



# BMI219 - Spring 2012

## **Statistical Methods for Array and Sequence Data**

CBMB Faculty

Course director: Dr. Mark Segal

Center for Bioinformatics & Molecular Biostatistics

UCSF Division of Biostatistics

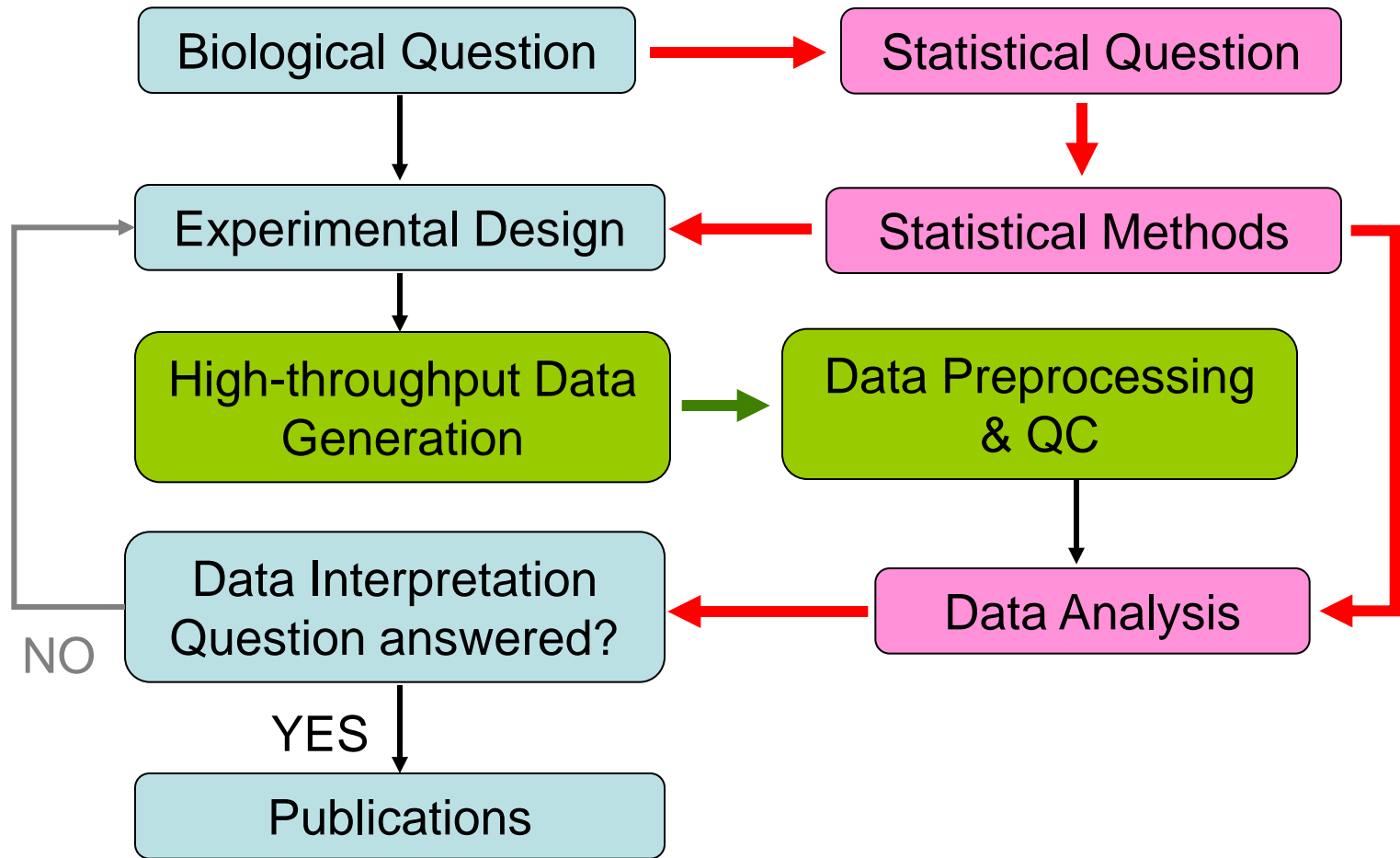
5/7/12 – 5/25/12

2-4pm Genentech Hall S227

# Announcements

- ❑ Lectures: 5/7/12 – 5/25/12 @ 2-4pm; Genentech Hall S227
- ❑ Website:
  - Short: <http://tiny.ucsf.edu/bmi219>
  - Long: <http://www.epibiostat.ucsf.edu/biostat/cbmb/bmi219/>
- ❑ Handouts: On website
- ❑ Course requirement: Attendance
- ❑ Please fill out your contact details (online form)

# Course Aims



# Statistical Issues

## ❑ **Experimental design:**

- Which array? How many samples? Which to co-hybridize?

## ❑ **Data preprocessing:**

- Quality assessment, image analysis, normalization

## ❑ **Combining replicates for differential expression:**

- Effect estimate, statistical significance, multiple testing

## ❑ **Annotation:**

- Shared characteristics (functional groups, motifs) among DE genes?

## ❑ **Data archive:**

- Data submission to public database

# Statistical Issues

## ***Technology dependent:***

- Preprocessing & QC

## ***Biological question dependent:***

- Hypothesis testing & Multiplicity
- Classification, prediction
- Clustering

## ***Both technology & biological question dependent:***

- Experimental design
- Meta-data integration

# Course Roadmap – Week 1

- ❑ May 7 (Lecture 1): [Dr. Henrik Bengtsson](#)  
Overview of statistical issues in bioinformatics.  
Introduction to microarray analysis.
- ❑ May 8 (Lecture 2): [Dr. Henrik Bengtsson](#)  
SNP array processing and copy-number analysis.
- ❑ May 9 (Lecture 3): [Dr. Katie Pollard](#)  
Clustering and multiple testing procedures.
- ❑ May 10 (Lecture 4): [Dr. Jun Song](#)  
Analysis of ChIP-Chip data.
- ❑ May 11 (Lecture 5): [Dr. Tom Hoffman](#)  
Genome wide association studies.

# Course Roadmap – Week 2

- ❑ May 14 (Lecture 6): [Dr. Mark Segal](#)  
Regression and survival analysis with array-based prediction.
- ❑ May 15 (Lecture 7): [Dr. Barry Taylor](#)  
Introduction to next generation sequencing.
- ❑ May 16 (Lecture 8): [Dr. Adam Olshen](#)  
Next generation sequencing - statistics.
- ❑ May 17 (Lectures 9): [Dr. Jun Song](#)  
Analysis of ChIP-Seq data.
- ❑ May 18 (Lecture 10): [Dr. Jun Song](#)  
Analysis of ChIP-Seq data.

# Course Roadmap – Week 3

- ❑ May 21 (Lecture 11): [Dr. Sean Thomas](#)  
Epigenomics.
- ❑ May 22 (Lecture 12): [Dr. Alisha Holloway](#)  
Analysis of RNA-Seq data.
- ❑ May 23 (Lecture 13): Drs. [Sean Thomas](#) & [Alisha Holloway](#)  
Emerging technologies.
- ❑ May 24 (Lecture 14): [Dr. Saunak Sen](#)  
High-throughput genetic screening.
- ❑ May 25 (Lecture 15): [Dr. Mark Segal](#)  
Sequence analysis case studies.