Where we are

- Supervised Learning
  - Rules/Trees
  - Ensembles/Boosting
    - Ex: Random Forests
  - Optimization with Gradient Descent
    - Regression
    - Classification
      - Logistic Regression and SVMs
- Unsupervised Learning

Unsupervised Learning

Four categories of machine learning

- Supervised learning: feedback supplied explicitly
- Reinforcement learning: feedback supplied by the environment

  Ex: Control theory, decision theory
- Game theory: feedback supplied by other actors in the system
- *Unsupervised learning:* *no feedback supplied*

> "Almost all work in unsupervised learning can be viewed in terms of learning a probabilistic model of the data."
>
> Zoubin Ghahramani, 2004

Applications of Unsupervised Learning

- Outlier detection or monitoring
  - "Is this normal?"
- Classification
  - "What group is this item most similar to?"
- Compression and Communication
  - send the string "abababababababab" or send "ab*7"

- In all cases, we identify patterns that describe the data and put them to use

Clustering

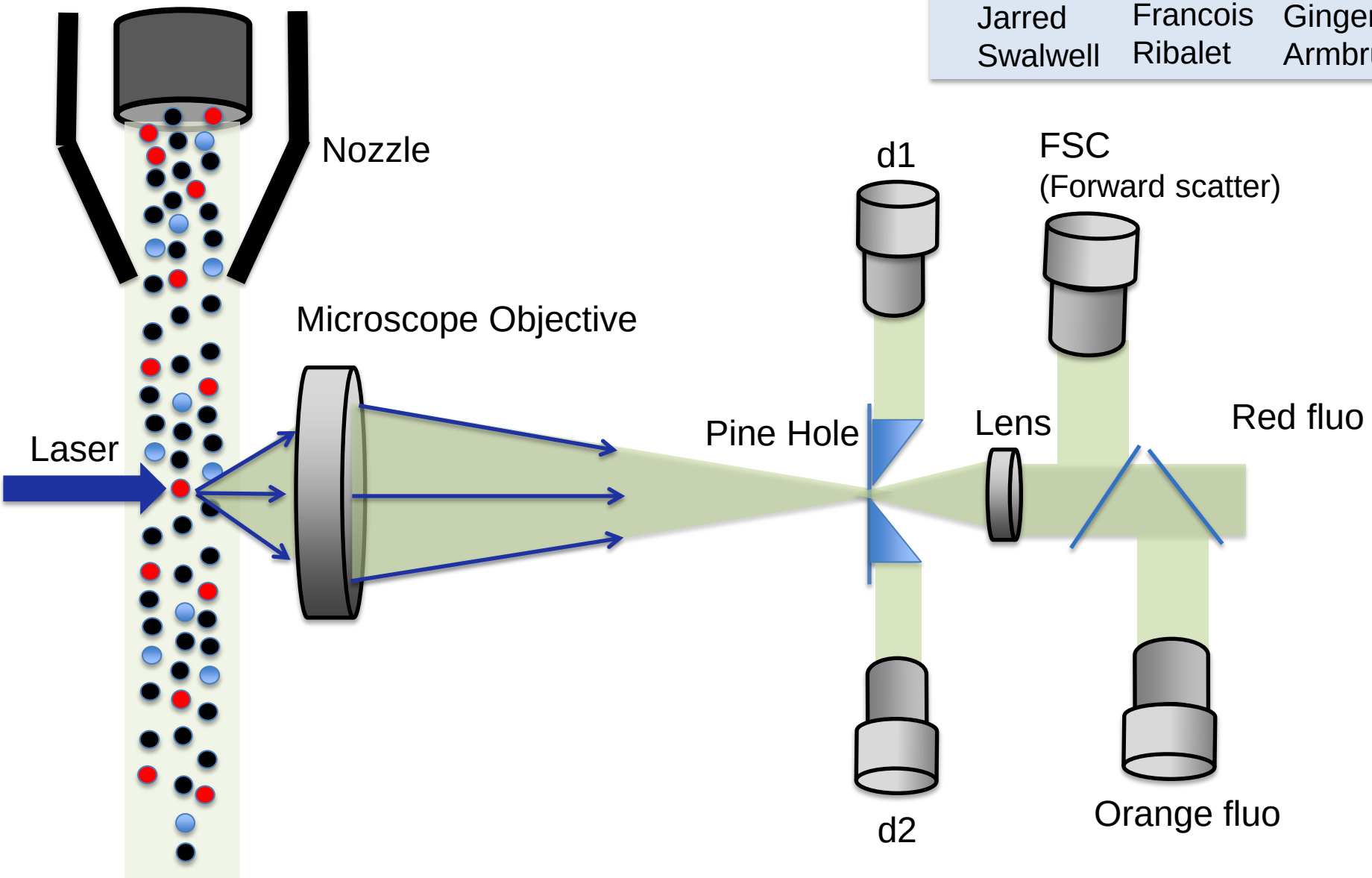
- No precise definition of a cluster
- Output is always a set of sets of items
- Items may be points in some multi-dimensional space
  - Ex: Find similar
- Items may be vertices in a graph
  - Ex: Community detection in social networks

SeaFlow


Jarred Swalwell


Francois Ribalet


Ginger Armbrust

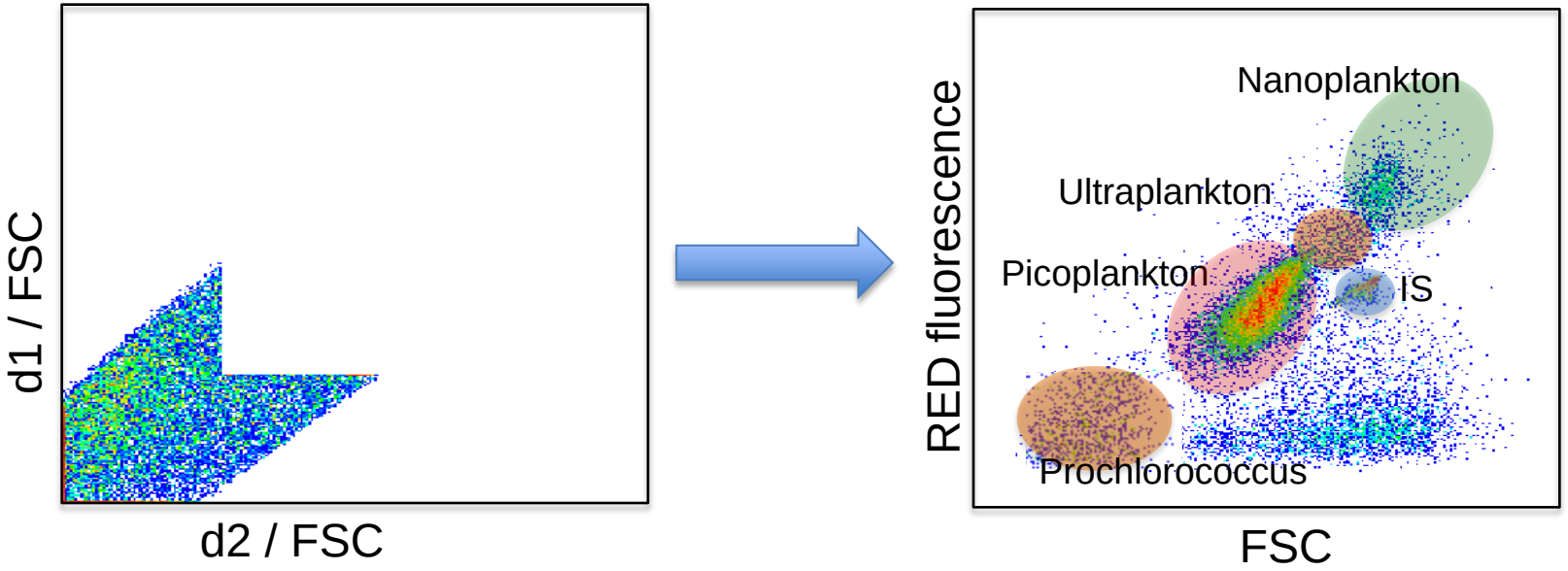


SeaFlow


Jarred Swalwell


Francois Ribalet


Ginger Armbrust



- Continuous observations of various phytoplankton groups from 1-20 μm in size
  - Based on RED fluo: *Prochlorococcus*, Pico-, Ultra- and Nanoplankton
  - Based on ORANGE fluo: *Synechococcus,* Cryptophytes
  - Based on FSC: Coccolithophores