A Snap Shot into the Life Cycle of Sequencing in Genomics Core Facilities

DNA Sequencing Research Group (DSRG)
Joanna Hamilton
Geisel School of Medicine at Dartmouth
Norris Cotton Cancer Center
DSRG Committee

• Robert Lyons (Univ. of Michigan)
• Stuart Levine (MIT)
• Charles Nicolet (USC)
• Savita Shanker (Univ. of Florida)
• Jyothi Thimmapurim (Purdue Univ)
• Chris Wright (Univ. of Illinois)
• Marie Adams (Univ. of Wisconsin)

• Tim Hunter (UVM), EB Liason

Thanks to those of you who completed the survey
Goals of the Survey

• Identify the widespread use of available sequencing technologies
• Which applications are most frequently employed?
• Begin to identify trends in instrumentation
• Help core facilities predict genomic trends
DSRG Survey Design

• Survey distributed by ABRF email blast (past and present members) and the ABRF list-serve to all member emails

• Survey sections
  – Sanger
  – NGS
  – Bioinformatics
  – Core management

• Within 24 hours we had ~80 responses
159 Responses from around the world
The majority of **Institutions** offer Sanger, NGS and Bioinformatics.

- **Sanger**: 1.9%
- **Sanger + Bioinformatics**: 3.1%
- **Sanger + NextGen**: 7.5%
- **NextGen**: 3.8%
- **NextGen + Bioinformatics**: 14.5%
- **Bioinformatics**: 1.3%
- **No Core**: 1.9%
- **Sanger+NextGen+Bioinformatics**: 66%
Sanger sequencing Section

• ~80% of responses still offer Sanger at their institutions, either alone or in combinations with other services

• Given the increase in NGS we wanted to determine if Sanger sequencing demand was decreasing
Does your Core Facility offer Sanger Sequencing?

- Business is Decreasing, 33%
- Business is Stable, 48%
- Business is Increasing, 7%
- Discontinued, 12%

~20% of responses did not offer Sanger Sequencing, or it was offered in a separate core.

Note: Business increasing from one respondent was due to low previous year.
How Many Sanger Reactions do you run per Year?
The Percentage of Successful Sanger Sequencing Reactions

60-90% of cores re-run samples for free
Next Generation Sequencing Section

- Platform distribution across cores
- Instrument Usage
- Frequency of NGS applications
- Technical Failure Rates
The **Number of Instruments** from each Platform in operation in Core Facilities

- **Illumina** – 66%
- **Ion Torrent** – 23%
- **Others** – 11%

80% of those with a GAIIx also have another Illumina instrument
What is the percentage of Cores with each Sequencing platform?

Sequencing Platforms in NGS cores

- illumina
- LifeTech
- Roche
- Solid
- PacBio
- Other

Illumina platform distribution

- MiSeq
- 1500/2500
- NextSeq
- 1000/2000
- GA
- 1T
- X10

Percentage of NGS cores

Percentage of Illumina labs
Do core facilities have multiple instruments of the same platform?
Several core facilities offer multiple sequencing platforms

*Illumina*

*LifeTech*

*Other*

*SOLiD, PacBio, and Roche*
Most “retired” instrument is the GAIIx.
NextSeq is the most likely to be purchased in the next 12 months

<table>
<thead>
<tr>
<th>Product</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illumina HiSeq 1500/2500</td>
<td>15%</td>
</tr>
<tr>
<td>Illumina MiSeq</td>
<td>5%</td>
</tr>
<tr>
<td>Illumina NextSeq</td>
<td>20%</td>
</tr>
<tr>
<td>Illumina 1T</td>
<td>15%</td>
</tr>
<tr>
<td>Illumina X10</td>
<td>10%</td>
</tr>
<tr>
<td>Ion Torrent Proton</td>
<td>5%</td>
</tr>
<tr>
<td>PacBio</td>
<td>20%</td>
</tr>
<tr>
<td>OTHERS??</td>
<td>5%</td>
</tr>
</tbody>
</table>

Which of the following 2014 releases are you most likely to purchase?

- Illumina NextSeq
- Oxford Nanopore MiniION
- Illumina X10
- Qiagen Gene Reader
- None

Source: Mizuho Securities USA and GenomeWeb survey. (n) = 55.

n=33 (yes and maybe)
What Percentage Operating Capacity do you run your currently owned sequencers?
How often do you experience technical failures?

- **Illumina (n=138)**
- **Ion Torrent (n=35)**
- **Others (n=19)**

<table>
<thead>
<tr>
<th>Operating Capacity</th>
<th>Percentage of responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;10%</td>
<td>70%</td>
</tr>
<tr>
<td>10-30%</td>
<td>60%</td>
</tr>
<tr>
<td>30-70%</td>
<td>50%</td>
</tr>
<tr>
<td>&gt;70%</td>
<td>40%</td>
</tr>
</tbody>
</table>
Library Preparation questions

• Which applications are performed most frequently
• Which applications are you most likely to add to your list of services for investigators
• Upstream sample prep methods
• Library QC
RNA-seq is the most often performed application for NGS.
Whole genome single cell and targeted RNA-seq are the applications most likely to be added to core services.
Does your core accept user generated library preparations?

- 9% of respondents do not accept user generated library preps
- 67% of cores accept user generated libraries
- 24% do so on a project by project basis

- QC of libraries is predominantly by Qubit, Bioanalyzer and qPCR

- Upstream of library preparation, only ~30% of cores offer DNA or RNA isolations, and that is ~20% for FFPE samples
Bioinformatics Overview

• Is analysis offered in the same core as the sequencing?
• How long is data stored for?
• Are additional services available, and are they included in the sequencing costs?
• How is data delivered to the customer?
Bioinformatics: Data Analysis (n=120)

- How is bioinformatics offered?
  - Same core: 44%
  - Different core: 43%
  - No informatics offered: 8%
  - No response: 5%

16% of respondents offered access to software tools, mostly (68%) in places where analysis was done in the core (do it yourself emphasis?)
When bioinformatics is offered in the same core.....

For same-core analysis (44%), majority offer secondary analysis

**Secondary analysis, alignments BAM file generation**

**Tertiary analysis, PCA, extensive DGE, assemblies, variant identification and interpretations**

**Other, File formatting, manipulations, GEO submission**

Most bioinformatic analysis is an add on cost
Bioinformatics: Data Storage Policies

High rate of non-response for these questions (30%)
Most cores provide short term storage as part of service
Many cores provided archival and long term storage as part of service
Many mechanisms exploited, majority of cores use multiple strategies

Email!!???  Must be Sanger contamination!
Core Management

• Are core facilities using a LIMS?
• Do cores charge for consultation?
• What problems do cores face, with samples, instruments etc.
Does your core use a Laboratory Information Management System (LIMS)?

- Yes: Home made (47%)
- Yes: Commercial (24%)
- No: 31%

Majority of responses of commercial systems used:
- ✓ Clarity
- ✓ iLAB
The majority of cores do not charge for consultations.

- Yes, we charge for all consultations – 1%
- Some – Initial consultation is free and then additional charges apply – 17%
- No – All consultations are free – 82%

Per hour charges ranged from $40-$100 per hour.
The majority of core financial support comes from customer charge backs.
97% of NGS sequencers are under vendor service contracts

Does the institution provide financial support for the service contracts?

We did not ask if these were the original contracts with the instrument purchase or a renewal...
A variety of problems are encountered in core facilities.
Conclusions

• Sanger sequencing business remains stable for ~50% of cores, with 12% no longer offering Sanger
• 67% of core facilities offer Illumina Sequencing, and 27% offer Ion Torrent
• RNA-seq is the most used application
• Watch out for single cell sequencing...
• Bioinformatics services available within ~40% of core, with another ~40% offering a separate core at their institution
Next DSRG Project

• Comparison of Enzymatic DNA Fragmentation library kits (Non-covaris)

• Several kits, some new to the market
  – Nextera XT
  – Kapa
  – Agilent
  – NEB
  – Homegrown

• Compare and contrast these different systems
  – ease of use
  – reproducibility

• Sequencing at a central core

We are looking for new members and study participants

Nextera chemistry simultaneously fragments and tags DNA in a single step. A simple PCR amplification then appends sequencing adapters and sample indexes to each fragment.
How do you QC libraries generated by your core or customer labs?
Which applications do you offer in your core facility?