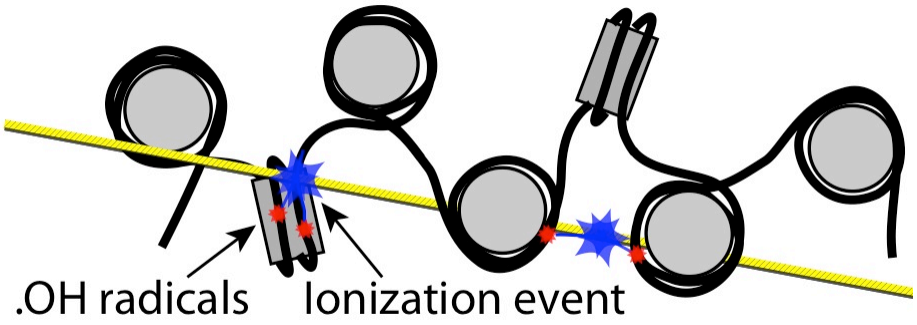
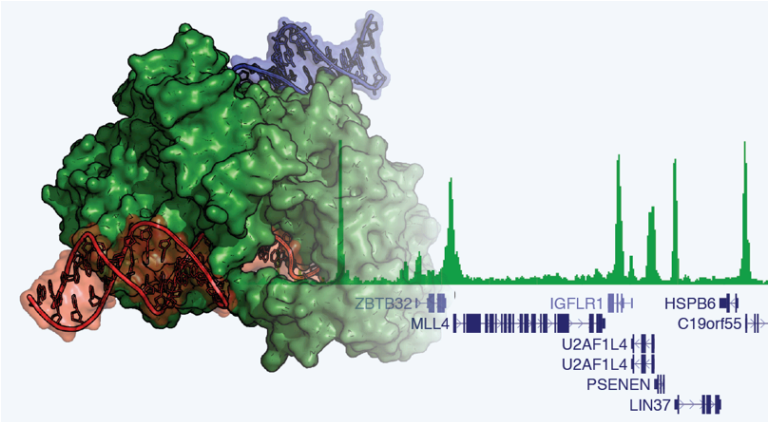




New directions for genomic data generation and functional validation

William Greenleaf | Stanford University

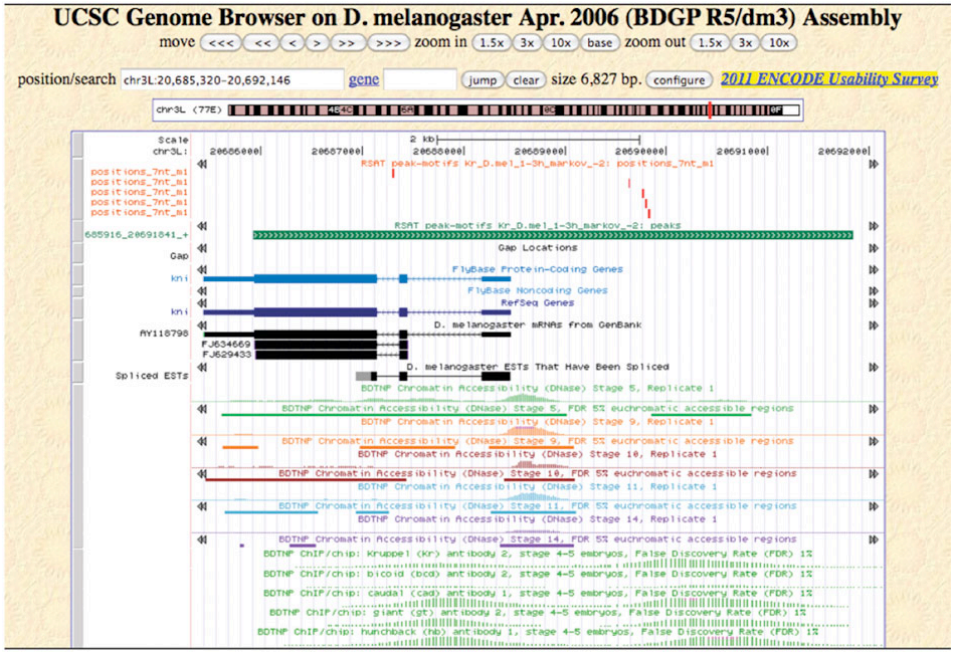
Greenleaf Lab

department of GENETICS

Conflicts: Epinomics, Eve Bio

Not so much an encyclopedia as an atlas

?



Brussels

Coordinates: 50°51′N 4°21′E

From Wikipedia, the free encyclopedia

Brussels (French: *Bruxelles*, [bʁyseɛl]; Dutch: *Brussel*, [ˈbrʏsəl]), officially the **Brussels-Capital Region**[5][6] (French: *Région de Bruxelles-Capitale*, Dutch: *Brussels Hoofdstedelijk Gewest*),[7] is the capital and largest city of Belgium and the *de facto* capital of the European Union (EU).[8][9] It is also the largest urban area in Belgium, comprising 19 municipalities, including the municipality of the City of Brussels, which *de jure* is the capital of Belgium, in addition to the seat of the French Community of Belgium and of the Flemish Community.[10]

Brussels has grown from a 10th-century fortress town founded by a descendant of Charlemagne to a sizeable city.[11] The city has a population of 1.2 million and a metropolitan area with a population of over 1.8 million, both of them the largest in Belgium.[12][13]

Since the end of the Second World War, Brussels has been a major centre for international politics. Hosting principal EU institutions,[14] the secretariat of the Benelux and the headquarters of the North Atlantic Treaty Organization (NATO), the city has become the polyglot home of numerous international organisations, politicians, diplomats and civil servants.[15]

Brussels is just a few miles north of the boundary between Belgium's language communities—French in the south, Dutch in the North. Historically a Dutch-speaking city, it has seen a major shift to French since Belgian independence in 1830. Today, although the majority language is French, the city is officially bilingual. All road signs, street names, and many advertisements and services are shown in both languages.[16] Brussels is increasingly becoming multilingual with increasing numbers of migrants, expatriates and minority groups speaking their own languages.

Brussels

Bruxelles

Brussel

Region of Belgium

Brussels-Capital Region

Région de Bruxelles-Capitale

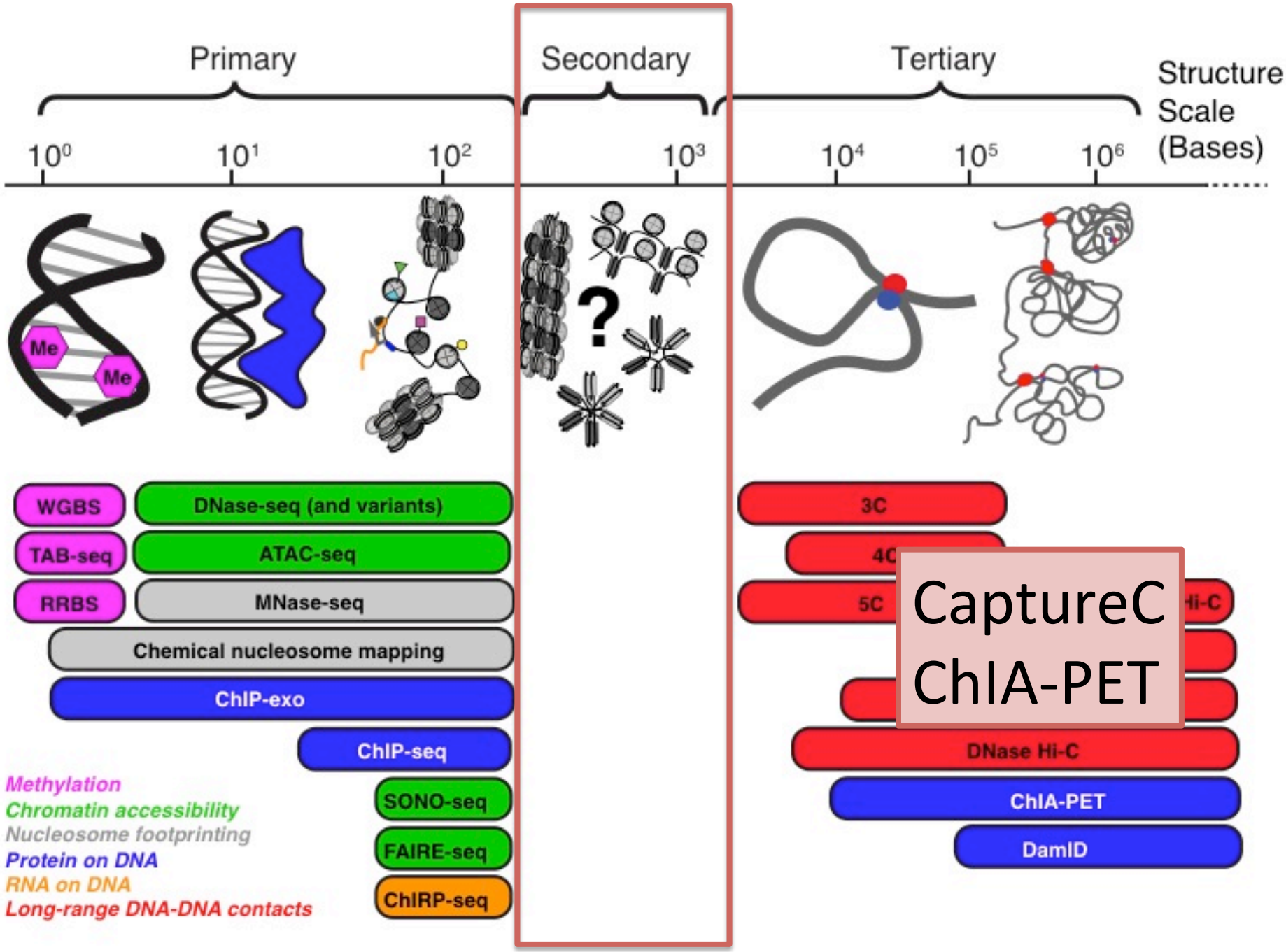
Brussels Hoofdstedelijk Gewest

A collage with several views of Brussels, Top: View of the Northern Quarter business district, 2nd left: Floral carpet event in the Grand Place, 2nd right: Brussels City Hall and Mont des Arts area, 3rd: Cinquantenaire Park, 4th left: Manneken Pis, 4th middle: St. Michael and St. Gudula Cathedral, 4th right: Congress Column, Bottom: Royal Palace of Brussels

Contents

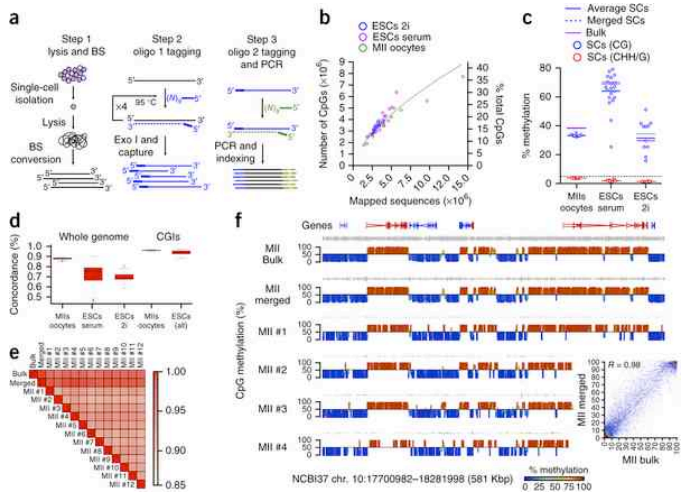
Outline

- Projects to provide 1) new types of genomic maps and/or 2) high-throughput, hypothesis-less functional validation and quantitative understanding:
  - Chromatin 2ndary structure (1)
  - Single-cell regulatory information (1 & 2)
  - Quantitative biochemical investigations of sequence preferences (1 & 2)
  - High throughput and combinatorial CRISPRa/i/KO in analogy to shRNA screens (2)


Primary
Secondary
Tertiary
Structure Scale (Bases)
10^0
10^1
10^2
10^3
10^4
10^5
10^6
Me
Me
?
WGBS
TAB-seq
RRBS
DNase-seq (and variants)
ATAC-seq
MNase-seq
Chemical nucleosome mapping
ChIP-exo
ChIP-seq
SONO-seq
FAIRE-seq
ChIRP-seq
3C
4C
5C
Hi-C
DNase Hi-C
ChIA-PET
DamID
CaptureC
ChIA-PET
Methylation
Chromatin accessibility
Nucleosome footprinting
Protein on DNA
RNA on DNA
Long-range DNA-DNA contacts

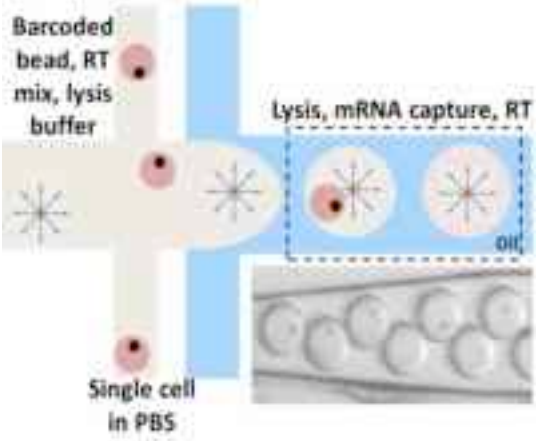
2) Understanding functional elements in single cells

Single-cell methylation

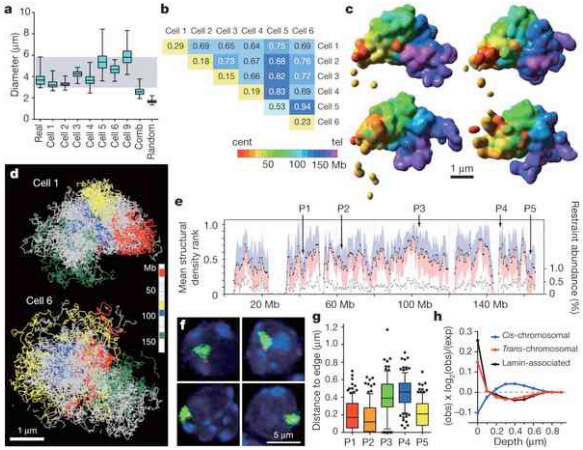


Smallwood et al. Nature methods 2014

Droplets for single-cell RNA/chromatin analysis

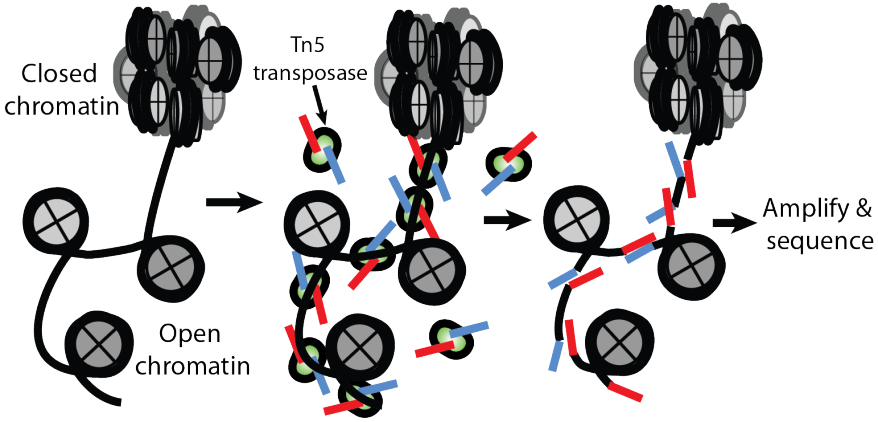


Single-cell Hi-C

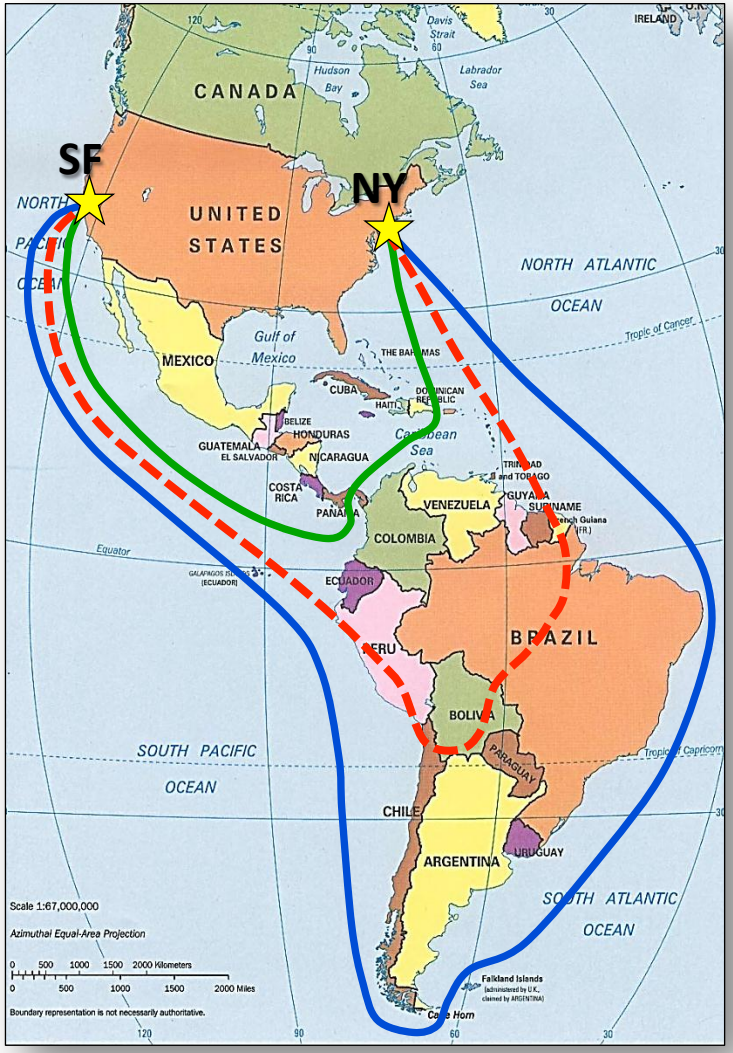


Nagano et al. Nature 2013

Single-cell ATAC-seq

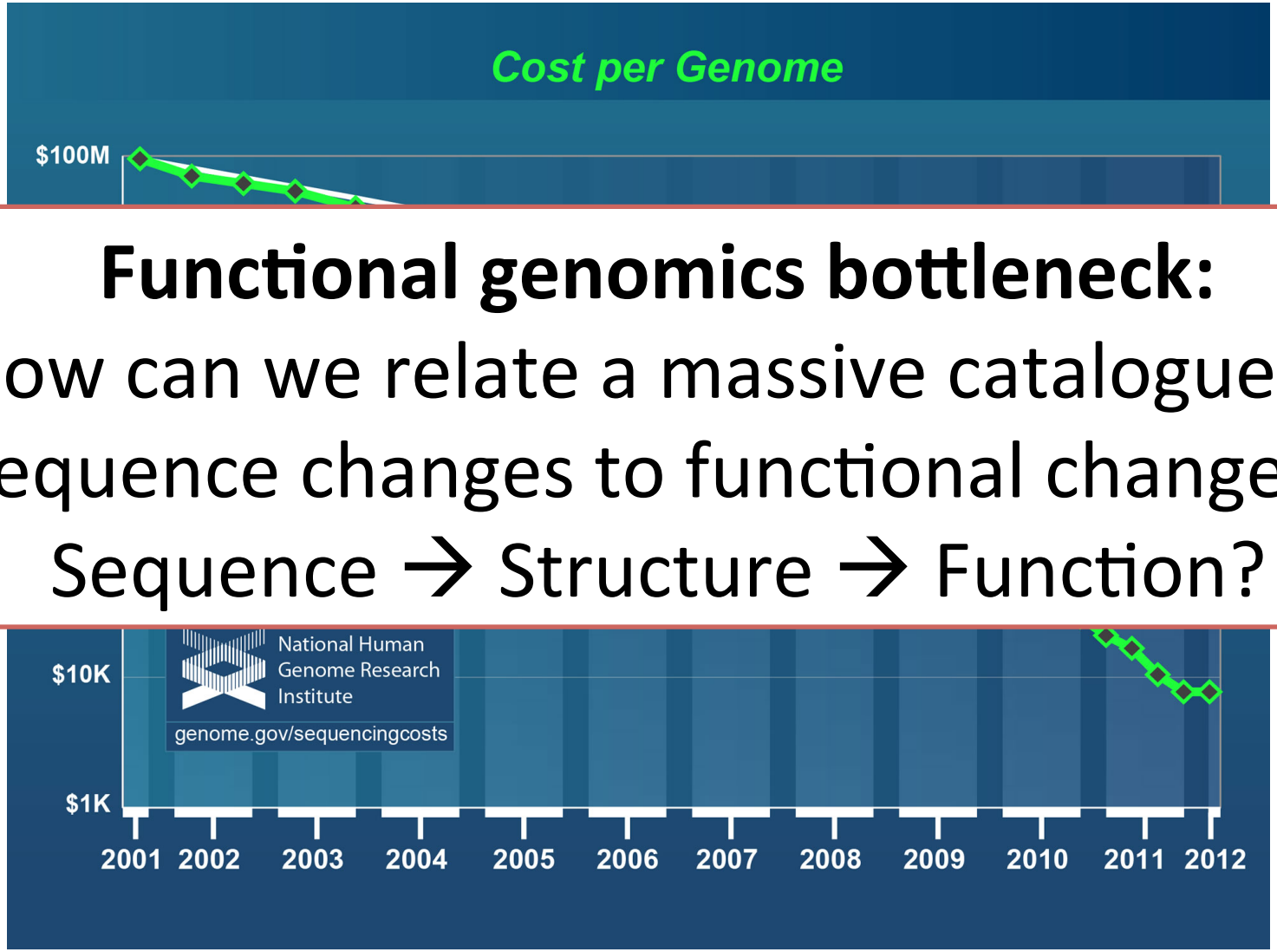


Why Single Cell?



Understanding “mechanism,” or establishing causality does not necessarily imply “hypothesis driven”

Our understanding of functional impacts of variation lags behind our ability to identify variation

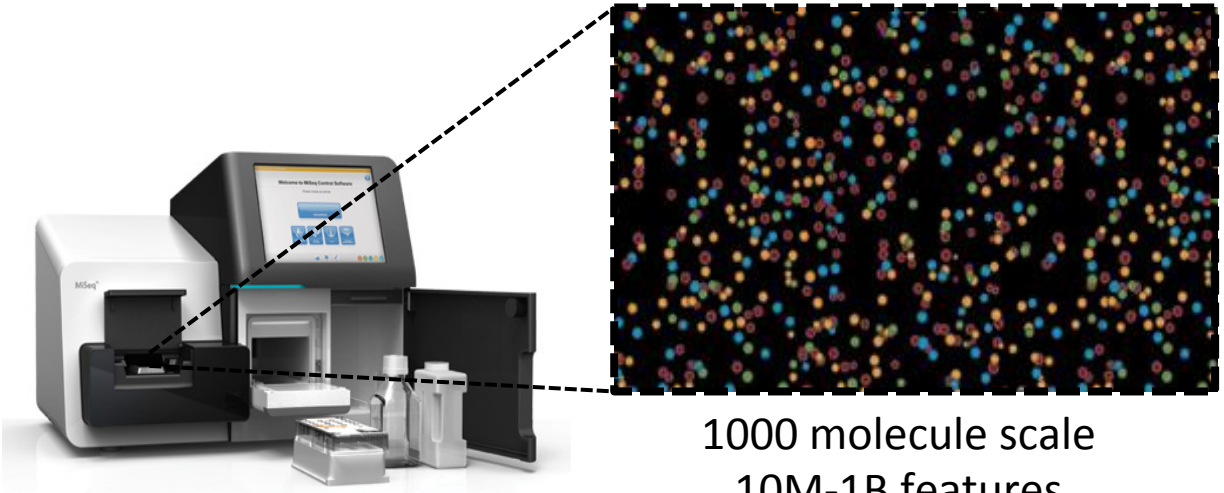


Functional genomics bottleneck:

How can we relate a massive catalogue of sequence changes to functional changes?

Sequence → Structure → Function?

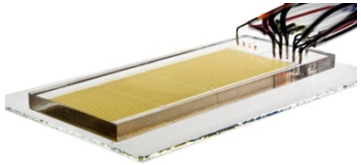
3) Ultra high-throughput quantitative biochemistry



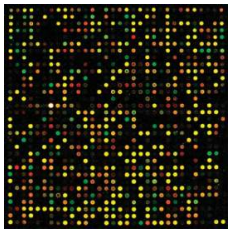
1000 molecule scale
10M-1B features

96

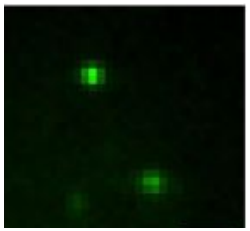
384



microfluidic
lab-on-a-chip

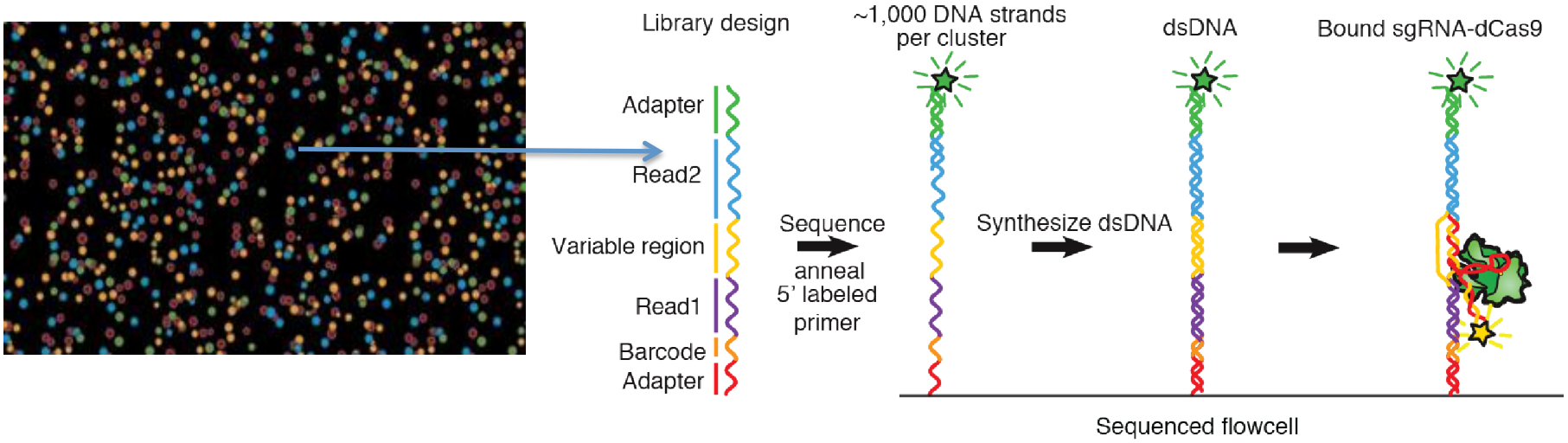


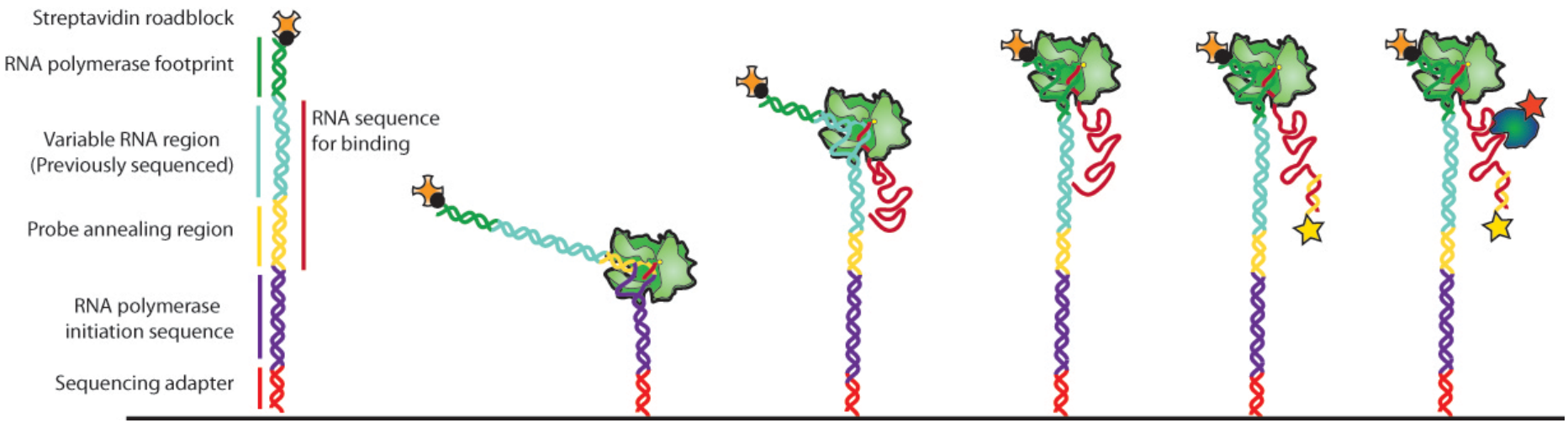
microarray



single
molecule

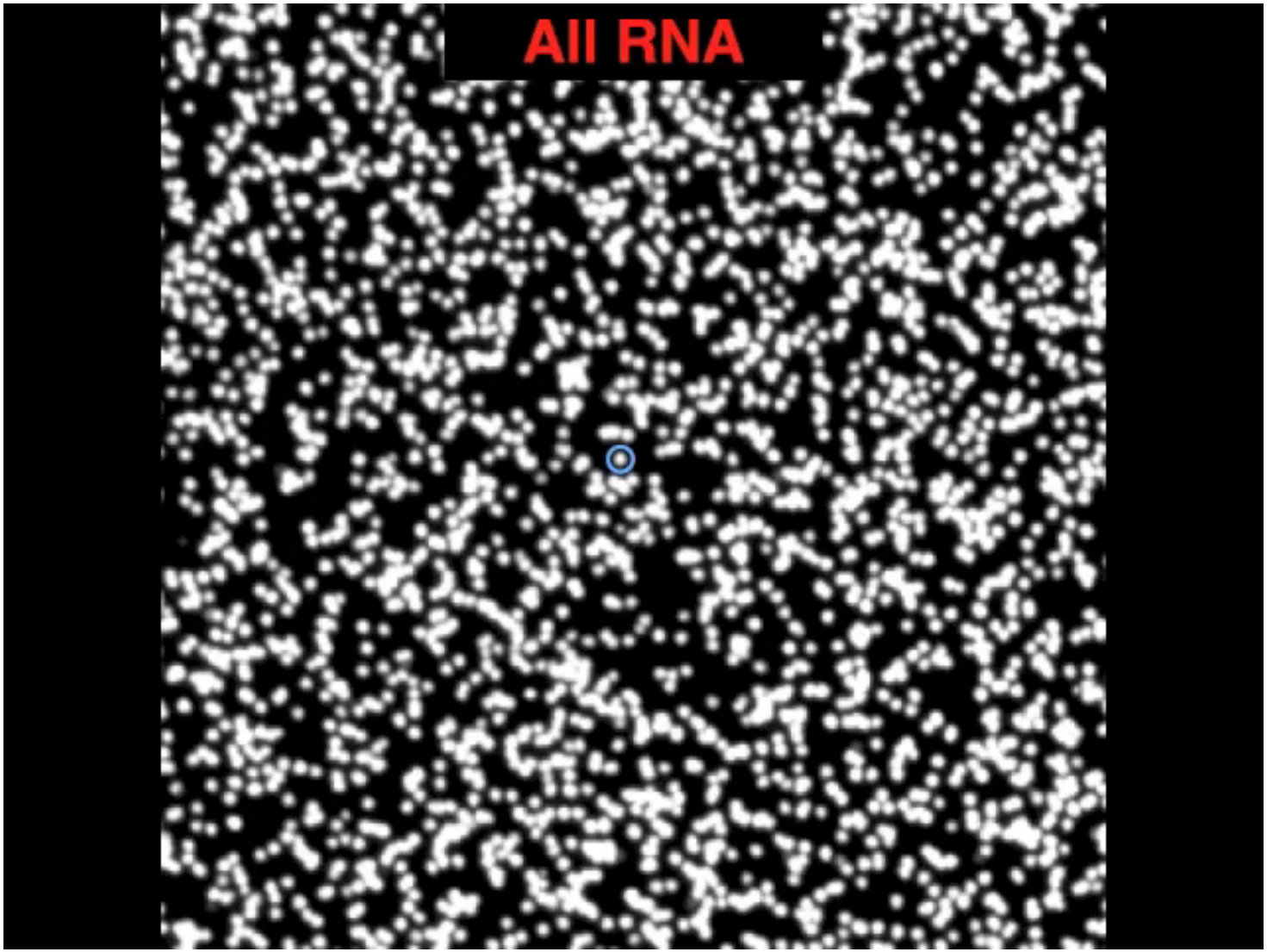
High-throughput functional investigations on a sequencer



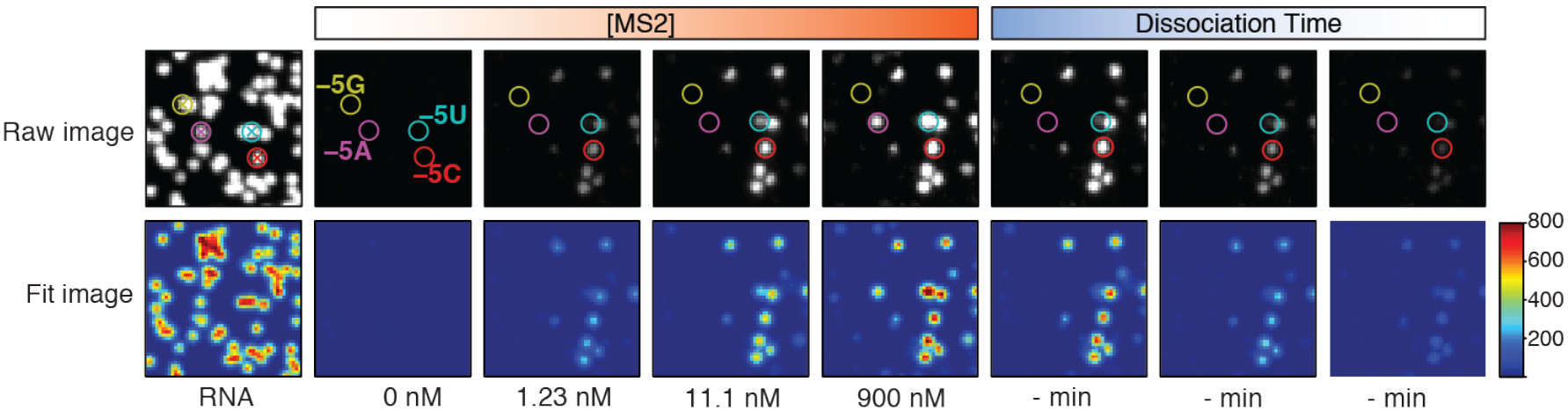


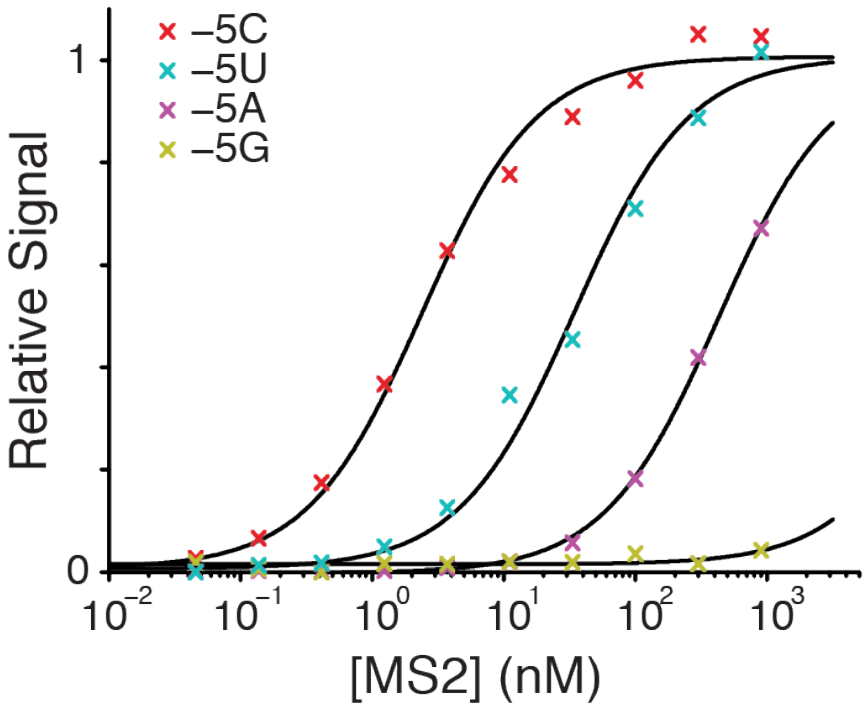
This work was pioneered by Chris Burge's lab [R. Nutiu *et al.*, *Nat Biotechnol* (2011)]

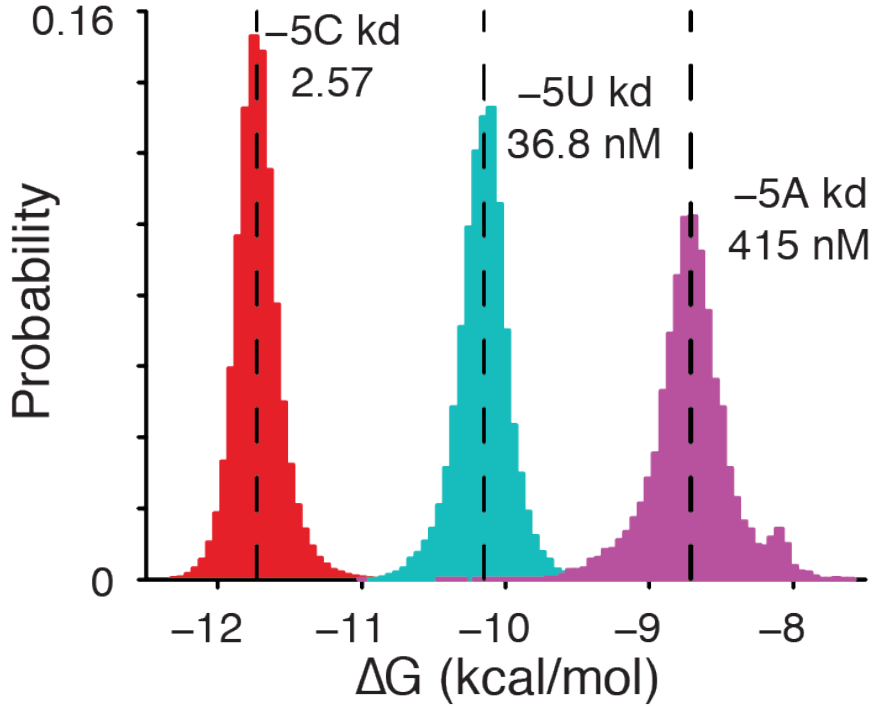
Equilibrium binding and off-rates (themodynamic and kinetic measurements)



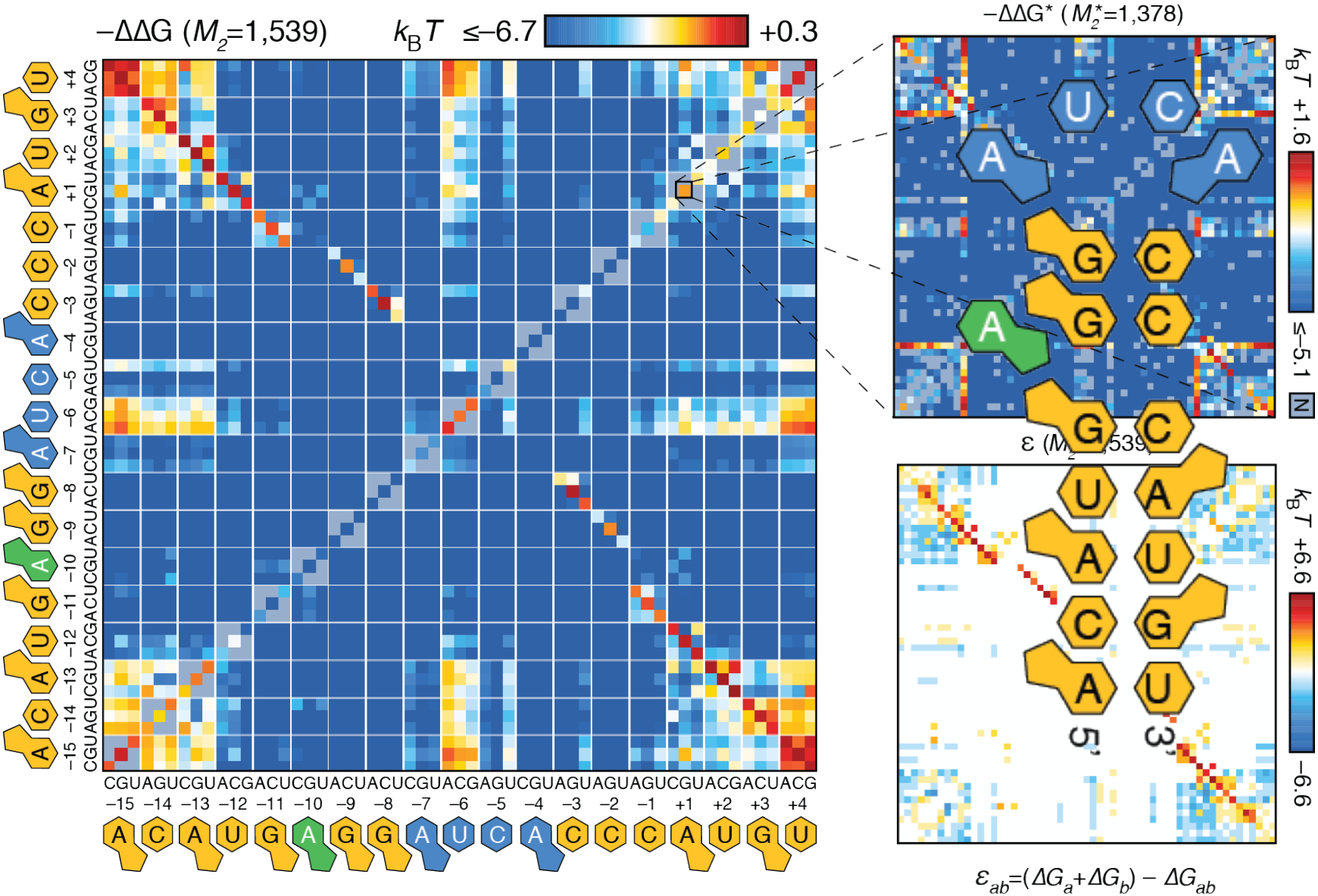
Binding and dissociation measurements



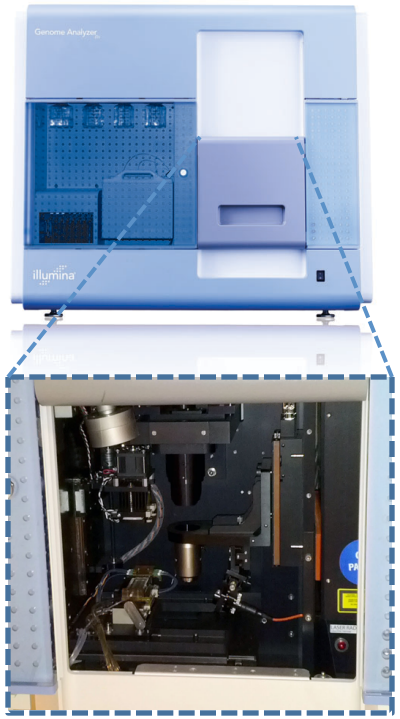




Reconstructing a comprehensive functional landscape for binding



Pirate-ing sequencing instruments is cheap and "easy"



GAIIx

- 6 years ago, $600k
- Now ~$0
- Pirate-able parts

best of both worlds +

Custom Imaging Station

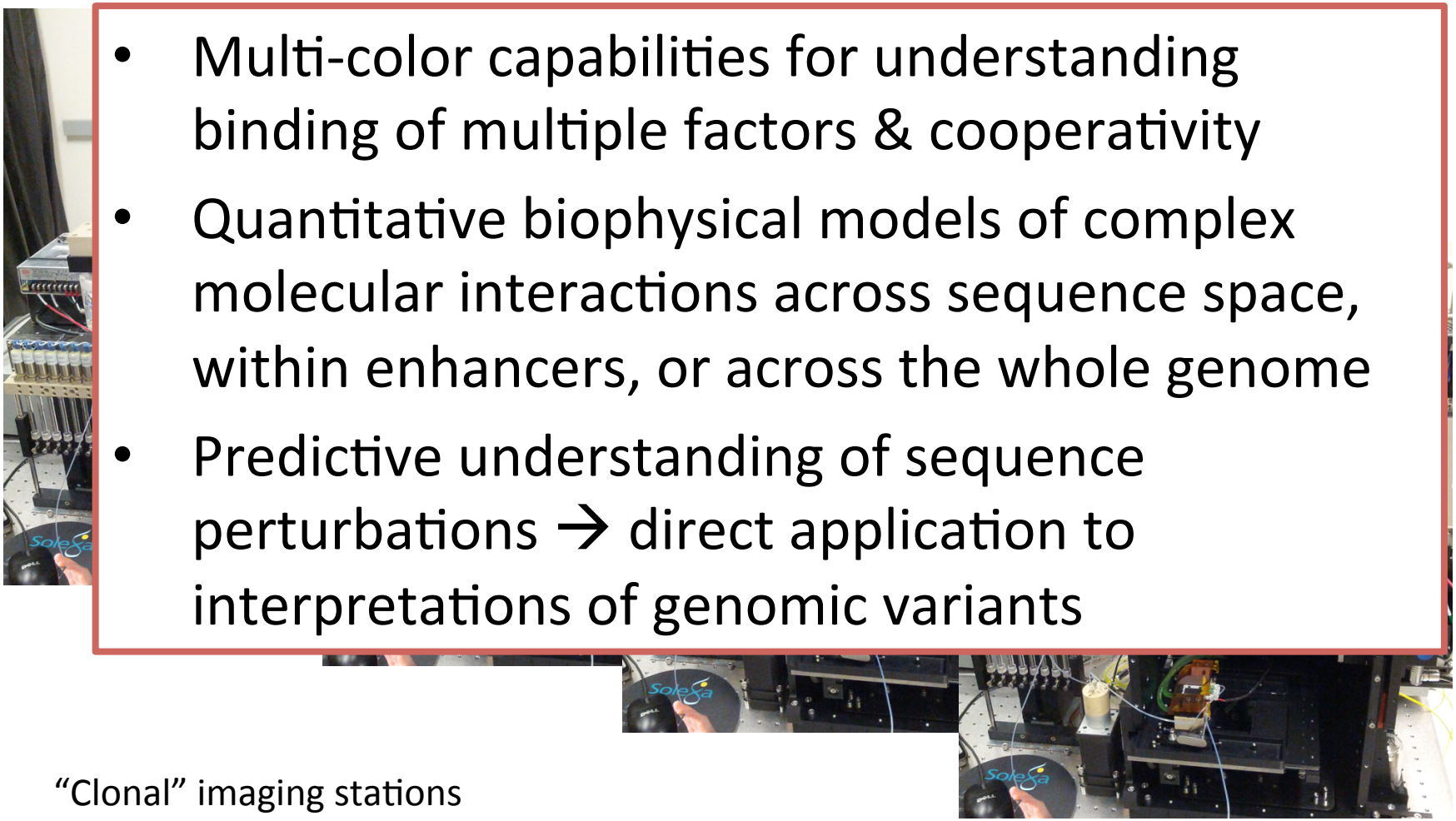
- All hardware under custom MATLAB software control
- Temperature control added
- Automated scripting for hands-free data collection

MiSeq

- ~20M reads
- 1 day / chip
- Simple workflow

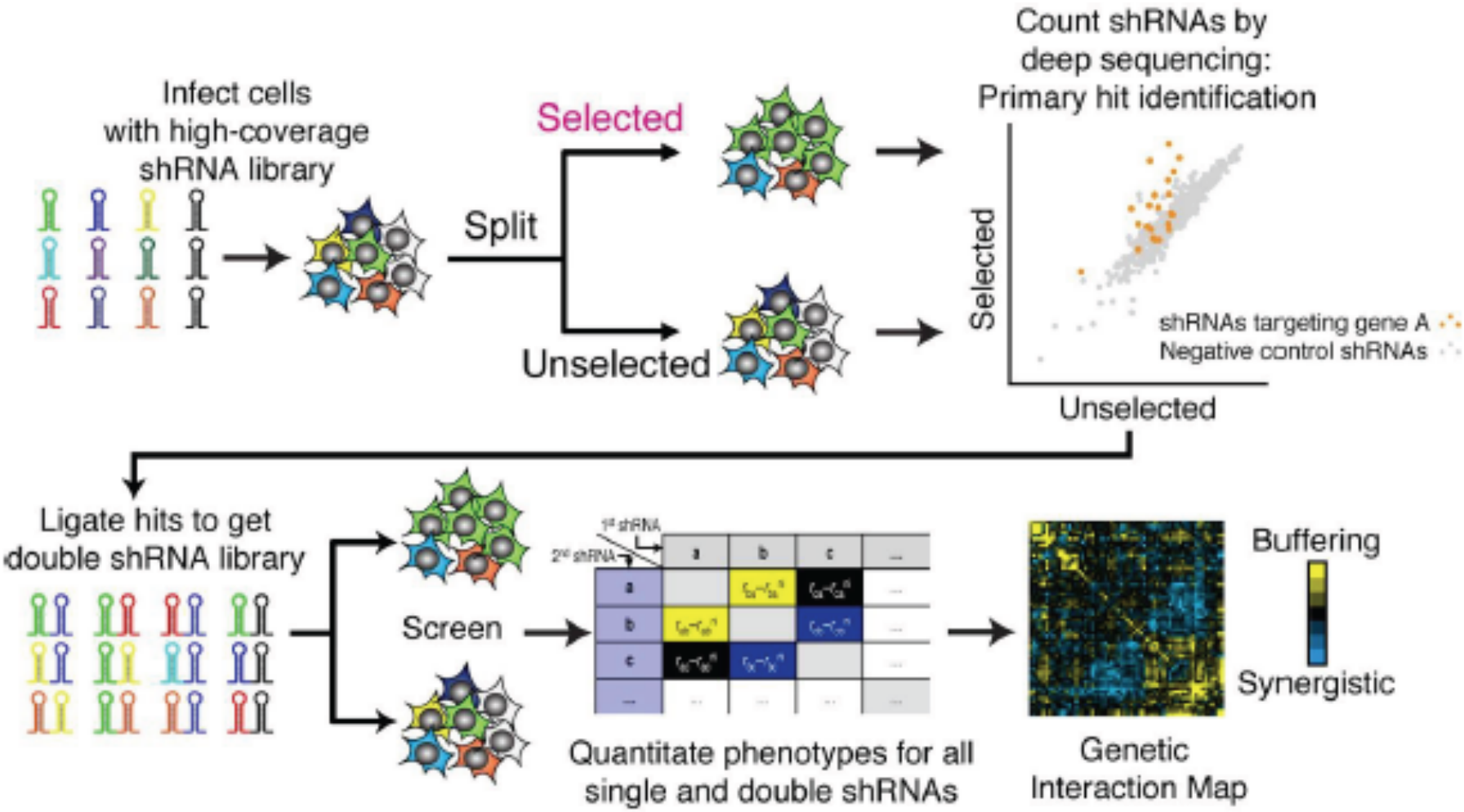
Marginal cost of imaging stations <$10k… perfect for creating a center

- Multi-color capabilities for understanding binding of multiple factors & cooperativity
- Quantitative biophysical models of complex molecular interactions across sequence space, within enhancers, or across the whole genome
- Predictive understanding of sequence perturbations → direct application to interpretations of genomic variants



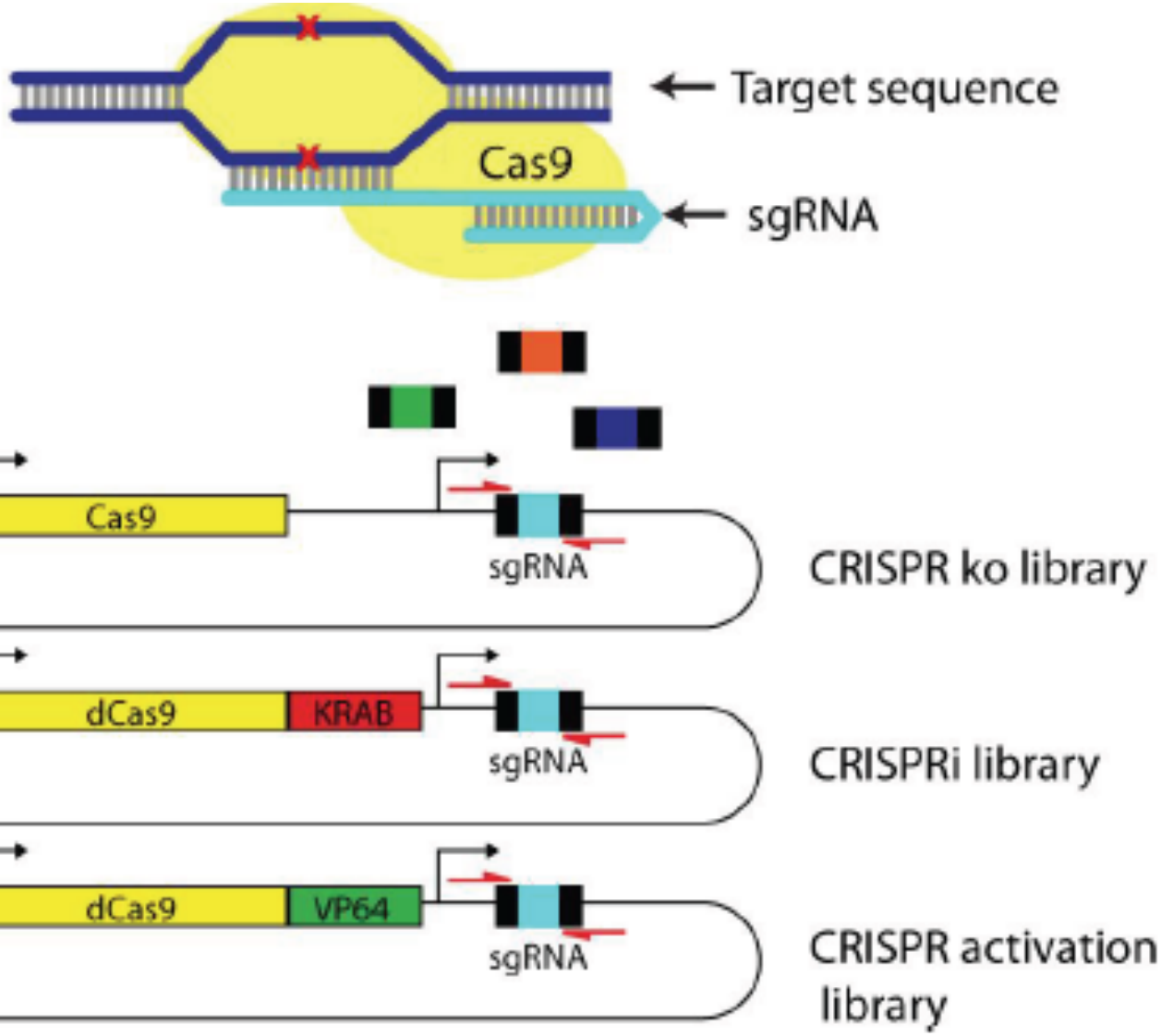
"Clonal" imaging stations

4) CRISPR-based high throughput functional screens in analogy to combinatorial shRNA screens



Figures courtesy of Michael Bassik

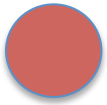
Combinatorial CRISPR libraries allow targeting of functional elements



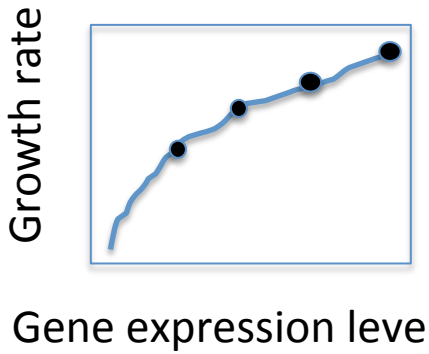
Figures courtesy of Michael Bassik

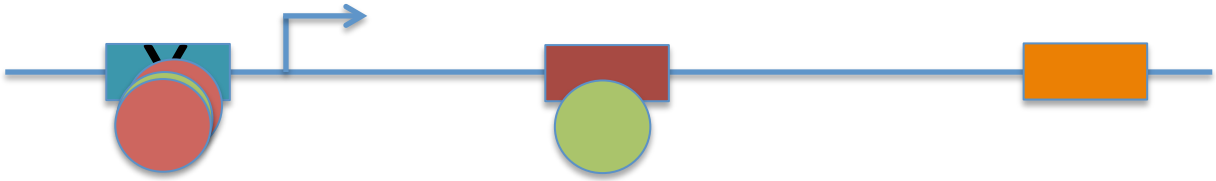
Experimental outline for gene *i* of *N*

X = CRISPER KO


= CRISPERi

= CRISPERa





- Needed: Mapping from gene expression of hundreds of genes to growth defects.
- CRISPRi/a/KO targeted to every putative potential regulatory element.
- All pairwise combinations of these perturbations of these elements.
- Potential for high-throughput quantitative picture of single element effects on gene expression in vivo, and synergy between these elements in tandem
- Provides comprehensive, quantitative measure of element function on (gene expression) and insight into additive regulatory logic.

Acknowledgements

Carlos Araya

Curtis Layton

Viviana Risca

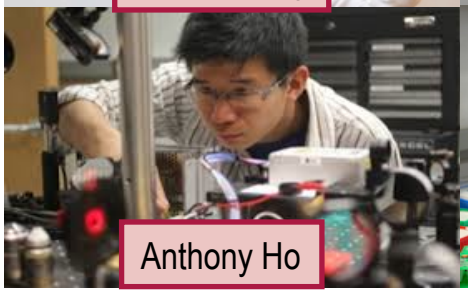
Johan Andreasson

David Zhang

Jason Buenrostro

Lauren Chircus

Sarah Denny


Anthony Ho

Beijing Wu

Alicia Schep

David Moskowitz

Funding: Rita Allen Foundation, NIH, NSF, Stanford

Howard Chang