3-D clustering to identify multiple oligomerization states by FRET

Kim Scott
Mentor: Henry Lester
SURF Seminar Day: October 17, 2009

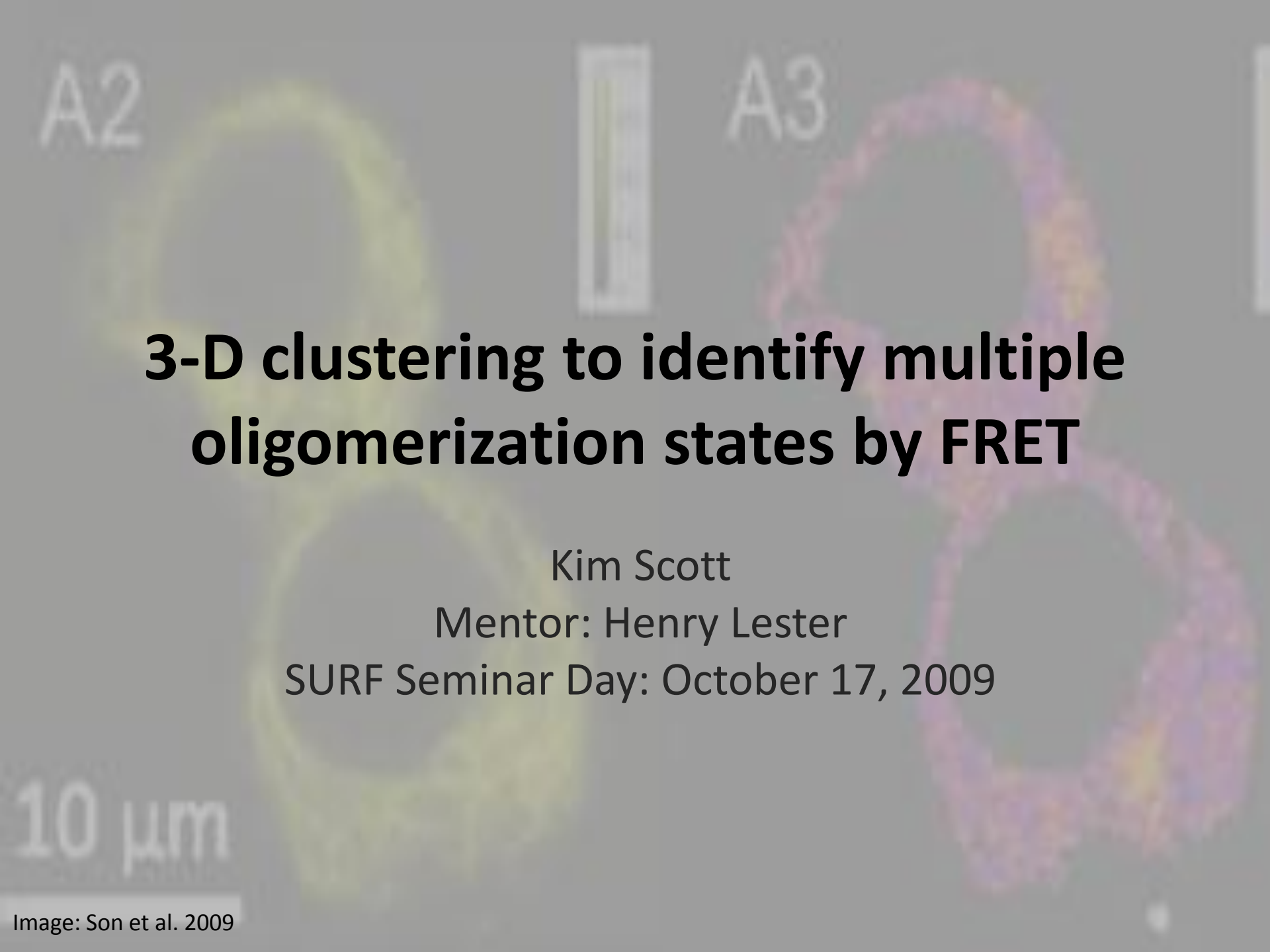
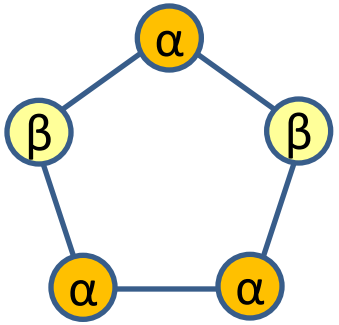
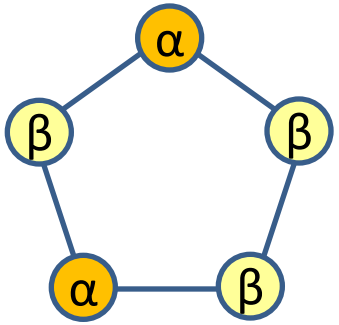


Image: Son et al. 2009

Nicotinic Acetylcholine Receptors (nAChRs)

$(\alpha4)_3(\beta2)_2$:
EC50 ~100µM

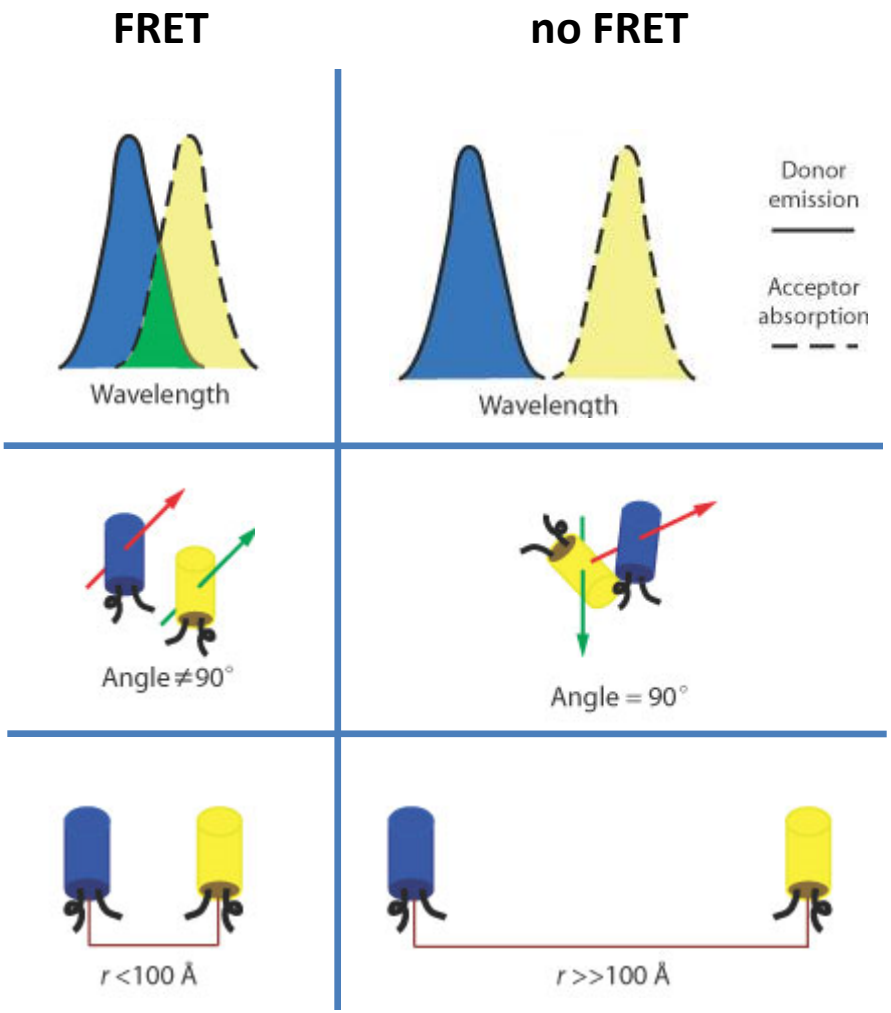




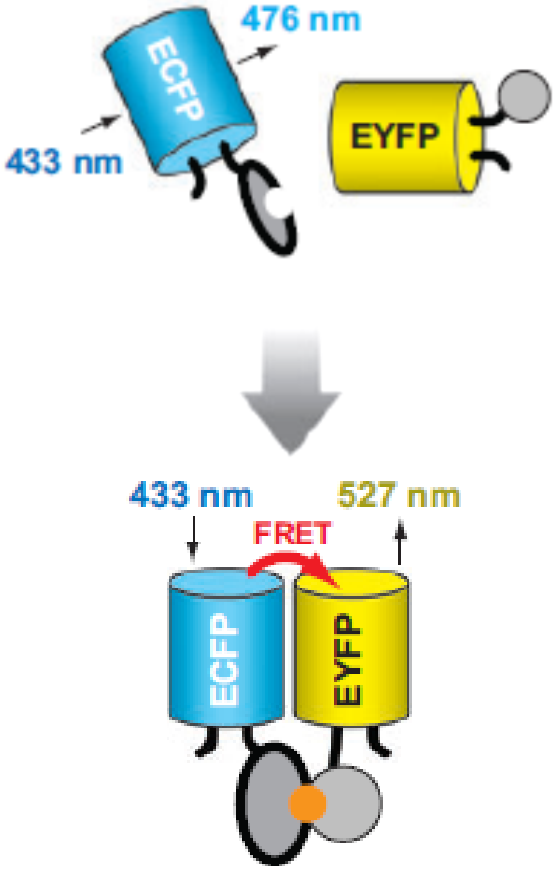
$(\alpha4)_2(\beta2)_3$:
100 times as sensitive!

- Pentameric ion channels found throughout the brain
- Composed of a variety of possible subunits in varying stoichiometries
- Open in response to acetylcholine (naturally), nicotine (much stronger!)
- Presumed to underlie the mechanisms of nicotine addiction, tolerance, & withdrawal… plus its protective effect against Parkinson's

FRET microscopy

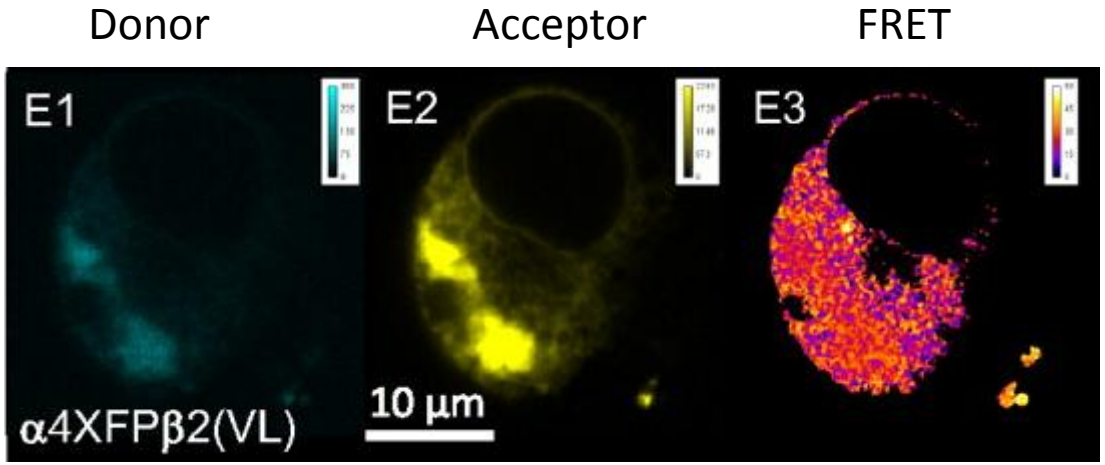


(Vogel et al., 2006)



(Wang et al., 2008)

Extending FRET to study stoichiometry



(Son et al., 2009)

Goal: to estimate the prevalence of distinct nAChR stoichiometries from the distributions of donor, acceptor, and net FRET pixel values.

Challenges:

- Multiple stoichiometries of assembled receptors, partially assembled receptors of unknown geometry, and unpaired donors and acceptors all present.
- Heterogeneous population even within single pixels
- Unknown subcellular localization of FRETing oligomers

Current NFRET histogram analysis

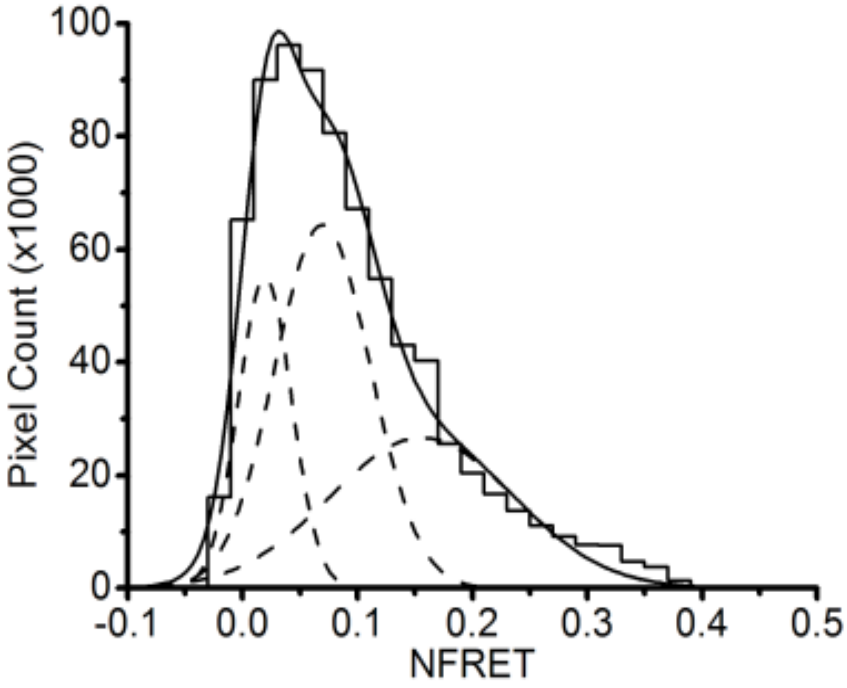
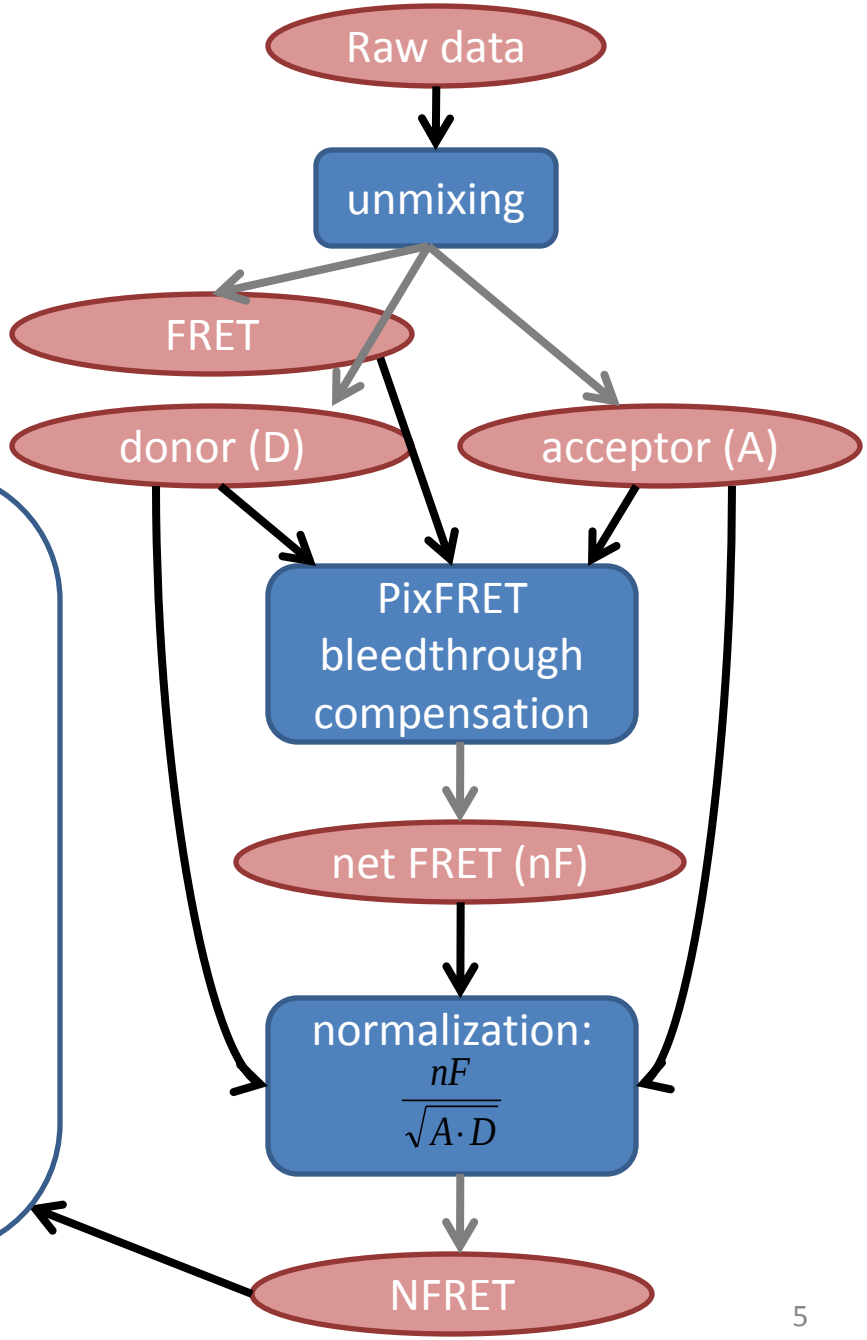


Fig 8H, Moss et al., submitted to JGP



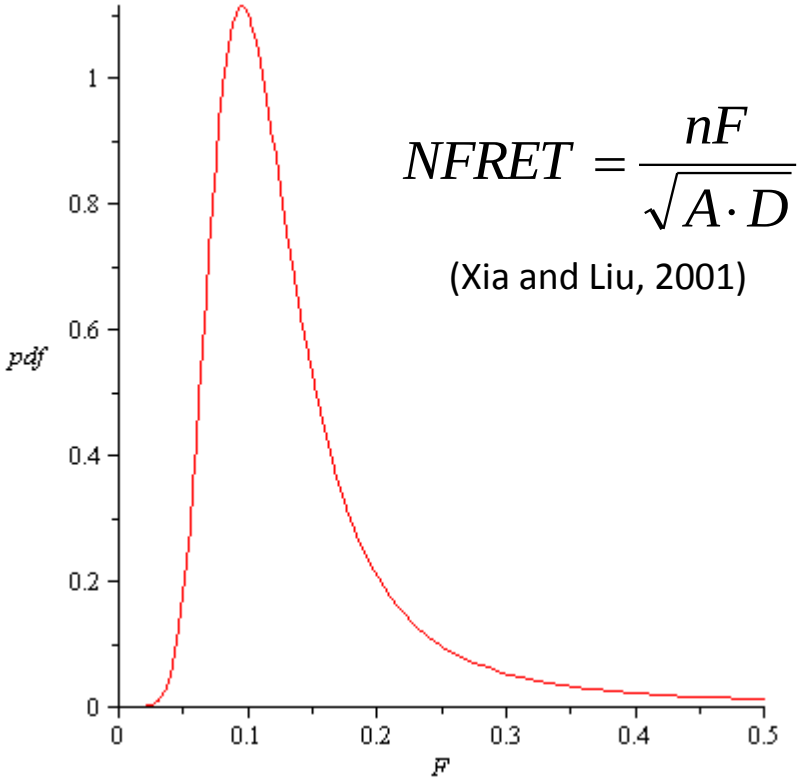
Dangers of fitting NFRET histograms

Single oligomer NFRET values with

nF ~ N(1, 0.2)

A ~ N(10, 4)

D ~ N(10, 4)



$$NFRET = \frac{nF}{\sqrt{A \cdot D}}$$

(Xia and Liu, 2001)

1) NFRET distribution from a single oligomer with varying nF, A, and D measurements is skew.

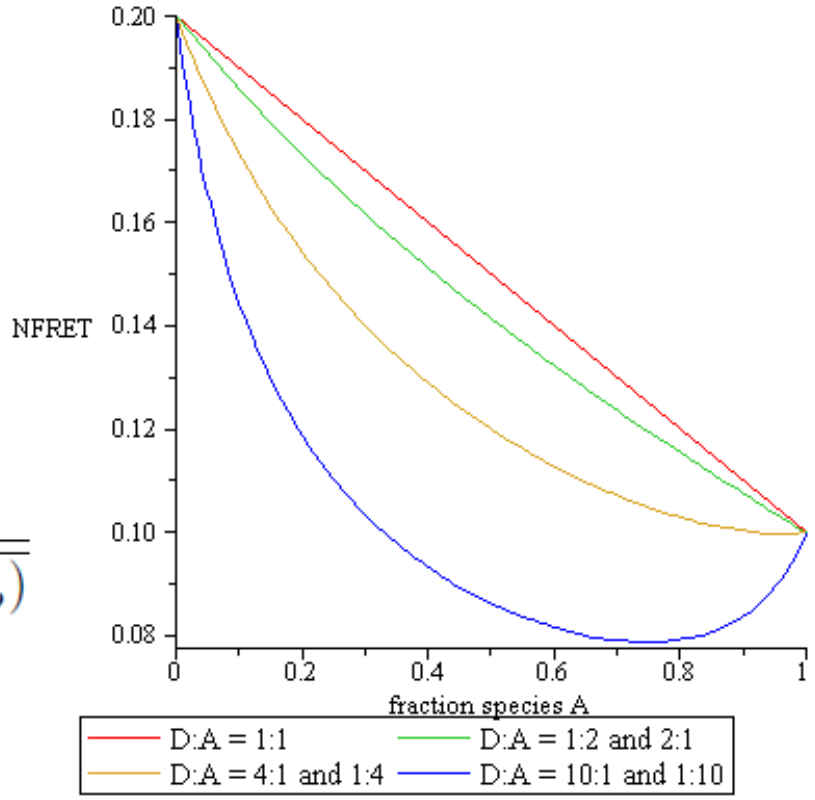
Dangers of fitting NFRET histograms

$$NFRET = \frac{nF}{\sqrt{A \cdot D}}$$

(Xia and Liu, 2001)

Given that a fraction *f* of the total FRETing constructs are of type A, the NFRET value

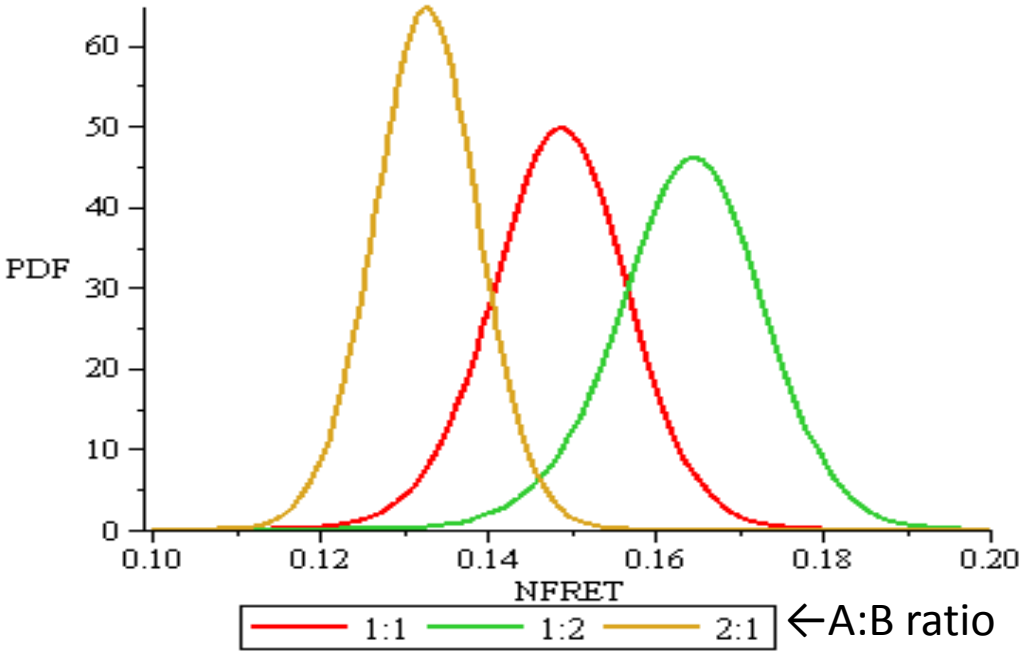
$$T(f) = \frac{f\,nF_a + (1-f)\,nF_b}{\sqrt{(f\,A_a + (1-f)\,A_b)\,(f\,D_a + (1-f)\,D_b)}}$$



2) NFRET from multiple species combines nonlinearly (sometimes non-monotonically)

Dangers of fitting NFRET histograms

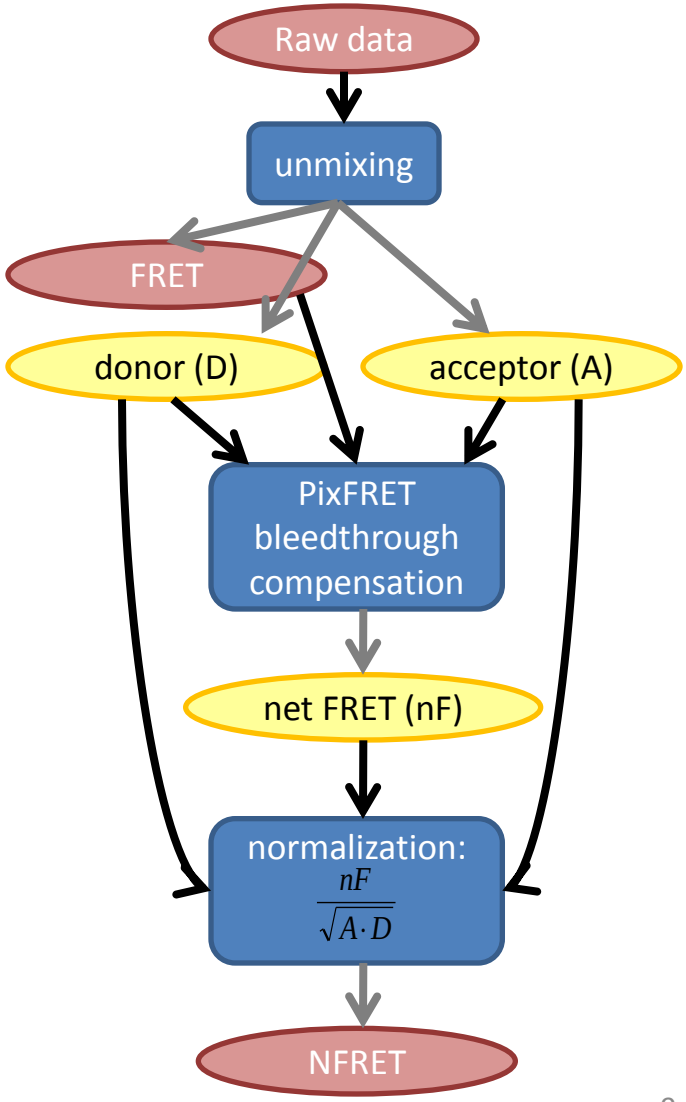
Two independently normally-distributed species with fixed nF, A, and D values per oligomer:



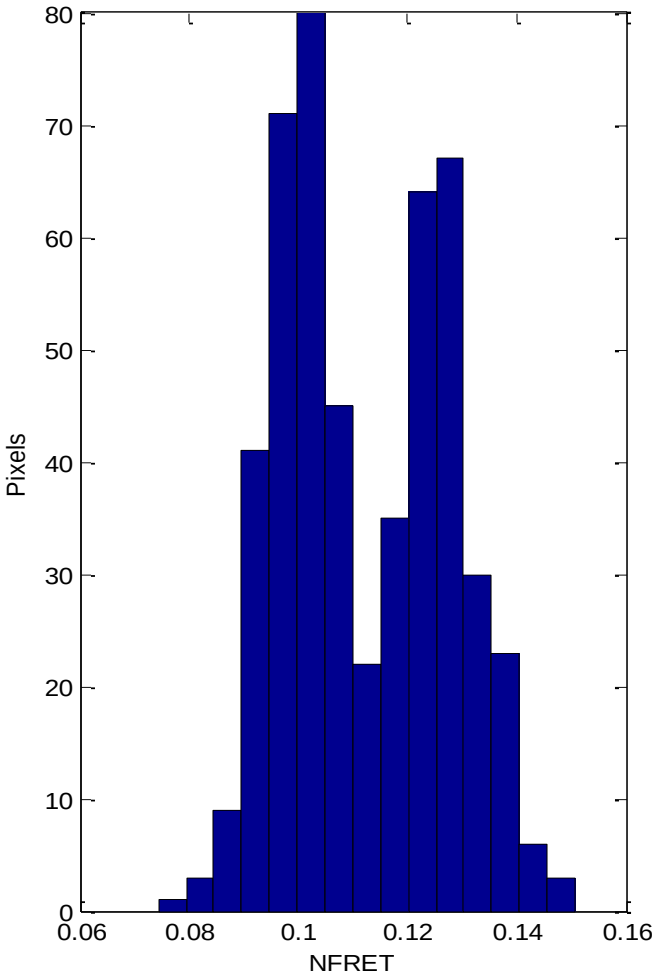
3) Even “ideal” situations (with no variation in nF, A, and D) give skew distributions of NFRET values.

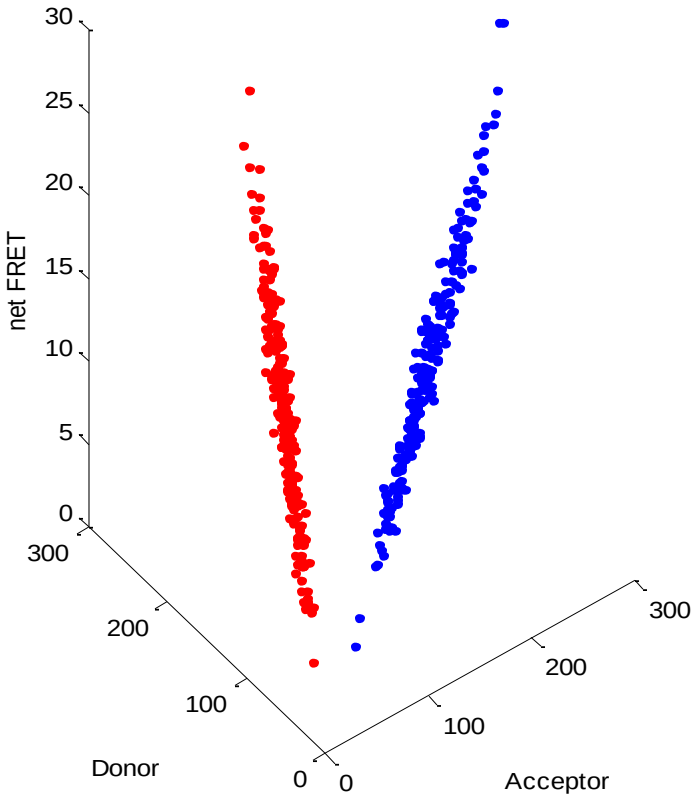
The case for direct clustering instead

- Why collapse 3D information to 1D unnecessarily?
- Clustering automatically assigns pixels to populations.
- Deals more readily with unpaired fluorescence.



Two pure population model

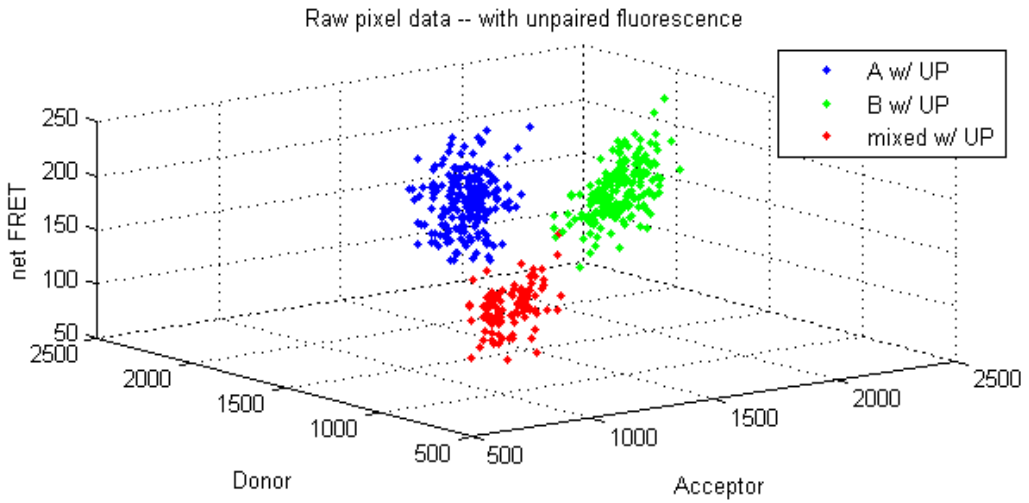


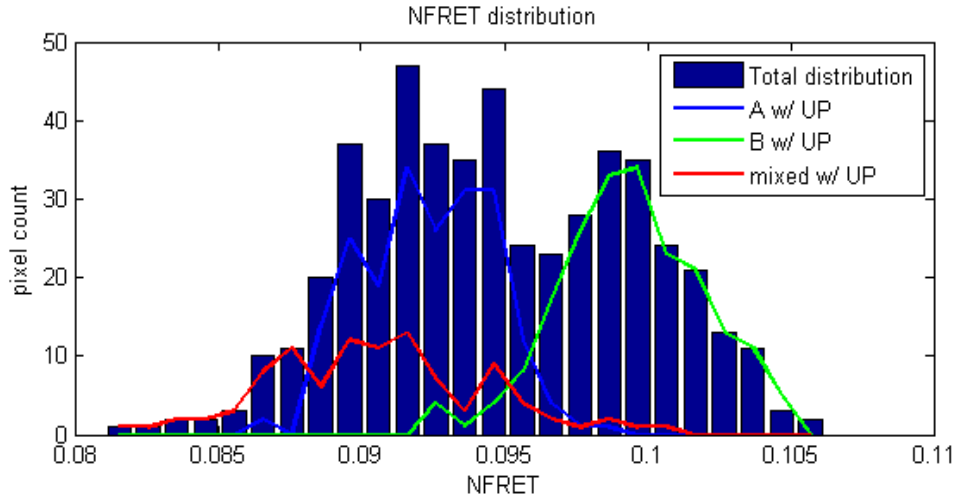


Species A, mean NFRET 0.1: nF ~ N(1, 0.2) A ~ N(7, 0.7) D ~ N(14, 1.4)

Species B, mean NFRET 0.125: nF ~ N(1.25, .25) A ~ N(14, 1.4) D ~ N(7, 0.7)

Segments with unpaired donor/acceptor





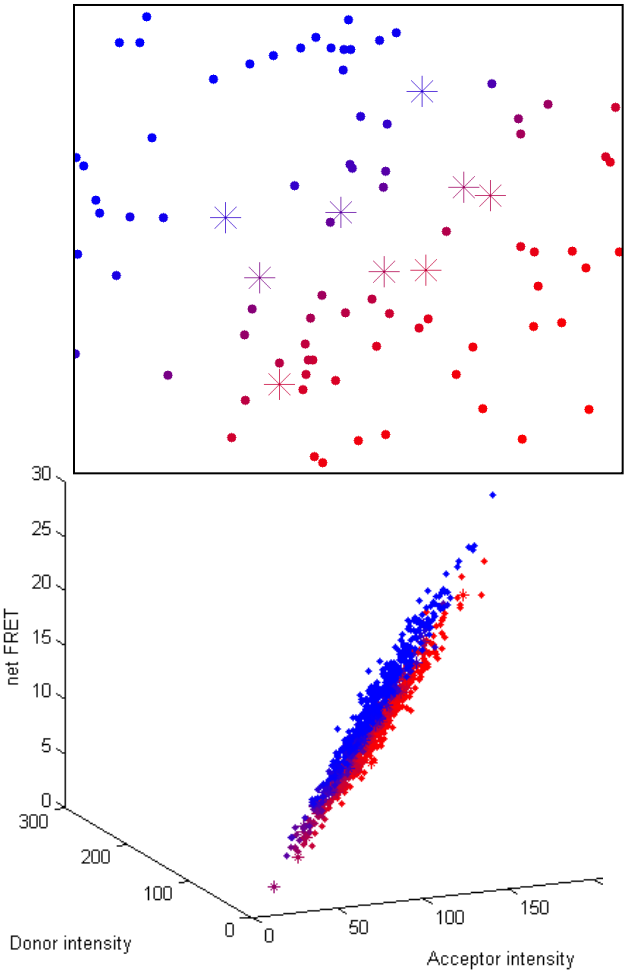
- Same species A and B, concentrations [100 50 50] and [50 100 50]
- Total unpaired concentration [25 25 25]
- Unpaired fluorophores have same properties as lesser of donor & acceptor in FRETing species

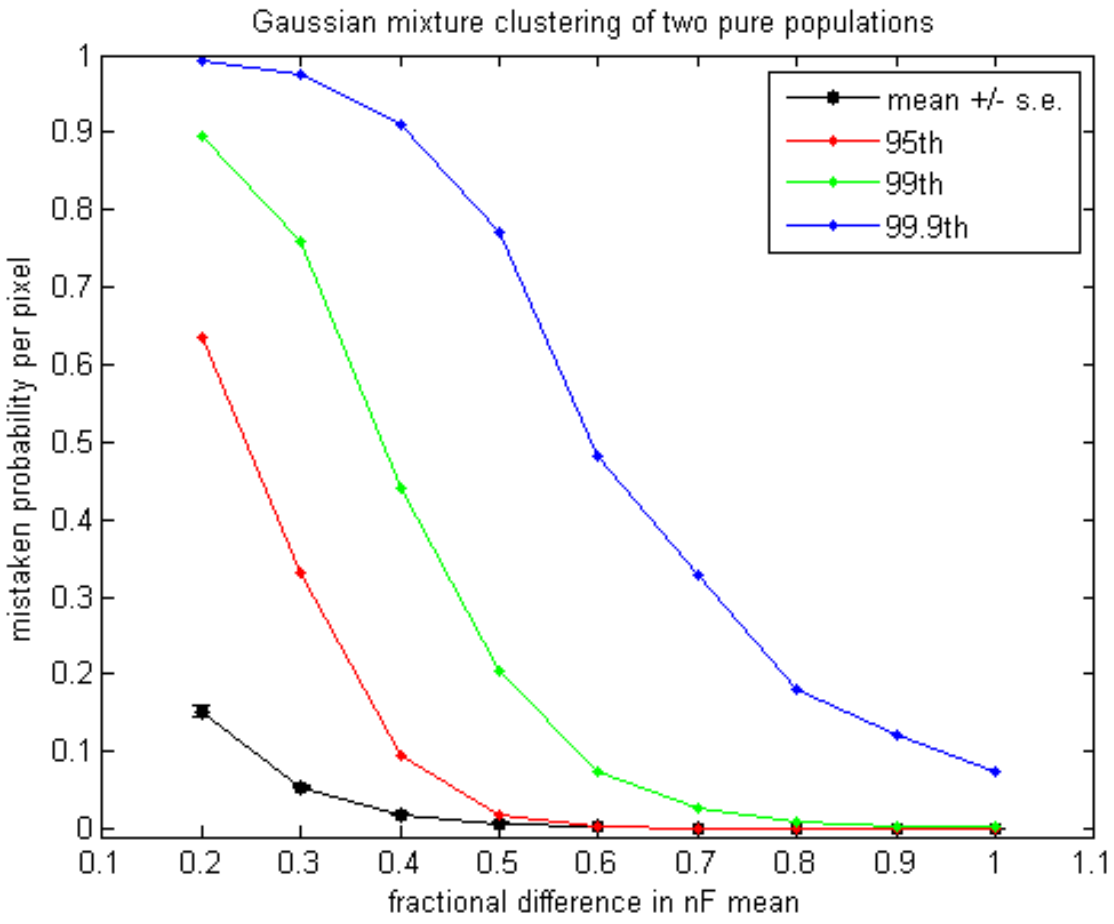
Choice of clustering algorithm

- Projective k-means
  - Clusters points along lines (representing varying concentrations of a single ratio of species, plus unpaired fluorescence)
  - Doesn't split high- and low-concentration regions
- **Gaussian mixture (GM) model**
  - Fits points to a set of Gaussian clusters
  - Doesn't ignore concentration
  - May be more robust to impure "segmentation"

Both easily extended to probabilistic clustering.

Performance of GM clustering





25 images each, 2000 pixels per image. 20% uncertainty in nF, 10% in A and D
Average concentration 10 oligomers (small) in both populations.

Performance of GM clustering

- **Accurately and reproducibly** clusters pixels from pure-population and segmented models, even with unpaired fluorescence
- Consistently **identifies the number of clusters** using Bayesian information criterion (introduces a parameter penalty to avoid overfitting)

Next steps

- Next focus is on clustering real data from two experiments: with three and one putative populations of nAChRs
- Use of membrane-specific and non-FRETing distributions to calibrate expected clusters
- Modeling varied transfection ratios and matching clusters across cells

Acknowledgments

Henry Lester

Fraser Moss

Rigo Pantoja

Rahul Srinivasan

Crystal Dilworth

Lester lab

Amgen Foundation

Caltech SFP office