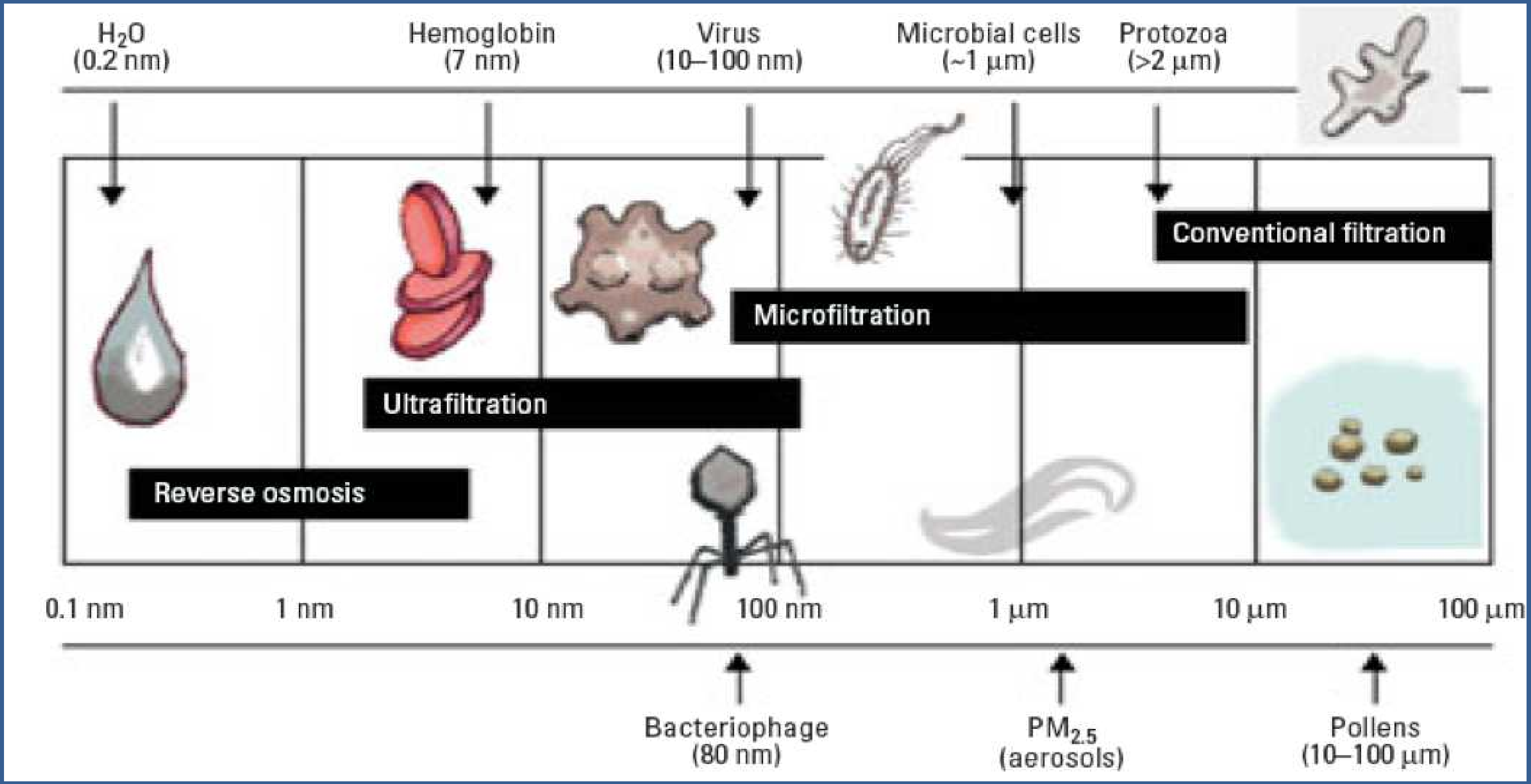
Goal

Develop rapid and cost-effective tools for the detection of threats to water quality

In situ sensors are ideally suited for rapid detection; however, they are often not very specific

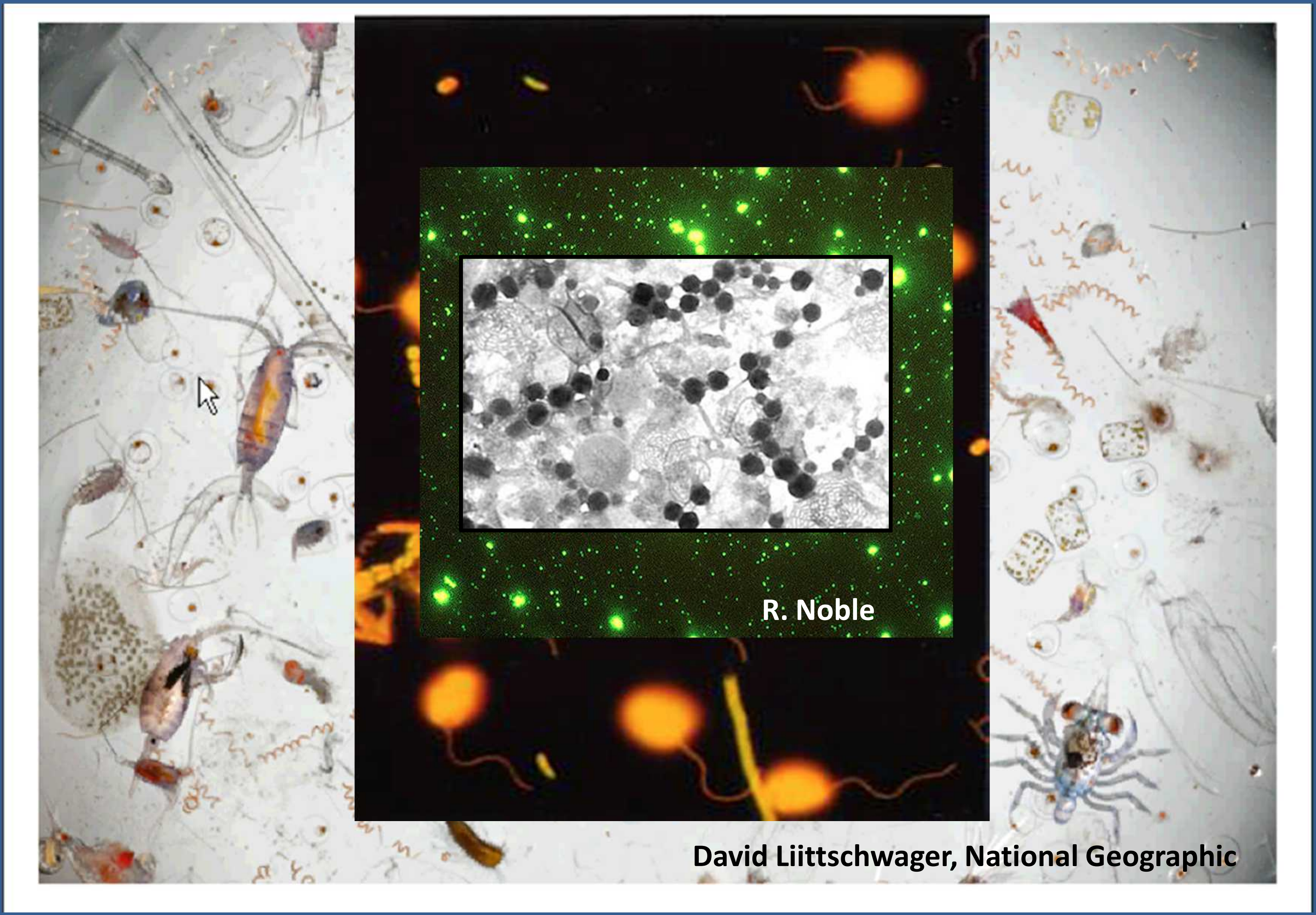
Specific techniques are often labor intensive and/or expensive to implement

Particles are closely tied to water quality



Masciangioli, T.; Zhang, W.-X. *Environ. Sci. Technol*. 2003, 37, 102A-108A).

A drop of seawater

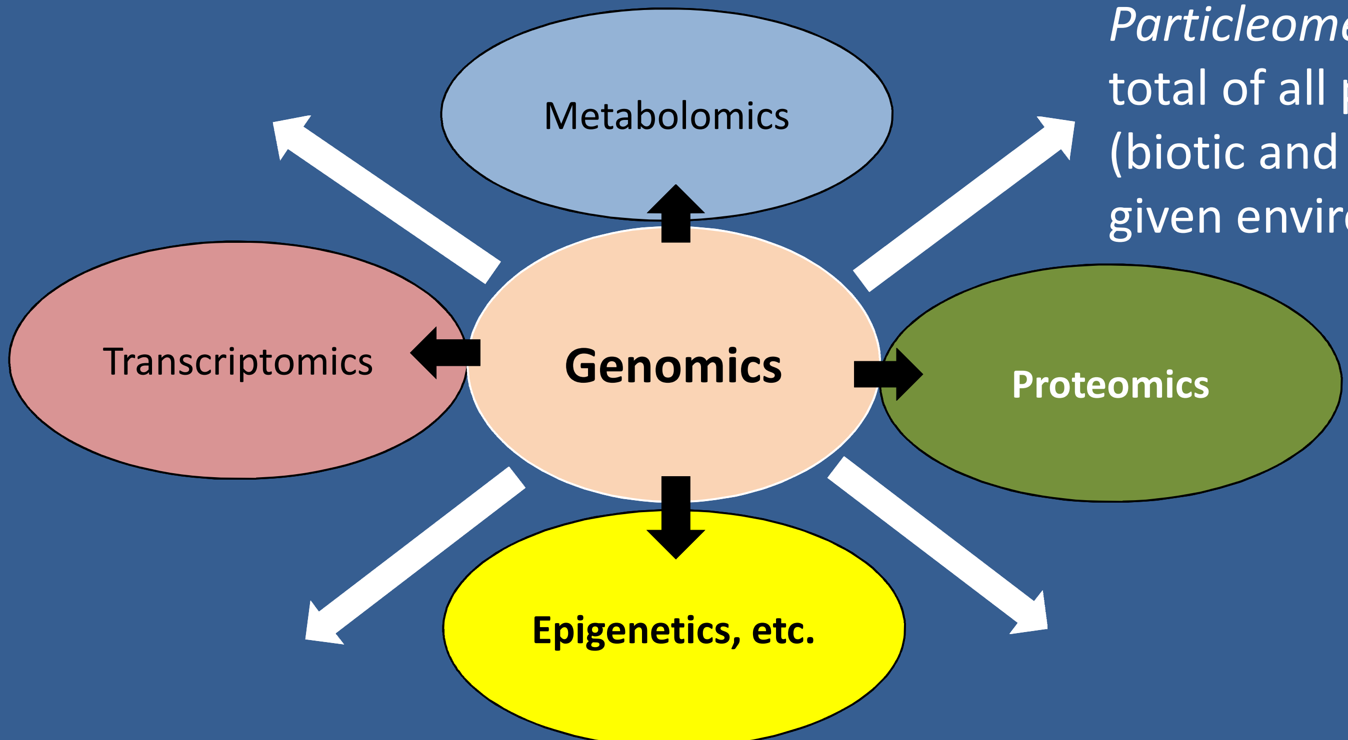


An –*omics* approach to particle characterization

Genome → the holistic characteristics of a set of genes corresponding to an organism

No one gene operates in isolation

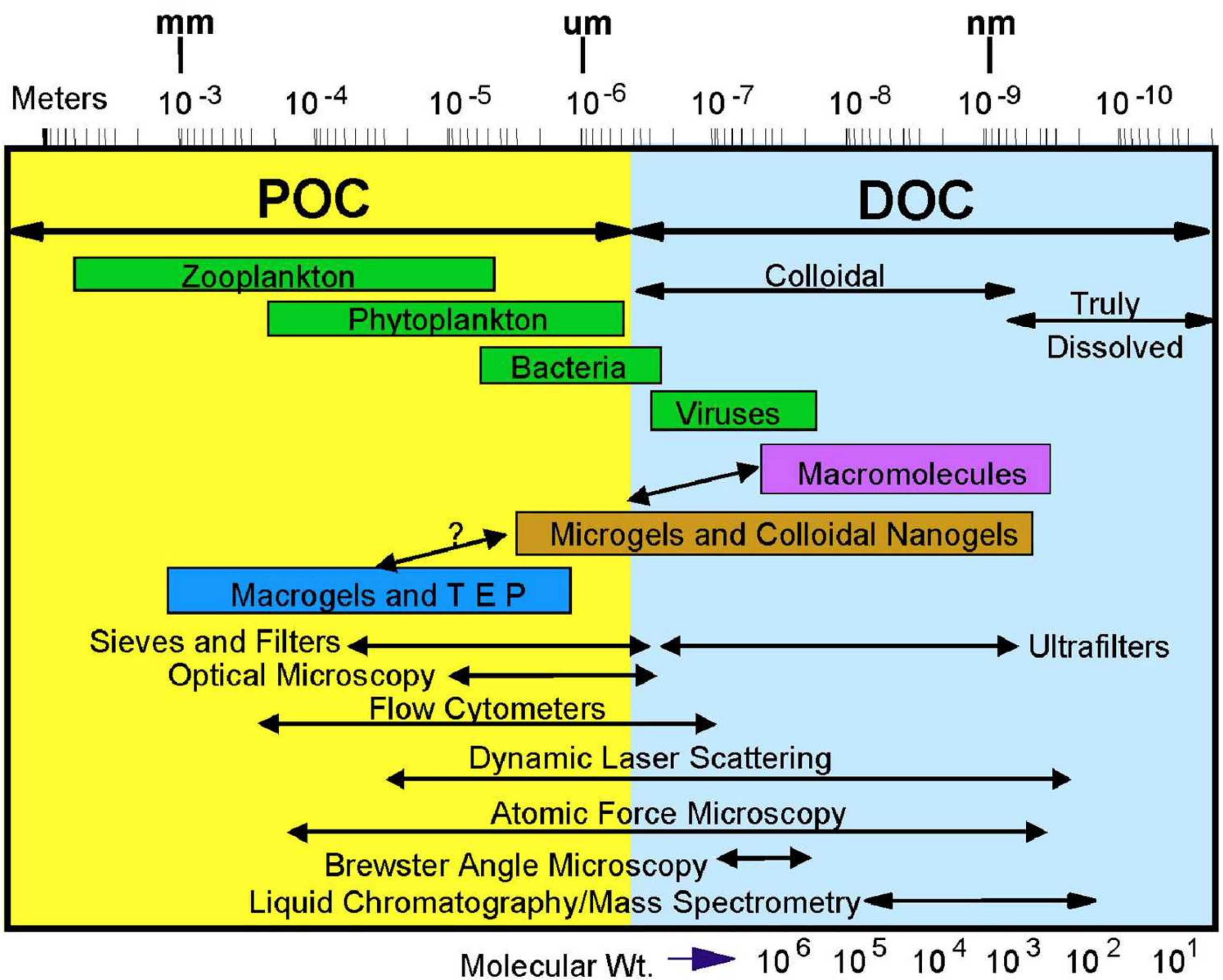
Particleome: The sum total of all particles (biotic and abiotic) in a given environment



Why use an –*omics* approach?

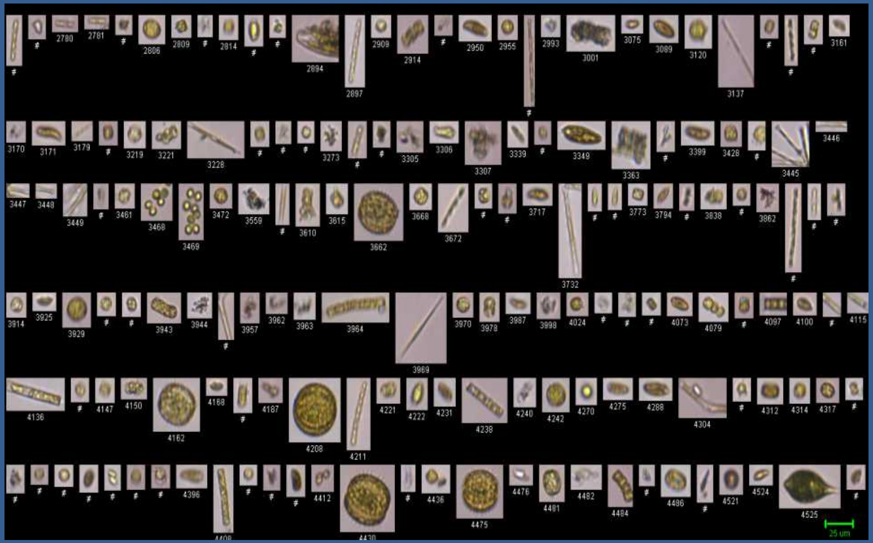
- Difficult to identify a single property that defines water quality
- Changes in holistic properties might be more sensitive and less expensive to track
- Move from a biomarker approach to a fingerprinting approach to identify threats to water quality
- Apply multivariate statistical techniques to real-time data, borrowing from bioinformatics

A variety of techniques are used to characterize particles



F. Azam, UCSD

FlowCAM



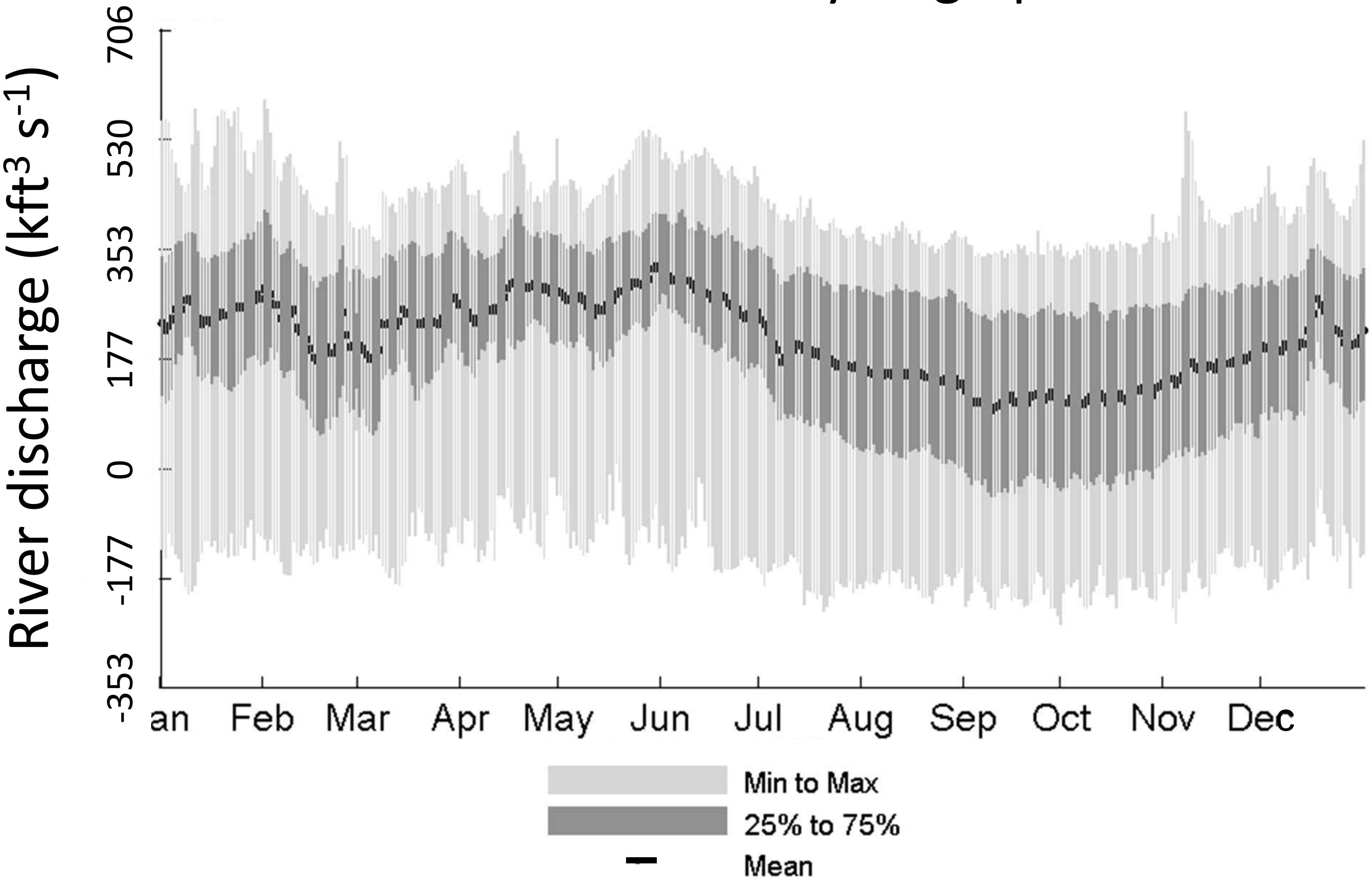
- Detects individual cells by scatter and/or fluorescence induced by laser light excitation;
- Rapidly counts and photographs individual particles;
- Distinguishes the shape and unique fluorescence properties of each cell in a sample;
- Provides a suite of particle properties (35 variables) as well as a set digital images of the particles.

Data Analysis

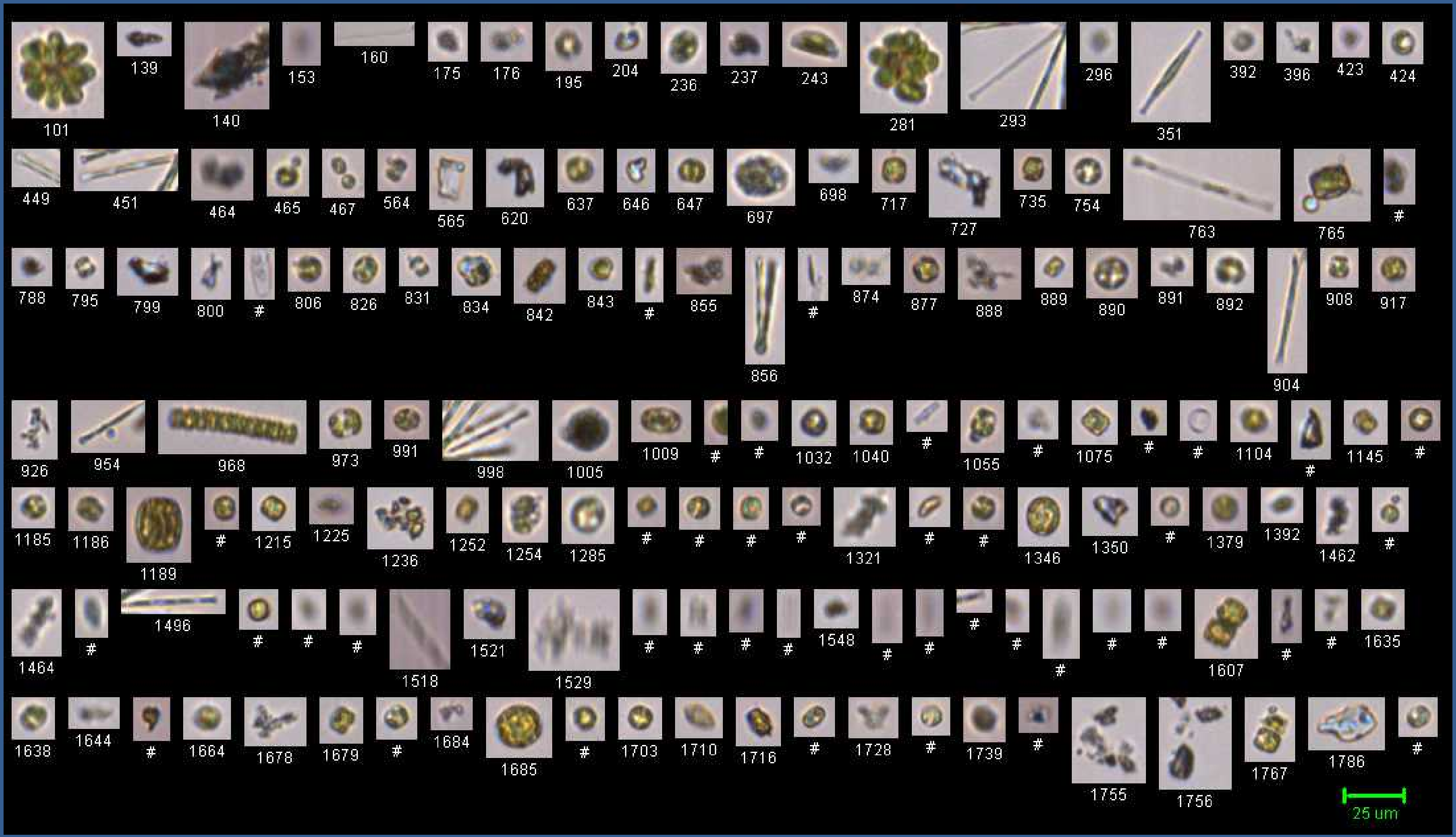
- Data reduction: Principal Component Analysis, (=Empirical Orthogonal Function analysis)
- Correlation analysis: correlate EOFs to environmental variables or to presence of pathogen of interest
- Example data set: Columbia River (Land Ocean Biogeochemical Observatory, LOBO)

Beaver Army Terminal
Map
Traffic
10 mi
20 km
Imagery ©2012 TerraMetrics -
Report a problem

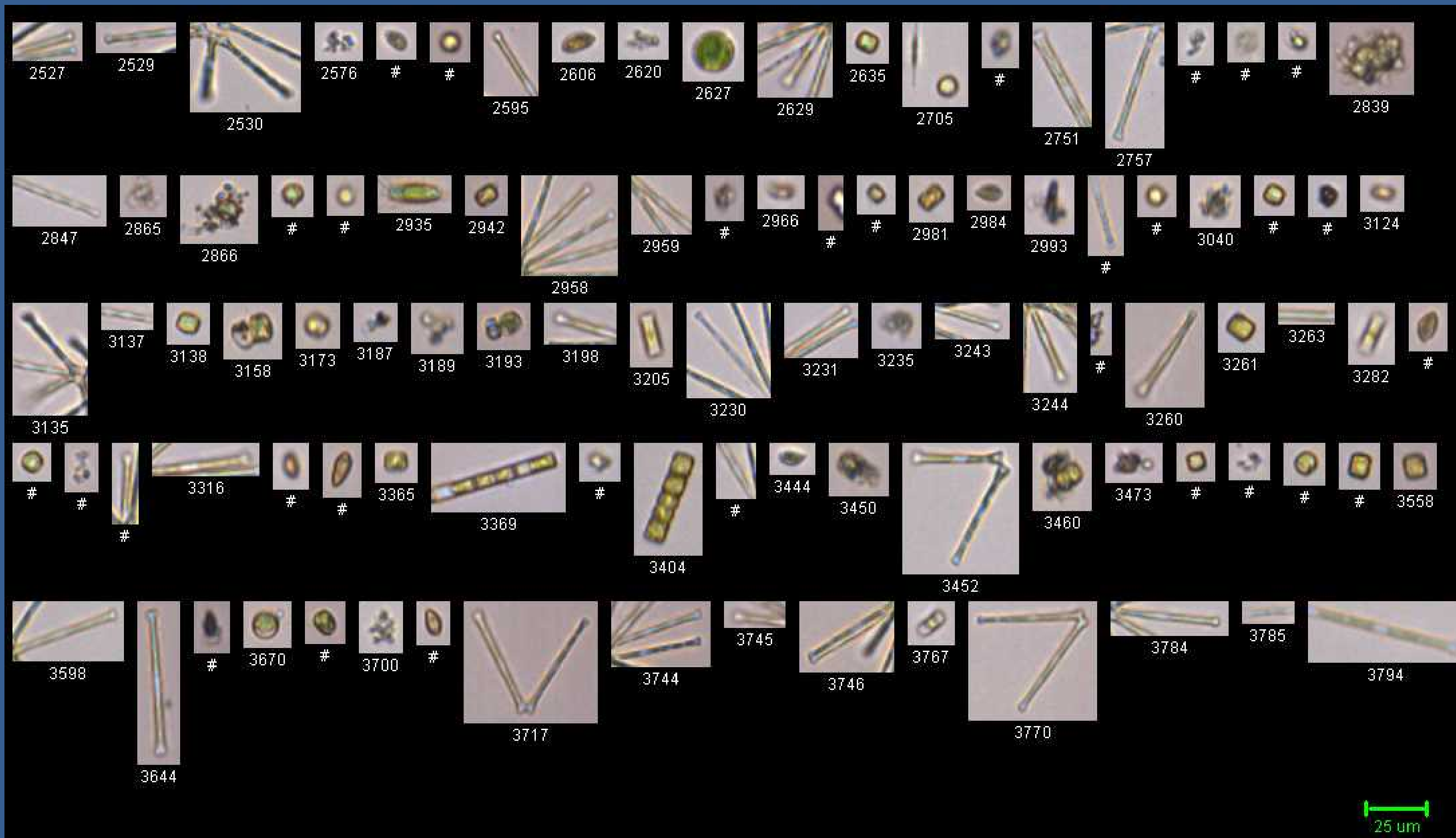
Columbia River Hydrograph



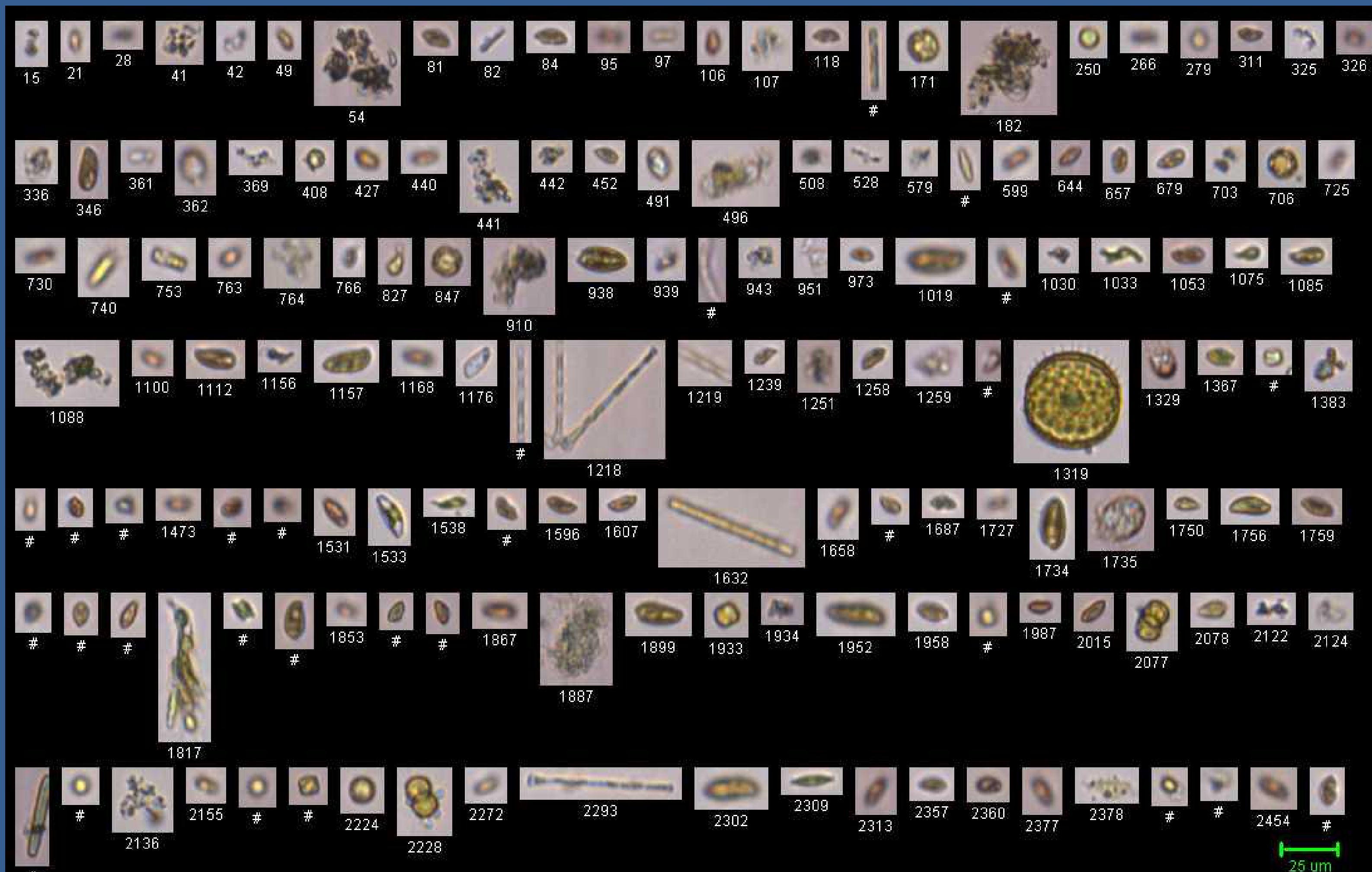
3/23/2011: Pre-freshet



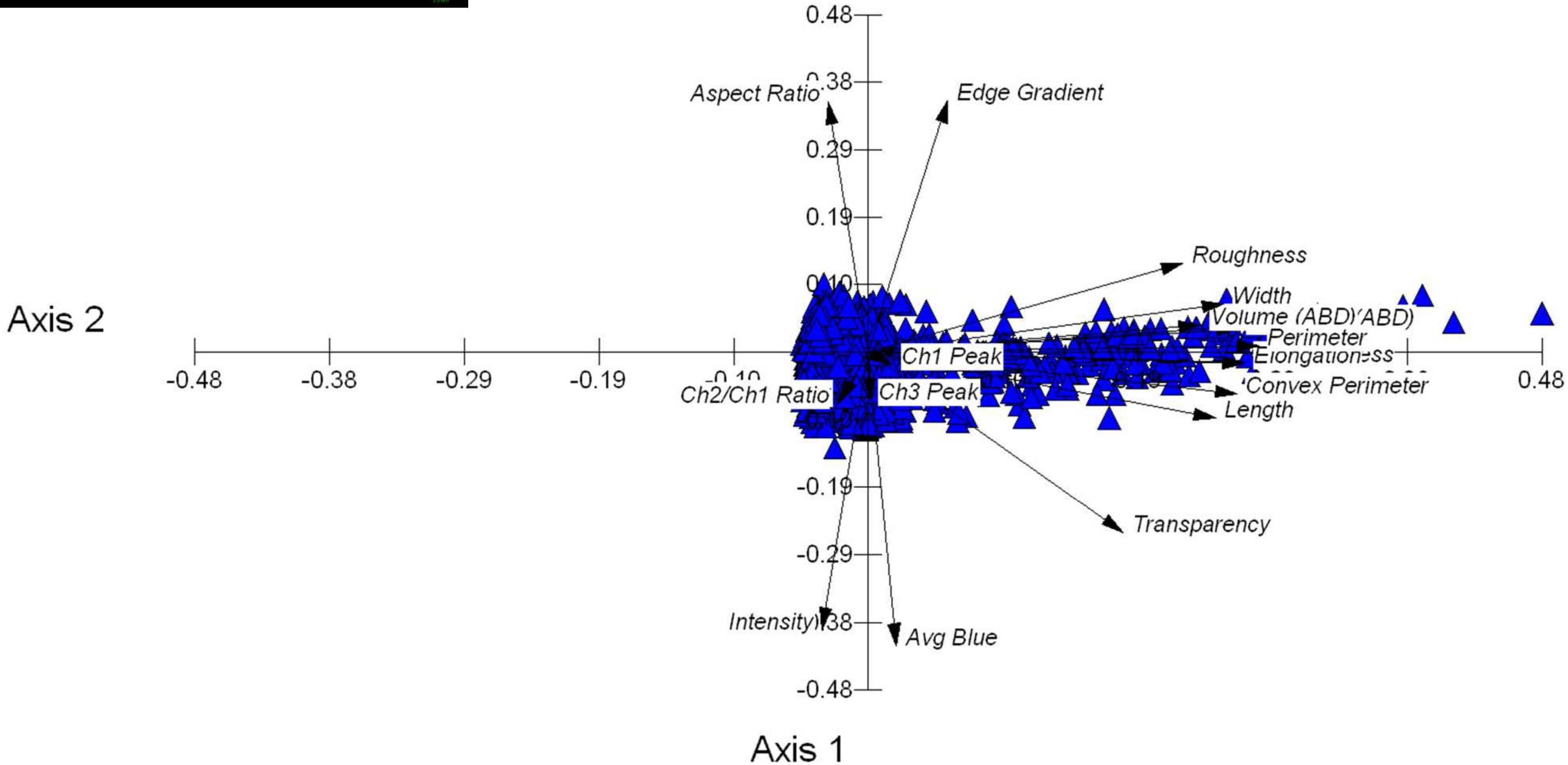
4/20/2011: spring bloom (pre-freshet)



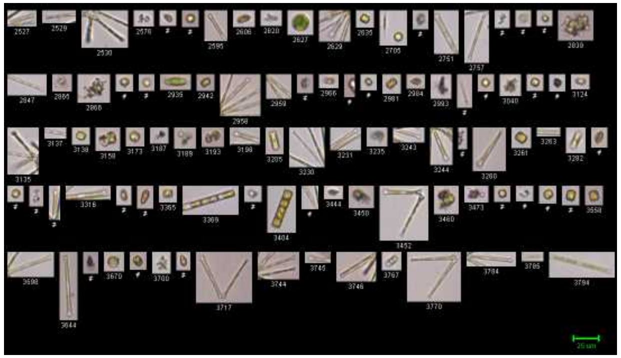
10/29/2010: fall/winter



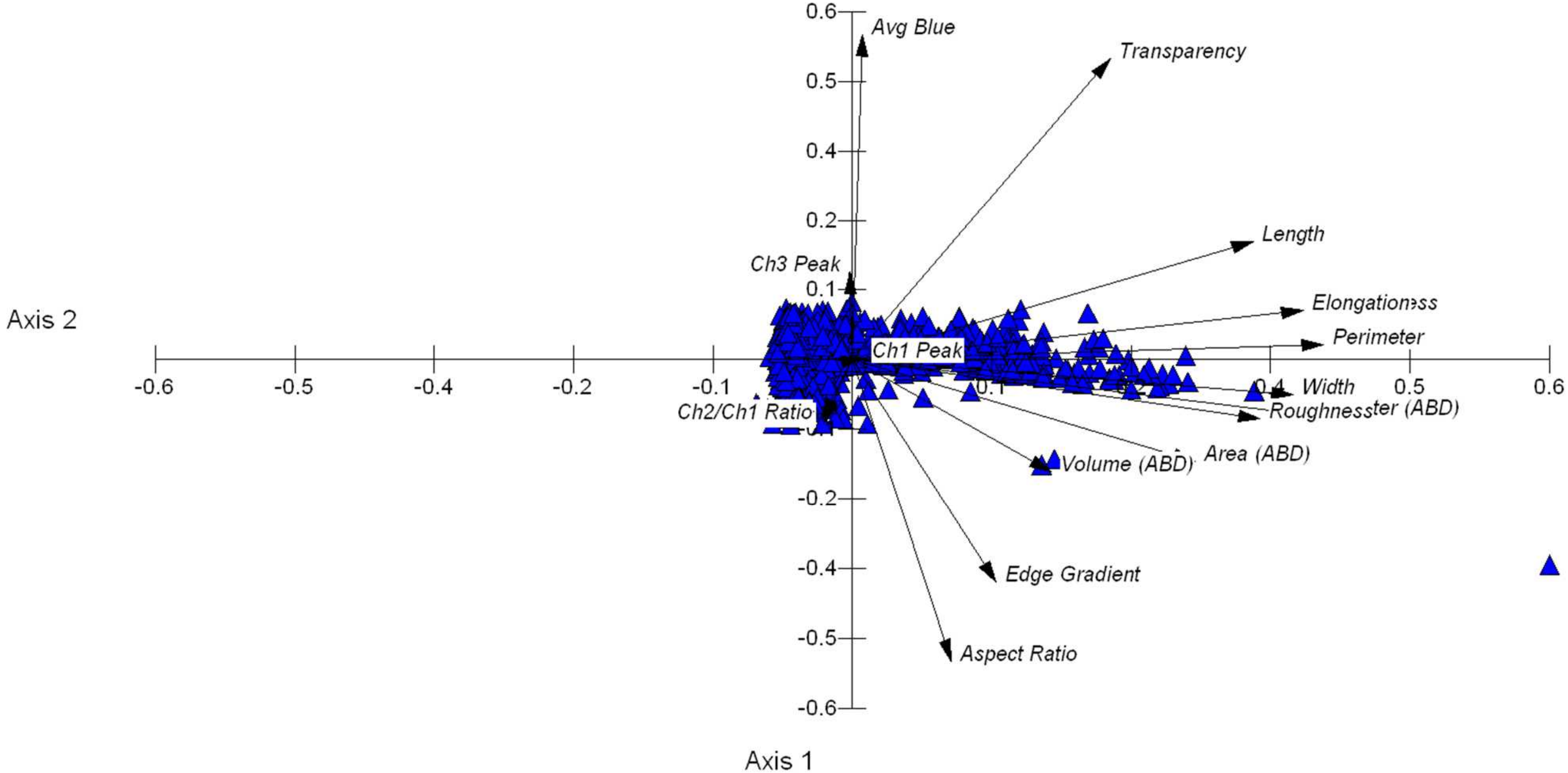
PCA case scores: 3/23/2011



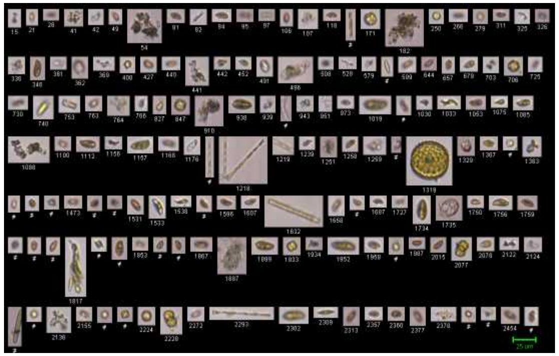
Vector scaling: 0.84



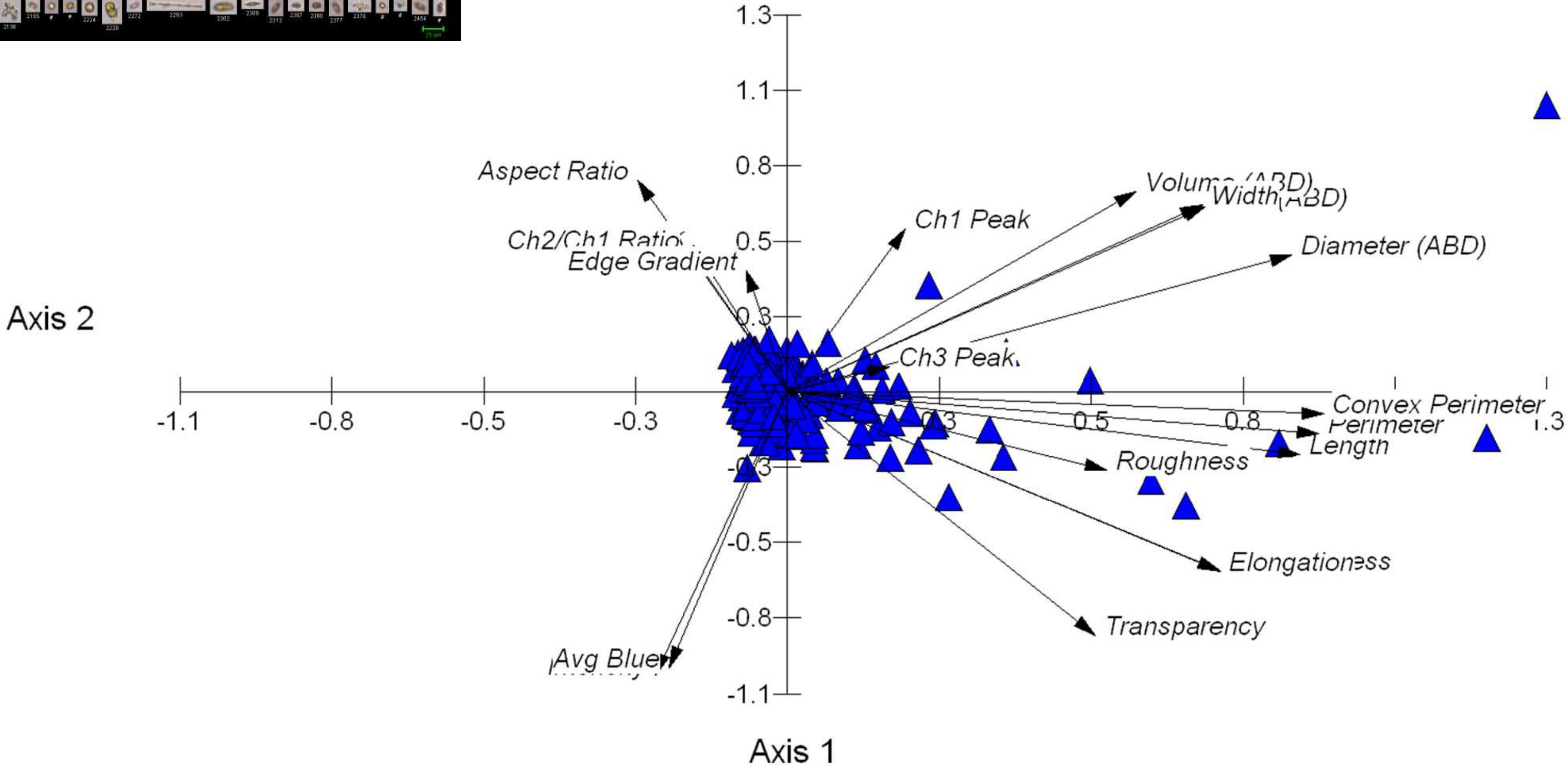
PCA case scores 4/20/2011



Vector scaling: 1.12

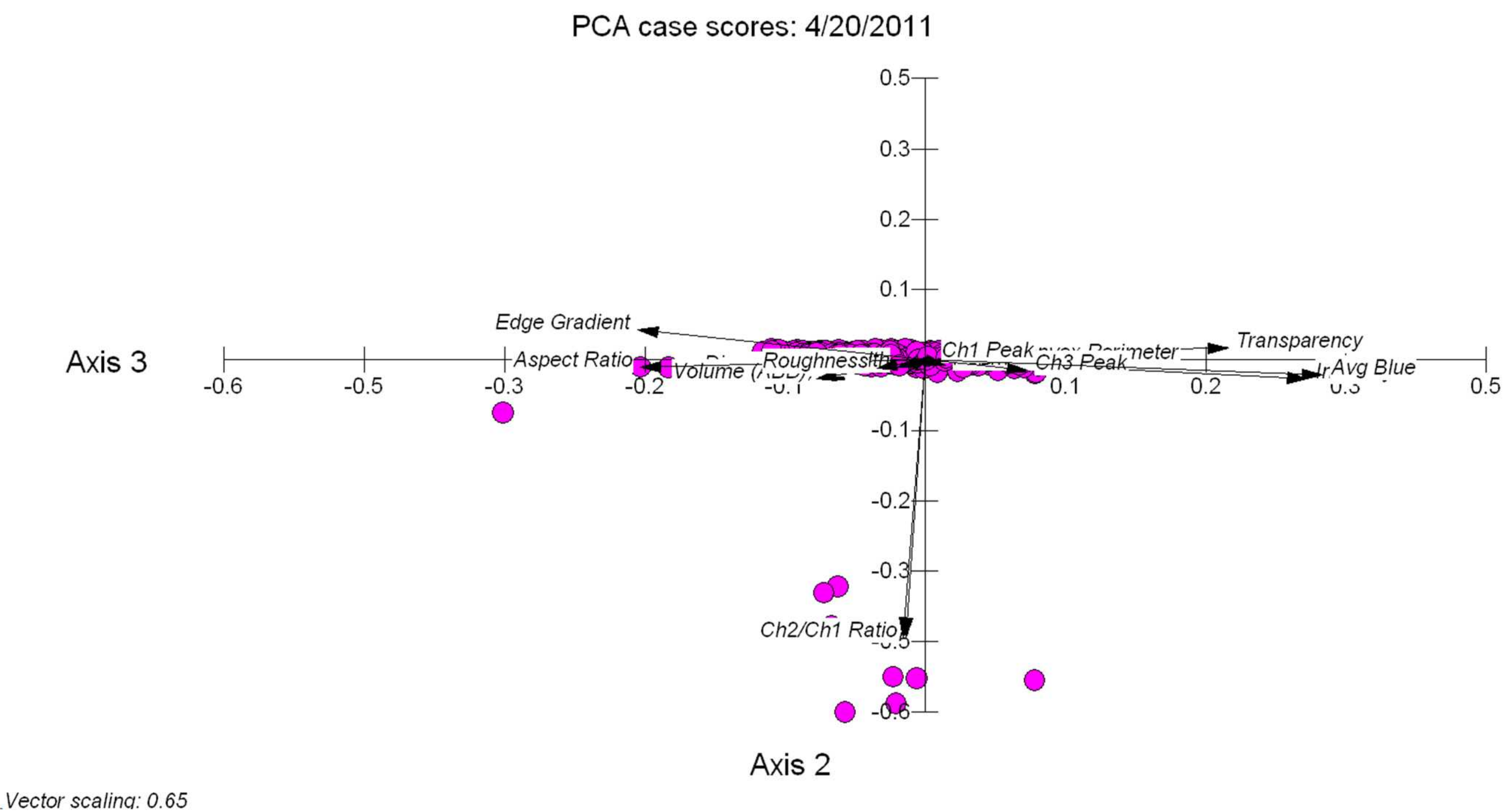


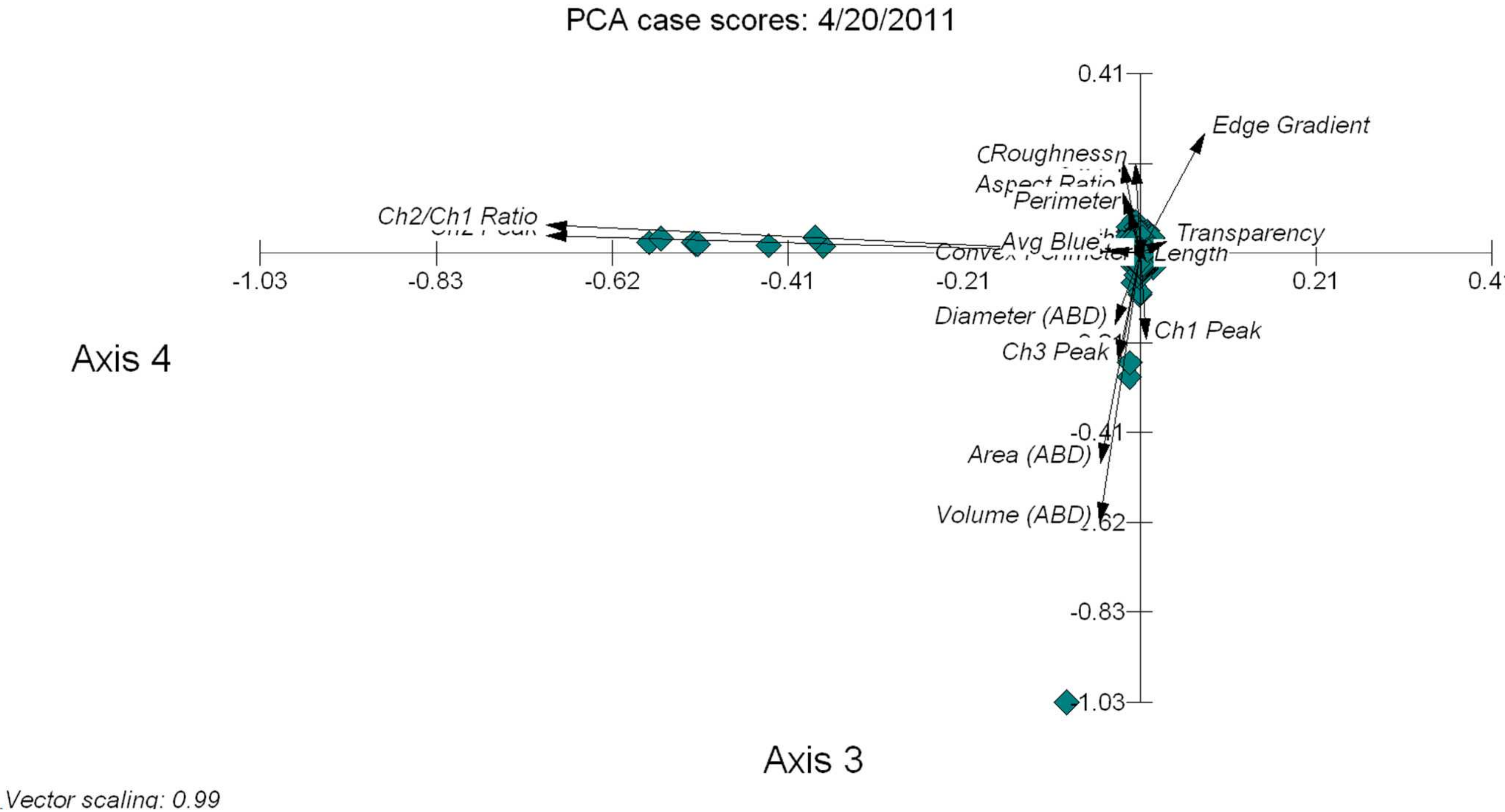
PCA case scores: 10/29/2010

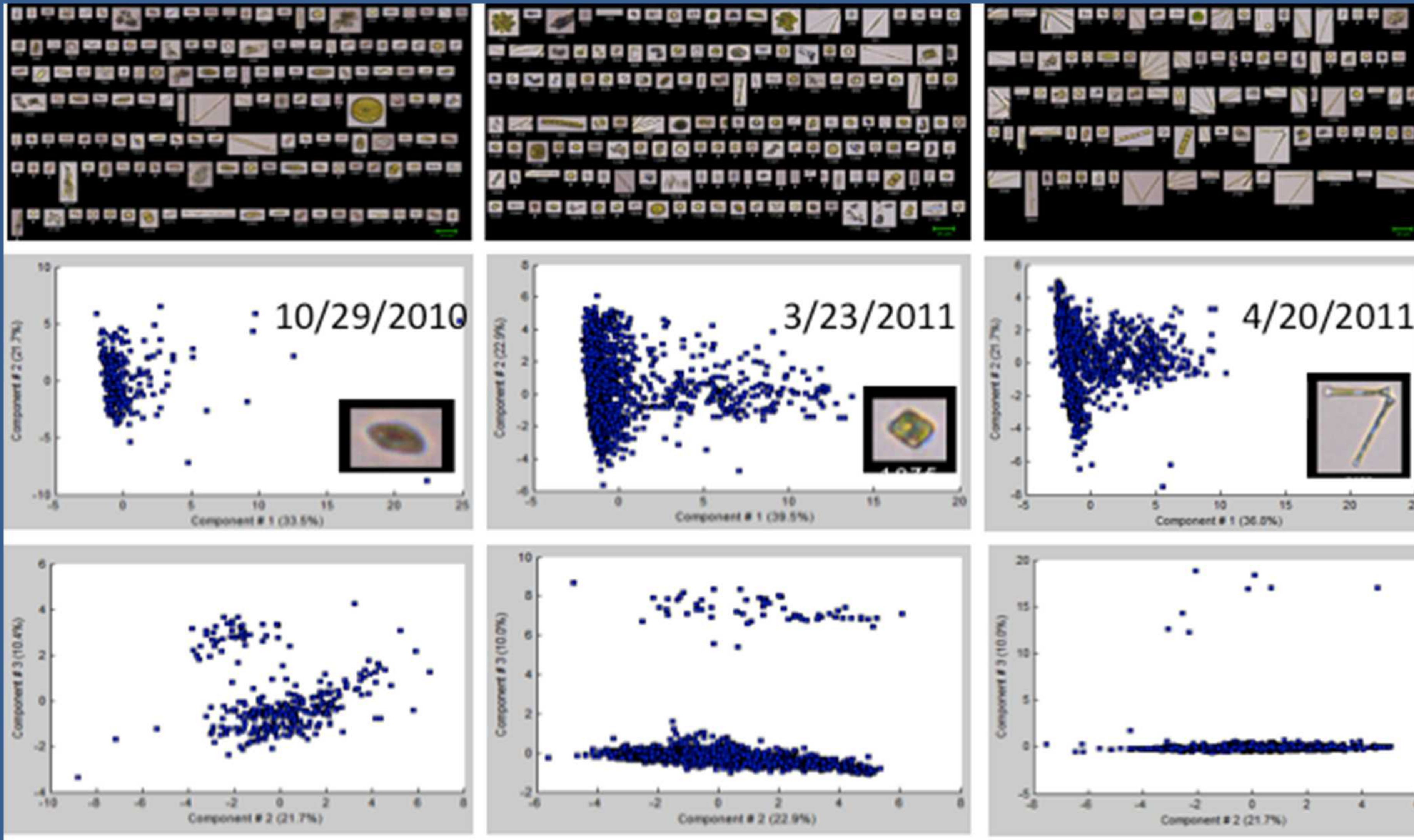


Vector scaling: 2.62

Axis 2 = PC 2: blue, transparency (hetero vs. autotroph?)
Axis 3 = phycoerythrin fluorescence/chl fluorescence



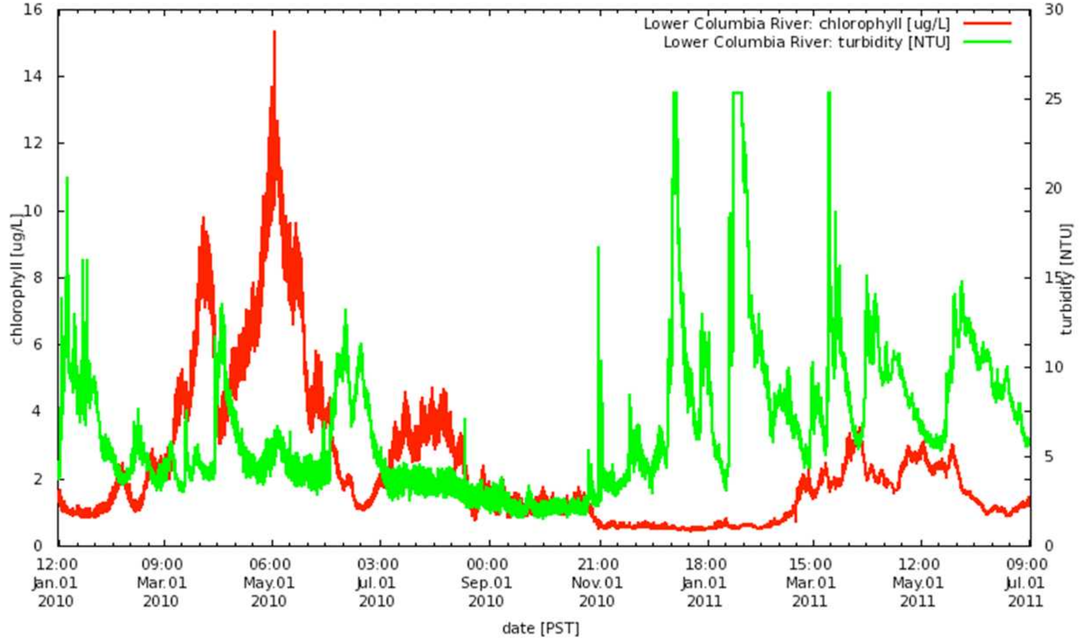

PCA case scores: 4/20/2011
Edge Gradient
Roughness
Aspect Ratio
Perimeter
Ch2/Ch1 Ratio
Transparency
Avg Blue
Length
Diameter (ABD)
Ch1 Peak
Ch3 Peak
Area (ABD)
Volume (ABD)
0.41
-0.41
-0.62
-0.83
-1.03
-1.03
-0.83
-0.62
-0.41
-0.21
0.21
0.41
Axis 4
Axis 3
Vector scaling: 0.99



10/29/2010
3/23/2011
4/20/2011
Component # 1 (33.5%)
Component # 2 (21.7%)
Component # 3 (10.4%)
Component # 1 (39.5%)
Component # 2 (22.9%)
Component # 3 (10.0%)
Component # 1 (36.0%)
Component # 2 (21.7%)
Component # 3 (10.0%)

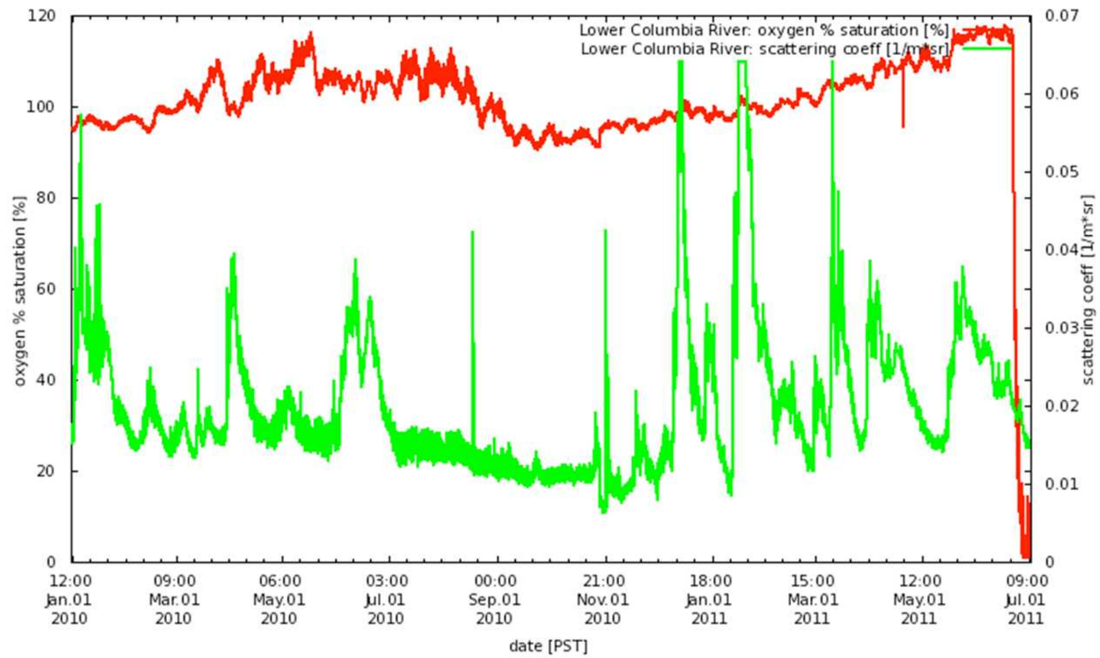
— chl

— turbidity

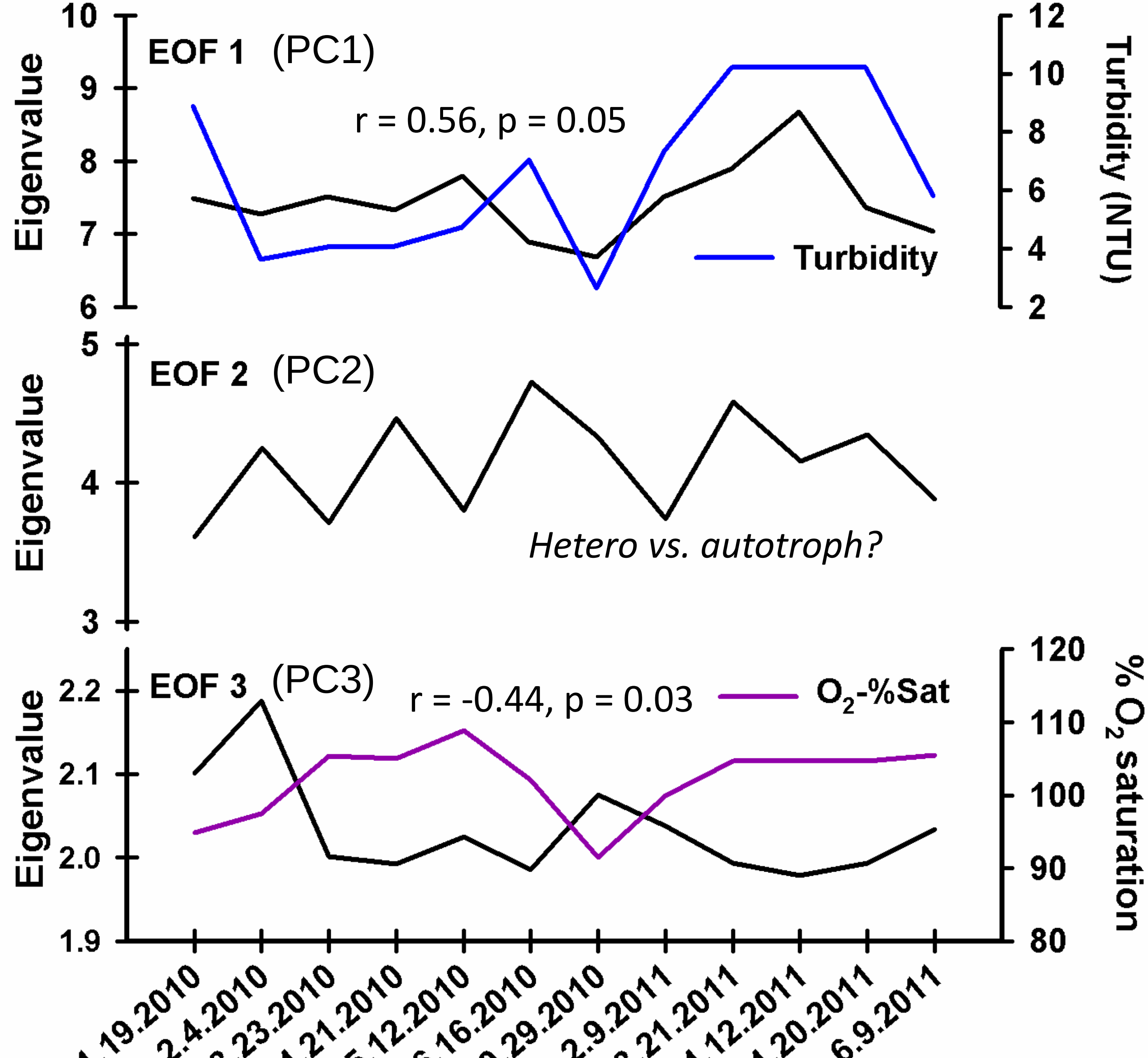


— $\%O_2$ sat

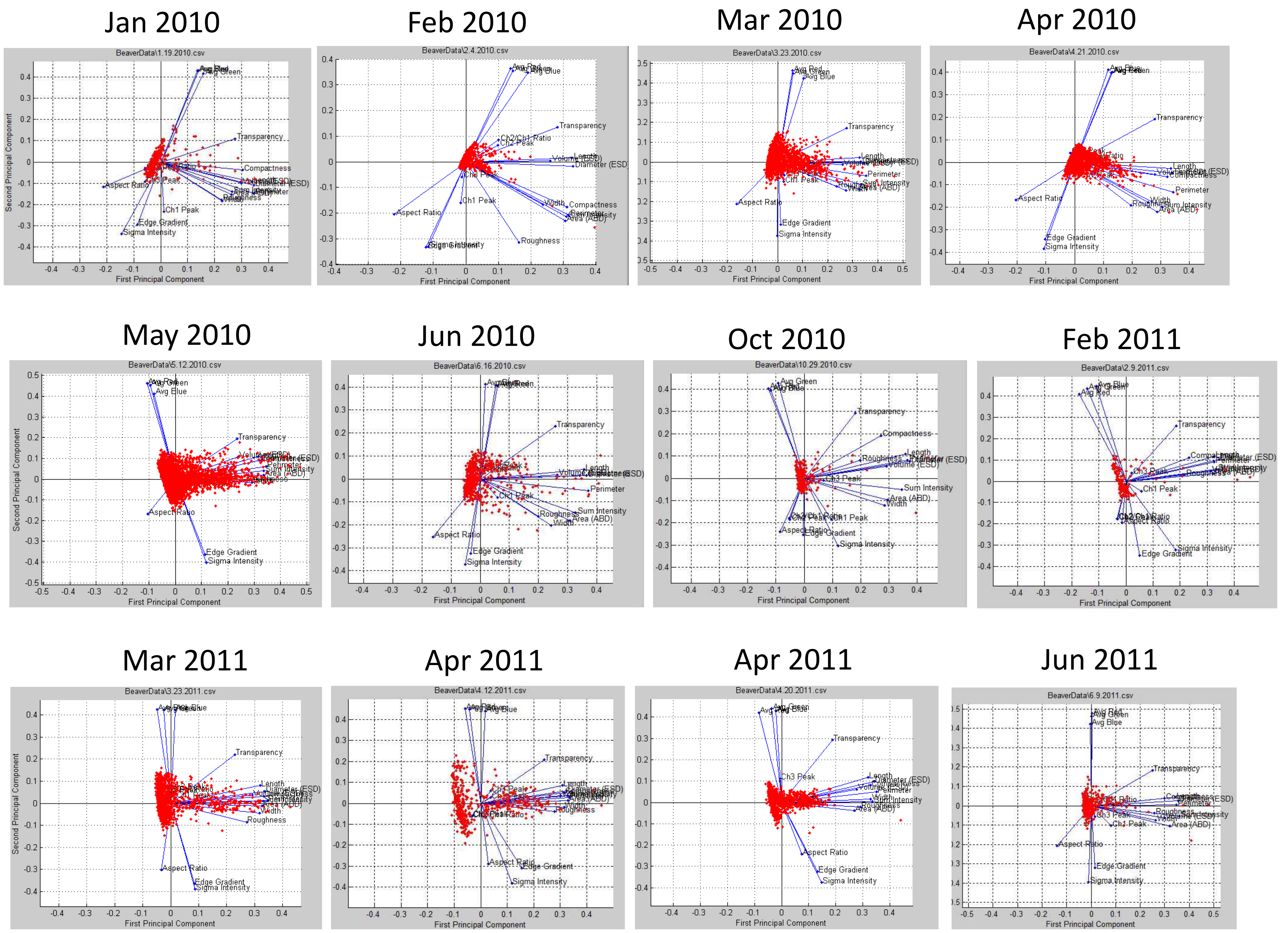
— scattering



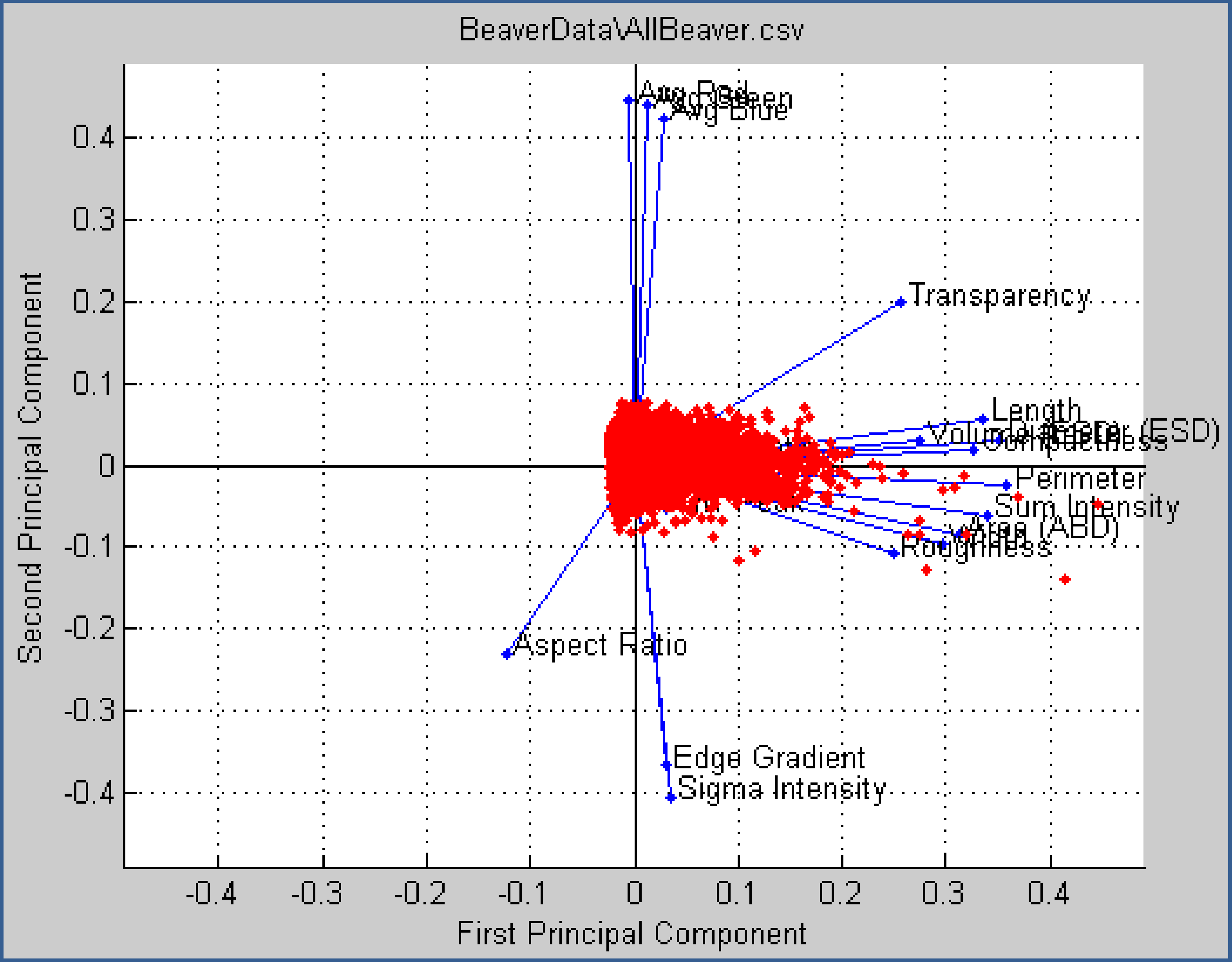
Variance



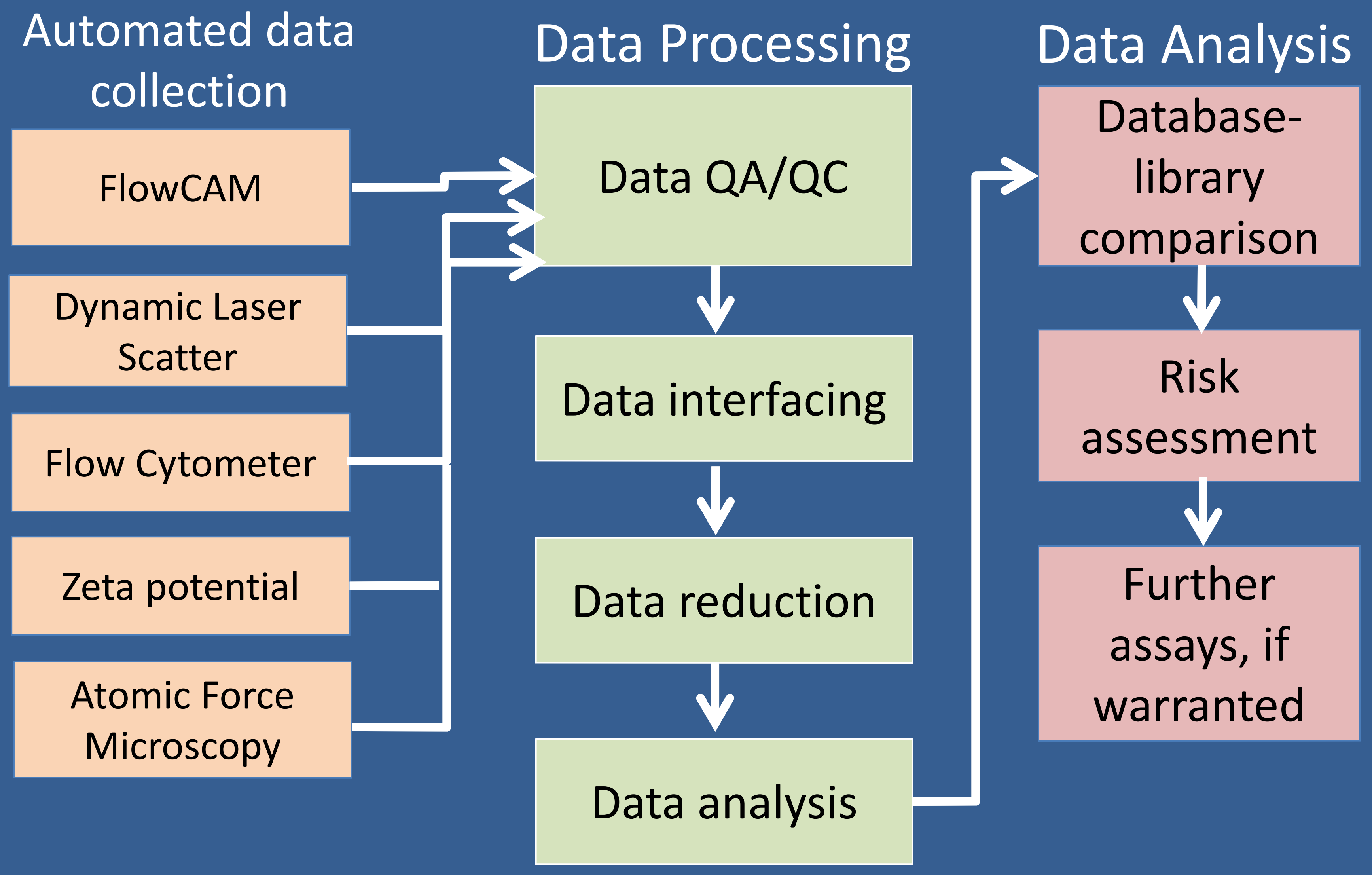
EOF 1 (PC1)
r = 0.56, p = 0.05
Turbidity
Eigenvalue
Turbidity (NTU)
10
9
8
7
6
12
10
8
6
4
2
EOF 2 (PC2)
Hetero vs. autotroph?
Eigenvalue
5
4
3
EOF 3 (PC3)
r = -0.44, p = 0.03
O2-%Sat
Eigenvalue
% O2 saturation
2.2
2.1
2.0
1.9
120
110
100
90
80
1.19.2010
2.4.2010
3.23.2010
4.21.2010
5.12.2010
6.16.2010
10.29.2010
2.9.2011
3.21.2011
4.12.2011
4.20.2011
6.9.2011


Jan 2010
BeaverData\1.19.2010.csv
Feb 2010
BeaverData\2.4.2010.csv
Mar 2010
BeaverData\3.23.2010.csv
Apr 2010
BeaverData\4.21.2010.csv
May 2010
BeaverData\5.12.2010.csv
Jun 2010
BeaverData\6.16.2010.csv
Oct 2010
BeaverData\10.29.2010.csv
Feb 2011
BeaverData\2.9.2011.csv
Mar 2011
BeaverData\3.23.2011.csv
Apr 2011
BeaverData\4.12.2011.csv
Apr 2011
BeaverData\4.20.2011.csv
Jun 2011
BeaverData\6.9.2011.csv
First Principal Component
Second Principal Component
Avg Red
Avg Green
Avg Blue
Transparency
Compactness
Length
Diameter (ESD)
Volume (ESD)
Perimeter
Sum Intensity
Area (ABD)
Width
Roughness
Ch1 Peak
Ch2 Peak
Ch3 Peak
Ch2/Ch1 Ratio
Aspect Ratio
Edge Gradient
Sigma Intensity

Global PCA


BeaverData\AllBeaver.csv
Second Principal Component
First Principal Component
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4
Transparency
Length
Perimeter
Sum Intensity
Roughness
Aspect Ratio
Edge Gradient
Sigma Intensity

Vision



Automated data collection
FlowCAM
Dynamic Laser Scatter
Flow Cytometer
Zeta potential
Atomic Force Microscopy
Data Processing
Data QA/QC
Data interfacing
Data reduction
Data analysis
Data Analysis
Database-library comparison
Risk assessment
Further assays, if warranted

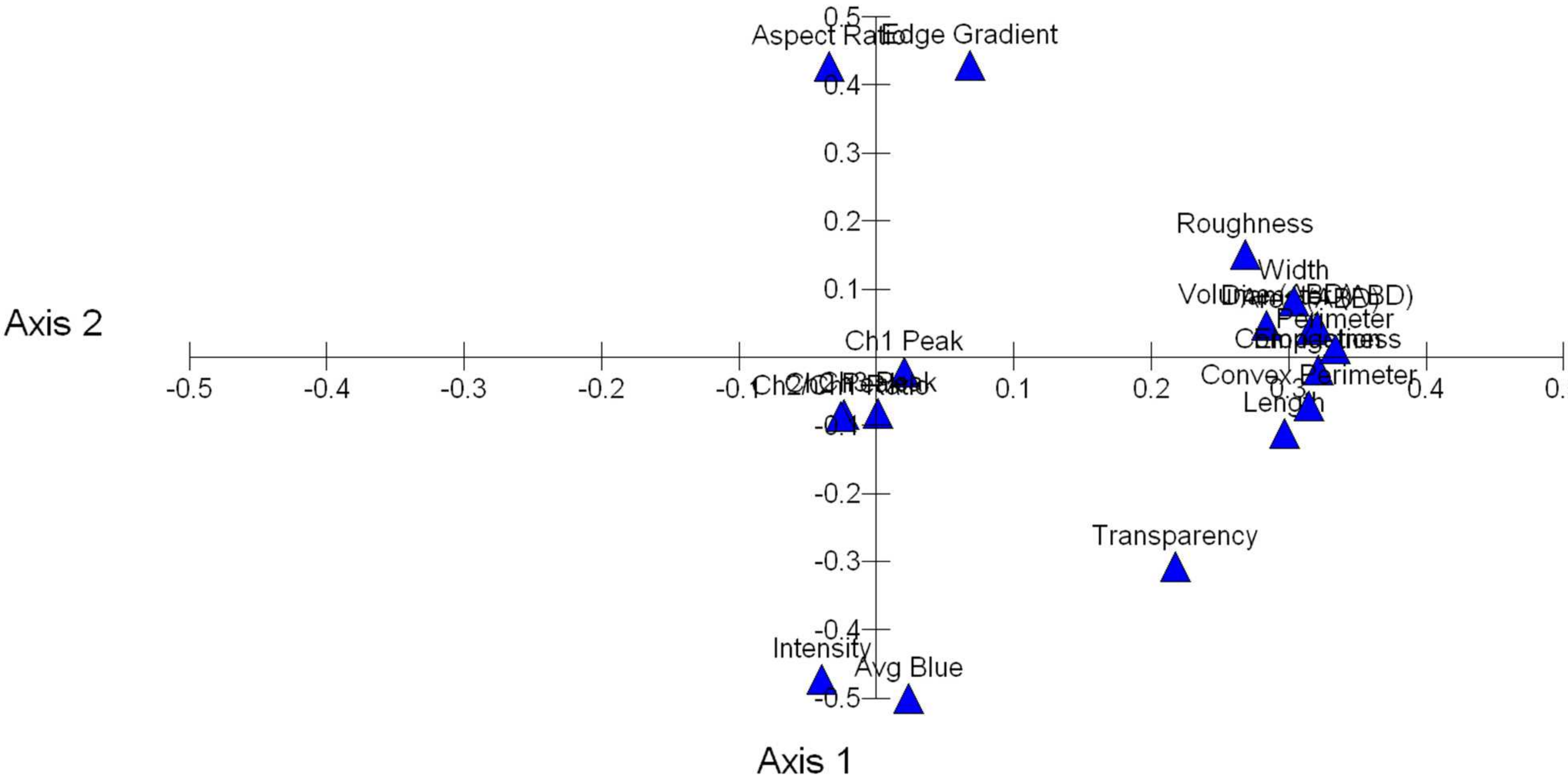
Challenges

- Large size range of particles in the natural environment
- Different size ranges often requires different technique to measure particles → Data interfacing
- Obtaining comprehensive and representative databases/libraries that provide meaningful diagnostics

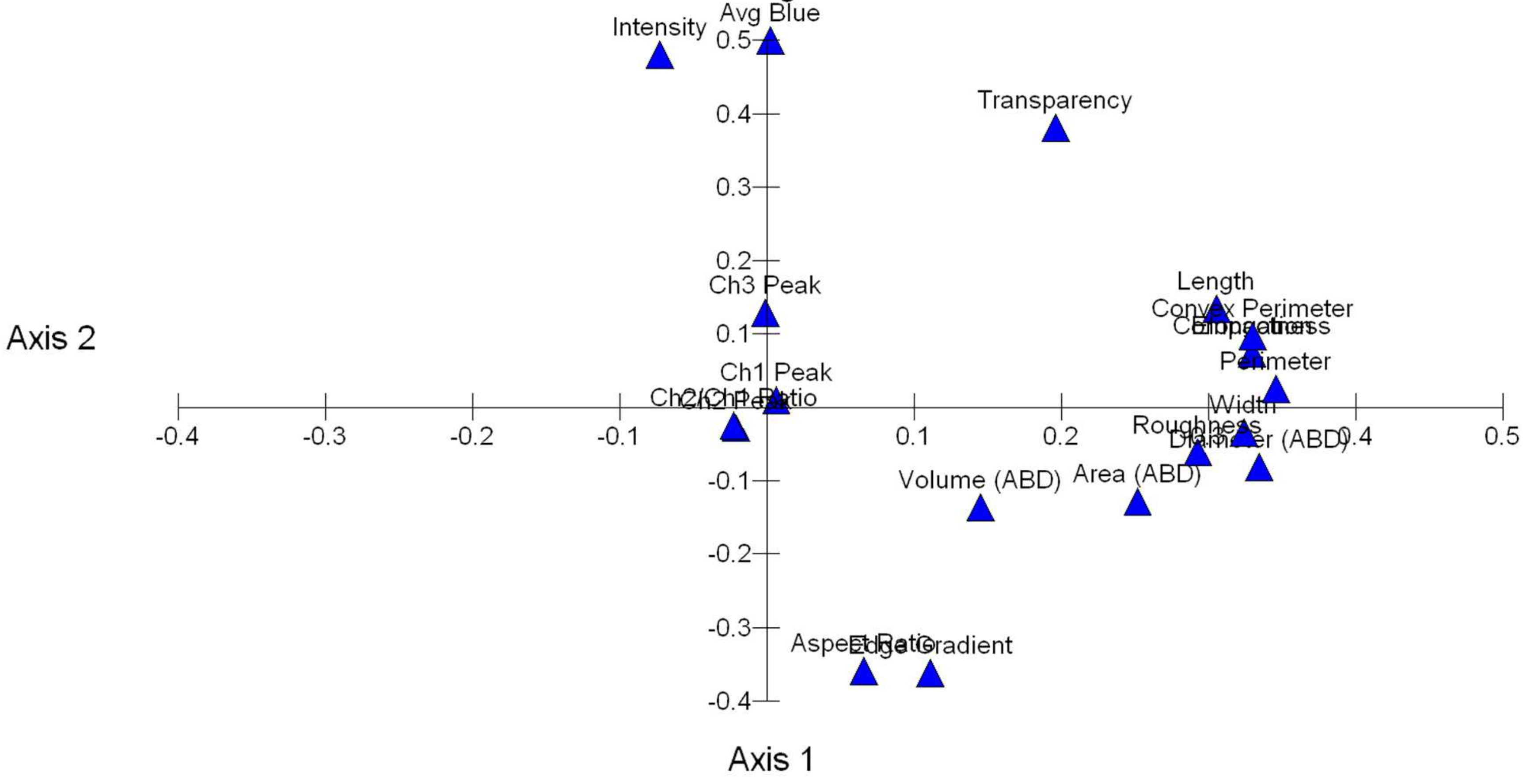
Summary and ongoing work

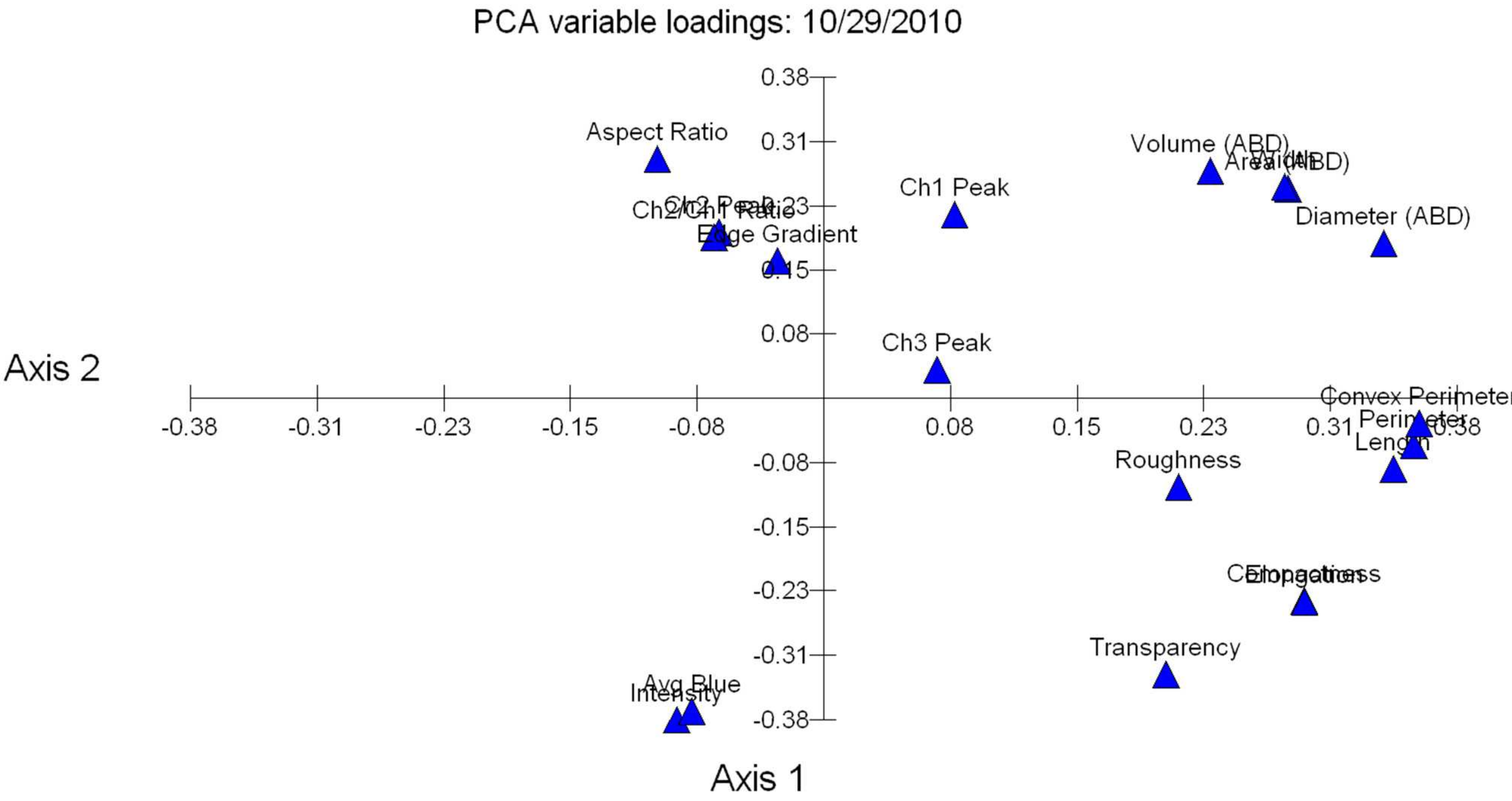
- Through particle-omics, we seek to characterize holistic properties of the particle load in water to identify meaningful changes or qualities
- Goal 1: perform laboratory tests using perturbations to assess sensitivity of the approach
- Goal 2: build a library of data and images for the lower Columbia River
- Goal 3: develop a standardized workflow for data processing

PCA variable loadings: 3/23/2011



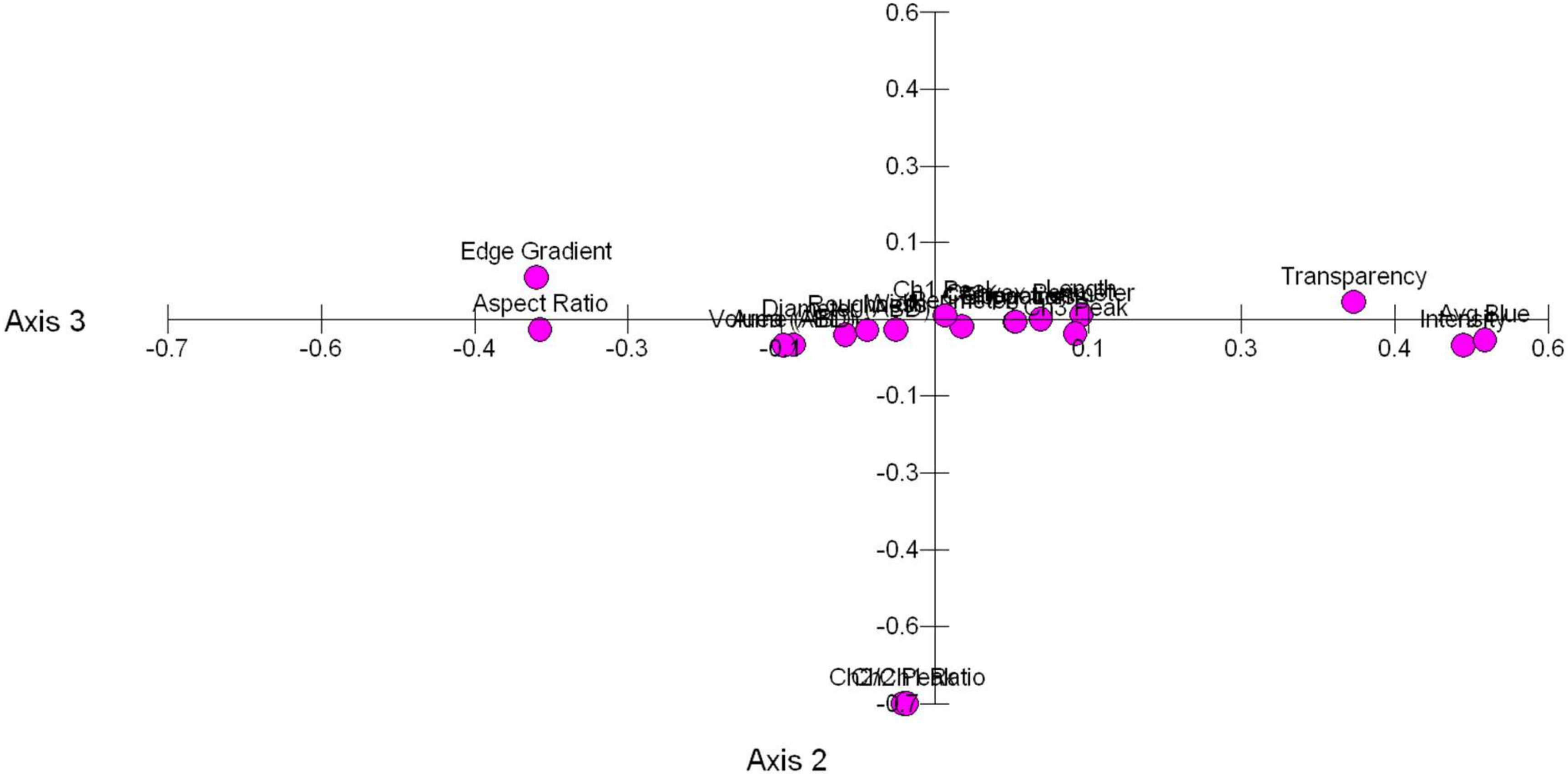
PCA variable loadings: 4/20/2011



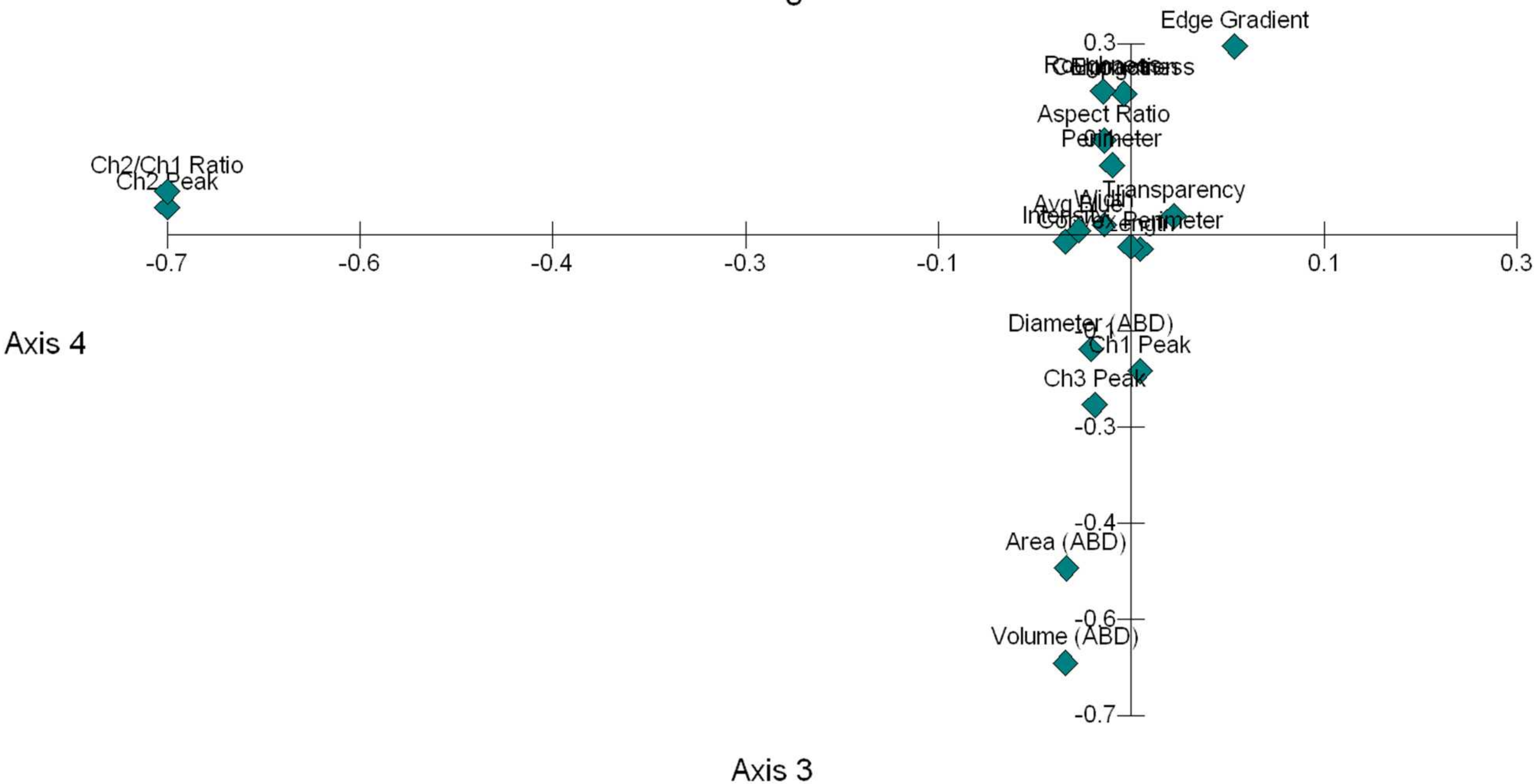


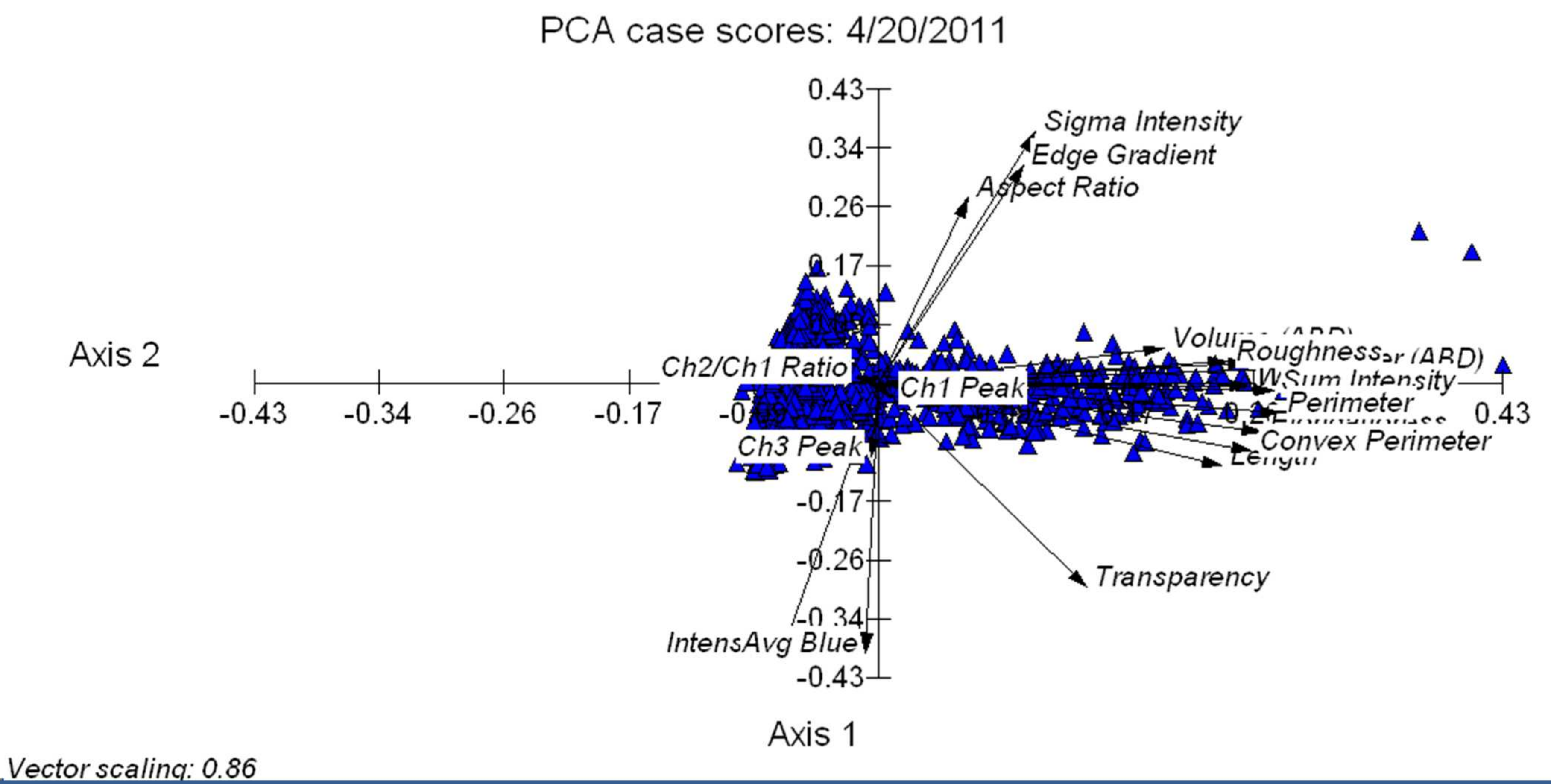
PCA variable loadings: 10/29/2010
Axis 2
Axis 1
0.38
0.31
0.23
0.15
0.08
-0.08
-0.15
-0.23
-0.31
-0.38
-0.38
-0.31
-0.23
-0.15
-0.08
0.08
0.15
0.23
0.31
0.38
Aspect Ratio
Ch2 Peak
Ch2/Ch1 Ratio
Edge Gradient
Ch1 Peak
Volume (ABD)
Area (ABD)
Width
Diameter (ABD)
Ch3 Peak
Convex Perimeter
Perimeter
Length
Roughness
Compactness
Elongation
Transparency
Avg Blue
Intensity

PCA variable loadings: 4/20/2011



PCA variable loadings: 4/20/2011




PCA case scores: 4/20/2011
Axis 2
Axis 1
0.43
0.34
0.26
0.17
-0.17
-0.26
-0.34
-0.43
-0.43
-0.34
-0.26
-0.17
0.43
Sigma Intensity
Edge Gradient
Aspect Ratio
Ch2/Ch1 Ratio
Ch1 Peak
Ch3 Peak
Roughness
Sum Intensity
Perimeter
Convex Perimeter
Transparency
IntensAvg Blue
Vector scaling: 0.86