CM581A2: NEXT GENERATION SEQUENCING PLATFORMS AND LIBRARY GENERATION

Fall 2015

Instructors:
Coordinator: Carol Wilusz, Associate Professor MIP, CMB
Instructor: Dan Sloan, Assistant Professor, Biology, CMB
Instructor: Mark Stenglein, Assistant Professor, MIP, CMB

Required Texts and Materials:
No required texts. Students will require access to LabArchives Electronic Lab Notebook which can be purchased for $15 per semester.

Description:
Graduate students from both biological and computational/mathematical fields will be trained in the theoretical and experimental aspects of next generation sequencing experiments with a focus on the Illumina platform.

Objectives:
Students will be able to:
• Tell the history of DNA sequencing, highlighting the technological advances that have allowed rapid, large-scale sequencing of genomes and transcriptomes.
• Describe the differences between Sanger sequencing, sequencing-by-synthesis (e.g. Illumina) and direct sequencing (e.g. PacBio, Oxford Nanopore)
• Describe in detail the steps to generate a library of DNA fragments for sequencing starting from isolation of genomic DNA including quality control, fragmentation, adapter ligation and amplification
• Highlight how preparation of libraries from total RNA differs from the protocols used for genomic DNA
• Explain how and why rRNA is removed from total RNA prior to RNA-seq
• Explain how and why strand-specific RNA sequencing is employed
• Tell how enrichment strategies and barcoding work, and how they can reduce the costs of DNA sequencing and analysis
• Give details on the steps of an Illumina sequencing reaction and tell how fluorescent output is generated and converted to sequence information
• List multiple applications of NGS sequencing and how library preparation should be modified for each
• Perform library preparation and submit samples for Illumina sequencing
• Troubleshoot common problems arising during library preparation and sequencing

Pre-requisites:
Life sciences BS degree with solid background in molecular biology, or CM000 Introduction to Nucleic Acids (experimental course) or equivalent.

Credit Hours: 1 Credit

Class Breakdown:
This is a 1CR lab class offered over 4 weeks with 8 hr per week of contact time. This course is designed to require an average of 4 hours per week of outside work (preparation, problem sheets, reading, online lectures, quizzes etc). Students with weaker biology backgrounds may need to spend more time preparing for class and/or completing homework.

Teaching Strategies:
Labs, presentations on theory behind experiments, quizzes, problem sets, online lectures, practice through problem solving.

Evaluation:

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laboratory Notebook</td>
<td>50%</td>
</tr>
<tr>
<td>Online Quizzes</td>
<td>10%</td>
</tr>
<tr>
<td>Problem Sets</td>
<td>15%</td>
</tr>
<tr>
<td>Attendance/Participation</td>
<td>5%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>20%</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

Grading Scale:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>90-100</td>
</tr>
<tr>
<td>B</td>
<td>80-89</td>
</tr>
<tr>
<td>C</td>
<td>70-79</td>
</tr>
<tr>
<td>D</td>
<td>60-69</td>
</tr>
<tr>
<td>F</td>
<td>59 &amp; below</td>
</tr>
</tbody>
</table>

This class does not issue grades of + or –
WEEK 1

Day 1

PRESENTATION:  Introduction to CM581A2 (30 min)
Overview, objectives, expectations, grading, academic integrity
Introduction to LabArchives Electronic Lab Notebook (20 min)
Introduction to NGS (50 min)
History of DNA Sequencing
  Sanger Sequencing (Gel, Capillary)
  Transition to NGS
Applications and Limitations of NGS
  WGS, RNA-Seq, CHIP-Seq, Metagenomics
Brief overview of Illumina Sequencing-by-Synthesis including library generation

LAB:  Isolation of genomic DNA from a prokaryote
Lysis, centrifugation, RNAse treatment, column purification (~ 1hr 40 min)

HOMEWORK:  LabArchives Set-Up (30 min)
Isolation and Quality Control of Genomic DNA (1 hr video lecture online)
  DNA isolation from prokaryotic, eukaryotic cells
    Lysis
    RNA removal
    Organic Extraction & Precipitation
  Quantification & Quality Control:
    Fluorometric
    qPCR
    Nanodrop (limitations)
Quiz 1 (30 min)

Day 2

LAB:  Quality control and quantification of genomic DNA (1 hr)
  Fluorometric quantification
  Nanodrop
DNA fragmentation and Quality Control (2hr)
  Covaris, Tapestation

PRESENTATION:  Library Generation for Illumina Genome Sequencing (1hr)
  Part 1:  DNA isolation & fragmentation (30 min)
    Covaris, Enzymatic, Nextera Transposons
    Quality Control
    TapeStation, Bioanalyzer, qPCR, nanodrop, fluorometry
  Part 2:  Adaptors, Amplification (30 min)
    Size Selection
    SPRI-beads, Gels, PippinPrep
    Pre-run QC and Quantification
    Gels, TapeStation, Bioanalyzer, qPCR, nanodrop, fluorometry
    Cloning and Sanger Sequencing

HOMEWORK:  LabArchives Recordkeeping (30 min)
Sequencing Platforms (1hr online video)
  Illumina Sequencing By Synthesis
    How it works
    Differences between MiSeq and HiSeq and NextSeq
  Other technologies – advantages and limitations
    Current technologies – Pacbio, Ion Torrent
WEEK 2

**Day 1**

**LAB:** Adapter Ligation, Amplification, Size Selection (2-3 hr)
- Nextera tagmentation reactions
- Post-Nextera cleanup
- Setup and run adapter build out PCRs

**PRESENTATION:** Optimizing your NGS Experiment (1 hr 40 min)
- Cost and Time Considerations
  - Library Prep
  - Sequencing
  - Analysis
- Making Mapping Reads Easier
  - Read length
  - Single end versus paired end reads
  - Mate pairs
- Controlling Coverage
  - Barcoding and Pooling Strategies
  - Moleculo
- Troubleshooting
  - Low complexity libraries
  - Read length
  - Errors
  - Amplification Bias
    - Polymerase choice
    - # of cycles
    - Avoiding amplification
  - Misidentification of barcodes
  - Double barcoding
  - Contamination

**HOMEWORK:** LabArchives Recordkeeping (30 min)
- Problem Set 1 (1 hr 30 min)

**Day 2**

**LAB:** Library size selection and pooling (3 hr)
- Size select libraries (by gel).
- Quantification of individual libraries for pooling (by qubit)
- Pool libraries. Cleanup pool.
- Additional PCR of final pool may be necessary depending on concentration...

**PRESENTATION:** RNA-Seq (50 min)
- RNA isolation and quality control
- rRNA depletion
  - poly(A) selection
  - Ribozero
- Fragmentation
  - Physical, Enzymatic, Chemical, Nextera
- Reverse transcription
- Adapter Ligation, Amplification, Size Selection, QC (same as for DNA, so brief)
Strand-specific RNA seq

**PRACTICE:**  
Problem Set 1 (50 min)

**HOMEWORK:**  
Quiz 3 (30 min)  
Problem Set 1 (30 min)  
Reading: (1hr)

PubMed PMID: 20111037.

---

**WEEK 3**

**Day 1**

**LAB:**  
Quality Control and quantification of final Libraries (2 hr)  
Kapa quant of final pooled library  
Gel of final pooled library (?)  
Preparation of sample sheet for MiSeq

**PRESENTATION:**  
Target Enrichment (1hr)  
Enrichment vs subtraction  
Enrichment methods  
Capture-based enrichment  
Exome sequencing  
[polyA capture]  
PCR-based enrichment (amplicon sequencing)  
16S rRNA seq and the microbiome  
Affinity-based enrichment  
RIP-Seq/ChIP-Seq  
Ribosome profiling  
Subtraction methods  
[RiboZero]  
RNaseH-mediated depletion  
DSN-seq/hydroxyapatite (for normalization)

**PRACTICE:**  
Problem Set 2 (30 min)

**HOMEWORK:**  
LabArchives Record Keeping (30 min)  
Quiz 4 (30 min)  
Problem Set 2 (1hr)

**Day 2**

**LAB:**  
Illumina MiSeq Laboratory Visit (2hr)

**PRESENTATION:**  
Experimental Design and Validation of Results (50 min)  
Importance of planning ahead - defining the experimental question and analysis strategy  
Controls – yes, you need them! Examples  
Replicates - Biological versus technical replicates  
Validation of NGS results

**PRACTICE:**  
Problem Set 2 (50 min)

**HOMEWORK:**  
Reading:  
Problem Set 2: (30 min)
Quiz 5: (30 min)

WEEK 4

Day 1

PRESENTATION: Applications of NGS (1hr 40 min)
- Variant detection
- Structural variation detection
- Microbial (16S) metagenomics
- Shotgun metagenomics
- Genotyping-by-sequencing
- Tn-Seq
- Host/Pathogen sequencing
- 3C-Seq
- RAD-seq & reduced representation
- Small RNA-seq

PRACTICE: Practice Exam Questions (1hr 40 min)

HOMEWORK: Prepare for Final Exam
- 1hr review lecture and quiz provided online

Day 2

FINAL EXAM: Open Book Final Examination (3hr)

Course Survey: (20 min)