

Challenges for smaller groups

The Donnelly Sequencing Center

Corey Nislow, Director

FGED, Toronto, Oct. 22, 2012



10 Rules for a successful sequencing facility

~~10 Rules for a successful sequencing facility~~

Lessons learned after 2 years in the trenches

- Equipment
- Vendor relationships
- DEMOs
- Scope of assays
- QC and SOPs
- Defining relationships, expectations
 - Collaborators
 - Staff
- Snapshot 2010 and 2012
- Storage

Hardware: You need more than a sequencer

- Get a room of your own- electrical and environmental specs
- Get 2 rooms- 1 wet, 1 dry
- FAS gets a key
- Accessory equipment- don't borrow anything
- Basic stuff- molecular bio lab
- Quantitation, concentration, visualization
- The latest greatest accessory.....

Vendor relationships

- It costs nothing to be nice and can be very expensive to be a jerk.
- Check in regularly, i.e even when things are working
- The hierarchy:
- Rep> Regional Manager> North American Honcho
- FAS
- Beta tests and reference sites

Be a DEMO artist



- Do you have the bandwidth for the DEMO?
- 25% chance that a device you test will be essential
- 75% chance it's a luxury or worse
- Invite other users in the area- good for networking

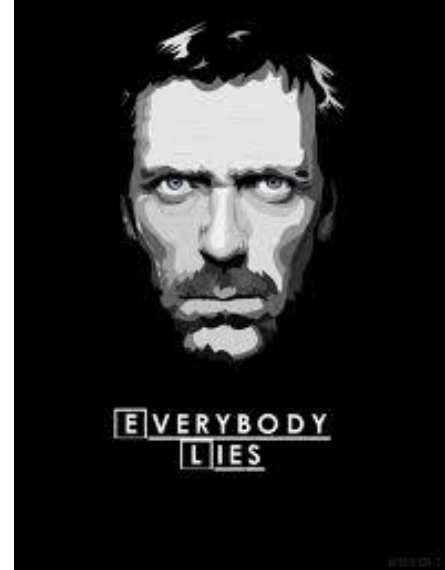
Offer every possible assay



- Educate your clients (reprogramming)
- Understand the experimental question
- Do they really need to sequence?
 - Don't underestimate Sanger QC
- Treat every new client similarly in the beginning. Experienced users will understand.

Do not QC clients samples

- Everybody lies about their samples and their timing
- Your notebook is your best defense
- e.g. DNase to “prevent cells from sticking”
- No gel, no quibit, no deal or no guarantees
- Connect with a full-service commercial shop for backup and as a foil



Every client is your friend. Your current collaborators are your special friends



- Defend your pricing structure
- Someone has to pay for the reagents
- And the time
- Work this out ahead of time

Overpromise and under deliver

- Don't betray your own calendar. Your staff will hate you. And the client will hate you. And your mom.

The DSC: a case study

- Doors open: Jan 2010
- Validation Run: April 21
- 4 months is a good rule-of-thumb for instrument delivery and ability to provide client data

Donnelly Sequencing Centre Stats

21-Apr-10	GAIIx validation run
10	Number of runs to-date
Types of runs to-date	UofT Lab
whole human genome	Dr. T. Hughes
whole human exome	Dr. C. Nislow
whole human transcriptome	Dr. Z. Zhang
whole <i>C. elegans</i> transcriptome	Dr. B. Blencowe
short-hairpin RNA	Dr. J. Moffat, <i>publication pending</i>
ChiP-Seq	Dr. G. Brown
<i>E.coli</i> barcodes	Dr. A. Emili
yeast barcodes	Dr. C. Nislow
whole yeast genome re-sequencing	Dr. C. Nislow
whole yeast genome <i>de novo</i> assembly	Dr. C. Nislow
whole yeast nucleosome	Dr. C. Nislow
Mate-paired libraries	<i>In progress</i>
directional mRNA libraries	<i>In progress</i>

Collaborations

Current internal collaborations: 12	Current external collaborations: 4
Dr. B. Andrews	Dr. Bjorn R. Olsen, Harvard University
Dr. C. Boone	Dr. Fritz Roth, Harvard University
Dr. G. Brown	Agilent Technologies Inc.
Dr. A. Emili	FlexGen B.V.
Dr. I. Ensminger	
Dr. A. Fraser	
Dr. J. Greenblatt	
Dr. T. Hughes	
Dr. J. Moffat	
Dr. C. Nislow	
Dr. D. Siddhu	
Dr. Z. Zhang	

The DSC: after 19 months

- Gone
 - GAllx
 - PGM
- Going
 - Caliper XT
 - Caliber Lab Chip
- New
 - Bioanalyzer
 - MiSeq (custodian)
 - 2500
- 21 GAllx runs
- 44 HiSeq runs
- 36 MiSeq runs

Private sector:

Mitomics, Ontario Public Health, Vertex Pharmaceuticals

North America:

SUNY Buffalo, Virginia State University

Ontario Community:

University of Toronto, McMaster University, Mt. Sinai, Toronto,
MaRS, UHN

International:

Jozef Stefan Institute, Riken Institute, Universitat d'Alacante

Assay Types:

Whole genomes, Total RNA, mRNA, ChIP-seq, Nuc-seq, Human
Exomes, Small RNAs, Diverse counting apps

Custom assays (DSN, BarSeq, integration site identification)

DSC Server Specs

DELL PowerEdge 910, 64Gb RAM, 48 Cores

DELL MD3200/1200 Storage, 60 Terabytes RAID5

Integrated w/CCBR Unified network

Primary tasks (storage, processing, distribution of seq. data)

1. Offloads data from the HISEQ
2. Off-line basecalling (Bustard) and/or basecall conversion
3. Data Distribution
4. Data Archiving
5. Limited downstream Analysis (Alignment, Counting etc)

Most analysis on secondary servers and the Banting cluster.

And how would you like your data?

- Oh I'm not picky filtered reads is fine
 - Well actually... could you map them?
 - To my own reference?
 - And find the SNPS? And CNV's
- From exome data?

Questions from the Organizers

- What recommendations do you have for better ways to share genomic data and ensuring reproducible research?
- What challenges have you met and/or steps have you taken with creating or using standards?

Thanks

- Dax Torti, Ph.D.
- Danica Leoung
- Larry Heisler, Ph.D.

- Brenda Andrews
- Jack Greenblatt

