Assay Automation and Quantitation -From Benchtop to 475



Jennifer Smith, PhD Caroline Shamu, PhD



How can laboratory automation improve my experiments?

How many experimental samples am I working with at a time?

Conditions

Cell lines

Dilution series

What number of replicates are needed for data analysis?

Technical

Biological

Which steps in the experiment contribute most to variability?

Each additional step adds variability

What steps are the slowest?

Reagent preparation

Conducting the experiment

Data acquisition

Data analysis

What steps have flexible timing?



What type of small molecule screen(s) will best address my biological question(s)?

What resources are available to me?

Compound libraries Laboratory automation

What am I interested in?

Target/pathway identification Novel tool compounds Early stage drug discovery



How large of a screening campaign am I interested in conducting?

Is a phenotypic (cell-based) or biochemical assay most appropriate?

What is the desired assay readout?

Plate reader Imaging FACS, qPCR, NMR

HARVARD MEDICAL SCHOOL http://iccb.med.harvard.edu

What functional genomics tools are (most) appropriate for my biological question(s)?

What resources are available to me?

Libraries

Laboratory automation

What am I interested in?

Knockdown (siRNA, shRNA, CRISPRi) Knockout (CRISPR) Activation (CRISPRa)



Can my cells be transfected? Infected?

What incubation time is required to quantitate phenotype? Is it compatible with the assay?

Short (siRNA, CRISPRi) – 3 to 5 days Long (shRNA, CRISPR) – 2+ weeks

What is the desired assay readout?

Selection (pooled library – shRNA, CRISPR) Imaging, plate reader (arrayed library – siRNA, CRISPR, shRNA, CRISPRi)



Cell perturbations: small molecules vs. functional genomics

	Small Molecules	Functional Genomics
Availability	Not available for every target	Available against every target
	Significant screening effort often require to obtain initial active small molecule	Options: siRNA, shRNA, CRISPR
		Multiple vendors
Mechanism	Inhibition, partial inhibition, activation of enzymes or other proteins	Gene editing
		RNA depletion
		Depletion or partial depletion of protein
Timing	Fast	Slower
	Often reversible	RNAi is gradual depletion
Specificity	Non-specific effects possible	Off-target effects possible



Comparing RNAi and CRISPR

	RNAi (s <i>iRNA</i>)	CRISPR
Modification	Knockdown Post-transcriptional RNA degradation	Knockout Transcription inhibition Transcription activation
Target	Transcript – coding and noncoding	DNA – ORF/promoter/enhancer (adjacent to PAM)
Site of action	Cytoplasm	Nucleus
Length of time	Transient – 2 to 6 days	Permanent (for knockout/knockin)
Off target effects?	Yes Strategies to address	Yes Fewer, <i>TBD</i>



What laboratory resources are available to me?

Fully automated platforms





http://iccb.med.harvard.edu

What laboratory resources are available to me?

Modular work stations







What laboratory resources are available to me?

Standard laboratory equipment – can conduct pooled screens

Hood, centrifuge, plate reader, qPCR instrument, microscope



NGS for quantitating enriched or depleted shRNAs/crRNAs Analytical chemistry resources for compound QC – LCMS, NMR



Liquid handling: bulk dispensers, automated pipettors



Labcyte Echo





Thermo Multidrop Combi



Tecan EVO150



Hewlett Packard D300



Hamilton STARlet



Pin transfer of compounds to library plate



384-pins (V&P Scientific)

Common assay volumes: 30 or 10 ul/well

Typical pin transfer: 33 nl, 100 nl, 300 nl

Compounds dissolved in DMSO

Screening concentration dependent on: Library concentration Dilution



Tip-based transfections for RNAi to library plates



```
http://iccb.med.harvard.edu
```

Creation of custom library plates

Perturbagens



96-well master plate with dilutions







384-well library screening plates





Assay readout: plate readers



PerkinElmer EnVision



Li-COR Aerius



Essen Biosciences IncuCyte FLR



Hamamatsu FDSS7000EX



LifeTechnologies QS7



Plate reader-based assay readouts

Uniform well read out

Absorbance Colorimetric enzyme assays

Luminescence

Cell viability - CellTiterGlo

Luciferase as reporter for transcription Luciferase as reporter for protein stability

Fluorescence

In-Cell Western Channel kinetics, membrane potential Fluorescence polarization, HTRF, AlphaLISA

Advantages

Fast (1-5 minutes/384-well plate) Output is numerical data, quantitative

Disadvantages

Limited information Reagent cost can be high









Screening by imaging: screening microscopes and image analysis



Molecular Devices IXM and IXM-C



GE InCell6000



Perkin Elmer Operetta



TTPLabTech Acumen eX3



Screening by imaging: screening microscopes and image analysis



Molecular Devices IXM and IXM-C

GE InCell6000









TTPLabTech Acumen eX3



http://iccb.med.harvard.edu

Image-based assay readouts

Ability to monitor multiple parameters on a per cell basis – high content screening

Cell number and viability

Presence/absence/intensity of probe

Changes in protein stabilization/localization

Changes in morphology – cell or organelle

Cell migration

Viral or bacterial replication, spread

Advantages

High content screening Potentially less expensive reagents



Disadvantages

Image capture on microscopes/scanner can be slow (10 min – 2+ hours/384-well plate)

Analysis can be challenging, time intensive



General high throughput screening workflow



http://iccb.med.harvard.edu

Timeline for small molecule assay development and screening



Timeline for small molecule assay development and screening



Timeline for RNAi/crRNA assay development & screening



Infrastructure

Reliable tools appropriate for scale of screen?

Informatics

- -Ability to track data? -Analyze data?
- -Visualize data?
- -Access to other screening results?

Choice of library

-Reagent type? -Single vs pooled? -Scale?





Scale:

How high throughput do assay and readout need to be? Microplate density– 96-well or 384-well? Higher? Automation requirements Controls Impact on data analysis





Considerations for assay development

Primary screen assay

-Asking the appropriate question?

-Adaptable to screening?

-Controls?

-Counter screen to focus hit selection?

-Strategy to minimize false negatives?



Secondary assays

-How will potential hits identified in primary screen be narrowed to those most relevant to question being asked?

-Orthogonal assays



Challenge – establish a robust assay that is physiologically relevant

Specific

Sensitive

Reproducible

Scalable

Biochemical assays:

-Is there a reliable and sufficient supply of protein?

-Timing

-Assay temperature

- Cell-based assays:

- -Cell lines vs primary cells
- -Are cells transfectable?

-Stable cells lines for reporter assays or bulk transfection/infection of reporter



http://iccb.med.harvard.edu

Challenge – establish a robust assay that is physiologically relevant

Assay readout

Plate type requirements

Optimize prior to including variable of compound/transfection/infection treatment:

- Protein concentrations
- Incubation times
- Antibody dilutions
- Washes
- Image analysis



Determine potential signal window

Use CRISPR to generate reporter cell line for screen

- Mutation or deletion to mimic disease
- Transcription reporter in endogenous context GFP, luciferase
- Tag endogenous protein with fluorescent marker



Compound screening – how does DMSO impact assay?

Functional genomics – essential to identify a negative control

Ideal – use different positive controls that have range of phenotypes in assay

Monitor dynamic range siRNA/crRNA/compound Genetic or conditional





Compound screening – how does DMSO impact assay?

Functional genomics – essential to identify a negative control

Ideal – use different positive controls that have range of phenotypes in assay

Monitor dynamic range

siRNA/crRNA/compound

Genetic or conditional







http://iccb.med.harvard.edu

Compound screening – how does DMSO impact assay?

Functional genomics – essential to identify a negative control

Ideal – use different positive controls that have range of phenotypes in assay

Monitor dynamic range

siRNA/crRNA/compound

Genetic or conditional





Compound screening – how does DMSO impact assay?

Functional genomics – essential to identify a negative control

Ideal – use different positive controls that have range of phenotypes in assay

Monitor dynamic range

siRNA/crRNA/compound

Genetic or conditional

siRNA library plate



Assay specific controls plate





http://iccb.med.harvard.edu

Measure of HTS robustness Opportunity to mimic automation utilized in screen



Z' factor =
$$1 - \frac{3(\sigma p + \sigma n)}{|\mu p - \mu n|}$$

Assumes normality Range 1 to $-\infty$

Biochemical small molecule > 0.75 ideal Cell-based small molecule > 0.6 ideal RNAi > 0.5 ideal

Zhang JH, et al. J Biomol Screen. 1999;4(2):67-73.



CB399 Nanocourse 2017

Assay optimization – pilot screen

How does assay perform when using a variety of compounds or siRNA/ crRNA?

Visualize and analyze data

Acceptable Z' factor? Potential hits? If siRNA, do potential hits make biological sense?

Good correlation of replicates?

Clear separation of positive and negative controls?

Any visible patterns or edge effects?

Utilize visualization software









A continuum of activity/phenotype is observed in the primary screen



Experimental compounds (or siRNAs)



A continuum of activity/phenotype is observed in the primary screen



Experimental compounds (or siRNAs)

Follow the positives!



Visualization of a full genome siRNA screen, with libraryand assay-specific controls included




Potential hits can be categorized by strength





Additional secondary high-throughput screens can limit and categorize hits



Plate:Well



What resources are available to me?

Software

Coding capability

How many data points am I working with?

Per experiment

Overall

What question am I trying to answer?

How do the data look across a specific experiment? Overall?

How reproducible are the data?

Are there positional effects? Biases?

Is the data normally distributed?

What is the potential hit rate? Within specific profile (multiple conditions)?

Does the analysis fit with what is anticipated from the raw data?

Are there trends in selected hits?

How will hits be identified?

HARVARD MEDICAL SCHOOL http://iccb.med.harvard.edu

What resources are available to me for data visualization?

Custom script- self, collaboration

The R Project for Statistical Computing https://www.r-project.org

MathWorks[®] MATLAB https:

https://www.mathworks.com

National Institute of Allergy and Infectious Diseases CARD Comprehensive Analysis of RNAi Data

https://card.niaid.nih.gov

spotfire.tibco.com



TIBC Spotfire

https://www.graphpad.com



What resources are available to me for data visualization?



Vortex

Intuitive and versatile scientific data visualization and analysis

Laurie Martensen laurie.martensen@dotmatics.com 619.306.3412



Data acquisition \longrightarrow analysis \longrightarrow visualization





Output file: csv





CB399 Nanocourse 2017

http://iccb.med.harvard.edu

Data acquisition \longrightarrow analysis \longrightarrow visualization

Formatted data file: Excel Output file: csv D the rank hand when the rank hand had had had been been been Exclude HTRF-Ch1_A HTRF-Ch2_A HTRF-Ratio A HTRE-Ch1 B RE-Ch2_B Plate Well Type. HTRE-Ratio B Line And Spin Facto March Starts March 2 21 10 888 Sold Starts Starts 1394 A01 394 A02 Plate well location 735) A03 A04 A05 Well identifier A06 394 A07 AOB 1 row/compound F1394 A09 And And Spin Facto Mart March. March 1 11 10 100 Social Science Spin-March 1 11 10 100 Social Science Spin-State 1 11 10 100 Social Science Spin-F:394 A10 A11 All values/well A12 A13 394 A14 7394 A15 All replicates/well 394 A16 A17 **F**394 A18 A19 394 A20 1394 A21 F1394 A22 7394 A23 A24 B02 Inter Mand Ngar Parts Mart Marth Mart 1 1 1 11 11 111 1111 F1394 394 **F**394 394 B09 394 B10 **F**394 B11 B12 Later Kand Spac Farm Maar Maaron Kalar S II III III HAR Kalar Sarbiya Sarbiya Kalar I II III III KAN Sarbiya Sarbiya Sarbiya F1394 B13 B15 394 B16 R11 **F394 B18** B19 B20 la → ▶ | 20161027-135939.xls / + / F394 C01 C02 2947(C03 C04 C05 1394 C06 Sent to ICCB-L data curator C07 C08 44 -Controls Custom script Screensaver Analysis software Visualization software

http://iccb.med.harvard.edu

How do the data look across a specific experiment?



Visualize using heat map function

Controls and potential hits should be clearly identifiable

Can be used to distinguish bad plates in a run or trends



How do the data look across a specific experiment?



Example: row trends, atypical hit rate



Max

How do the data look across a specific experiment? Overall?



Confirmed what investigator had suspected – first 16 plates different from rest of screen (45 plates total)



How do the data look across a specific experiment? Overall?



Highlighted issues quantitating cell numbers at the beginning of the screen. Screener paused, re-optimized, then continued.



How do the data look across a specific experiment? Overall?



Striping on two plates Experimental wells on one plate atypical



Basic scatterplot



plate-well

Most informative: basic scatter plot with a unique ID (plate-well, reagent ID) vs average data value. Color set by well 'type'

Enables visualization of: separation of controls, dynamic range potential positives positional effects trends in data (positional, time)



Basic scatterplot – problems!



Positional trend, periodicity



Time trend, drift



Low signal:background

Excessive scatter Unusually high or low plates Issues with control Possible saturation



Basic scatterplot – *big* problems!





Basic scatterplot for biochemical screen





How reproducible are the data?





Goal: R² > 0.8

Note: controls can skew data, artificially inflating R². Determine correlation +/- controls





Needed solution to average raw data by well location across all plates in a run

Statistics feature: mean of raw data using a 'group by function of row and column' Generates new data column "Z" **Graph: X = column, Y = row, Z = size**

Goal: all wells (except positive control) are the same size





Example: Clear trends in rows C/D and G/H

Technical issue with Combi?

Paired row trends frequently consequence of peristaltic pump manifold utilized to fill assay plates because 1 nozzle fills 2 rows (A/B, C/D, *etc.*)





```
http://iccb.med.harvard.edu
```



Edge effect – temperature gradient, evaporation



Are the data normally distributed?



Ave Viability (50 bins)

Built using bar graph

Ideal is bell-shaped curve, with both tails visible

Some assays only display one tail (signal issues, starting point of assay)

Distribution Important to know – influences data analysis methods



1. Triage

2. Normalization

Performed per plate Removes systematic data effects Options: sample median, control median, log transformation, scores May have to exclude some plates/wells from analysis Enables comparison and combination of data from different plates

- 3. Quality check
- 4. Hit selection statistical analyses



How will hits be identified?

Z-score

Assumes normality Parametric (score implies a probability) Scale?

Robust Z-score

Recommended for RNAi Robust to outliers Nonparametric

B-score

Can take into account positional effects

Percent of control

Normalized percent inhibition/activation

Rank product

SSMD – Strictly Standardized Mean Difference

RSA – Redundant siRNA Activity





What is the potential hit rate?



Ave Alpha_bin_stddev_0.5,1,2,3 Built using bar graph, bin by standard deviation option

Breakdown of experimental wells by standard deviations away from the average

Goal: observe majority of wells having minimal effect (green and yellow) and a few potential positives (red)

Provides estimated number of hits, how many enhancers and suppressors

Accept there will be both false positives and false negatives

Follow the positives

HARVARD MEDICAL SCHOOL http://iccb.med.harvard.edu

What is the potential hit rate? Within specific profile (multiple conditions)?

Is the assay capable of identifying potential hits with specified profile?

Focused siRNA screen



2 endpoints were measured: Phenotype score (screen) Cell number (counter screen)

Goal: identify genes that result in low phenotypic score (y-axis), but are not generally toxic (x-axis)

Ideal: low on y-axis, high on the x-axis and red CB399 Nanocourse 2017



Primary small molecule screen with 2 conditions



Same compound library is being run against wild-type cell and mutant cell line

Goal: identify compounds that kill mutant but not wt strain

Ideal: low on y-axis, high on x-axis



Primary small molecule screen with 3 conditions



plate-well

Same compound library is being run against wild-type cell and 2 different genetically modified strains

Goal: identify compounds that hit in 1 modified strain, but not the other or wild-

type Wild-type: color by average Abs Strain 1: position on y-axis by average Abs Strain 2: size by average Abs

Ideal: low on y-axis, in blue, large size



Quality check – does the analysis fit?

An investigator came back with an analyzed data set and specific questions:

Does the math fit?

Check algorithm for hit selection to see if it visually looks as intended

How does the counter screen (viability) change the data set?

Can Vortex be used to spot errors or trends in the selected hits?

Note: cell-based assay with multiplexed known bioactive compounds – high primary hit rate

Systematic Identification of Synergistic Drug Pairs Targeting HIV

Xu Tan¹, Long Hu^{2,#}, Lovelace J. Luquette III^{3,#}, Geng Gao¹, Yifang Liu², Hongjing Qu¹, Ruibin Xi³, Zhi John Lu^{2,*}, Peter J. Park^{3,*}, and Stephen J. Elledge¹ *Nat Biotechnol.* 2012 November ; 30(11): 1125–1130.





Quality check – does the analysis fit?



1. Does my algorithm for picking hits match what I would pick visually on the calculated data?

13230 number of experimental wells

Potential hits categorized as Strong (333), Medium (211) and Weak (284) based upon z-score



Do the analyzed and raw data correspond?



2. Does my algorithm for picking hits match what I would pick visually on the RAW DATA?



How does the counter screen change potential hits?



3. What does hit selection on the raw data look like after the counter screen filter (toxicity) is applied?



How does the counter screen change potential hits?



Counter screen enables follow up to be on the most relevant potential hits



Are there trends in the potential hits?

Which wells are my potential hits in?



Well_bin_distinct

Utilize distribution graph to check which wells are picked and how often

This example: wells in row E (E14) frequently represented, periodicity



Are there trends in the potential hits?

Which library plates are my potential hits in?



Plate_bin_distinct

Example: clear indication that the first 10 plates have a large number of hits

Library clustering of similar compounds, target type? May justify more stringent criteria



Practical

Discovery of biological research tools Probes of pathways Discovery/validation of new 'druggable' targets note: target ID can be challenging!

More Ambitious

Early-stage therapeutic lead discovery Requires significant medicinal chemistry Animal studies required to evaluate efficacy, PK/PD, toxicity....


Practical

Gene/target discovery in specific cell pathway or process

Small molecule target identification or elucidation of mechanism Perform in parallel with small molecule screen Screen for enhancers/suppressors of small molecule-induced phenotype

More Ambitious

Systematic annotation of a whole genome or gene family

Examples of high throughput screening projects

HOME / RESOURCES /

http://iccb.med.harvard.edu/small-molecule-publications

Small Molecule Publications

Below is a list of publicatio notify us of your publicatio

Khan AS, Murray MJ, Ho (GlaxoSmithKline Protein Replication that Prevent: ahead of print:doi:10.1099

Beelontally R, Wilkie GS, I Strang BL. Identification of IE2 proteins. Antiviral F

Balasubramanian A, Manz the identification of smal

Zervantonakis IK, Arvanitis acoustofluidic 3D tumor

RNAi Publications

http://iccb.med.harvard.edu/rnai-publications

Below is a list of publications that have resulted from RNAi screening efforts at ICCB-Longwood. Please notify us of your publications that involve use of the ICCB-Longwood Screening Facility.

Nelms B, Dalomba NF, Lencer W. A targeted RNAi screen identifies factors affecting diverse stages of receptor-mediated transcytosis. J Cell Biol 2017;

Simpson KJ, Smith JA. Knocking down the obstacles to functional genomics data sharing. Sci Data 2017;4:170019 doi:10.1038/sdata.2017.19.

Dutta B, Azhir A, Merino L-H, Guo Y, Revanur S, Madhamshettiwar PB, Germain RN, Smith JA, Simpson KJ, Martin SE, Beuhler E, Fraser LDC. An interactive web-based application for Comprehensive Analysis of RNAi-screen Data. Nat Commun 2016;7:10578.

Polachek WS, Moshrif HF, Franti M, Coen DM, Streenu VB, Strang BL. High-Throughput Small Interfering RNA Screening Identifies Phosphatidylinositol 3-Kinase Class II Alpha as Important for Production of Human Cytomegalovirus Virions. J Virol 2016;90(18):8360-71.

Savidis G, McDougall WM, Meraner P, Perreira JM, Portmann JM, Trinucci G, John SP, Aker AM, Renzette N, Robbins DR, Guo Z, Green S, Kowalik TF, Brass AL. Identification of Zika Virus and Dengue Virus Dependency Factors using Functional Genomics. Cell Reports 2016;16(1):232-46.

http://iccb.med.harvard.edu

Fluorescence Polarization (FP) Assays: protein-protein interactions



Figure 1 Schematic depicting the basic principle of fluorescence polarization. When a small peptide or nucleic acid ligand (dark circle) with a fluorescent label attached (white circle) is excited by polarized light at the excitation wavelength of the fluorophore, the ligand reorients to a significant degree due to molecular tumbling during the excited state lifetime of the fluorophore. This causes the emitted light to be largely depolarized. If the ligand is bound to a protein (gray ellipse), the resulting complex tumbles much slower, and the emitted light retains its polarization. Moerke (2009) Current Protocols in Chem Biol 1:1

Example ICCB-L FP screens:

Frey et al. (2006) PNAS *103*:13938– HIV gp41 conformational change Moerke at al. (2007) Cell *128*:257– eIF4E/eIF4G interactions



Compound 4EGI-1 is a competitive inhibitor of the eIF4E/eIF4G_interaction



Figure 1. A FP Assay Identifies the Compound 4EGI-1 as a Competitive Inhibitor of the eIF4E/eIF4G Interaction

http://iccb.med.harvard.edu

4EGI-1 Disrupts eIF4F Complex Formation and Inhibits Cap-Dependent Translation



-retic lysate binding assay

-purified protein binding

-cap-dependent translation assay

Figure 3. 4EGI-1 Disrupts eIF4F Complex Formation and Inhibits Cap-Dependent Translation

Moerke at al. (2007) Cell 128:257



4EGI-1 Disrupts the eIF4F Complex and Inhibits Expression of Oncogenic Proteins in Mammalian Cells



Moerke at al. (2007) Cell 128:257



4EGI-1 Has Proapoptotic Activity and Inhibits the Growth of Multiple Cancer Cell Lines



Figure 5. 4EGI-1 Has Proapoptotic Activity and Inhibits the Growth of Multiple Cancer Cell Lines

Moerke at al. (2007) Cell 128:257 CB399 Nanocourse 2017



Functional genomics example - Abraham Brass

Identification of Zika Virus and Dengue Virus Dependency Factors using Functional Genomics

Cell Reports 16, 232-246, June 28, 2016

George Savidis,^{1,5} William M. McDougall,^{1,5} Paul Meraner,^{1,5} Jill M. Perreira,¹ Jocelyn M. Portmann,¹ Gaia Trincucci,¹ Sinu P. John,² Aaron M. Aker,¹ Nicholas Renzette,¹ Douglas R. Robbins,¹ Zhiru Guo,³ Sharone Green,³ Timothy F. Kowalik,¹ and Abraham L. Brass^{1,4,*}

MORR – Multiple Orthologous RNAi Reagent – Dharmacon, Ambion





Functional genomics example - Abraham Brass

Pooled CRISPR/Cas9 survival enrichment screen

Cell Reports 16, 232-246, June 28, 2016



CB399 Nanocourse 2017

http://iccb.med.harvard.edu

Functional genomics example - Abraham Brass

Cell Reports 16, 232-246, June 28, 2016



