Principles of Measurement & Imaging

Including an Overview of Analytical Tools for Synthetic Biology

by Evan R. Daugharthy

30 March, 2021

Introduction Part I:

Why do we need analytical tools in synthetic biology?

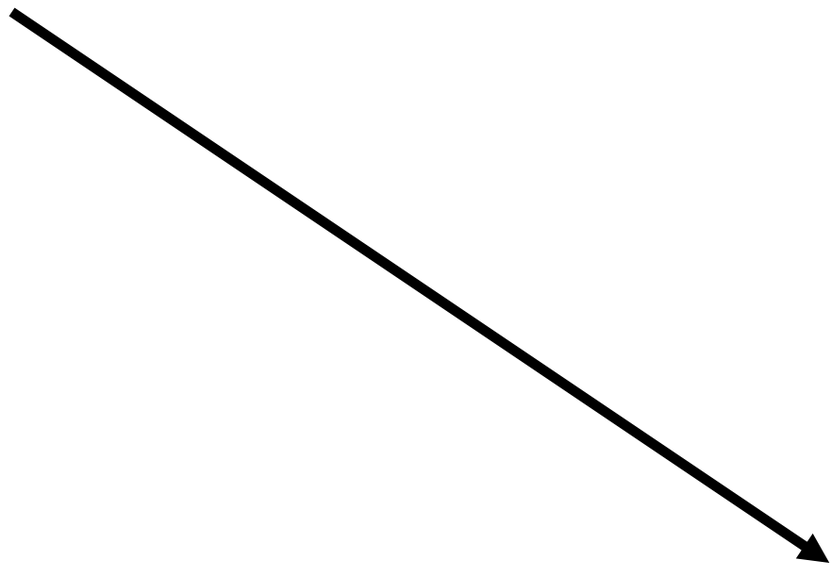
Introduction

Synthetic Biology Goal

Make a molecule via biosynthesis
Engineer an organism
Use DNA for nano-scale assembly
etc.

Introduction

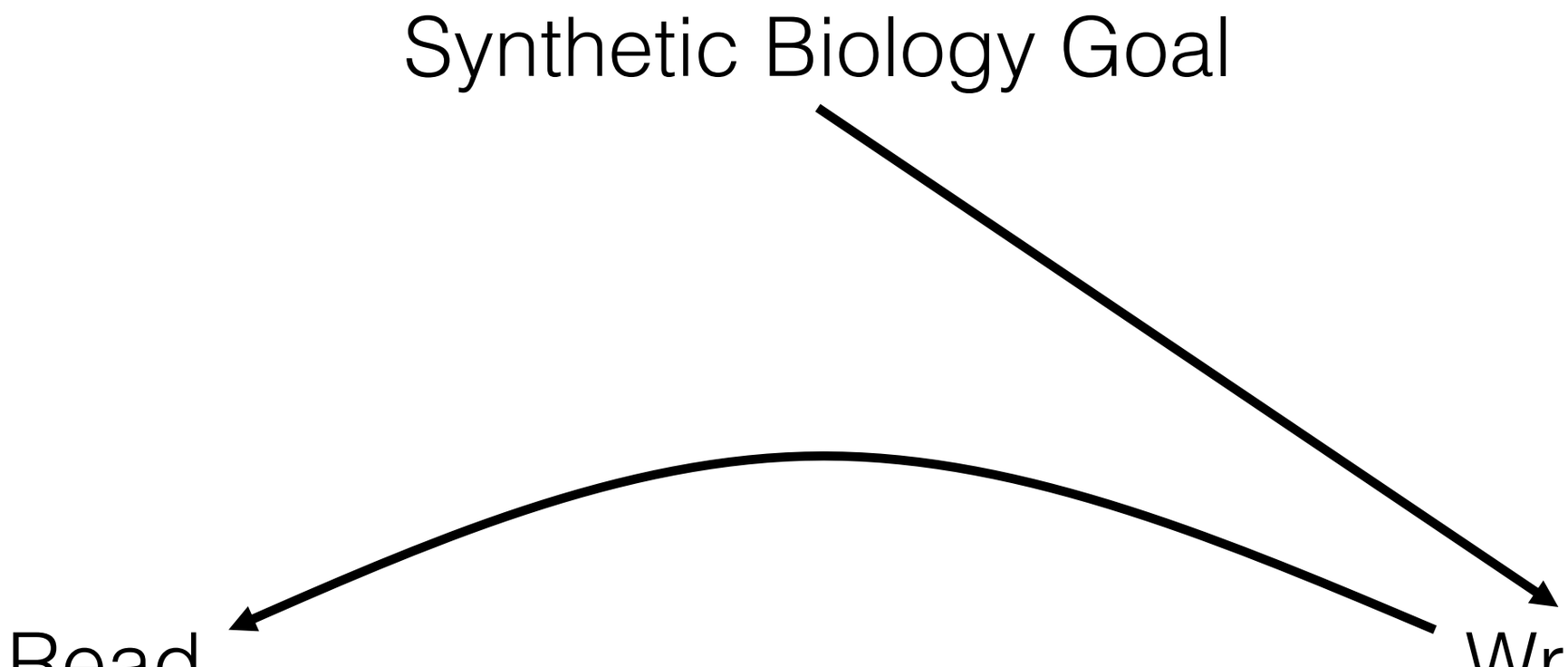
Synthetic Biology Goal



Write

DNA synthesis
CRISPR
DNA origami
metabolic engineering
directed evolution
cell-free system

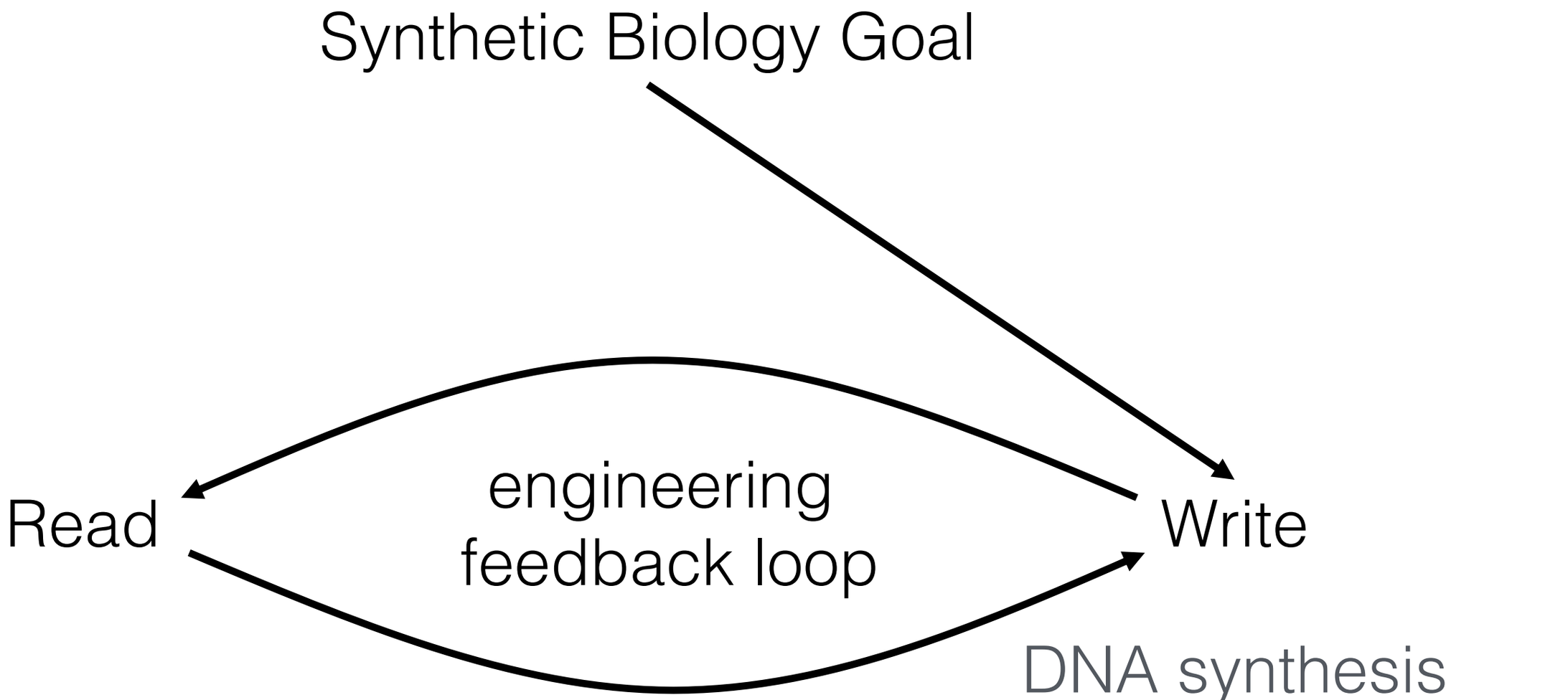
Introduction



proteomics
transcriptomics
genomics
functional assays
sensors

DNA synthesis
CRISPR
DNA origami
metabolic engineering
directed evolution
cell-free system

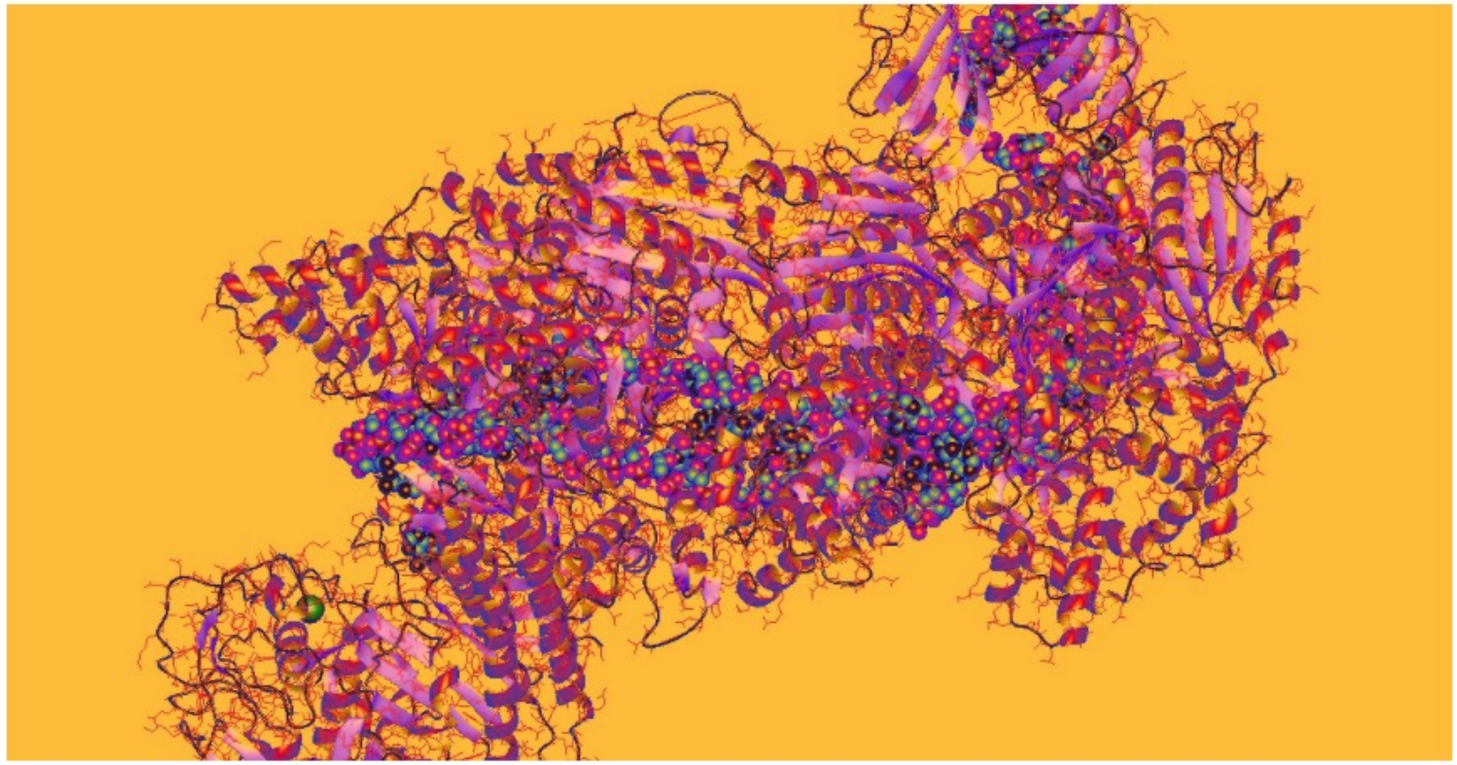
Introduction



proteomics
transcriptomics
genomics
functional assays
sensors

DNA synthesis
CRISPR
DNA origami
metabolic engineering
directed evolution
cell-free system

THE US HAS OFFICIALLY STARTED USING CRISPR ON HUMANS



ASTROJAN/VICTOR TANGERMANN

Evan Daugharthy
evan.daugharthy@gmail.com

De-extinction

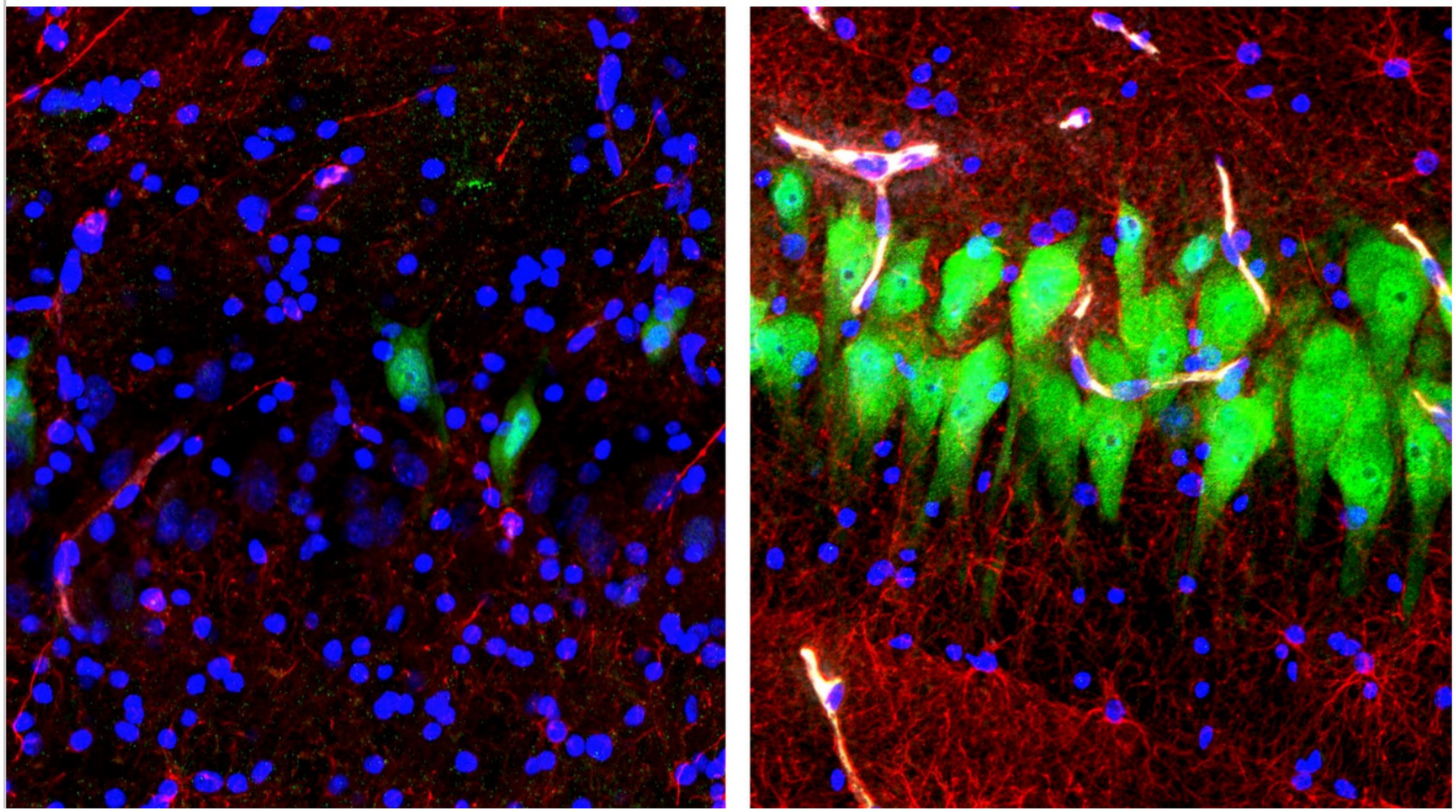
Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Cellular therapies

Scientists Restore Some Function In The Brains Of Dead Pigs

Nell Greenfieldboyce • April 17, 20191:01 PM ET



Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Can measurement save us?

Synthetic biology leap does not have to lead to monstrous outcomes

These technologies must be kept out of the hands of people who would misuse them

about 9 hours ago

Dick Ahlstrom

Follow

0

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Introduction Part II:

Brief history of molecules in biology

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR

WYSS INSTITUTE

Brief history of molecules in biology

1868-71

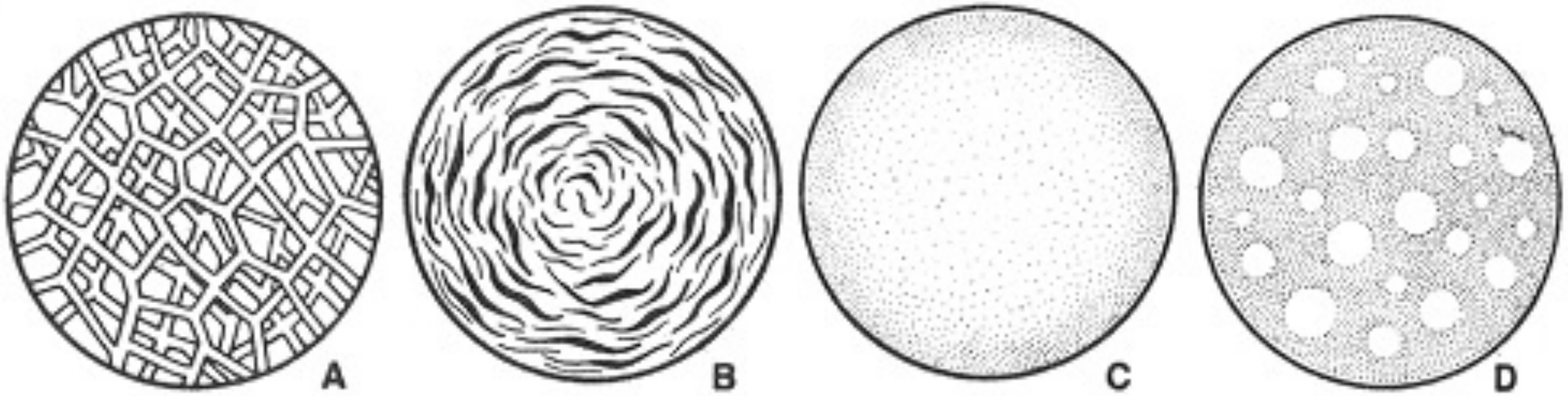
Nucleic Acid

1789

Protein

Brief history of molecules in biology

1880s: Types of biological “matter”



A. Reticular, B. Fibrillar, C. Granular, D. Alveolar

1789

Protein

~~macromolecule~~

1900-1930s: “colloidal theory”

Brief history of molecules in biology

1789

Protein

~~macromolecule~~

1900-1930s: “colloidal theory”

Brief history of molecules in biology

~1940: Two distinct nucleic acids with different properties

1868-71 1789

DNA **Nucleic Acid** **RNA** **Protein**

Brief history of molecules in biology

~1940: Two distinct nucleic acids with different properties


PENICILLIN,
New
Wonder Drug
from Mold

By IRMIS JOHNSON

A GREENISH BLUE mold like the one that grows on stale bread, or lends aroma and flavor to Roquefort cheese, now promises to be an important ally in helping wounded soldiers fight their way back to health.

Brief history of molecules in biology

1938
Fruit fly
Genome
Map

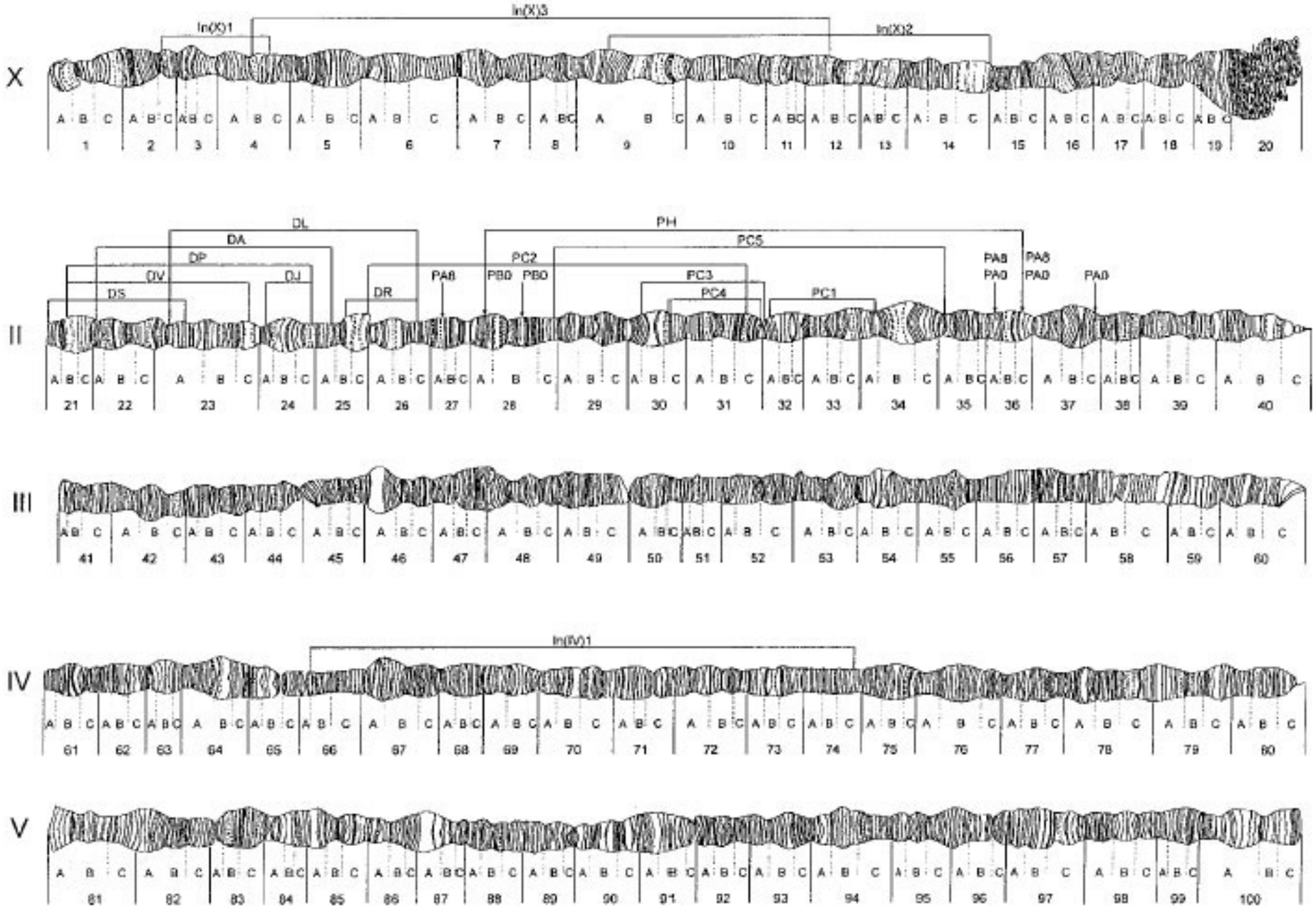
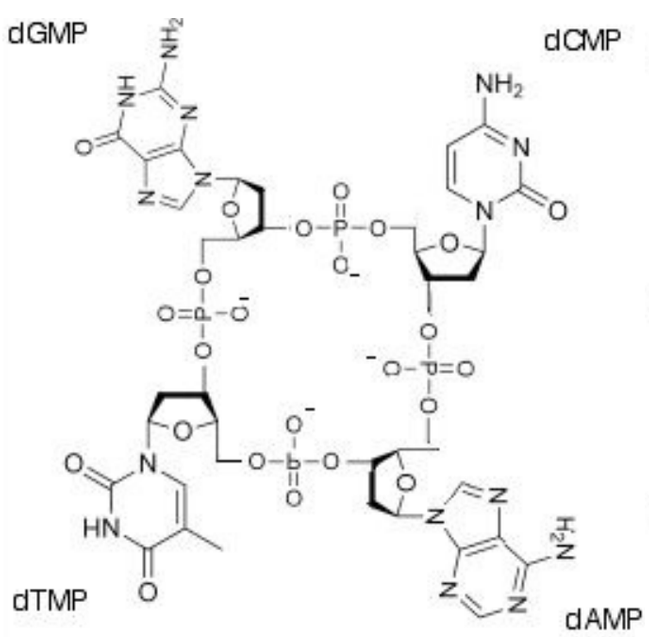


Fig. 1: polytene chromosome map of *Drosophila mediopunctata* with inversion breakpoints presented. Centromeres are shown to the right, telomeres to the left.

Brief history of molecules in biology

Tetranucleotide Hypothesis $(AGCT)_n$

DNA



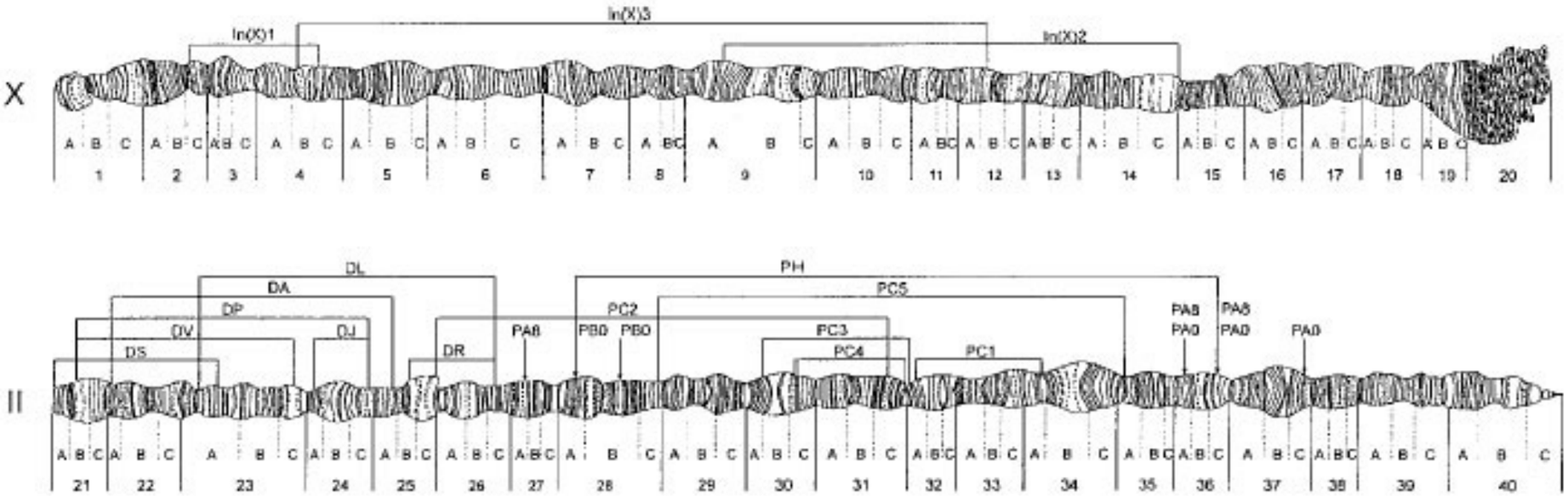


Fig. 1: polytene chromosome map of *Drosophila mediopunctata* with inversion breakpoints presented. Centromeres are shown to the right, telomeres to the left.

Brief history of molecules in biology

1944: Avery, MacLeod, McCarty
"the transforming activity… is actually an inherent property of the nucleic acid"

DNA

Elementary Chemical Analysis of Purified Preparations of the Transforming Substance

Preparation No.	Carbon	Hydrogen	Nitrogen	Phosphorus	N/P ratio
	per cent	*per cent*	*per cent*	*per cent*	
37	34.27	3.89	14.21	8.57	1.66
38B	—	—	15.93	9.09	1.75
42	35.50	3.76	15.36	9.04	1.69
44	—	—	13.40	8.45	1.58
Theory for sodium desoxyribonucleate.....	34.20	3.21	15.32	9.05	1.69

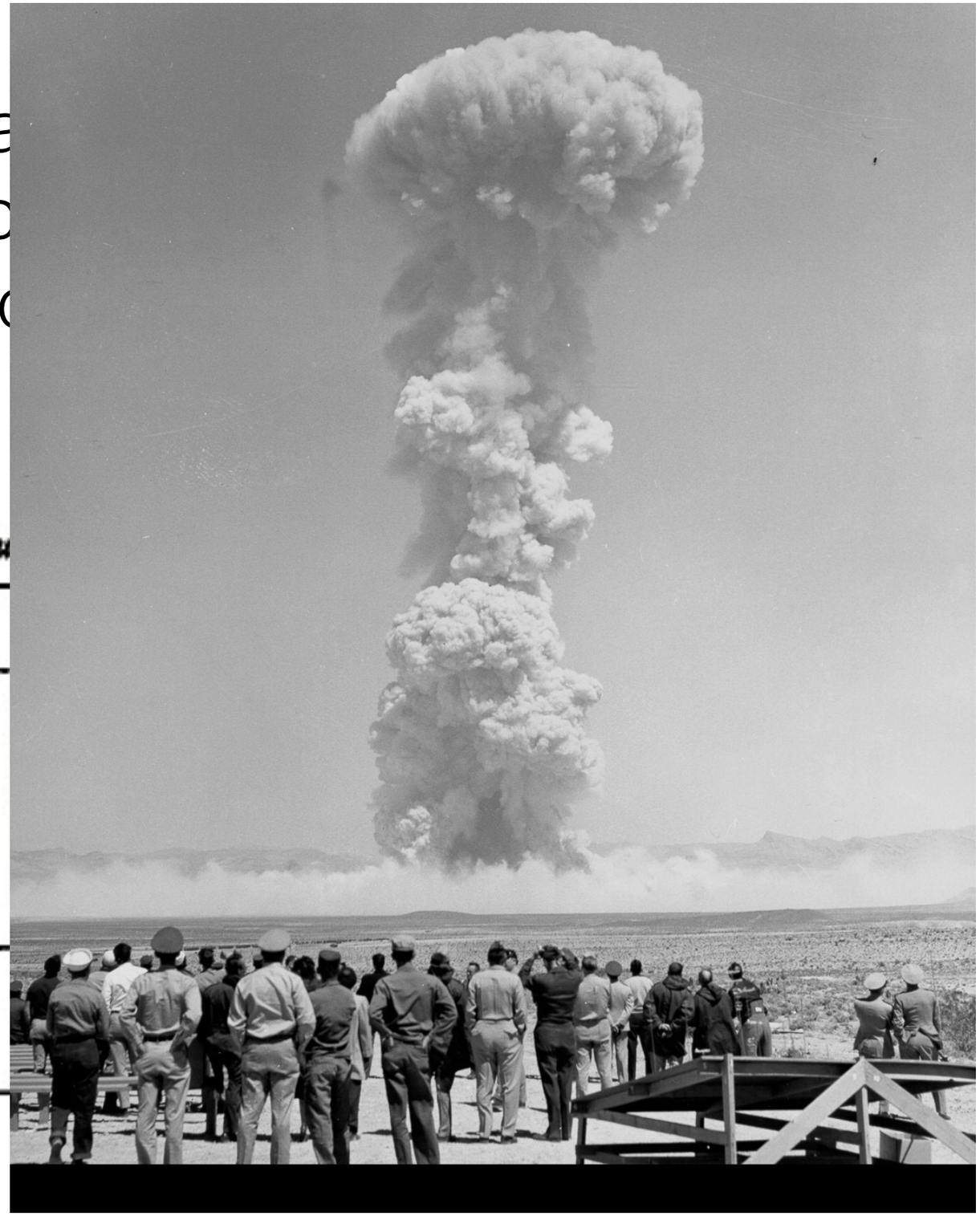
Brief history of molecules in biology

1944: Avery, Ma
"the transforming ac
inherent property o

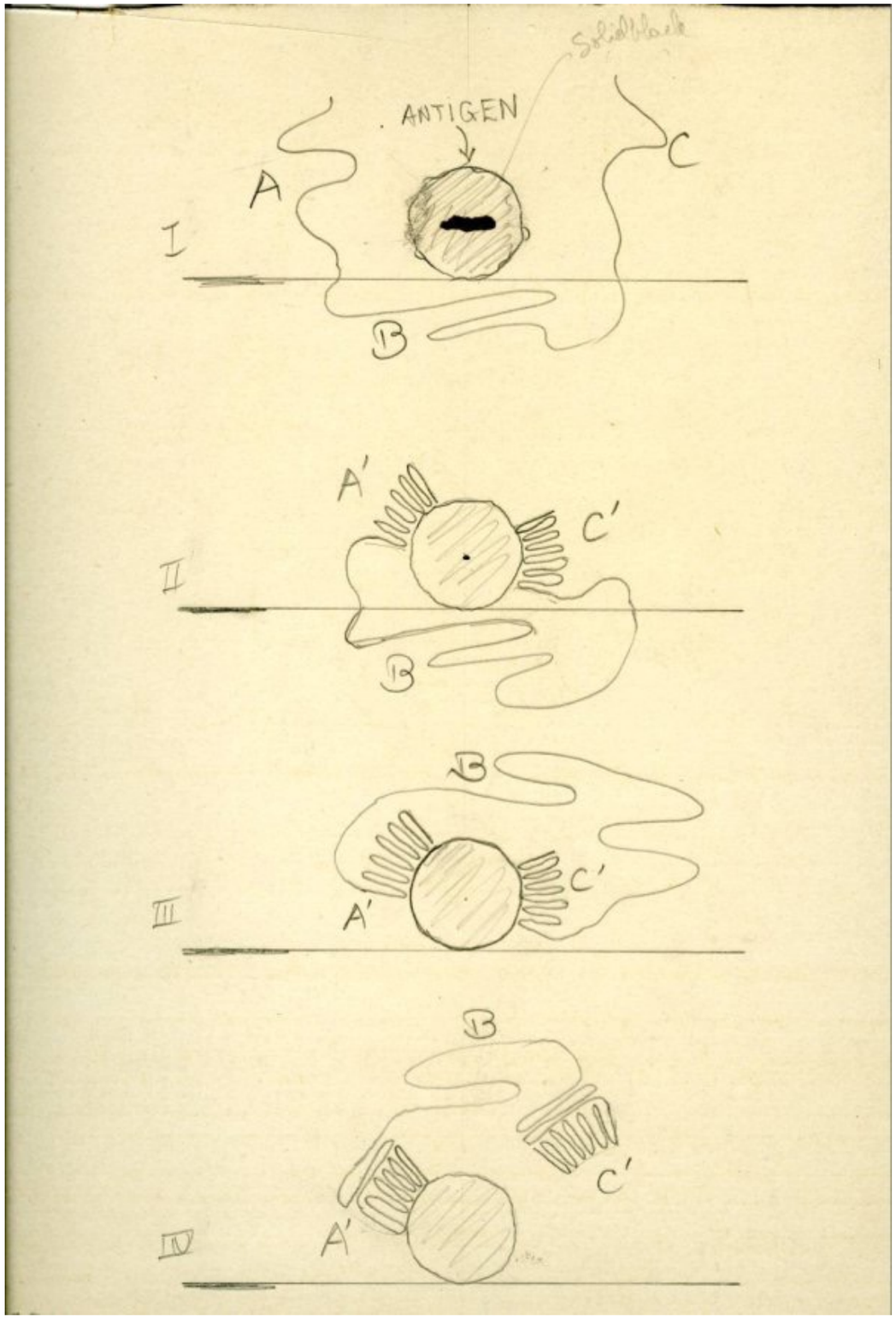
DNA

Elementary Chemical Analysis of Pu

Preparation No.	Carbon
	per cent
37	34.27
38B	—
42	35.50
44	—
Theory for sodium desoxyribonucleate.....	34.20



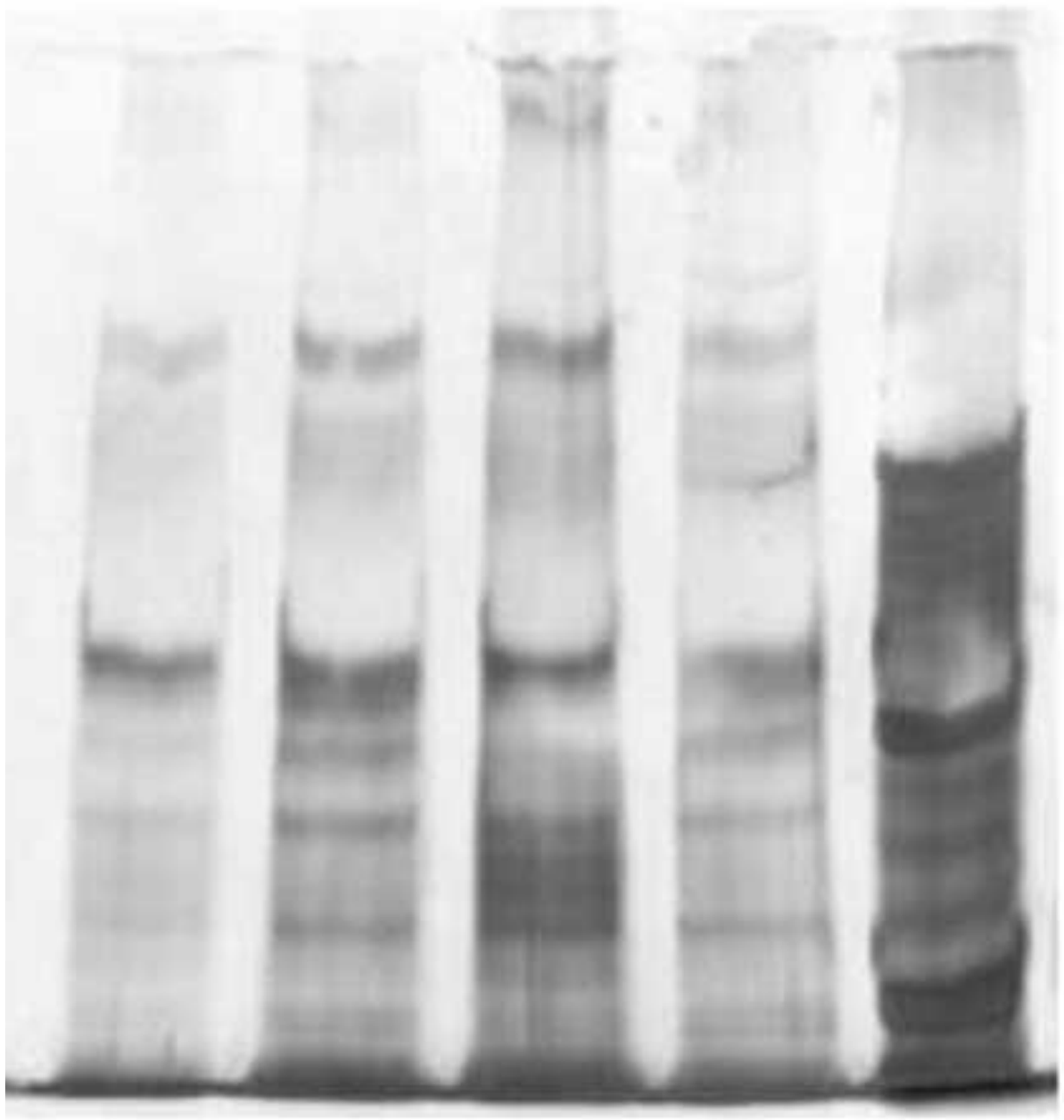
Brief history of molecules in biology



1940: Pauling hypothesizes all antibodies have same sequence

Protein

Brief history of molecules in biology



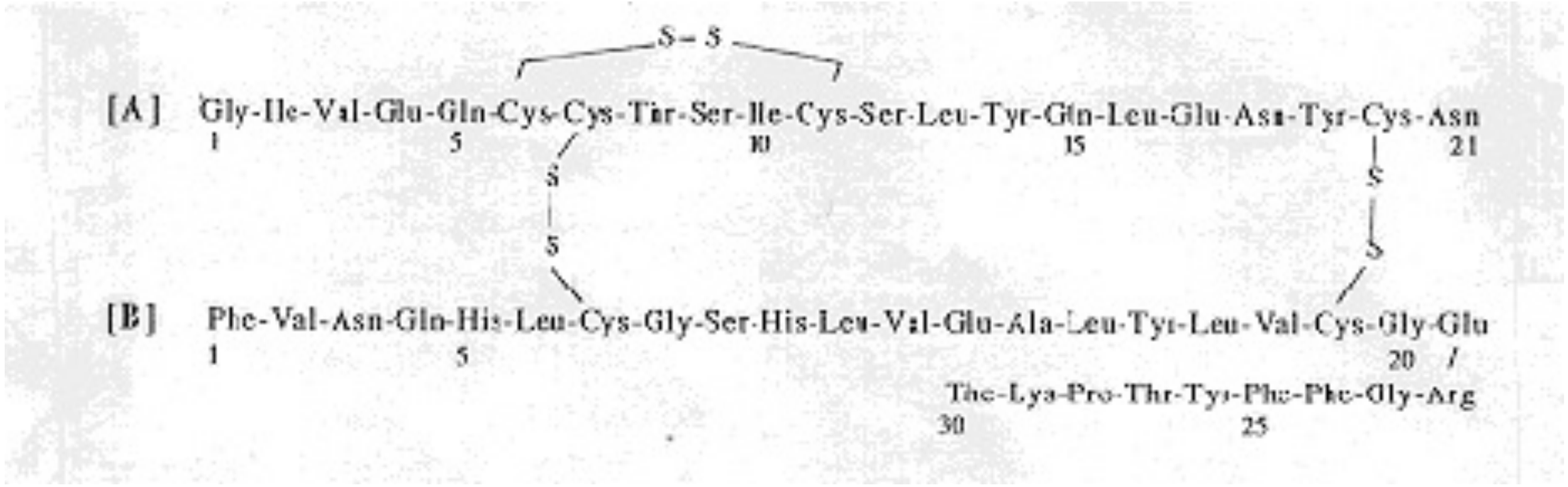
1948: Tiselius
"… substances are more complex than was originally supposed."

Protein

Brief history of molecules in biology

1951-53: Sanger sequences insulin

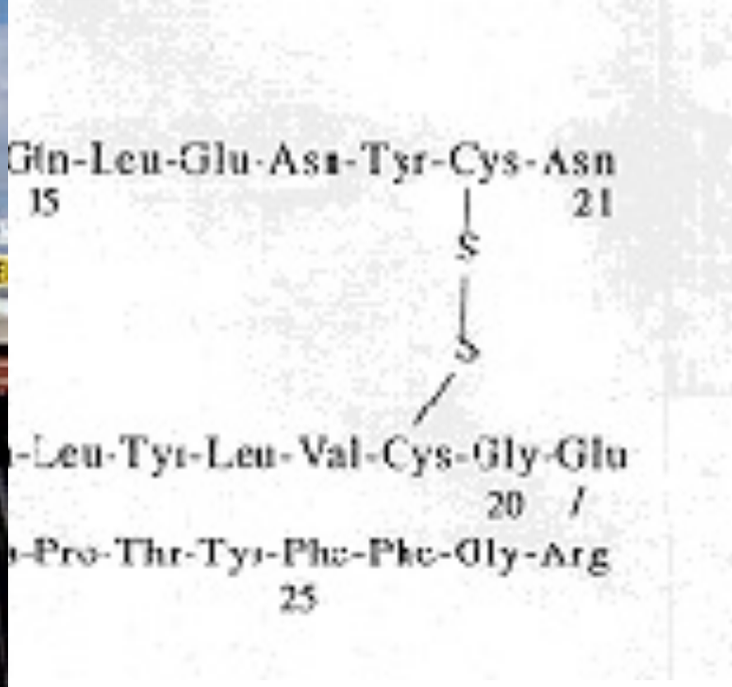
Protein



Brief history of molecules in biology

1951-53: Sanger sequences insulin

Protein

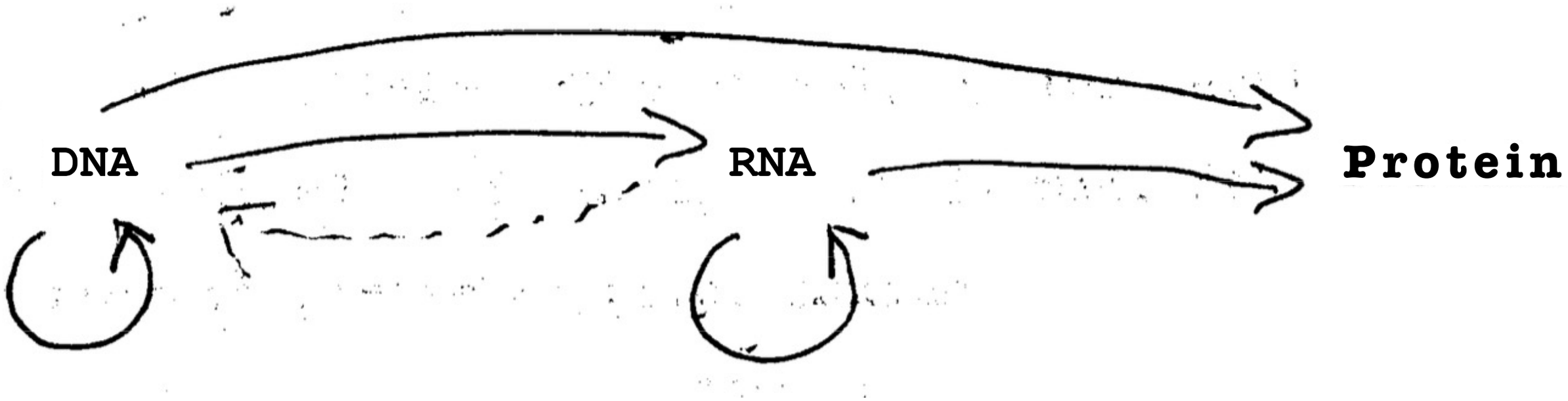


Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

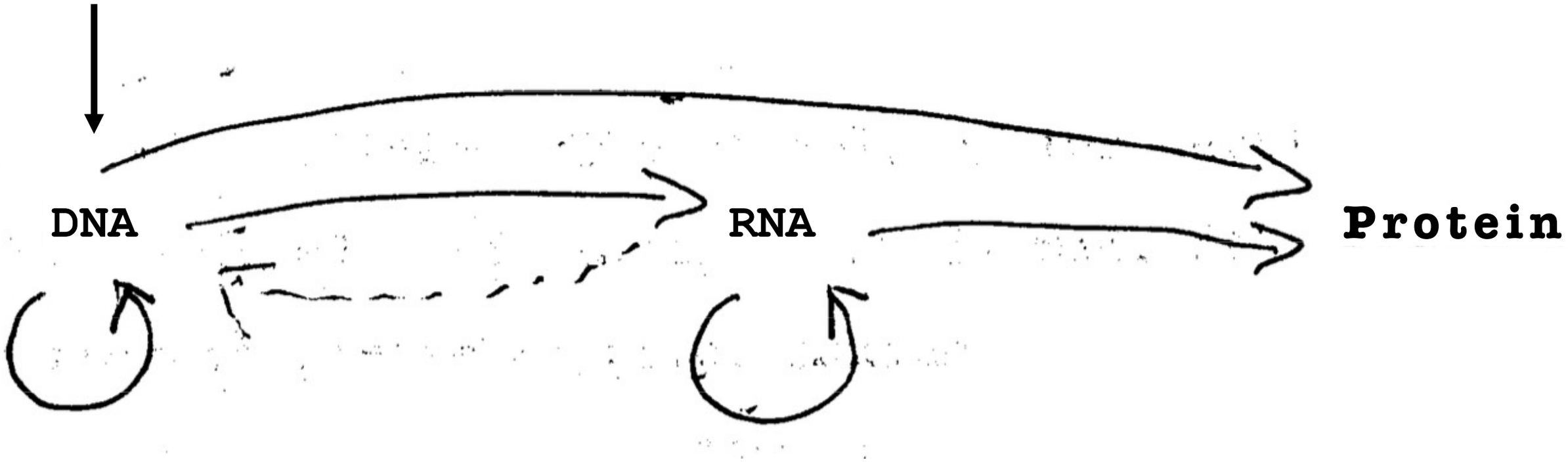
Brief history of molecules in biology

1956-60s: Central dogma of molecular biology



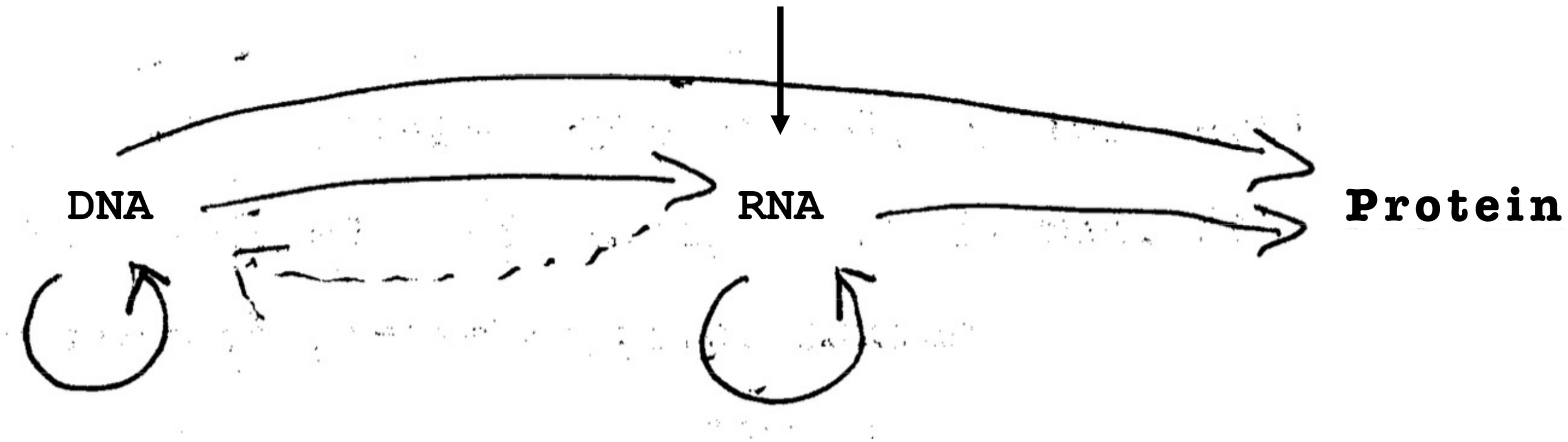
Brief history of molecules in biology

1970-80s: Histone modifications on chromatin
2000: Strahl and Allis' "Histone Code"



Brief history of molecules in biology

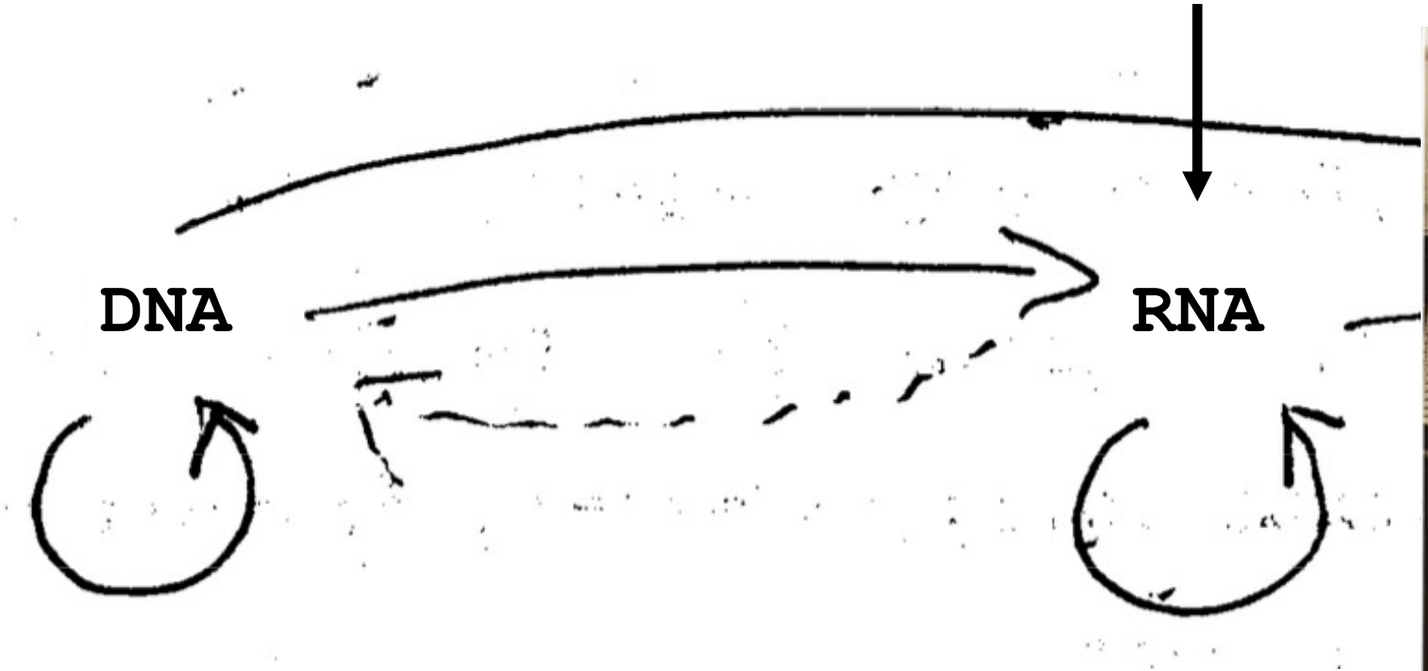
1977: Introns & RNA splicing
1986: RNA editing
1993: miRNA



Early 2000s: Whole genome transcription

Brief history of molecules in biology

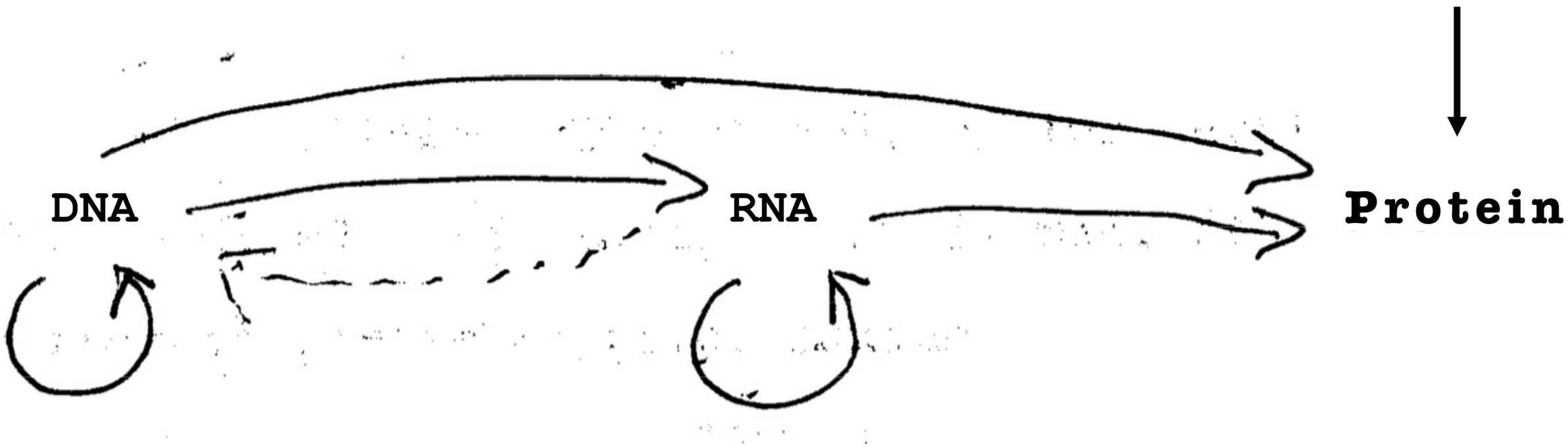
1977: Introns & RNA splicing
1986: RNA editing
1993: miRNA



Early 2000s: Whole genome transcription

Brief history of molecules in biology

1970-80s: Phosphorylation appreciated
Last 30 years >200 PTMs



Considerations for measurement

Biological systems are composed of many different molecules, all organized in space

Genome	Sequence (3 billion bp) Modifications (100 million methylation sites) Nucleosome (160 million histones) Organization (chromatin state, 2-20 Kb)
Transcriptome	Transcripts (75,000 species) Number (expression level) Splicing (70,000 splice junctions) Editing (2,000 A>I) Localization (regulation)
Proteome	Genes (20,000, >1m protein species) Number (expression level) Modifications (>20, millions of sites) Localization (function)

Towards Perfect Molecular Measurement

What is a measurement?

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR

WYSS INSTITUTE

Measurement Theory

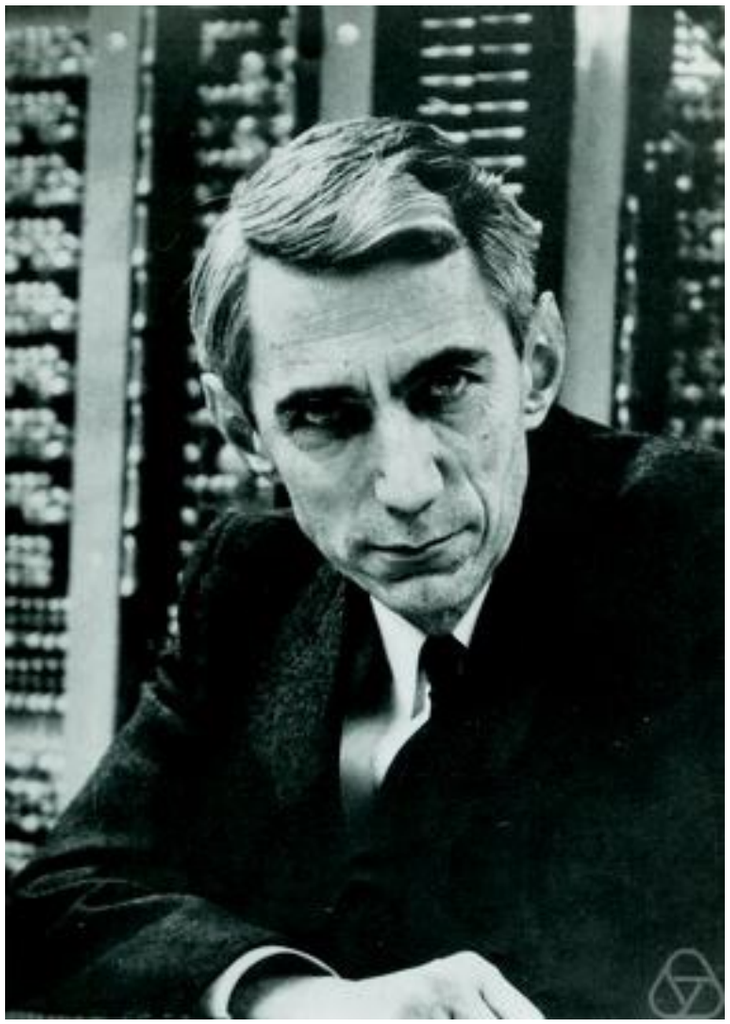
Measurement is an activity that involves interaction with a system with the aim of representing aspects of that system in abstract terms (classes, numbers)

Evan Daugharthy
evan.daugharthy@gmail.com

Measurement Theory

Information-Theoretic Account of Measurement

Claude Shannon 1916-2001



Entropy of an information source

$$H = -\sum_i p_i \log_2(p_i)$$

Evan Daugharthy
evan.daugharthy@gmail.com

Measurement Theory

Information-Theoretic Account of Measurement

Biological System

Information

Shannon: Source of information is anything with more than one state that can be realized.

Biological information = physical composition and localization of (all the) molecules

Evan Daugharthy
evan.daugharthy@gmail.com

Measurement Theory

Information-Theoretic Account of Measurement

Biological System

Shannon: Source of information is anything with more than one state that can be realized.

Biological information = physical composition and localization of (all the) molecules

“Analyte”
The entity to be subjected to measurement (in the sample)

“Measurand”
The quantity you intend to measure

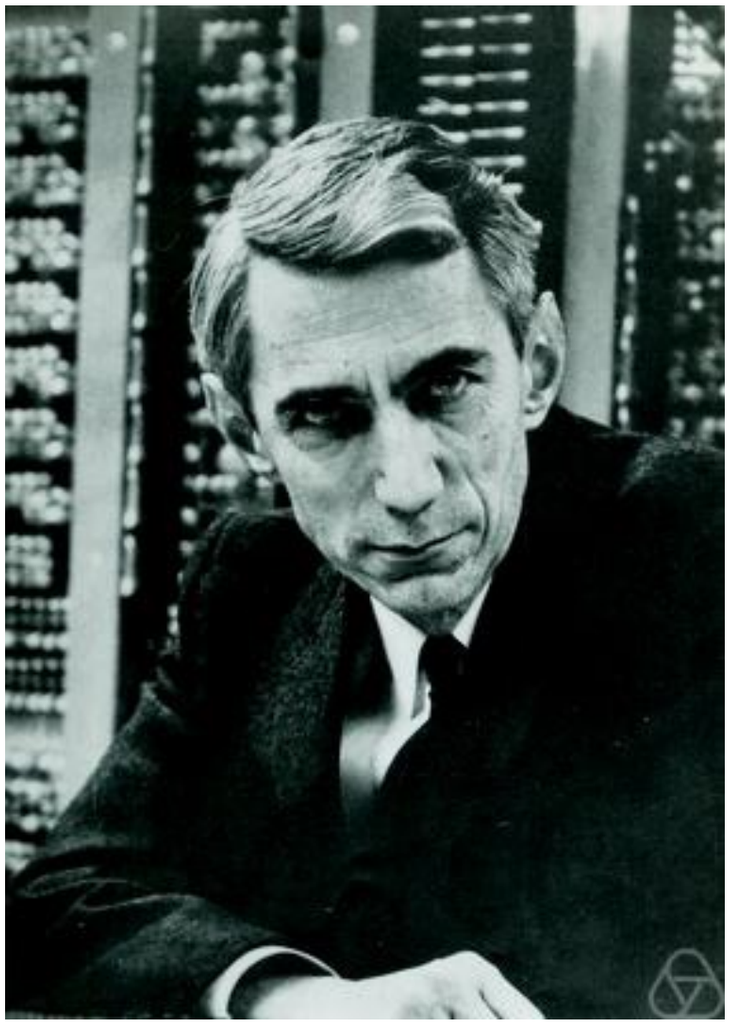
Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Measurement Theory

Information-Theoretic Account of Measurement

Claude Shannon 1916-2001



Mutual information

$$I(x', y') = \log_2 \frac{P(x'y')}{P(x')P(y')}$$

$$= \log_2 \frac{P(x'|y')}{P(x')}$$

$$= \log_2 \frac{P(y'|x')}{P(y')}$$

Measurement technologies are "information machines"

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Measurement Theory

"True Value"
Biological System

Information

Measurement Technology

Encoder

Message

Transmission

Decoder

Error
Resolution
Sensitivity
Specificity

"Measured Value"
Scientist

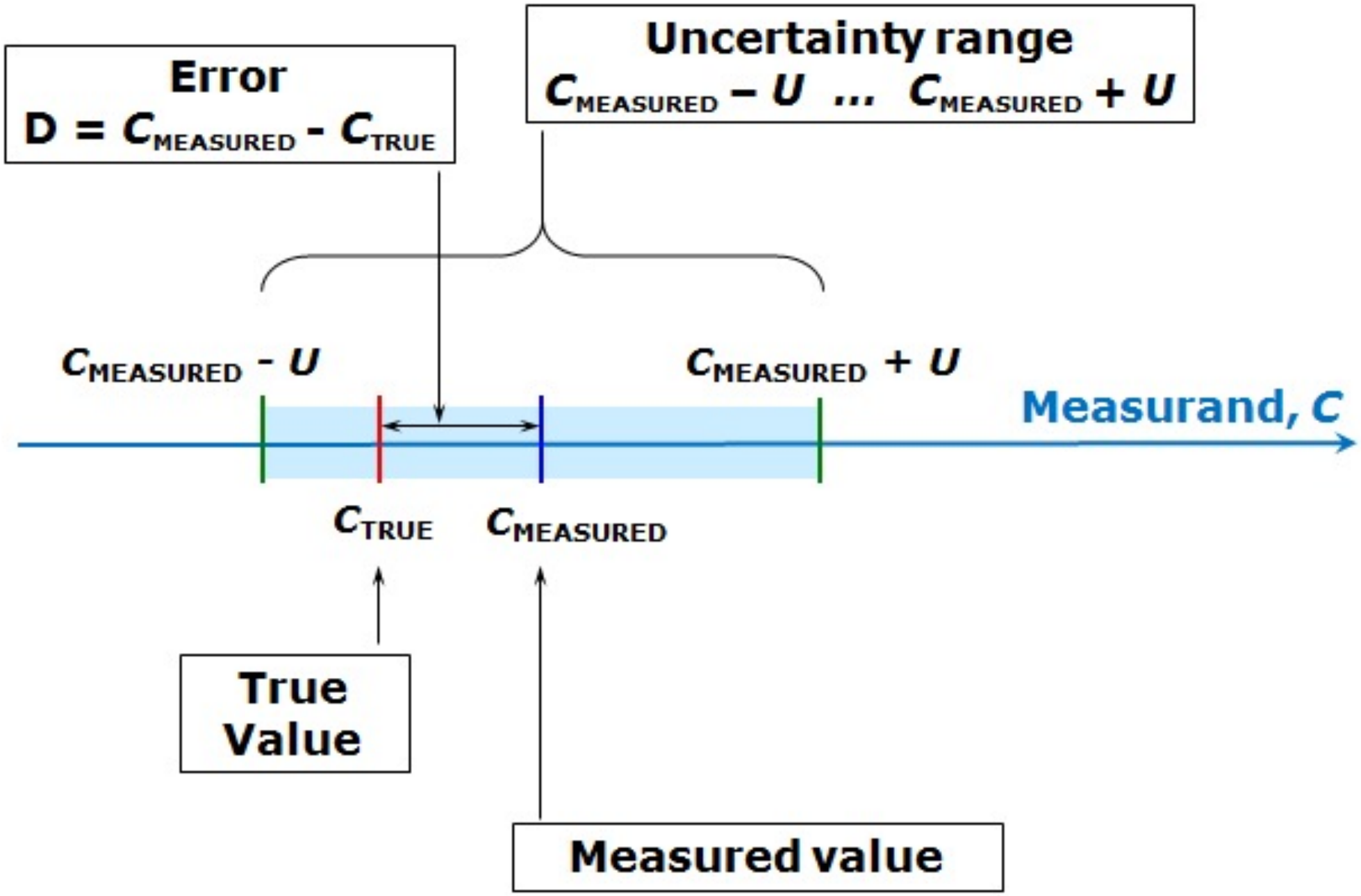
Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR

WYSS INSTITUTE

Measurement Theory

The difference between the measured value and the true value is called error. Error can have either positive or negative sign.



Evan Daugharthy
evan.daugharthy@gmail.com

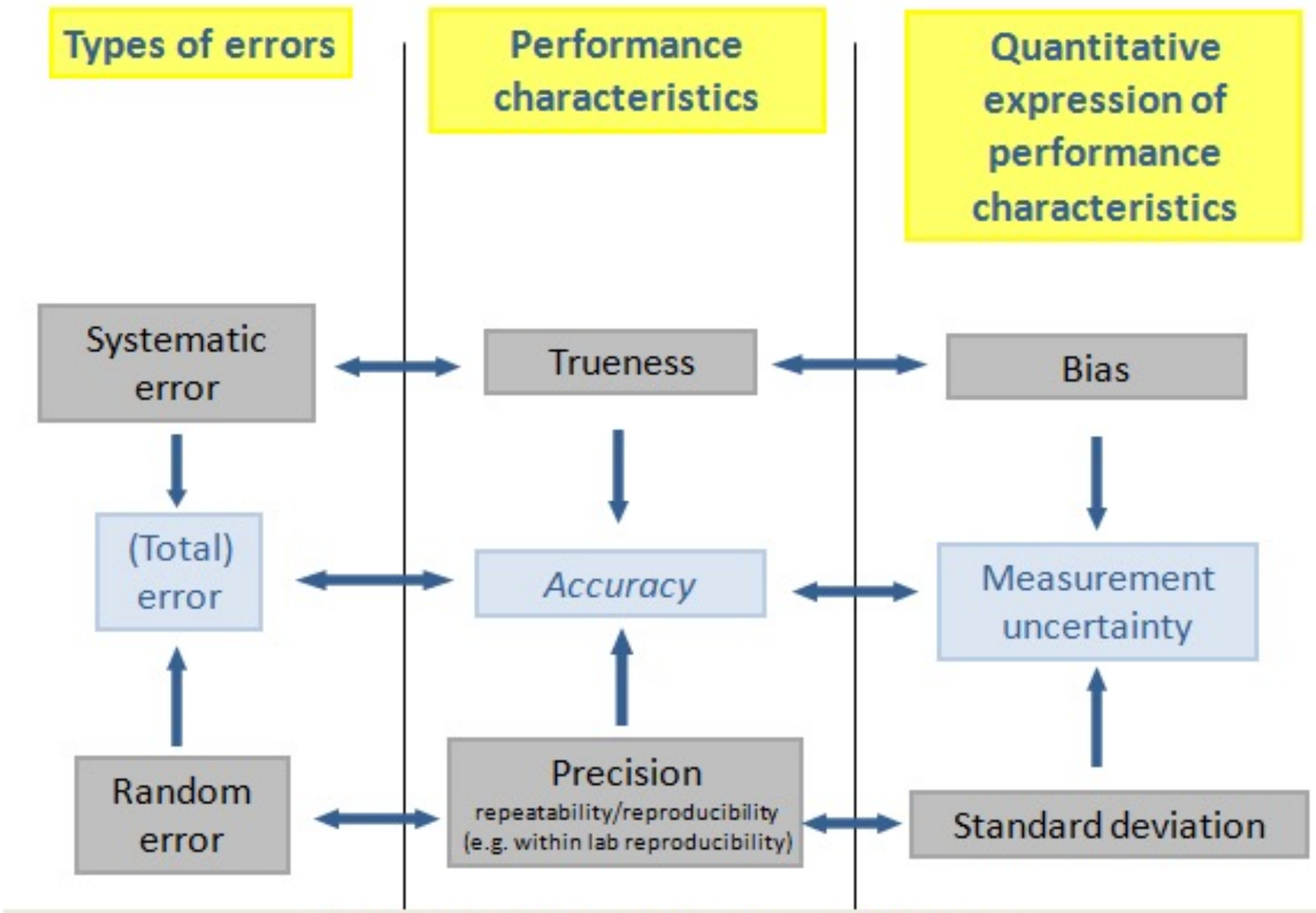
READCOOR WYSS INSTITUTE

Errors

Error can be divided into two parts:

Random error - having different magnitude and sign in the case of repeated measurements

Systematic error - having the same or systematically changing magnitude and sign in the case of repeated measurements



A. Menditto, et al *Accred. Qual. Assur.* 2006, *12*, 45

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Errors

There are several approaches to estimating measurement uncertainty

Guide to the expression of uncertainty in measurement (GUM)

Within-lab validation (Nordtest)

Lots of rules to follow in order to estimate & handle errors correctly, which are worth learning!

Evan Daugharthy
evan.daugharthy@gmail.com

Measurement Theory

"True Value"
Biological System

Information

Measurement Technology

Encoder

Message

Transmission

Decoder

Error
Resolution
Sensitivity
Specificity

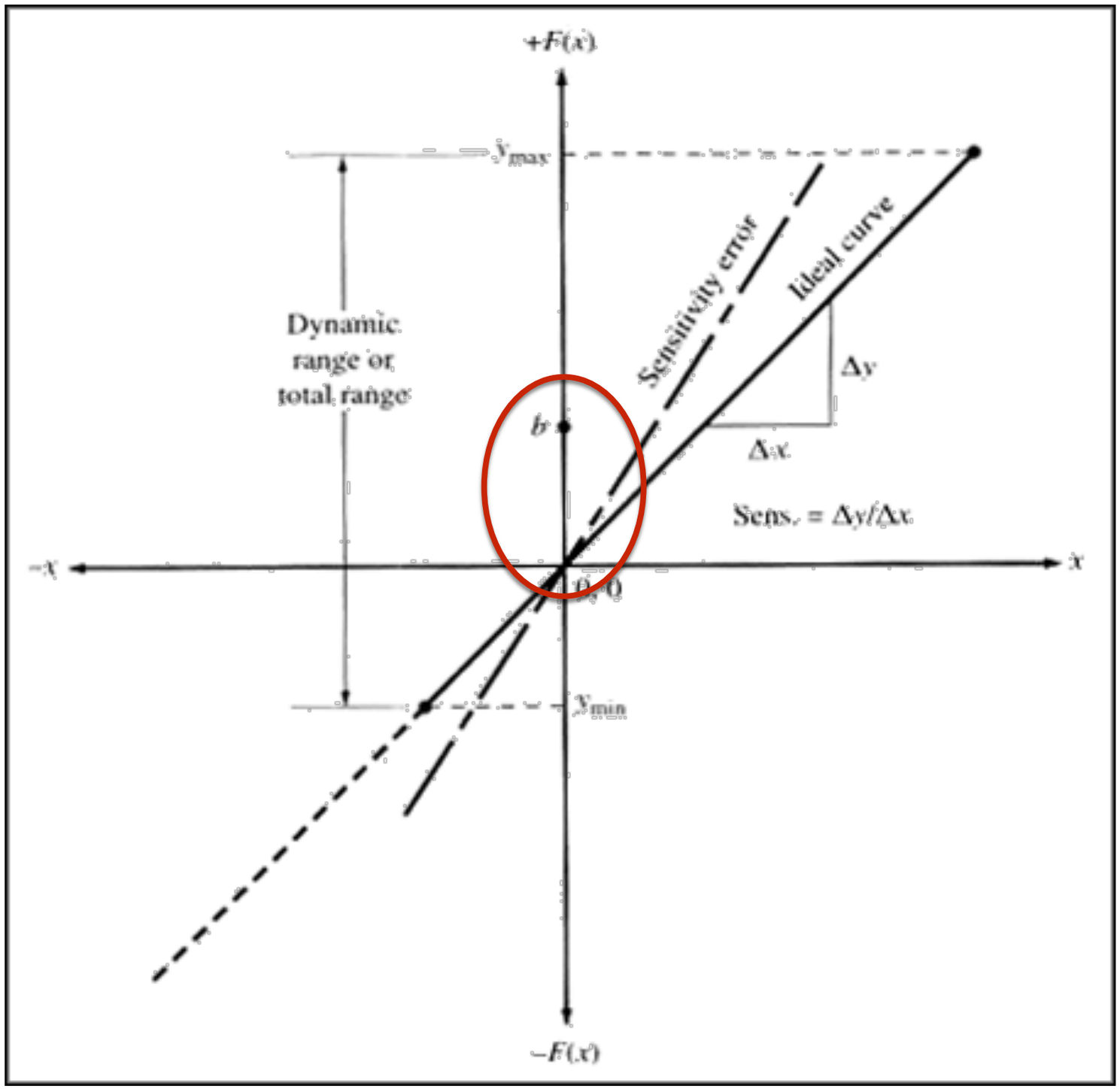
"Measured Value"
Scientist

Evan Daugharthy
evan.daugharthy@gmail.com

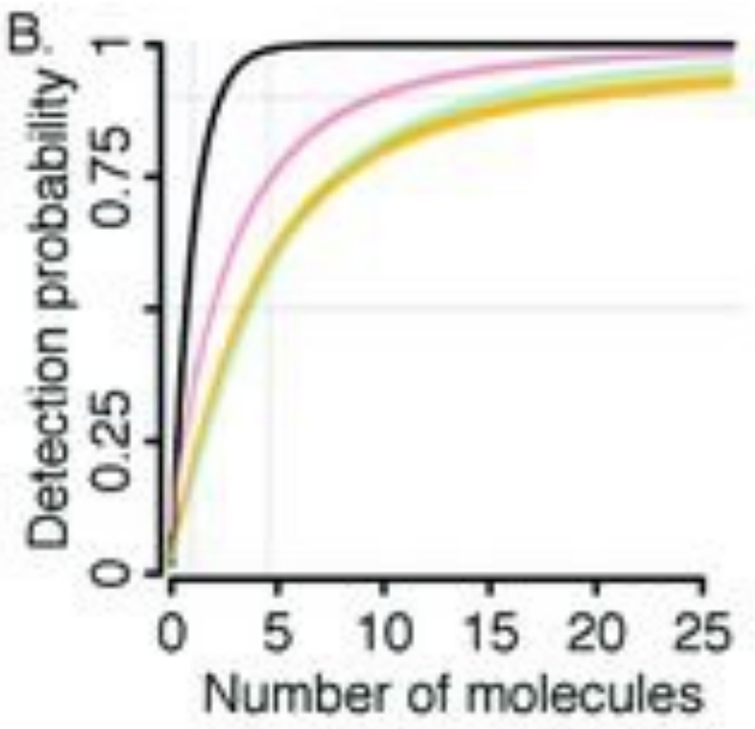
READCOOR

WYSS INSTITUTE

Resolution & Sensitivity



Sensitivity is defined by a response curve relating the input to output



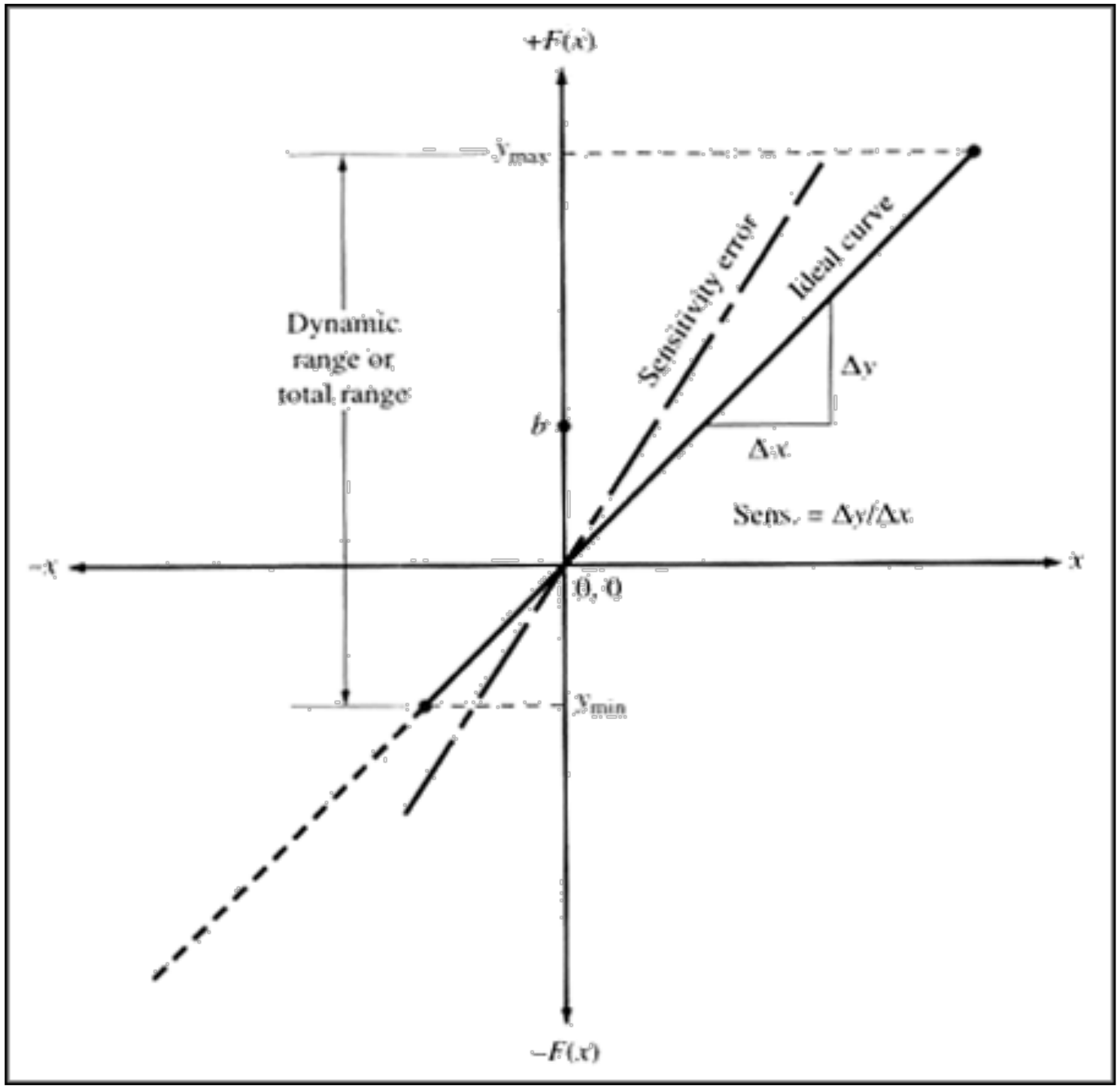
Dueck, Hannah R., et al. "Assessing the measurement transfer function of single-cell RNA sequencing." *bioRxiv*(2016): 045450.

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR

WYSS INSTITUTE

Resolution & Sensitivity



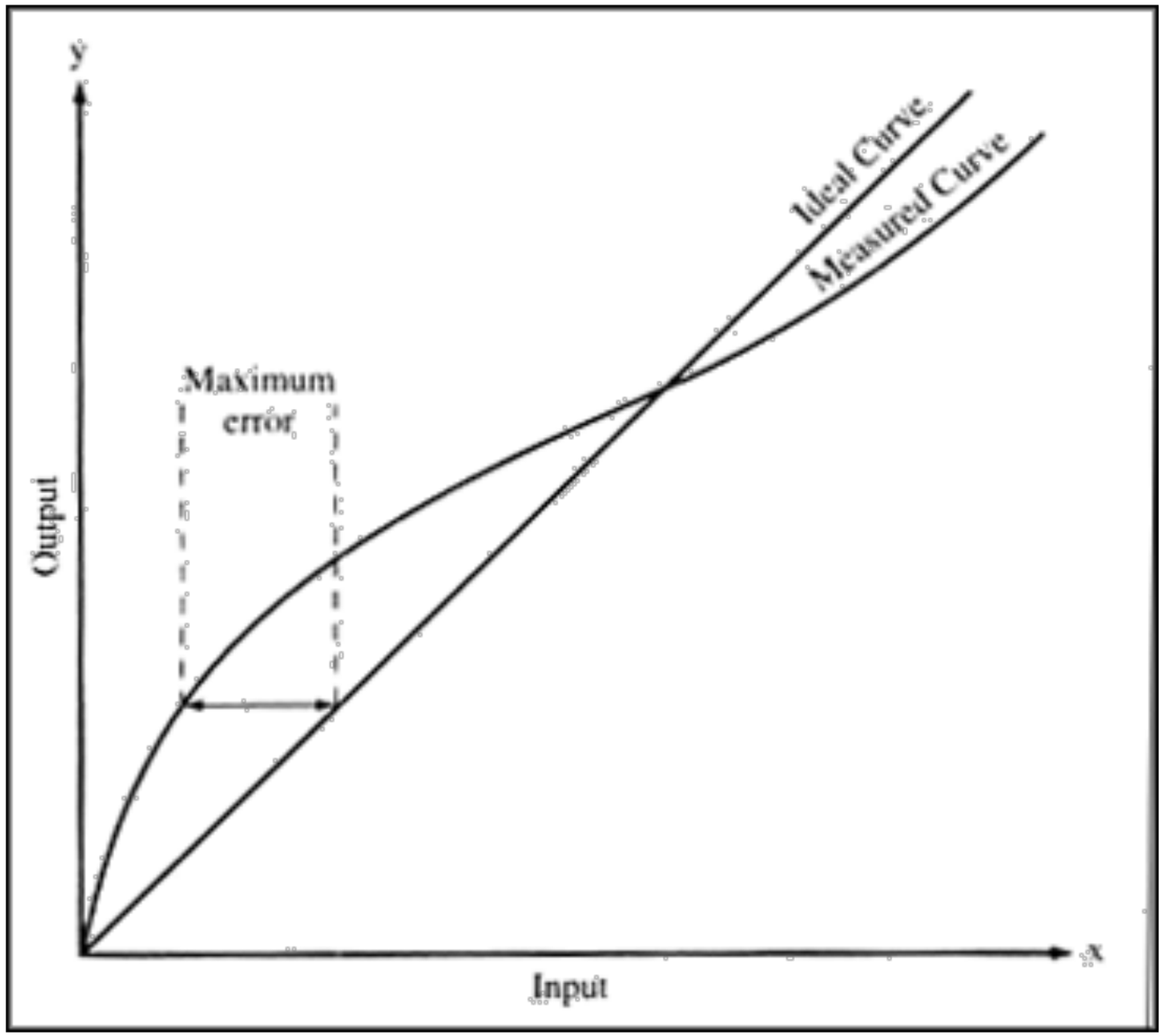
Dynamic Range is the total range of detection

Resolution is the smallest detectable increment change of input that can be detected in output

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Other factors

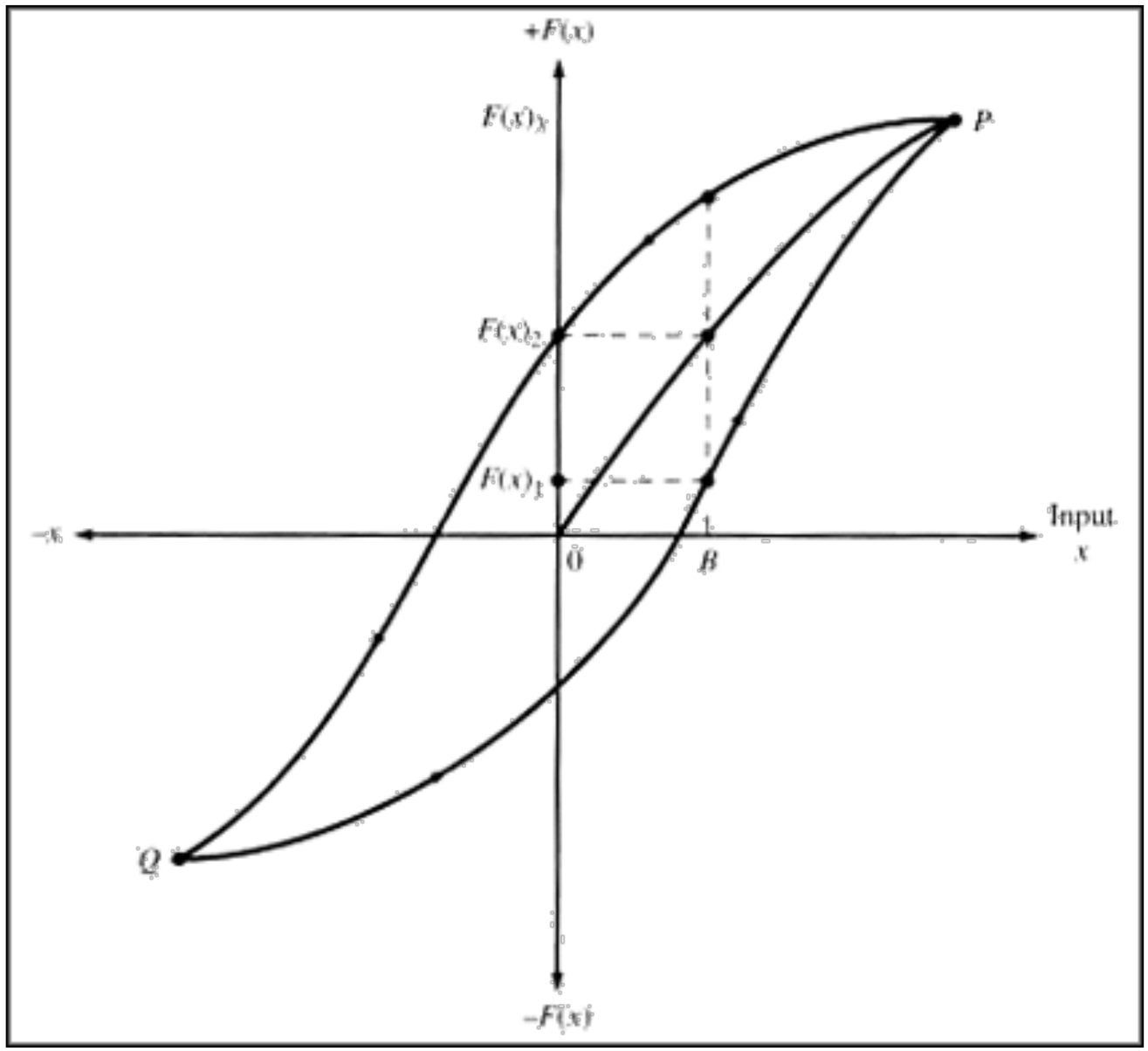


Linearity – are changes to input & output related in a linear function?

Hysteresis – dependence of a state on its history, e.g., direction of change

Evan Daugharthy
evan.daugharthy@gmail.com

Other factors

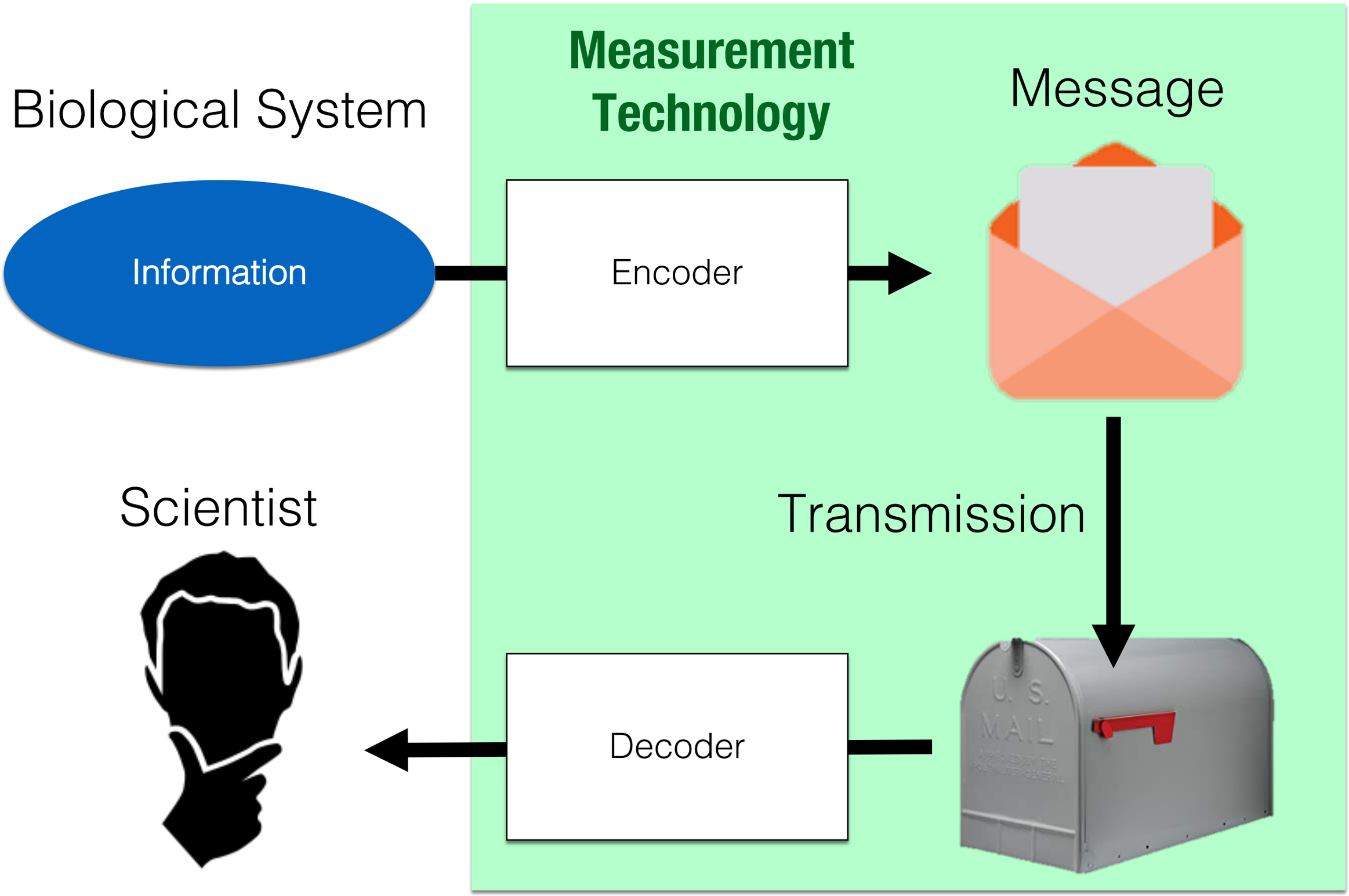


Linearity – are changes to input & output related in a linear function

Hysteresis – dependence of a state on its history, e.g., direction of change

Evan Daugharthy
evan.daugharthy@gmail.com

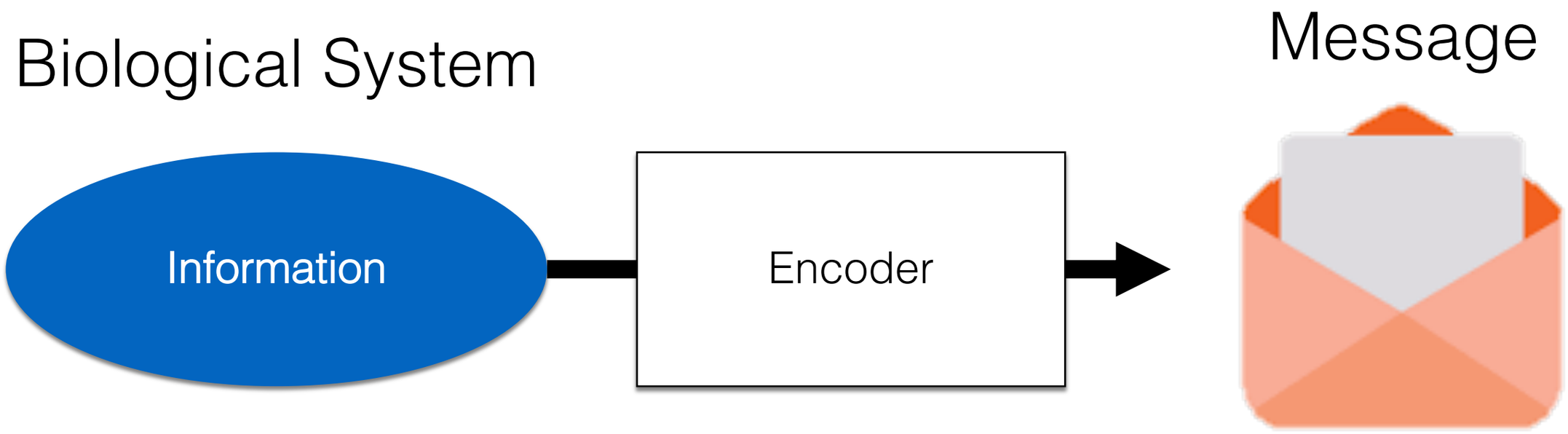
How to Build a Measurement Technology



Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

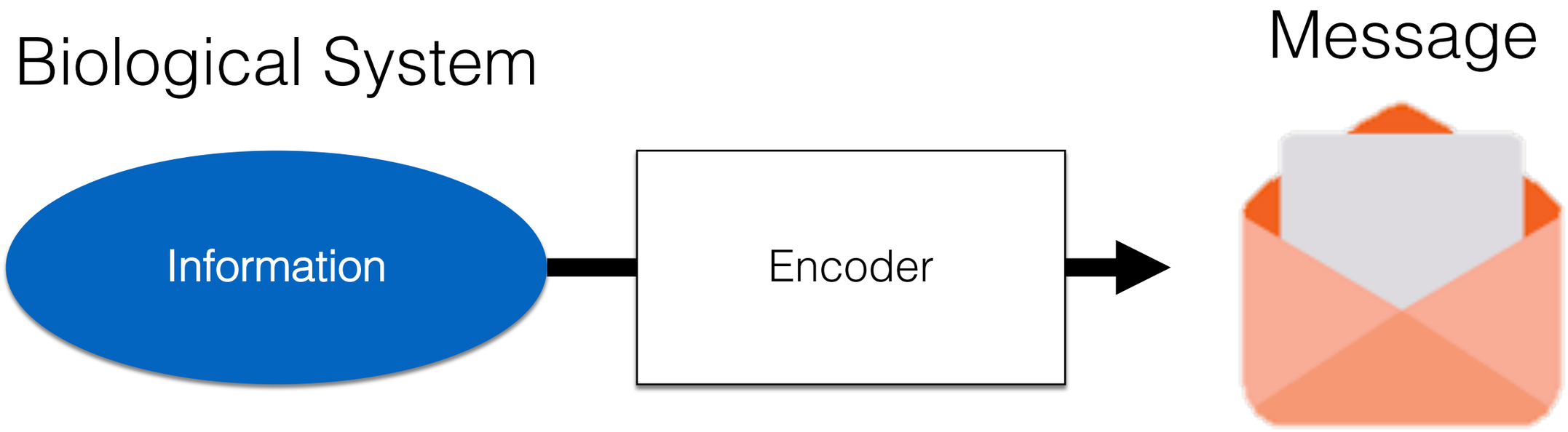
How to compose a message



Sometimes the information is the message!
For example, the information contained in the arrangements of bonds and chemical groups gives rise to physical signals, such as by the interaction with light.

Evan Daugharthy
evan.daugharthy@gmail.com

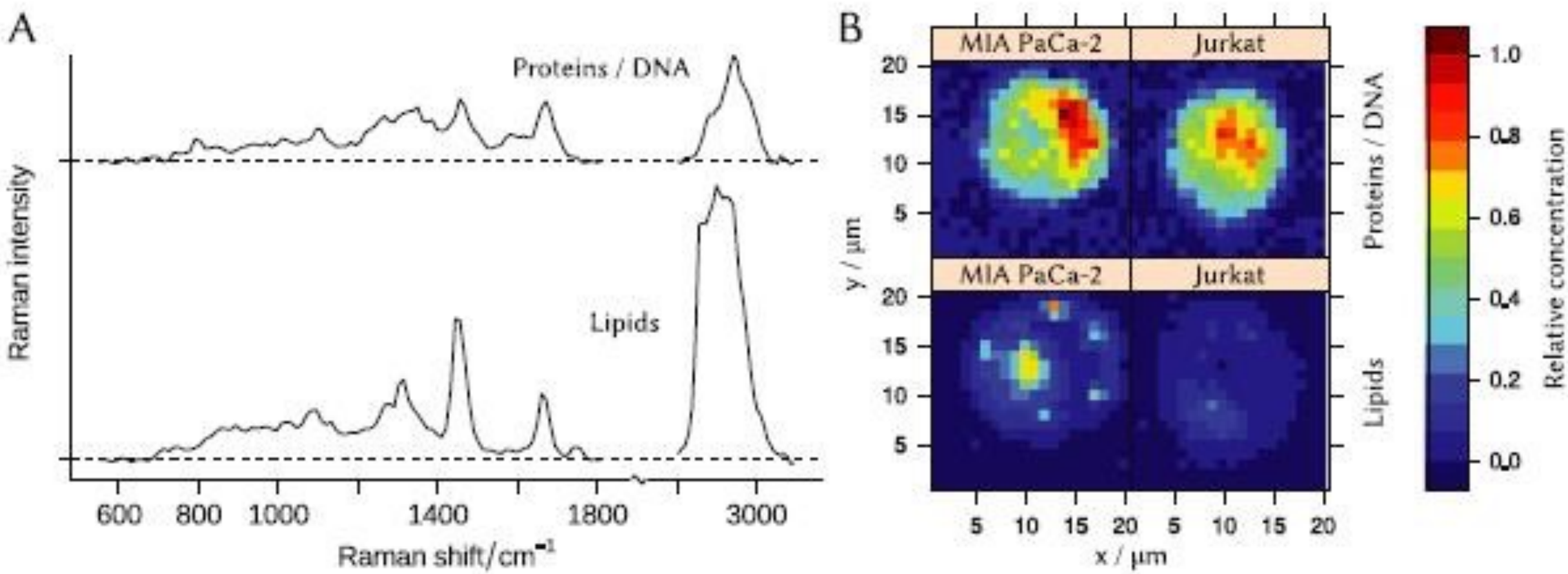
How to compose a message



- Detection of intrinsic size, weight, or charge
- Raman spectroscopy (intrinsic vibrational frequencies of chemical bonds)
- Electrical conductance

Evan Daugharthy
evan.daugharthy@gmail.com

How to compose a message

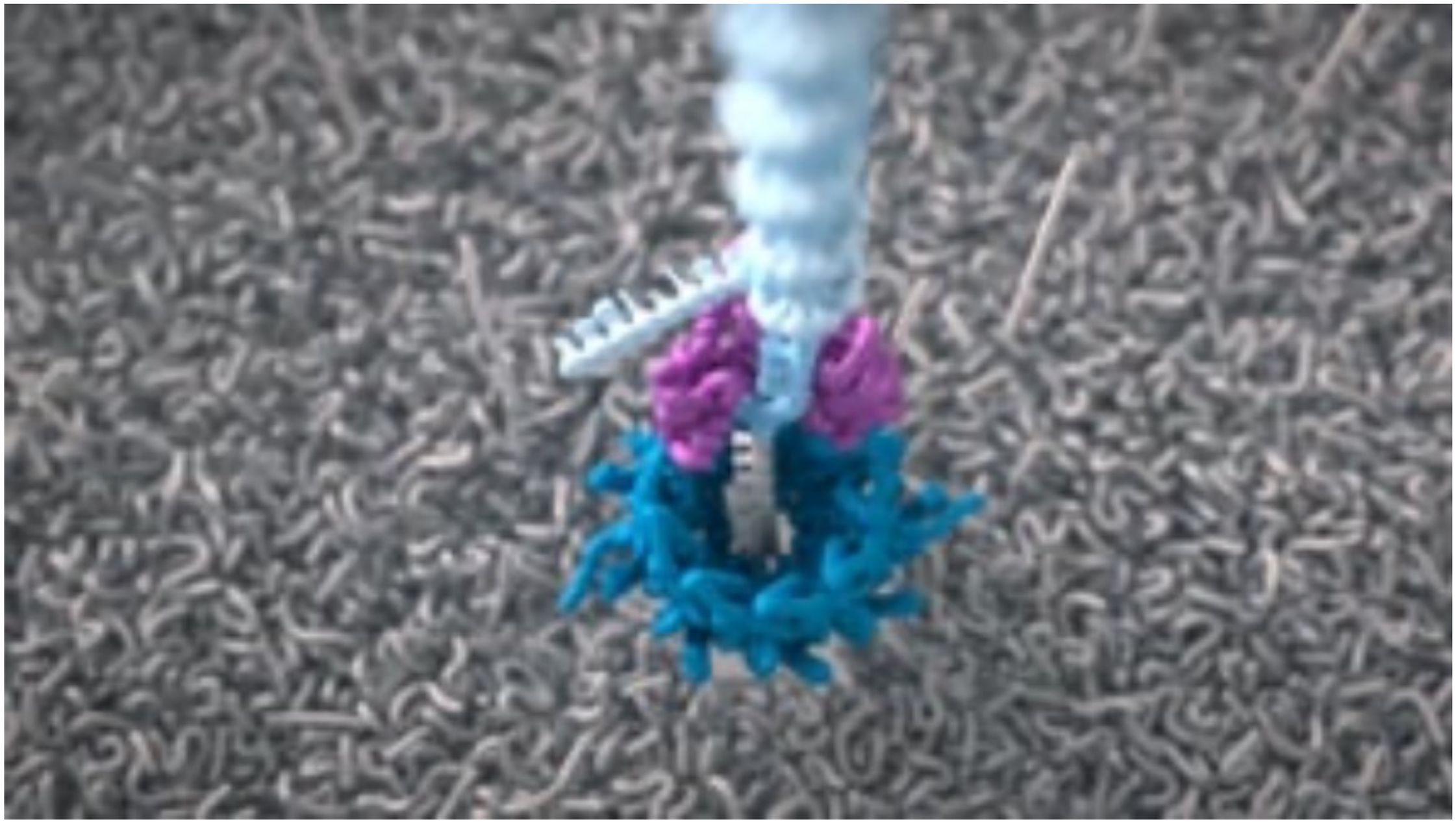


Unfortunately, most biomolecules are composed of a very limited set of particular bonds & chemical groups, which limits specificity & multiplexing

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Example of Direct Measurement: Nanopore Sequencing

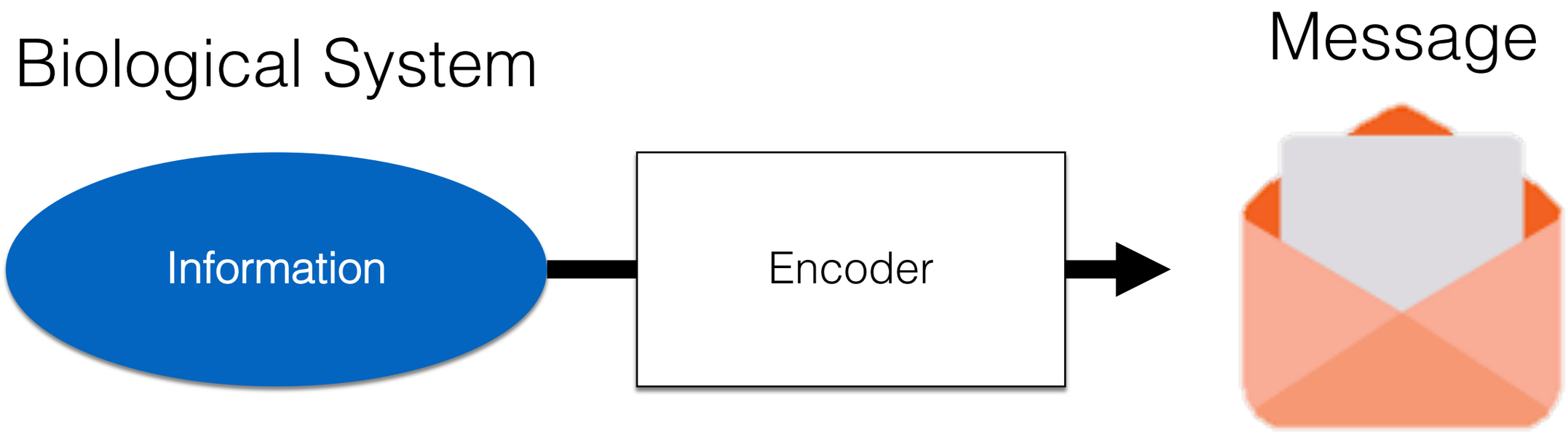


Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR

WYSS INSTITUTE

How to compose a message



For other types of measurement technologies, we have to actively form the message through experimentation

These approaches fall broadly into two groups:
Affinity & Reactivity

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

How to detect molecules by Affinity

- Affinity refers to the weak chemical interactions between biomolecules, such as hydrogen bonding, hydrophobic and hydrophilic interactions, electrostatic interactions, as well as the steric compatibility of biomolecular interfaces that enable these weak chemical interactions.
- Any ligand that exhibits a non-random binding pattern for other biomolecules under any conditions is capable of forming an informatic message from the underlying biological information.
- However, the more specific the affinity interaction is to a particular biomolecular composition, conformation, or spatiotemporal organization, the more information is transferred into the message.
- Message construction can also utilize either the formation or disruption of these weak interactions.

Evan Daugharthy
evan.daugharthy@gmail.com

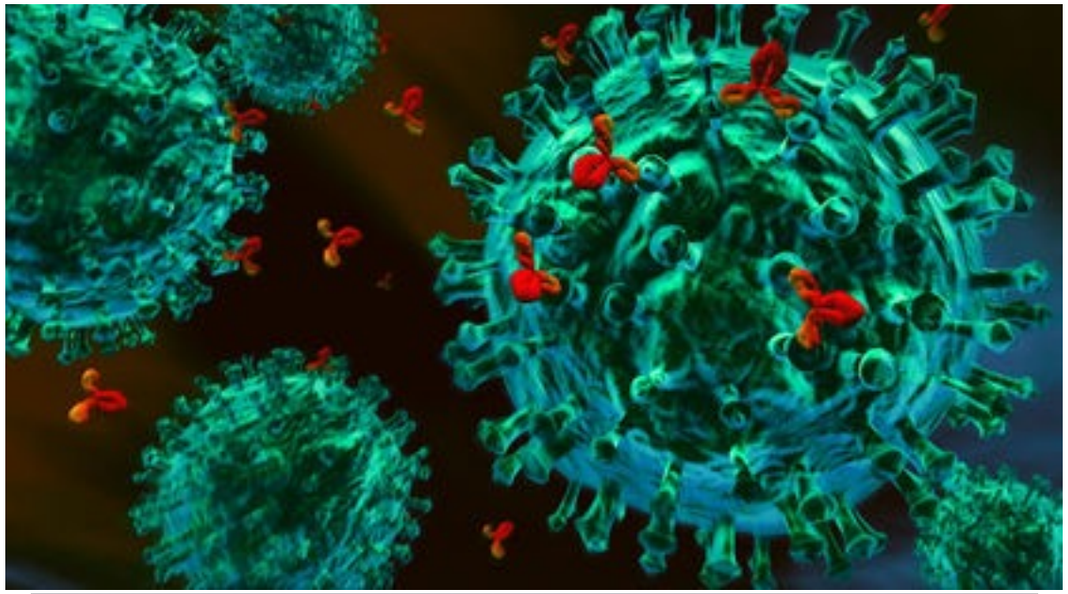
How to detect molecules by Affinity

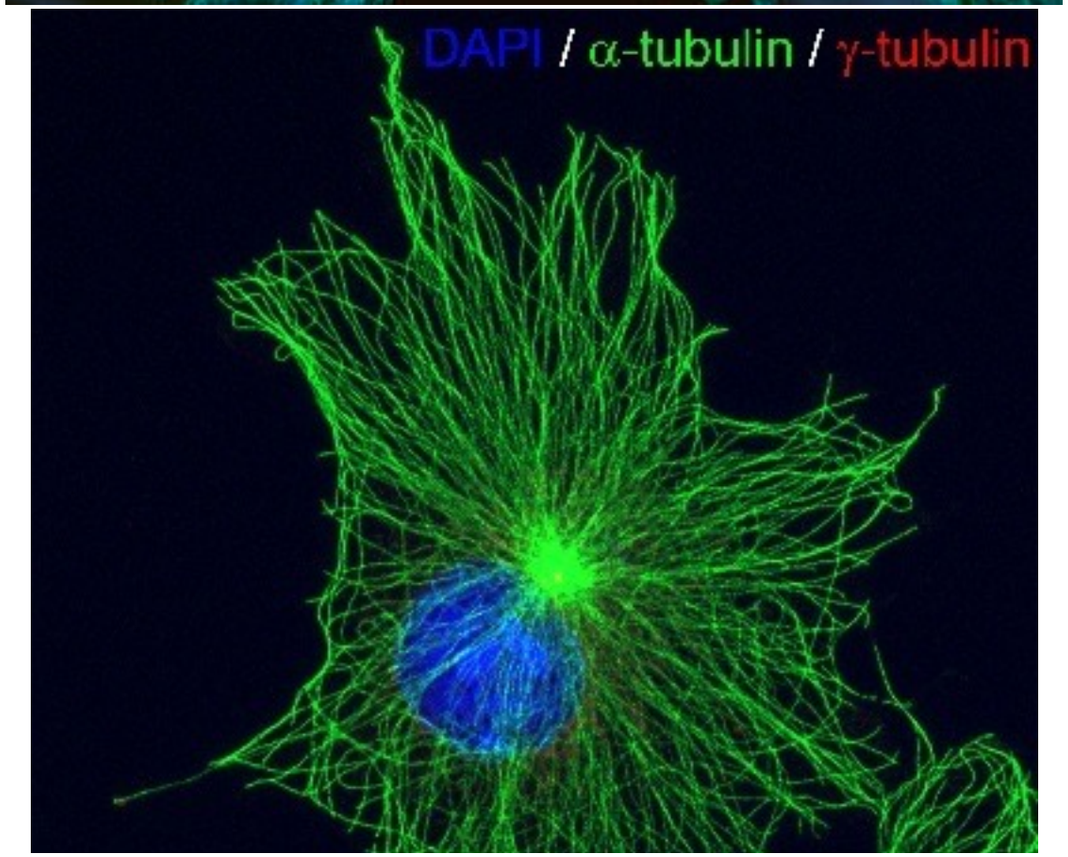
- Nucleic acid or nucleic acid analog hybridization probes
- Immunological proteins and immune-derived peptide fragments, such as antibodies, nanobodies, single chain variable fragments, and phage-display particles
- Aptamers, including those formed from nucleic acids, nucleic acid analogs, and polypeptides
- Proteins, such as lectins, which bind certain carbohydrate analytes
- Nucleic acid-guided nucleic acid binding proteins, such as by binding dCas9
- Heat or chemical denaturant treatment to disrupt weak interactions, e.g., DNA duplex melting curves

Evan Daugharthy
evan.daugharthy@gmail.com

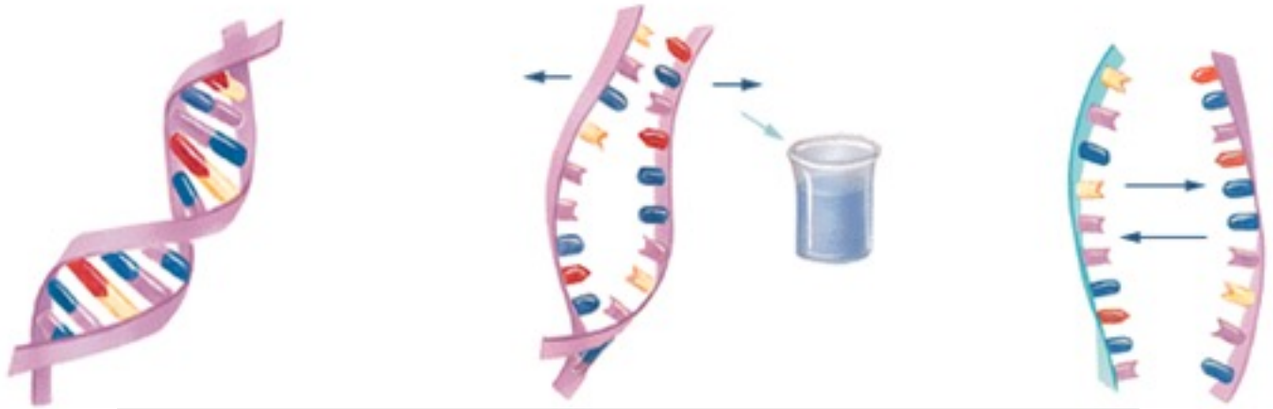
How to detect molecules by Affinity

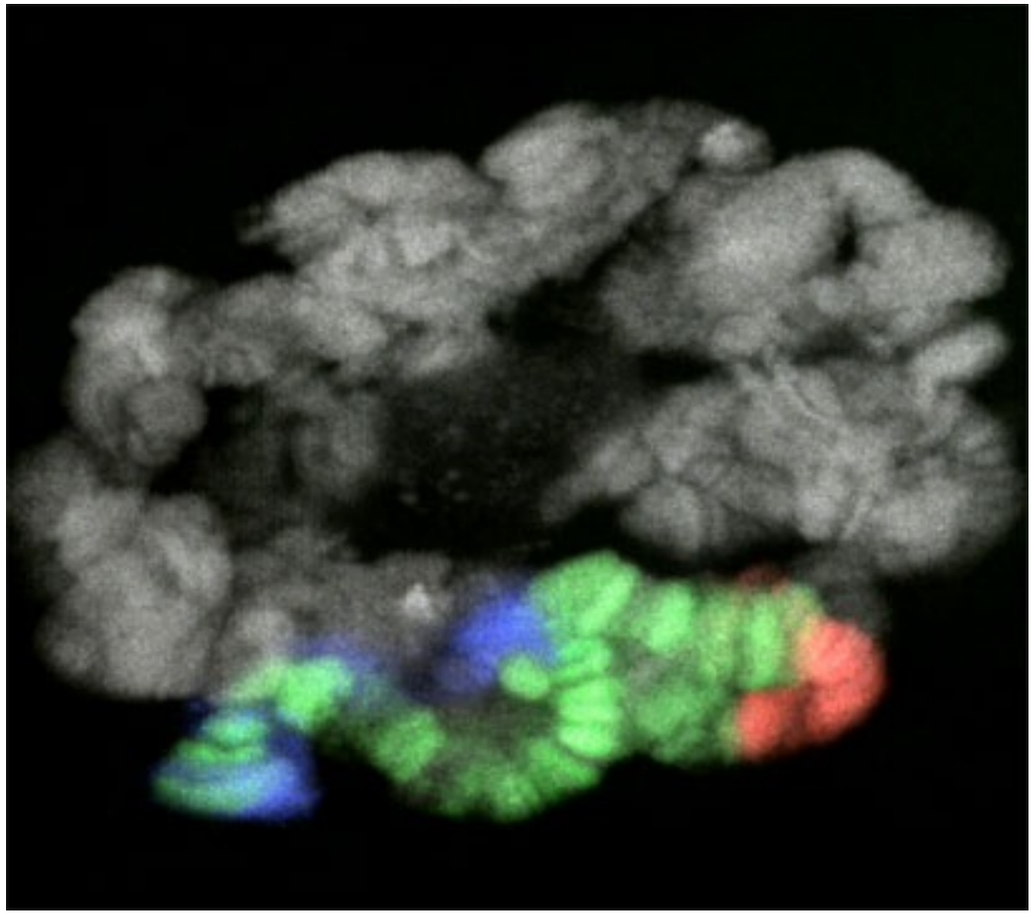
Antibody binding





Nucleic acid hybridization





Example: Akoya CODEX

Evan Daugharthy
evan.daugharthy@gmail.com

How to detect molecules by Reactivity

- Reactivity refers to formation or destruction of covalent or ionic chemical bonds.
- Again, in theory any chemical reaction that exhibits a non-random reactivity profile with biomolecules is capable of forming the message.
- The most common uses of reactivity stem from natural biochemical processes, since the reactions occurring inside living systems are generally highly specific.

Evan Daugharthy
evan.daugharthy@gmail.com

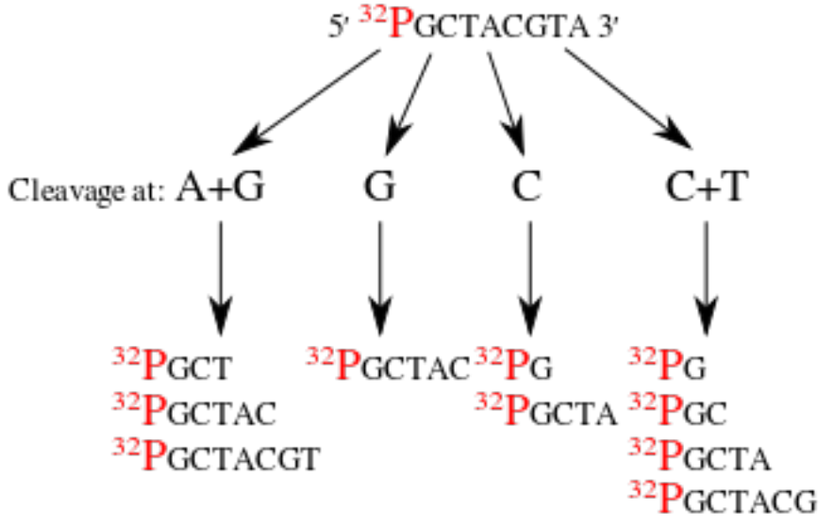
How to detect molecules by Reactivity

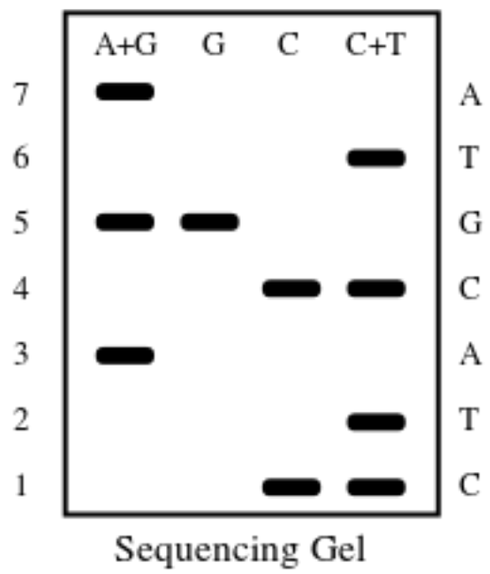
- Endonuclease digestion of nucleic acids to generate restriction fragments
- Protease digestion of peptides
- Blunt ended or single stranded ligation
- Nucleic acid synthesis, such as by a polymerase
- Bisulfite reaction with methylated DNA
- Nucleic acid-guided nucleic acid binding protein nuclease activity, such as Cas9

Evan Daugharthy
evan.daugharthy@gmail.com

How to detect molecules by Reactivity

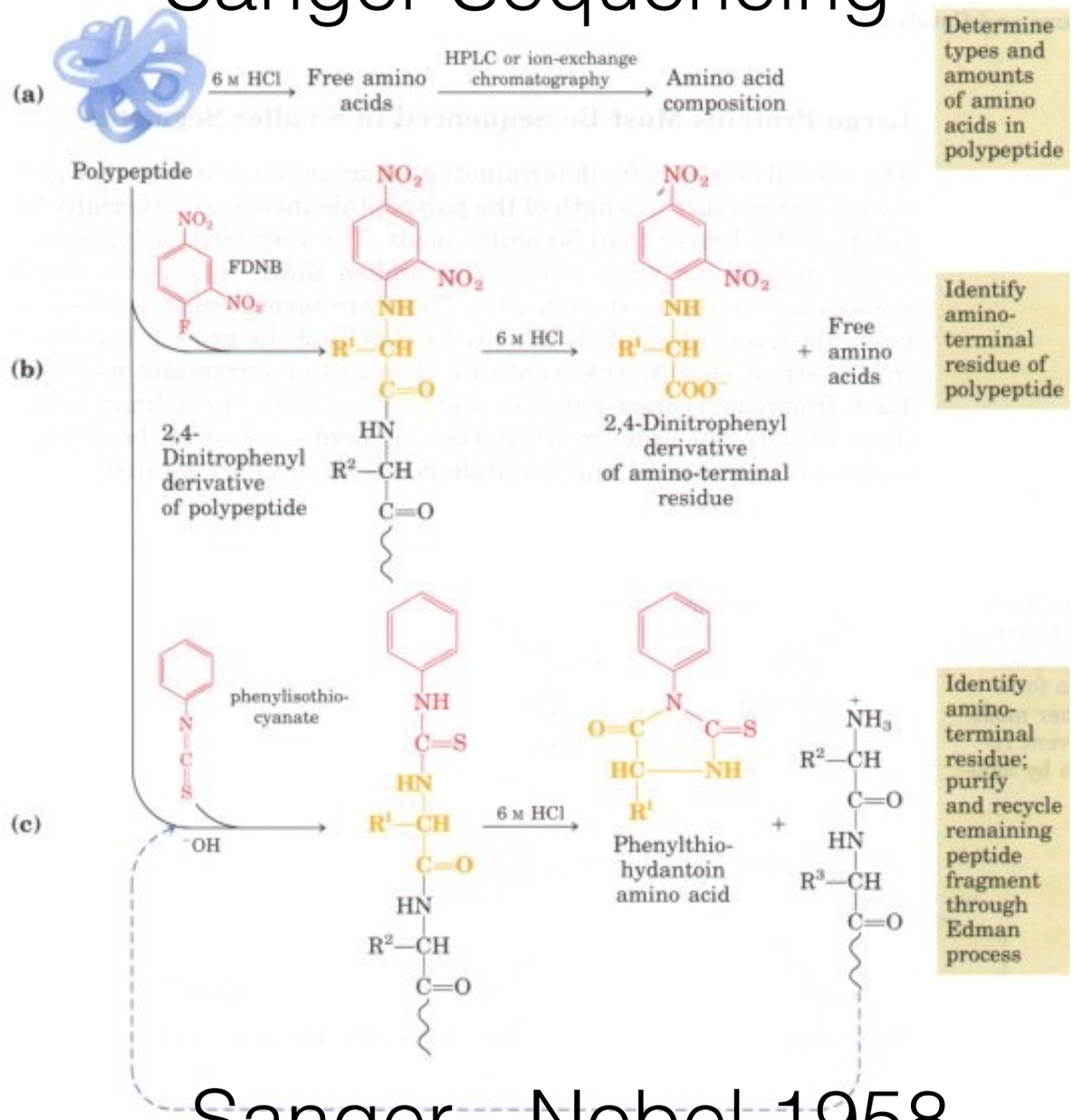
Maxam-Gilbert Sequencing





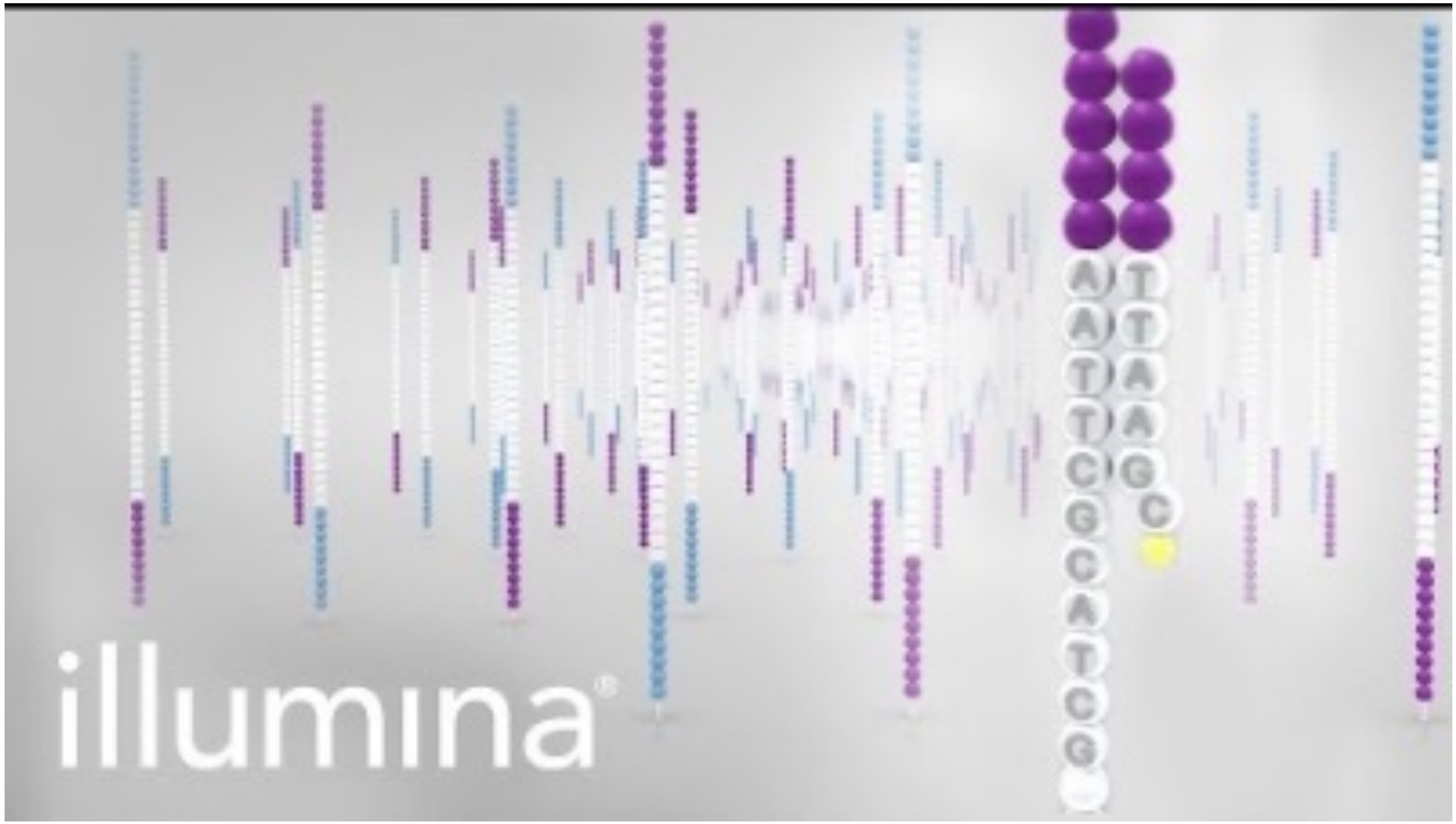
Gilbert, Sanger & Berg
Nobel 1980

Edman Reaction/ Sanger Sequencing



Sanger - Nobel 1958
(Insulin sequence)

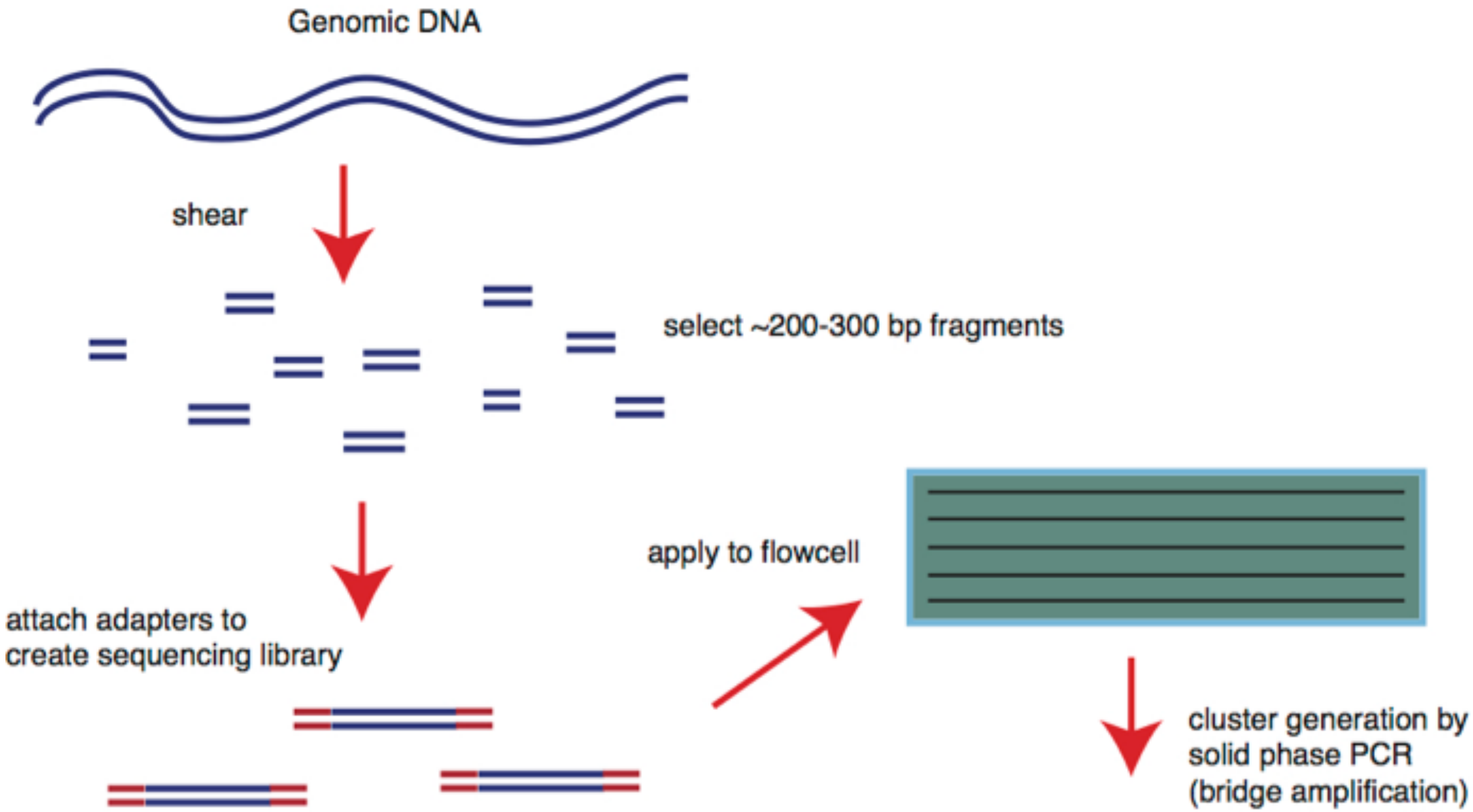
Example: Illumina



Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

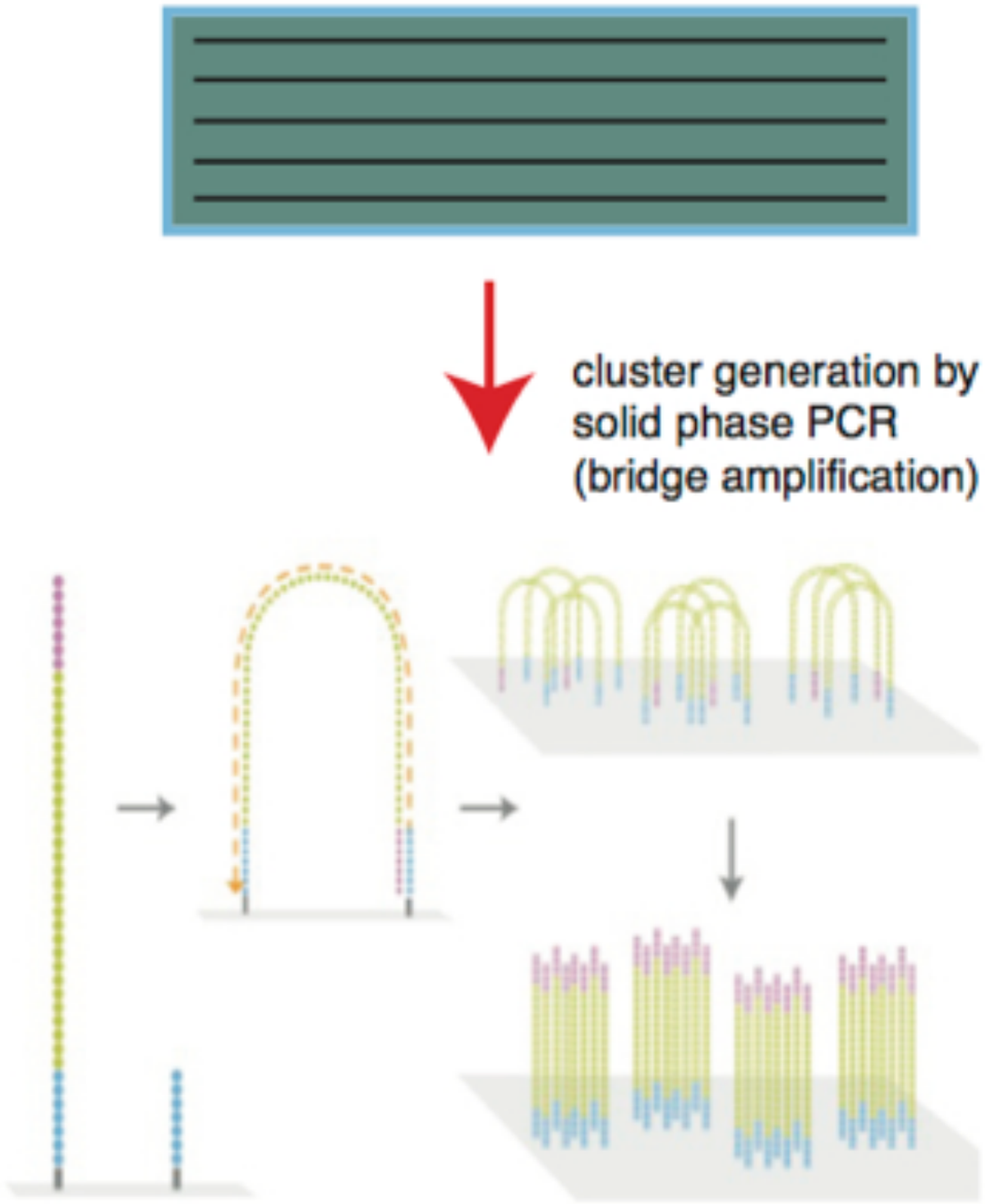
How to detect molecules by Reactivity



Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

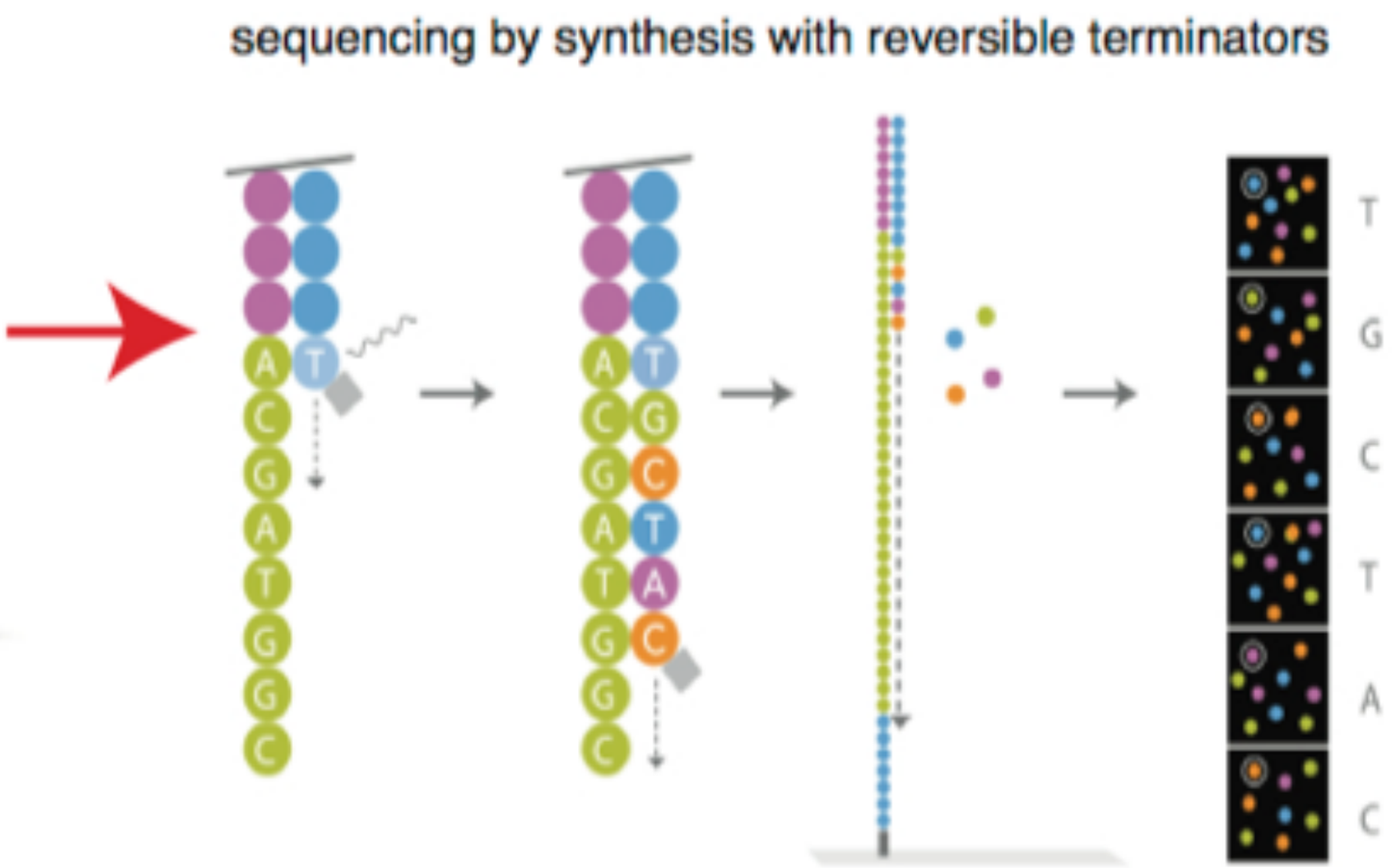
How to detect molecules by Reactivity



Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

How to detect molecules by Reactivity



Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR

WYSS INSTITUTE

Summary on Message Composition

- Most measurement technologies combine multiple methods to achieve a final assay
- E.g. FISSEQ technology has *many* steps utilizing both affinity and reactivity to convert the original specimen into an ”encoded message” suitable for transmission/detection and decoding
- Some subset of “all the information” in a sample is encoded into the message – it could be the information necessary to fingerprint to identify a molecule, or other types of information, like spatial information
- E.g., single-cell sequencing – each cell is loaded into a droplet and barcoded during message encoding – a cell barcode is incorporated into the message to tell the sequencer which cell the particular RNA or DNA molecule originated from (kind of like an IP/MAC address embedded in an internet data packet)

Towards Perfect Molecular Measurement

What is a perfect measurement?

Having all the required or desirable characteristics

vs

Being as good or complete as possible

Daugharthy, E., 2016. *Towards Perfect Molecular Measurement* (Doctoral dissertation).

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

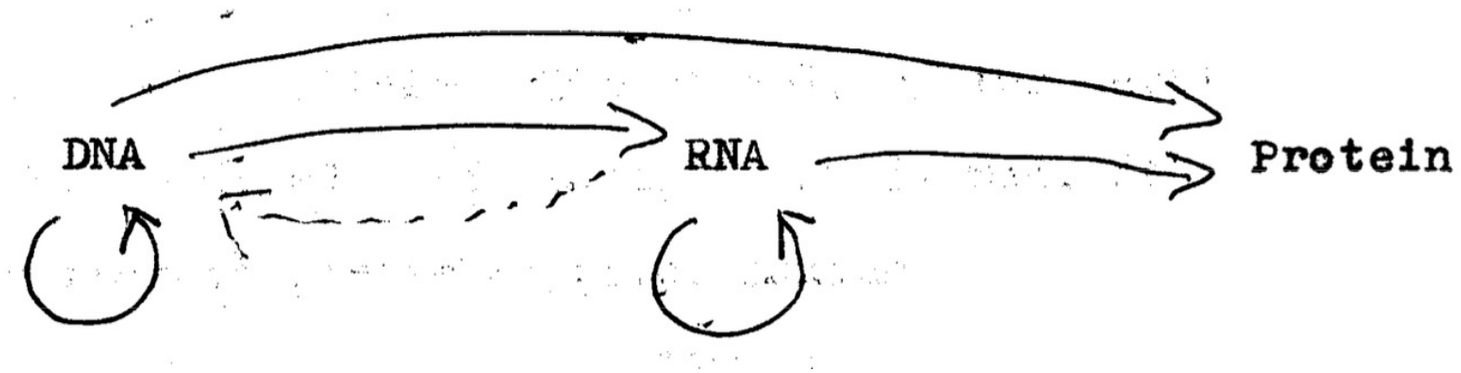
Towards Perfect Molecular Measurement

1. What questions are being asked about biological systems?

2. What types of information will provide these answers?

3. What types of observations and measurements will provide this information?

- Harold Morowitz, 1955



Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Towards Perfect Molecular Measurement

1. What questions are being asked about biological systems?

2. What types of information will provide these answers?

Molecular Mechanisms

Mechanisms are composed of entities and activities

A > B > C

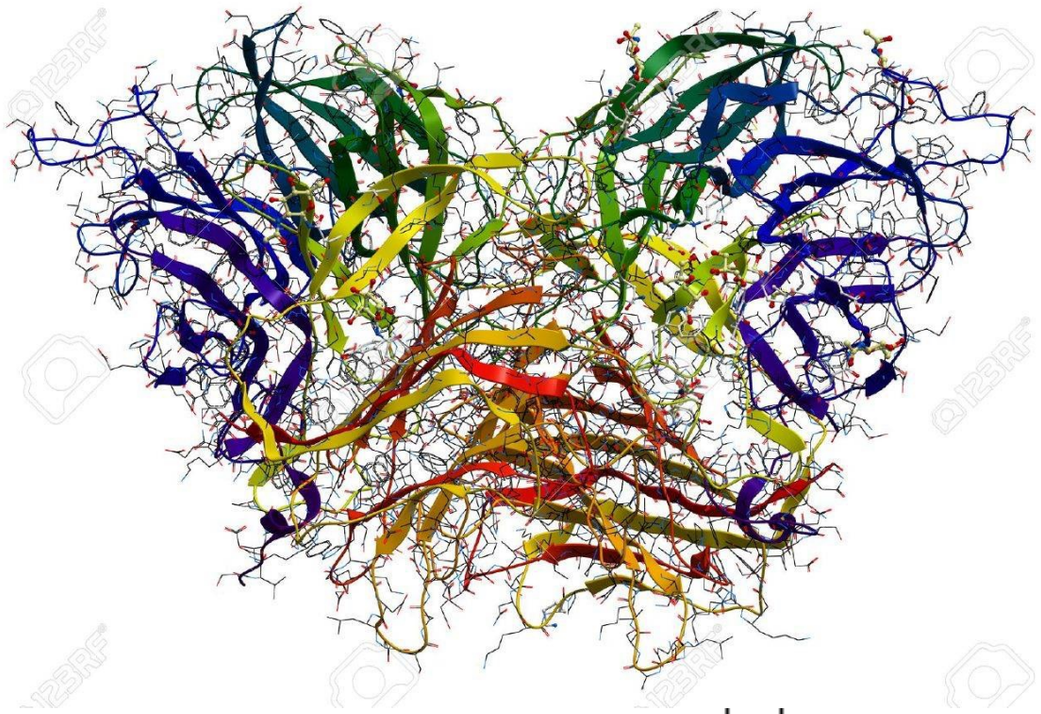
Philosophers of science Machamer, Darden, & Craver

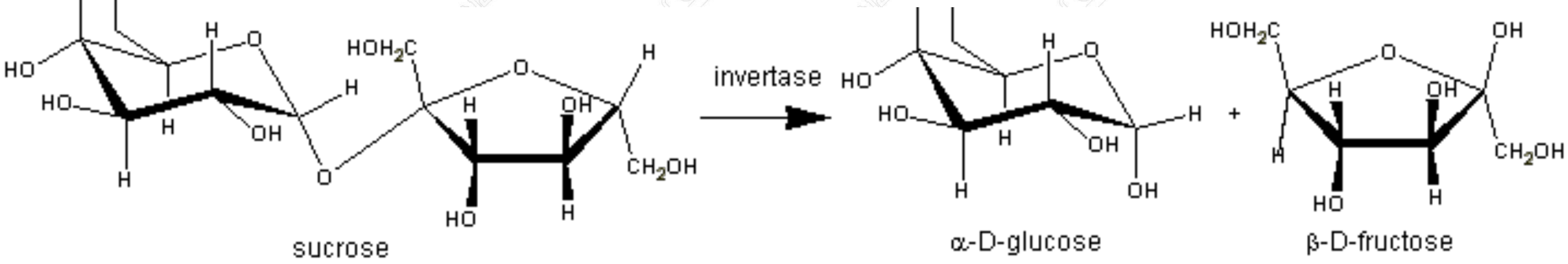
Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR

WYSS INSTITUTE

Mechanism





Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR

WYSS INSTITUTE

What about more complex phenomena?

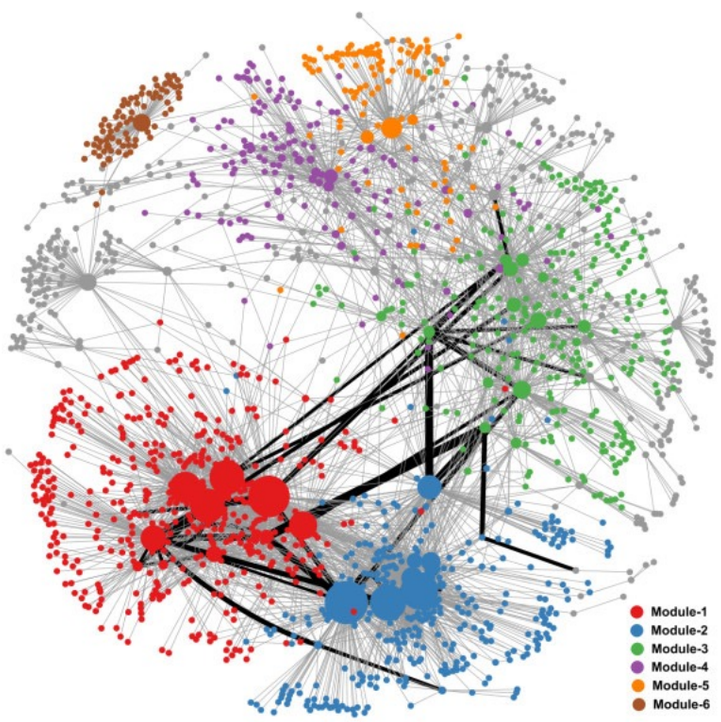
Evan Daugharthy
evan.daugharthy@gmail.com

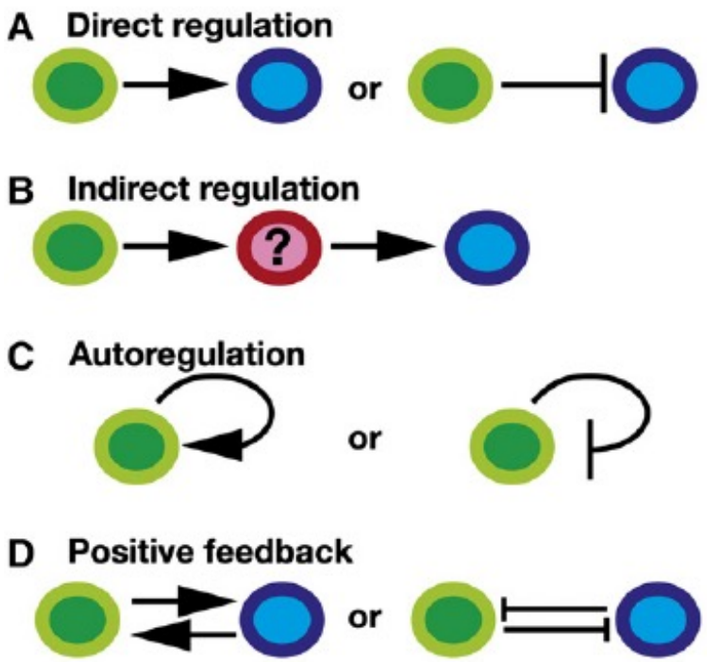
READCOOR WYSS INSTITUTE

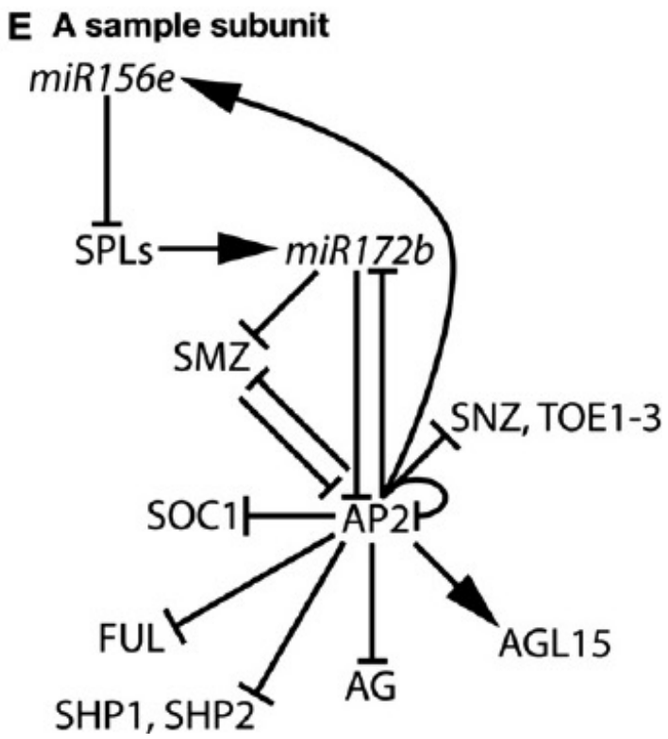
Towards Perfect Molecular Measurement

3. What types of observations and measurements will provide this information?

Systems biology is characterized by observing biological systems, experimentally perturbed or in their naturally dynamic states, using quantitative multidimensional and multiplex measurements, and then integrating measurement data and functional observations using mathematical and computational models.



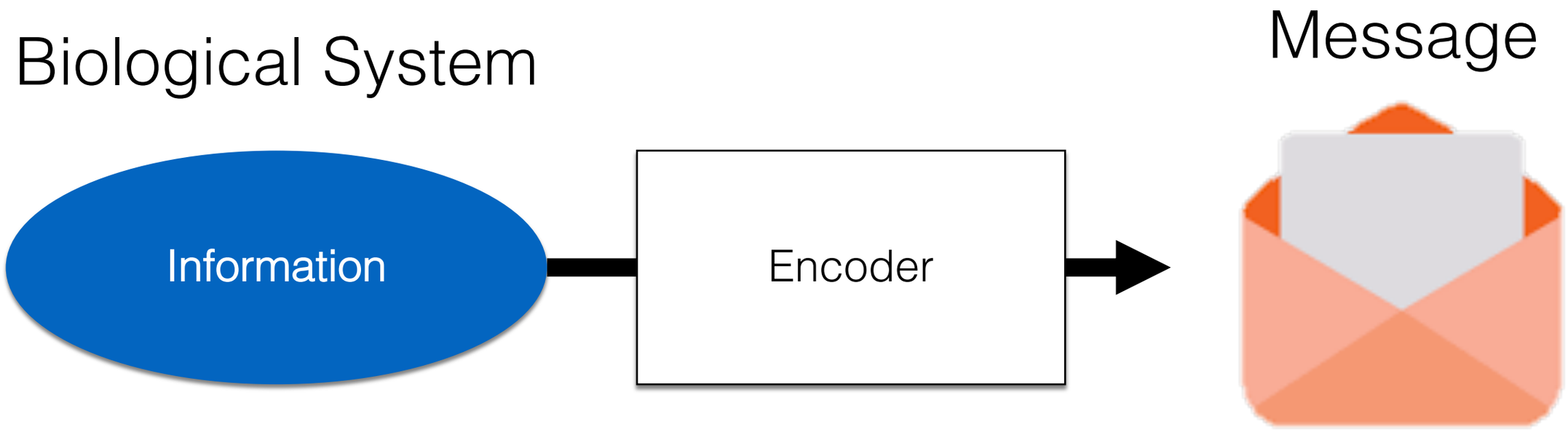




Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Considerations for measurement



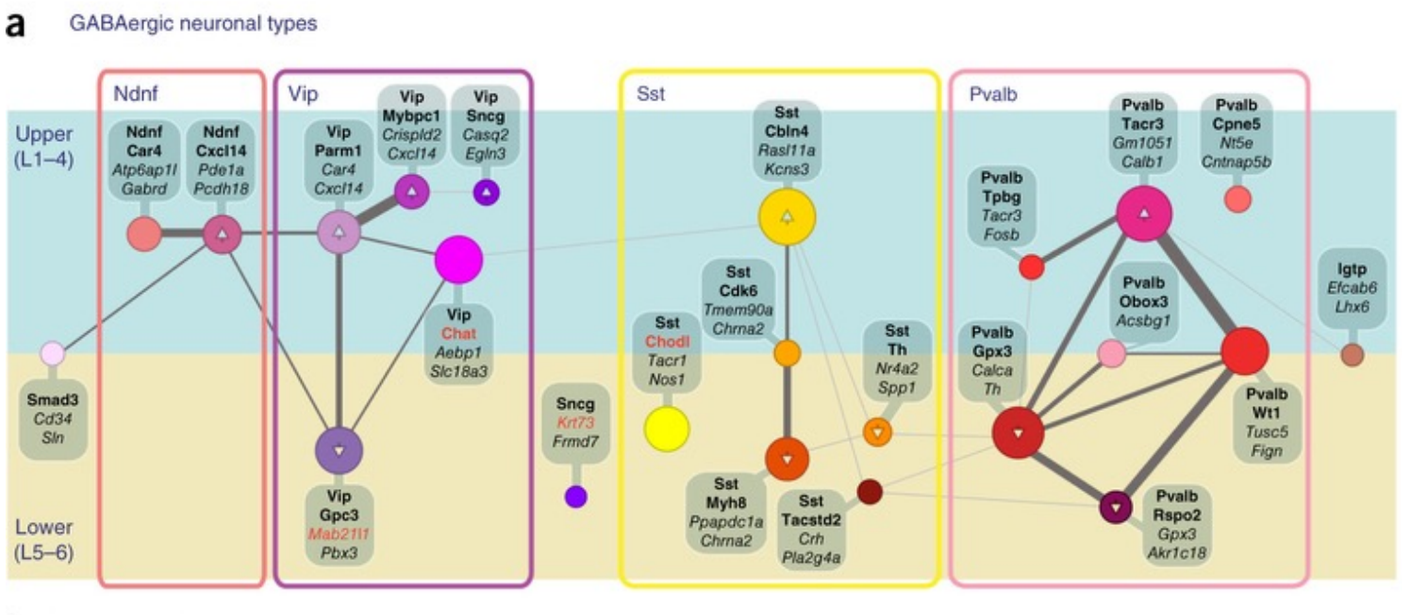
Cells vary in many dimensions simultaneously, need massively multiplex measurements to see the whole picture, i.e. simultaneously encode many species and types of molecules

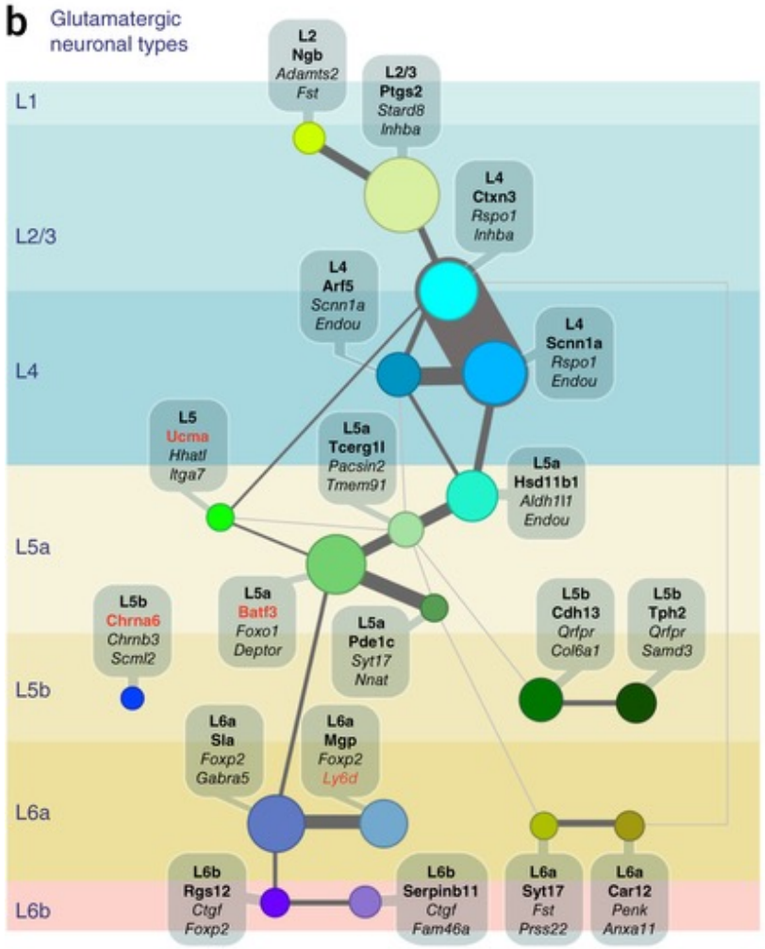
Brain FISSEQ

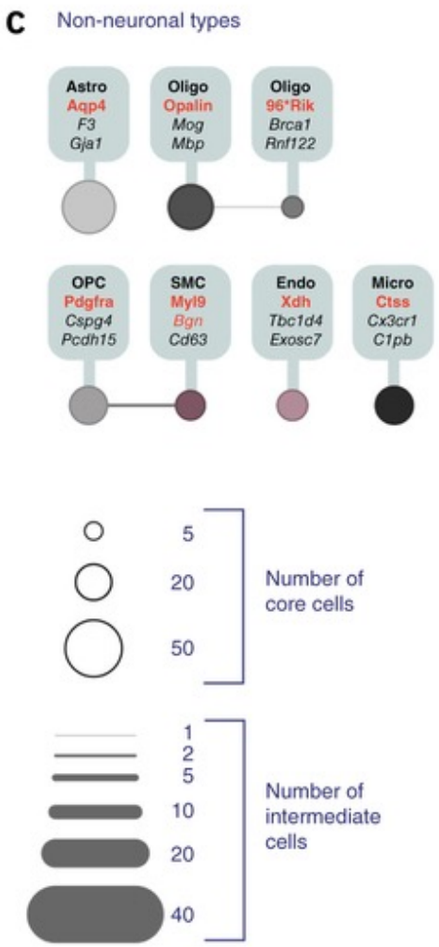
1989	2006	2015
210 cell types	145 neuron types	53 single neurons sequenced
1 neuron		
Alberts, Molecular Biology of the Cell	Vikaryous, Human cell type…	Dueck, Deep sequencing…

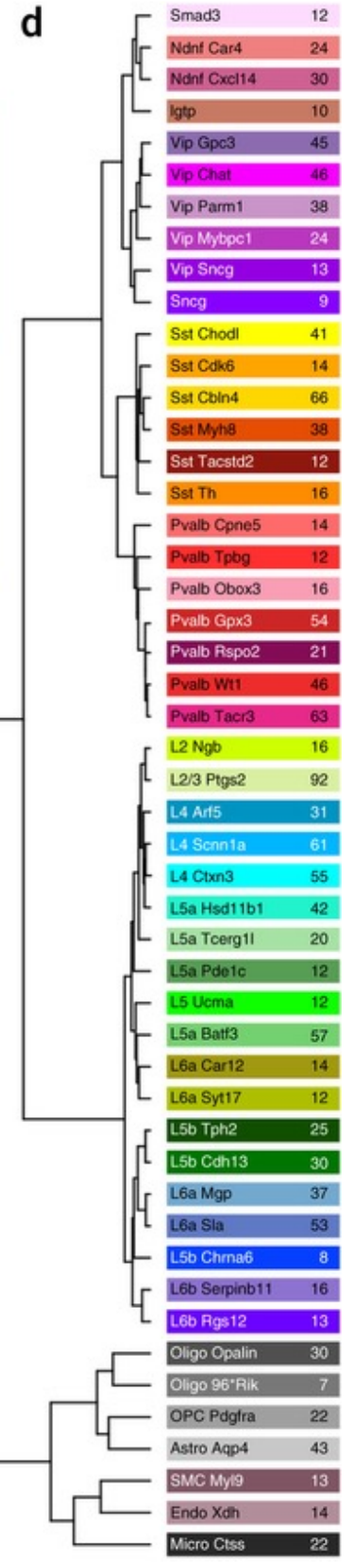
mouse hippocampus

Cells are high-dimensional entities





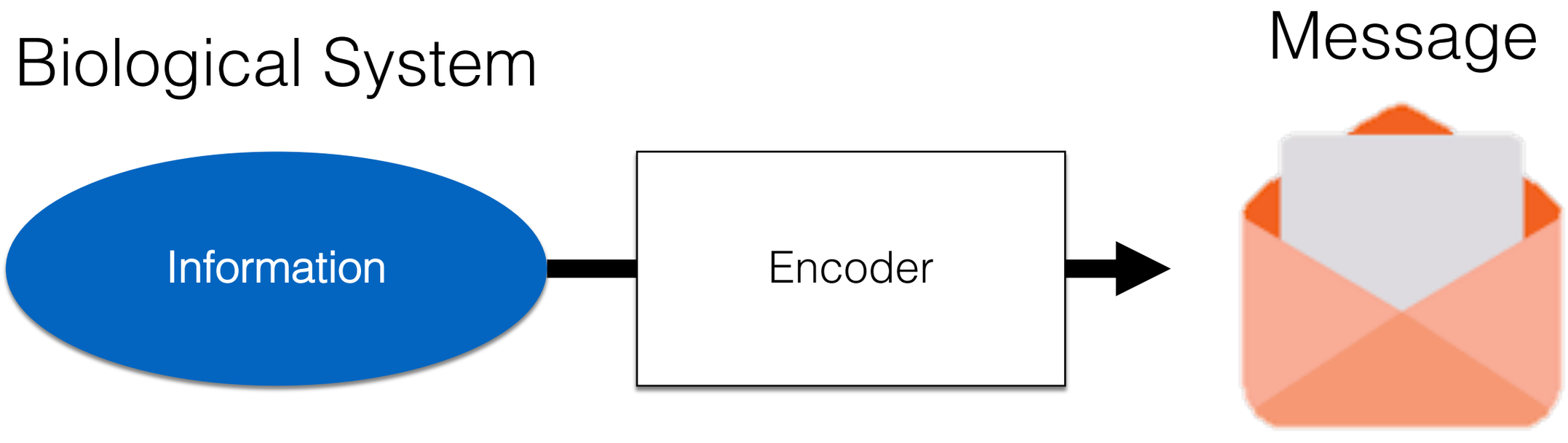




Adult mouse cortical cell taxonomy revealed by single cell transcriptomics

Tasic *Nature Neuroscience* (2016)

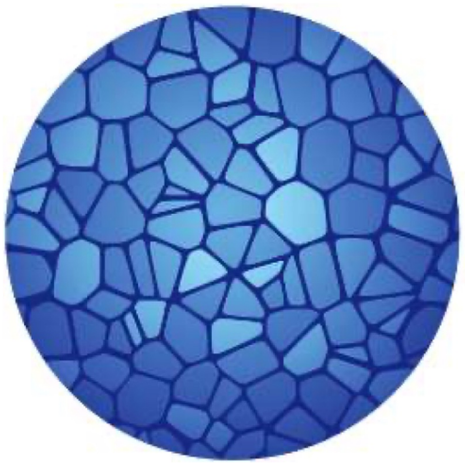
Considerations for measurement



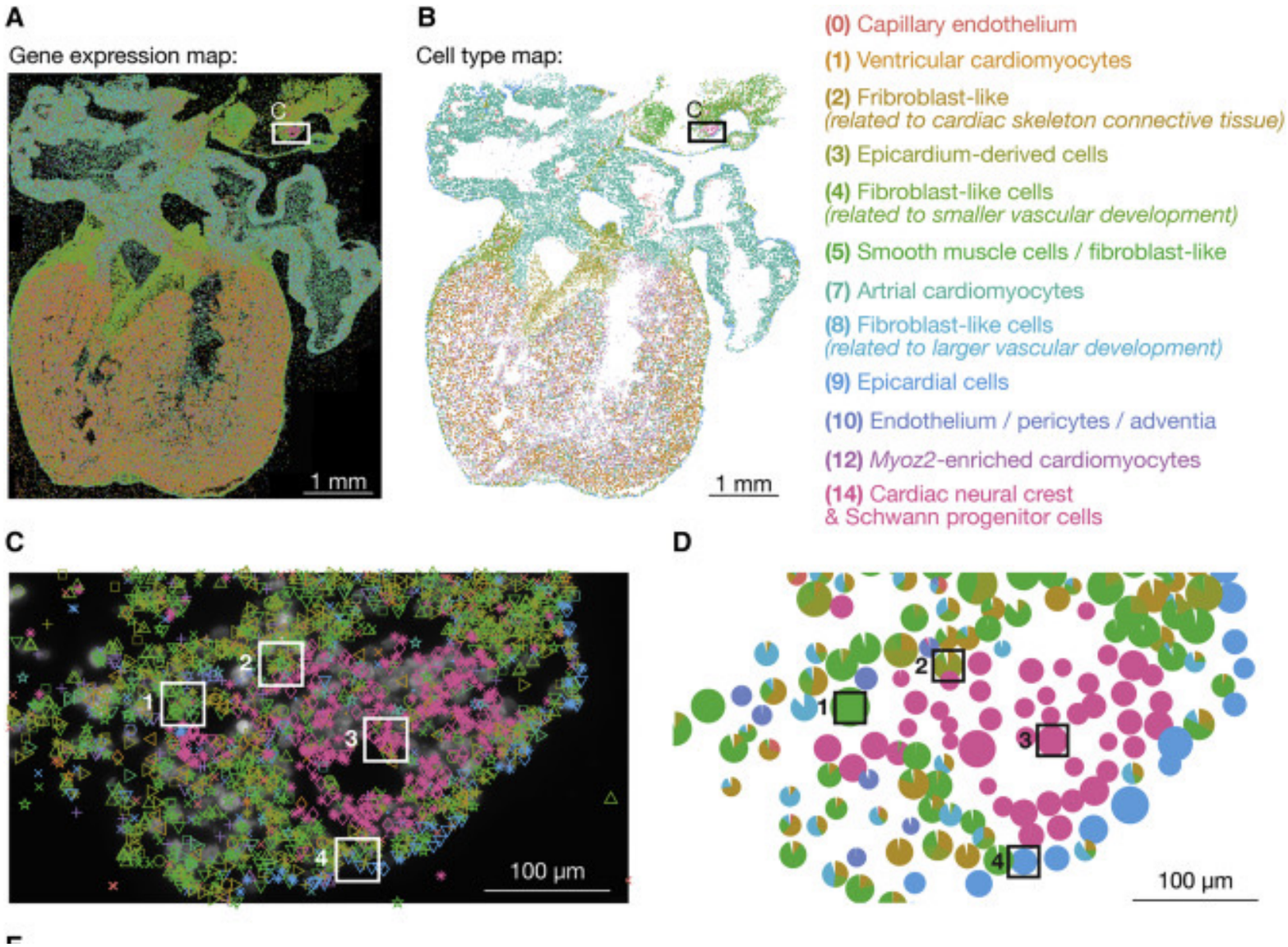
Biological systems vary over all spatial scales;
Measurement lacking resolution reduces sensitivity and can lead to incorrect conclusions

bulk **vs** tissue region
cell type
single cell
single molecule

Cell Atlas Project



THE HUMAN CELL ATLAS

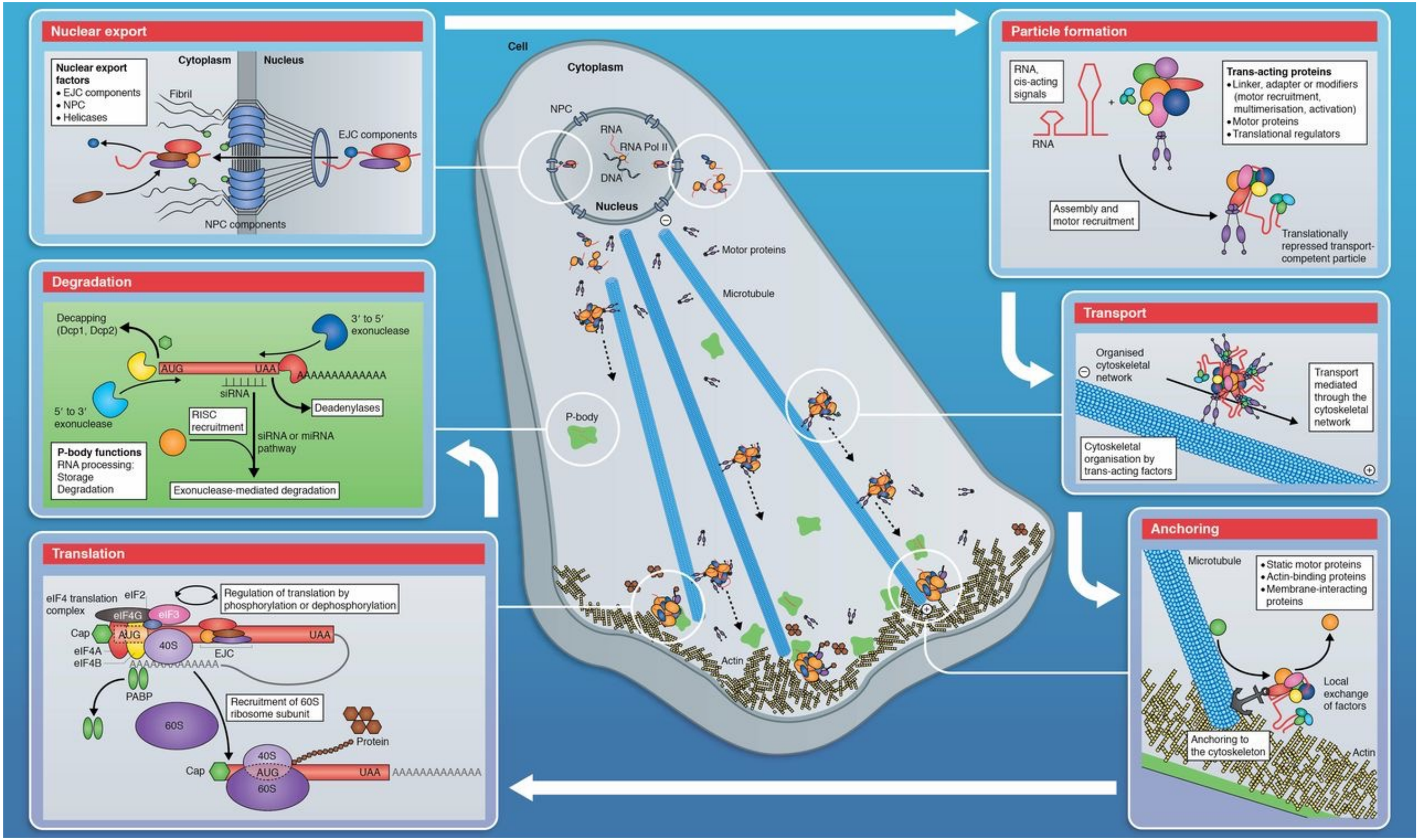


Asp, Michaela, et al. "A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart." *Cell* 179.7 (2019): 1647-1660.

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

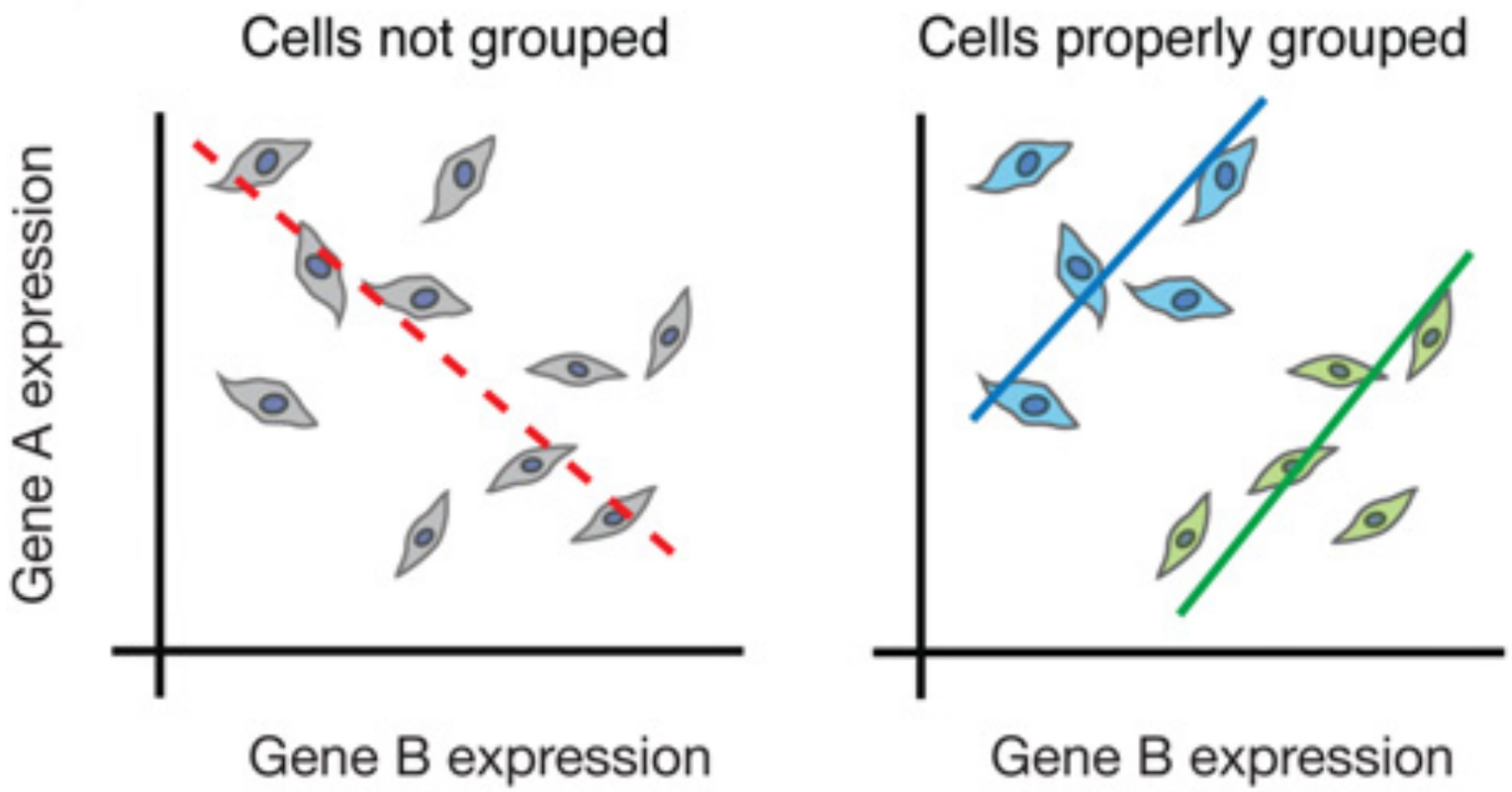
Meaningful variation at sub-cellular scale



Parton *et al. Cell Science* (2014)

Impact of bulk measurement technology

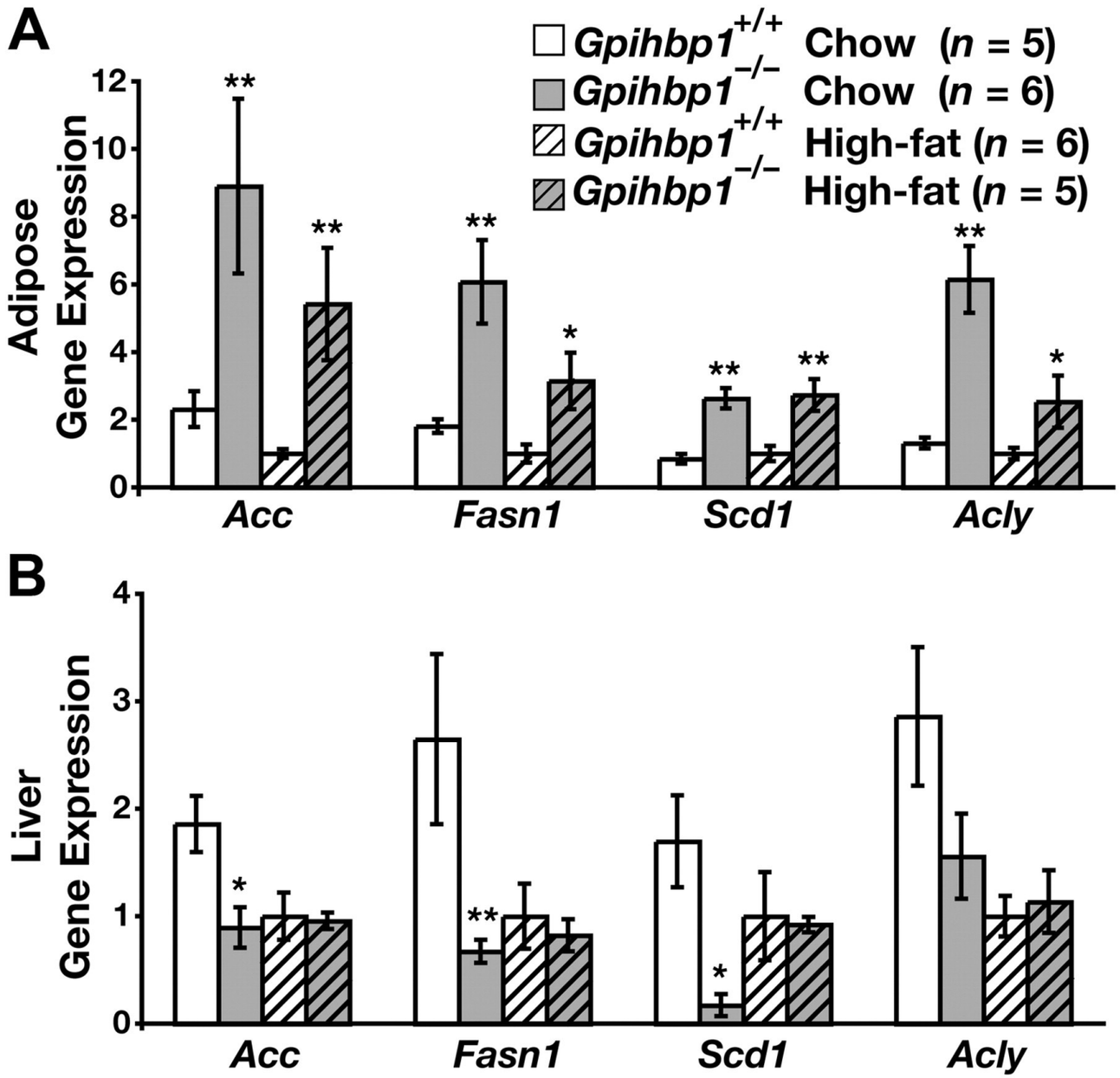
Simpson's paradox



bulk measurements yield qualitatively incorrect conclusions

Trapnell *Genome Research* (2015)

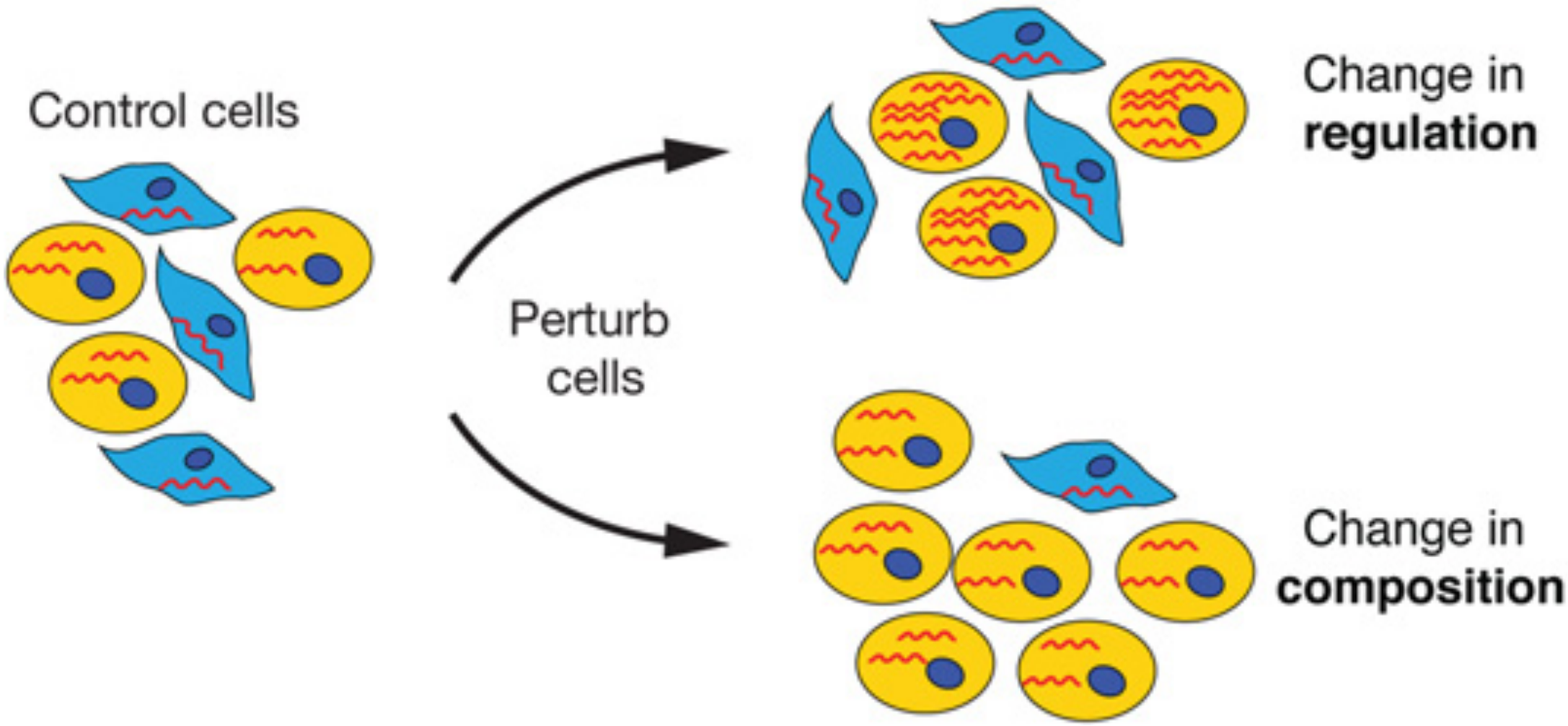
Meaningful variation at tissue scale



Reciprocal metabolic perturbations in liver and adipose tissue in the setting of defective lipolysis

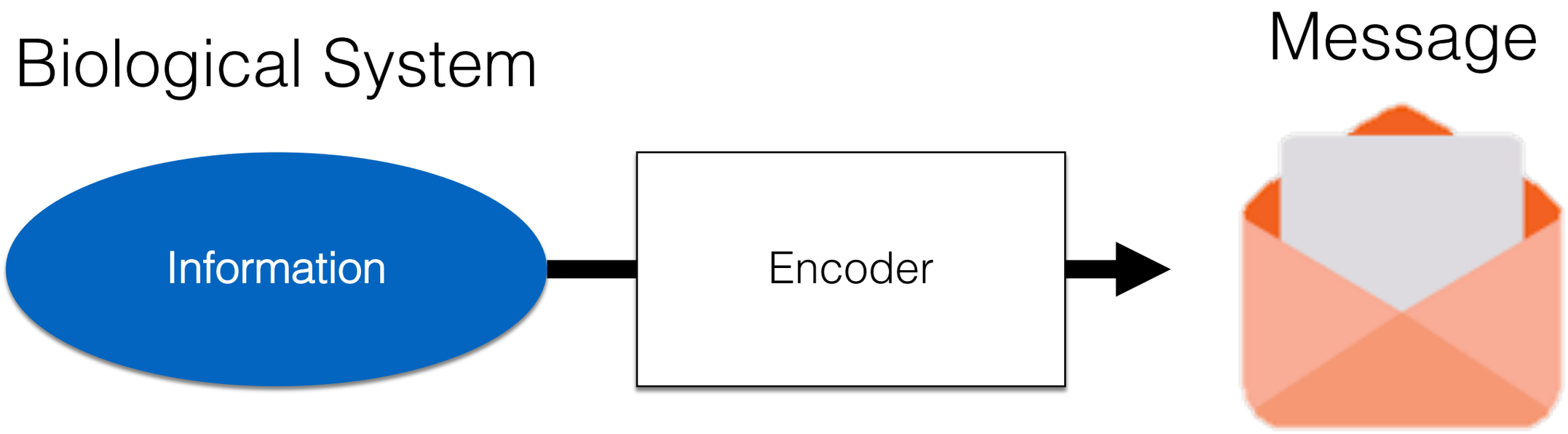
Weinstein *et al*. *ATVBAHA* (2012)

Other sources of variation confounding measurement



Trapnell *Genome Research* (2015)

Conclusion to message construction



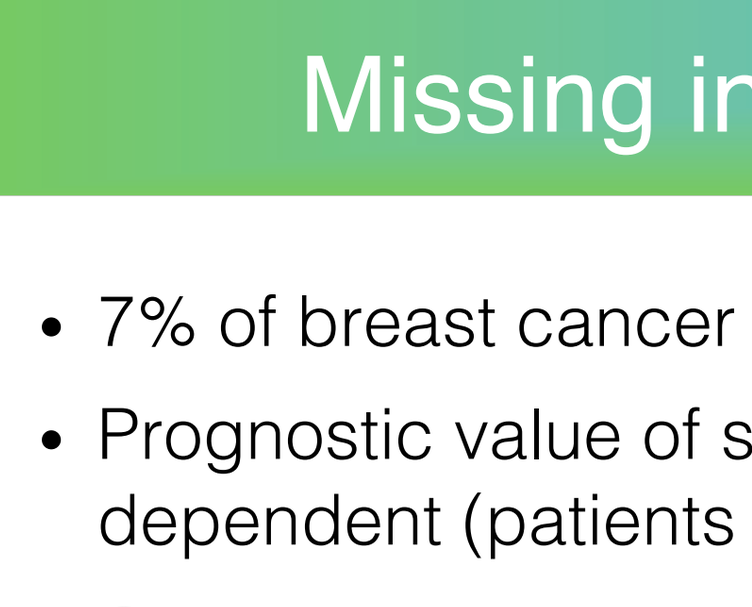
- We need to measure a lot of things simultaneously
- We need to keep the measurements confined spatially

Evan Daugharthy
evan.daugharthy@gmail.com

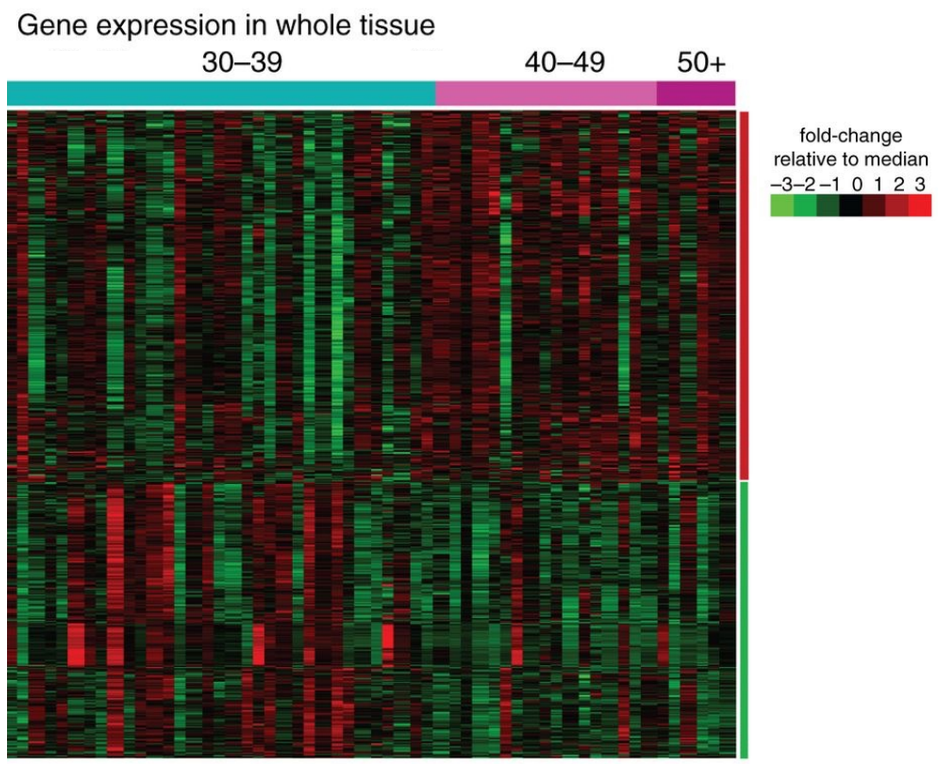
READCOOR WYSS INSTITUTE

Missing information in bulk measurement

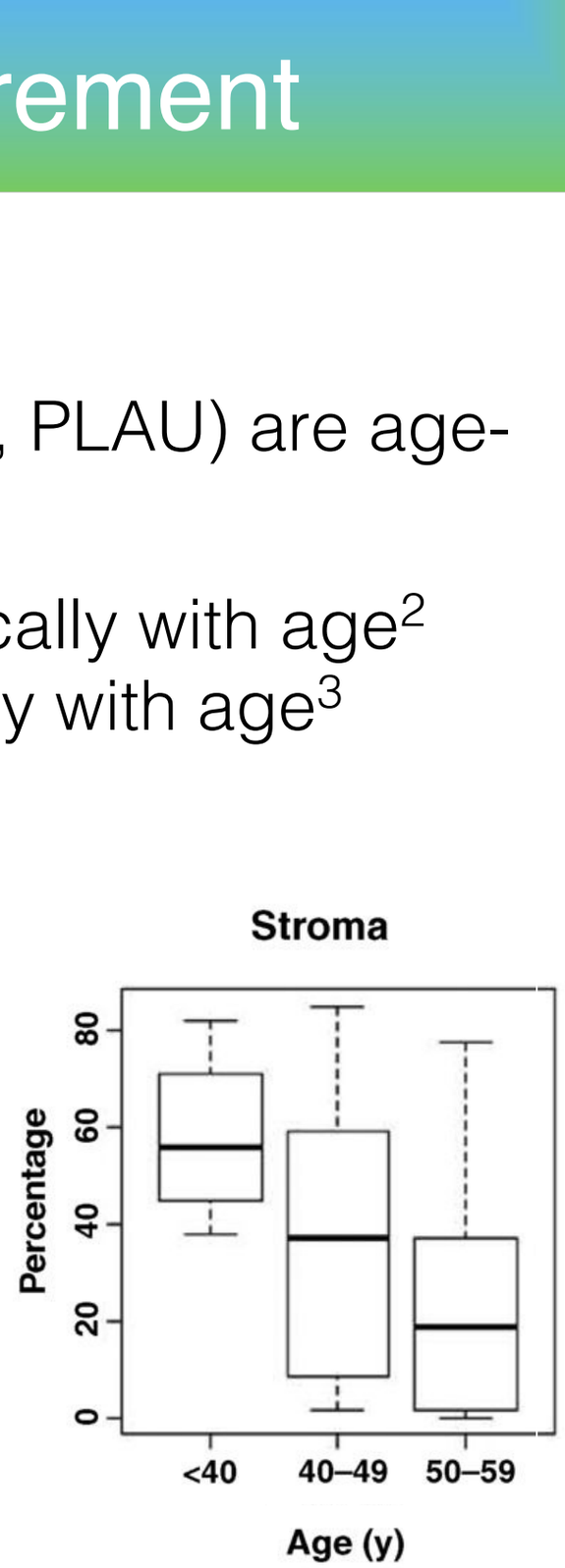
- 7% of breast cancer in patients age <40
- Prognostic value of stroma-related gene signatures (DCN, PLAU) are age-dependent (patients <40) for the $ER^-/HER2^-$ subtype[1]
- Gene expression in whole breast tissue changes dramatically with age[2]
- Cellular composition of breast tissue changes dramatically with age[3]



1 Azim *et al. Clin Cancer Res* (2012)

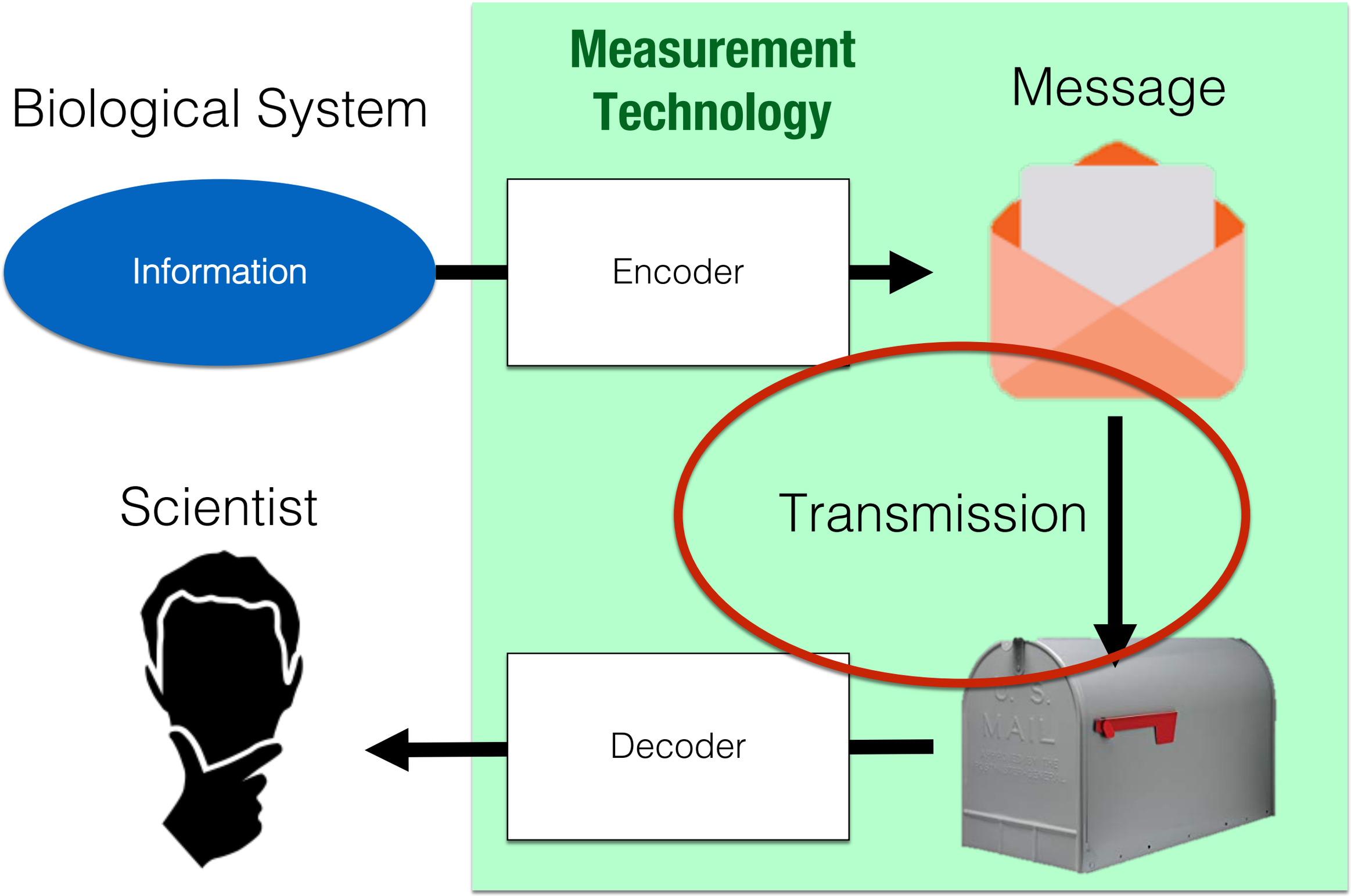


2 Pirone *et al. Cancer Epidemiol Biomarkers Prev* (2012)



3 Sun *et al. Clin Cancer Res* (2013)

How to transmit a message

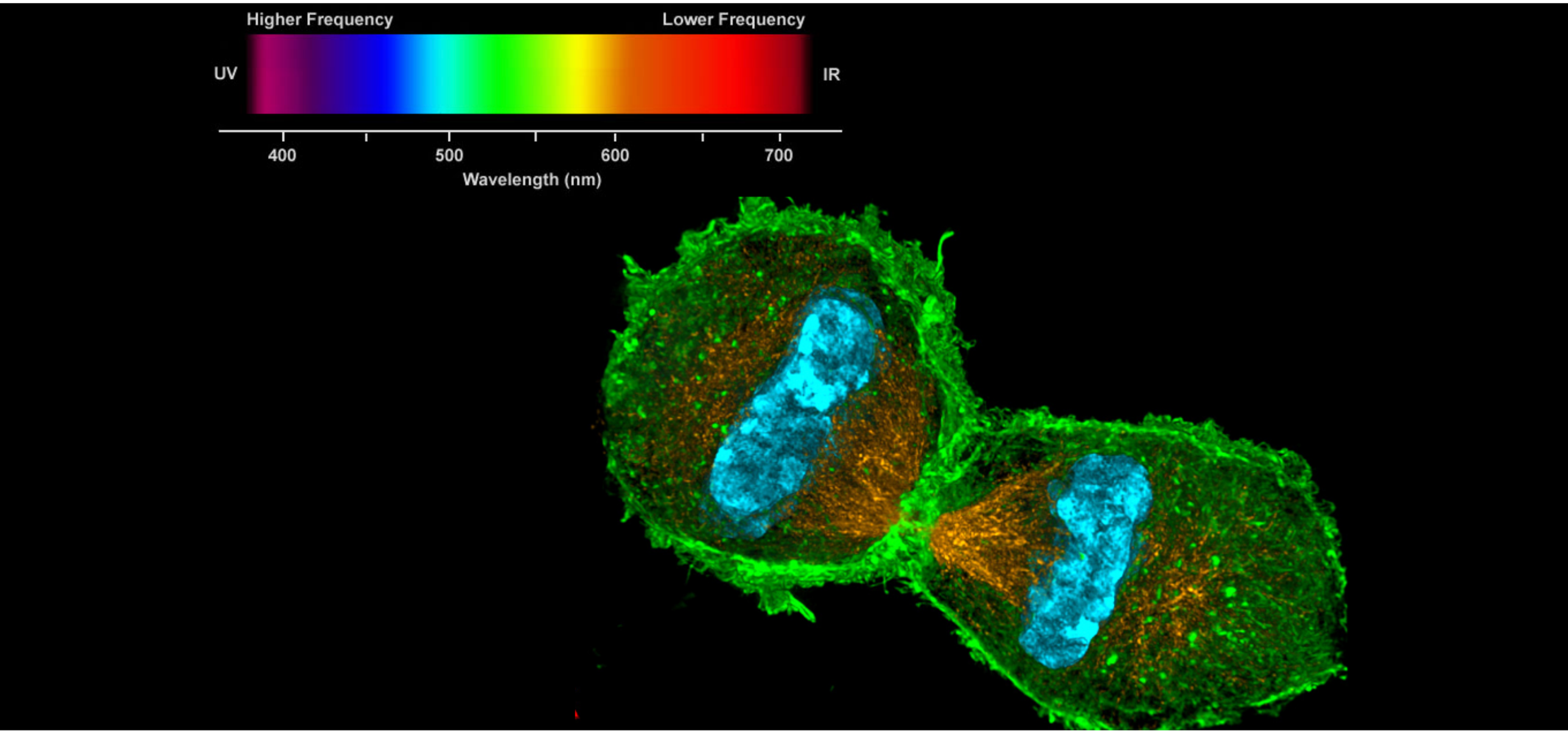


Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

How to transmit a message

The medium of information transmission in biological measurements is typically **light**



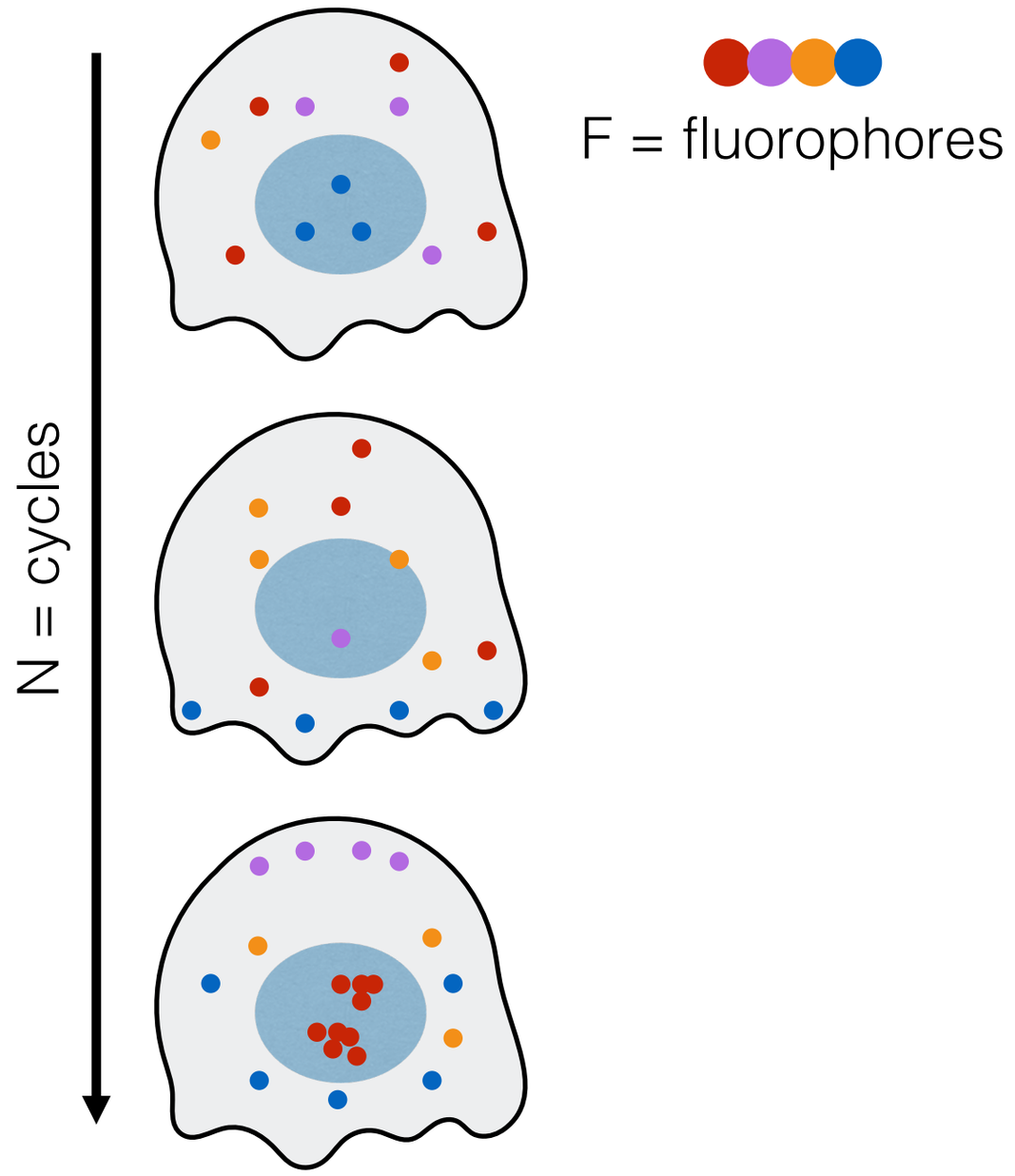
Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

The FISSEQ approach

FISSEQ approach to massively multiplex *in situ* molecular detection is **sequencing**

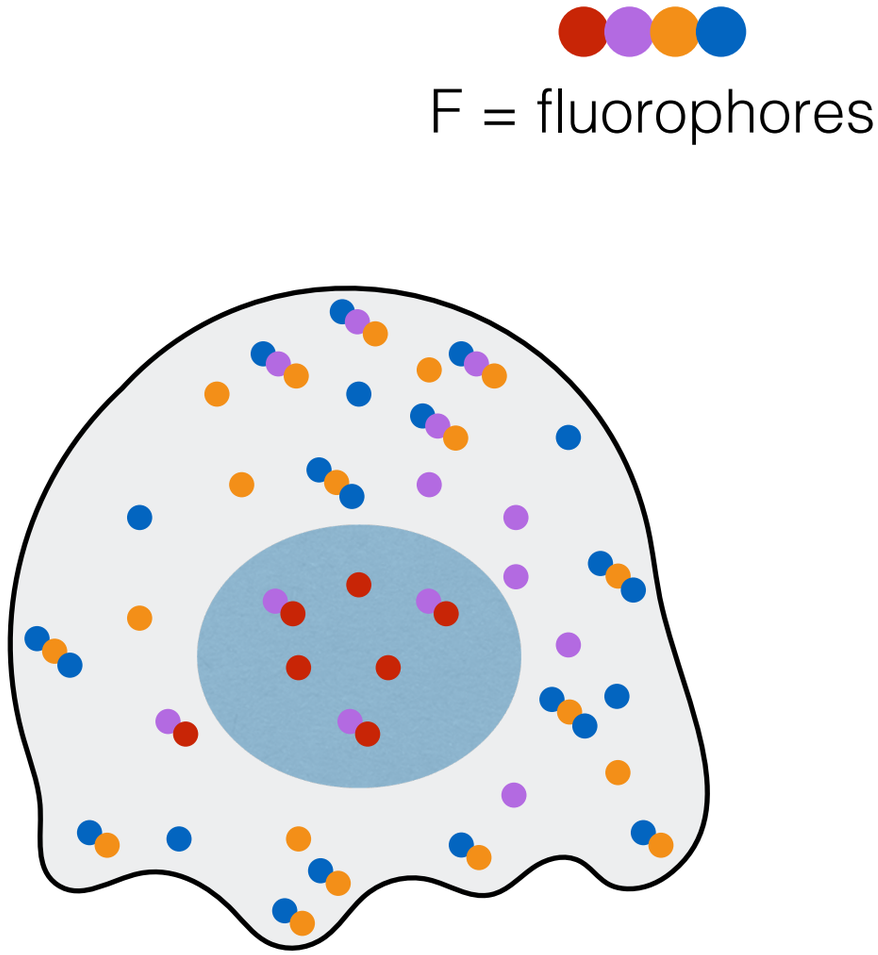
	Theoretical Multiplexity
Serial labeling	F×N



The FISSEQ approach

FISSEQ approach to massively multiplex *in situ* molecular detection is **sequencing**

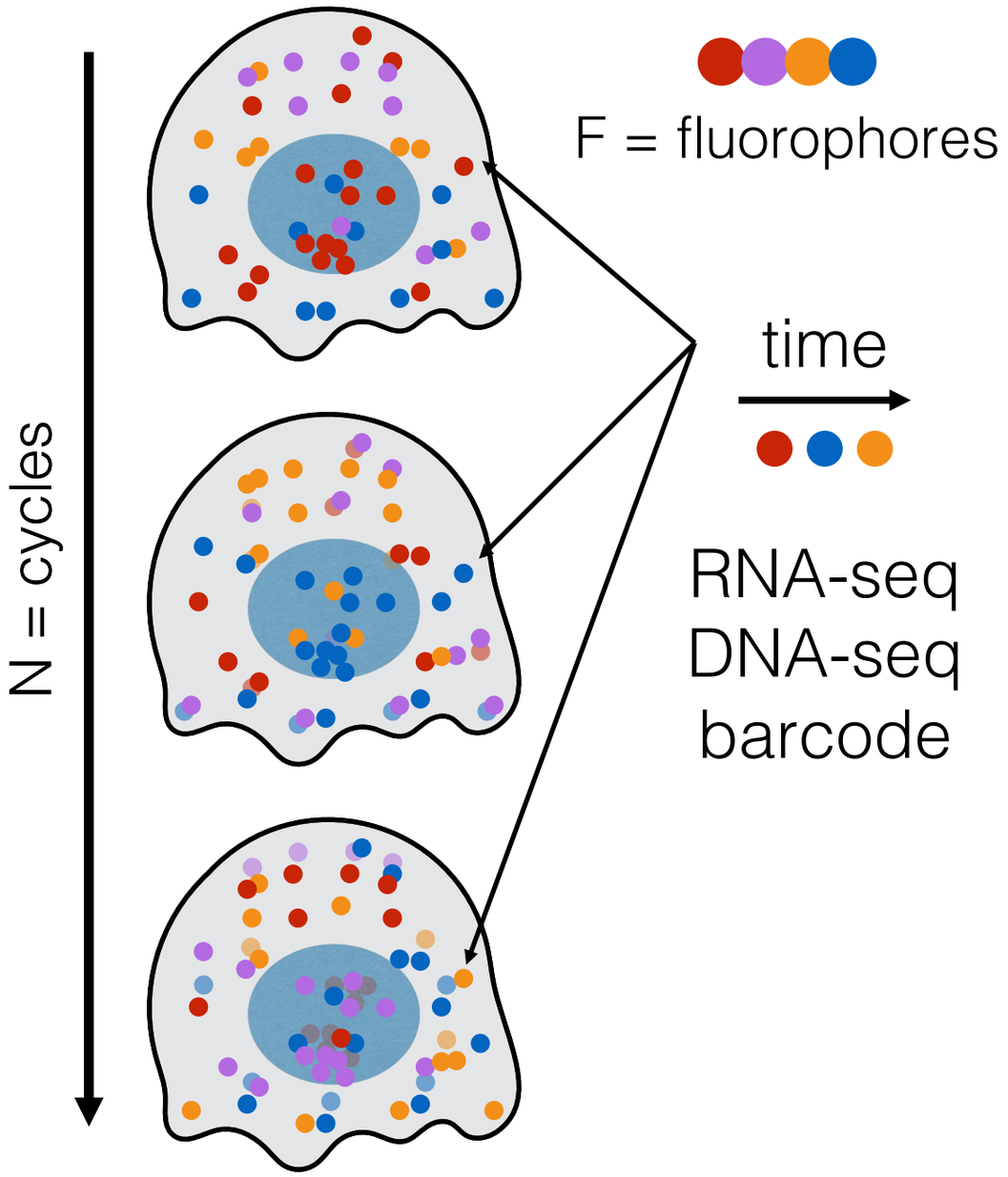
	Theoretical Multiplexity
Serial labeling	F×N
"Colorimetric" labeling	2^F-1



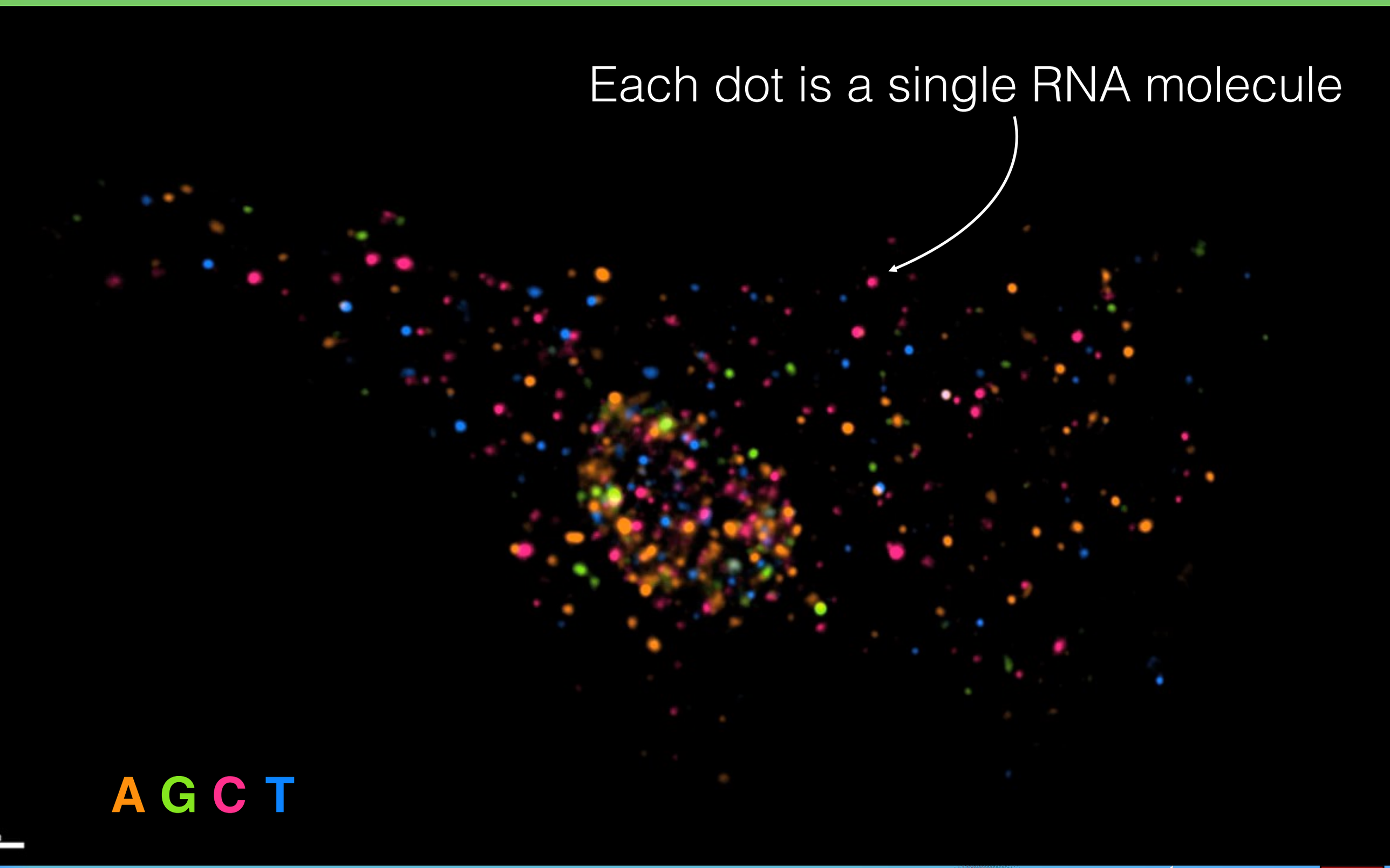
The FISSEQ approach

FISSEQ approach to massively multiplex *in situ* molecular detection is **sequencing**

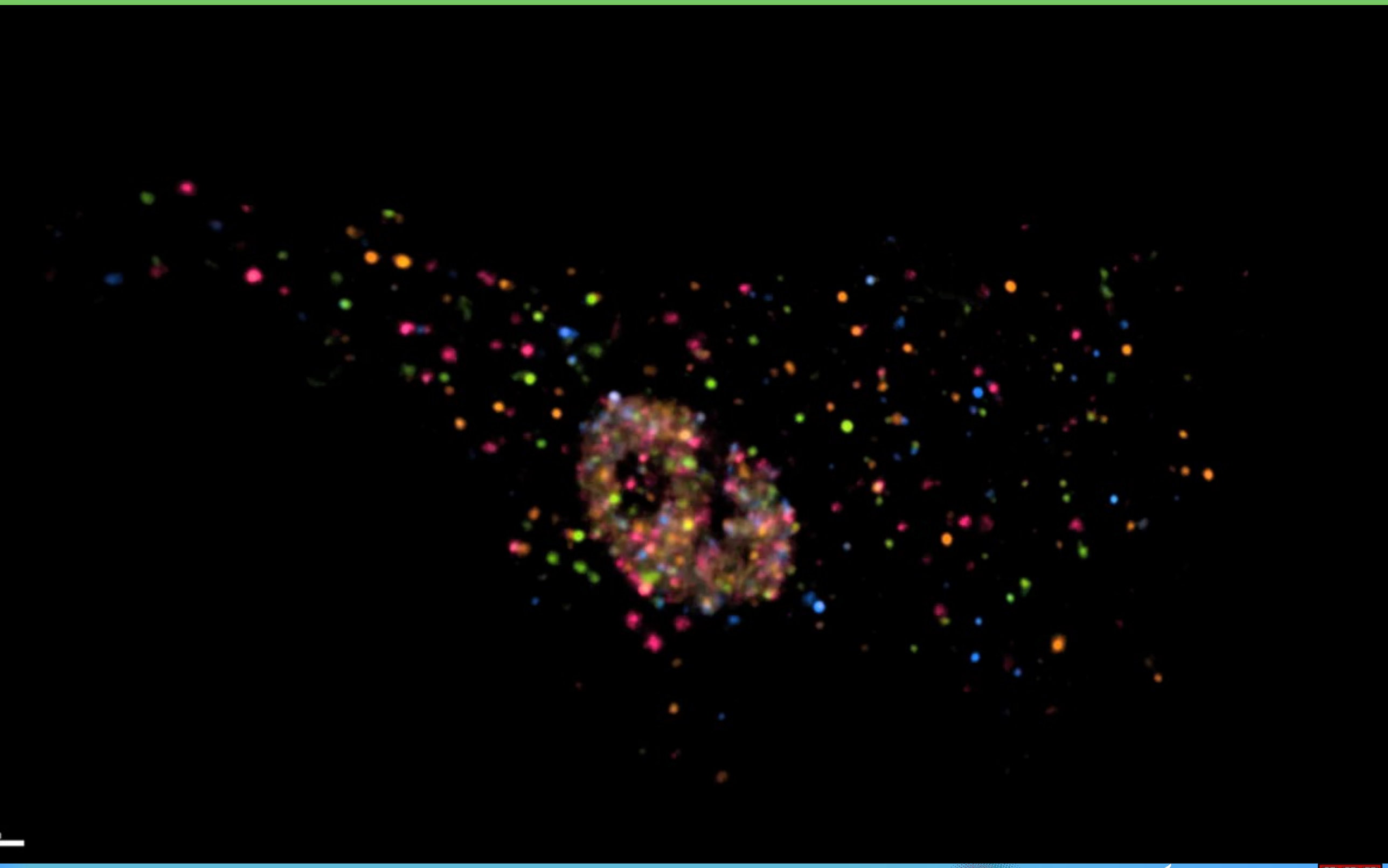
	Theoretical Multiplexity
Serial labeling	F×N
"Colorimetric" labeling	2^F-1
Sequencing	F^N



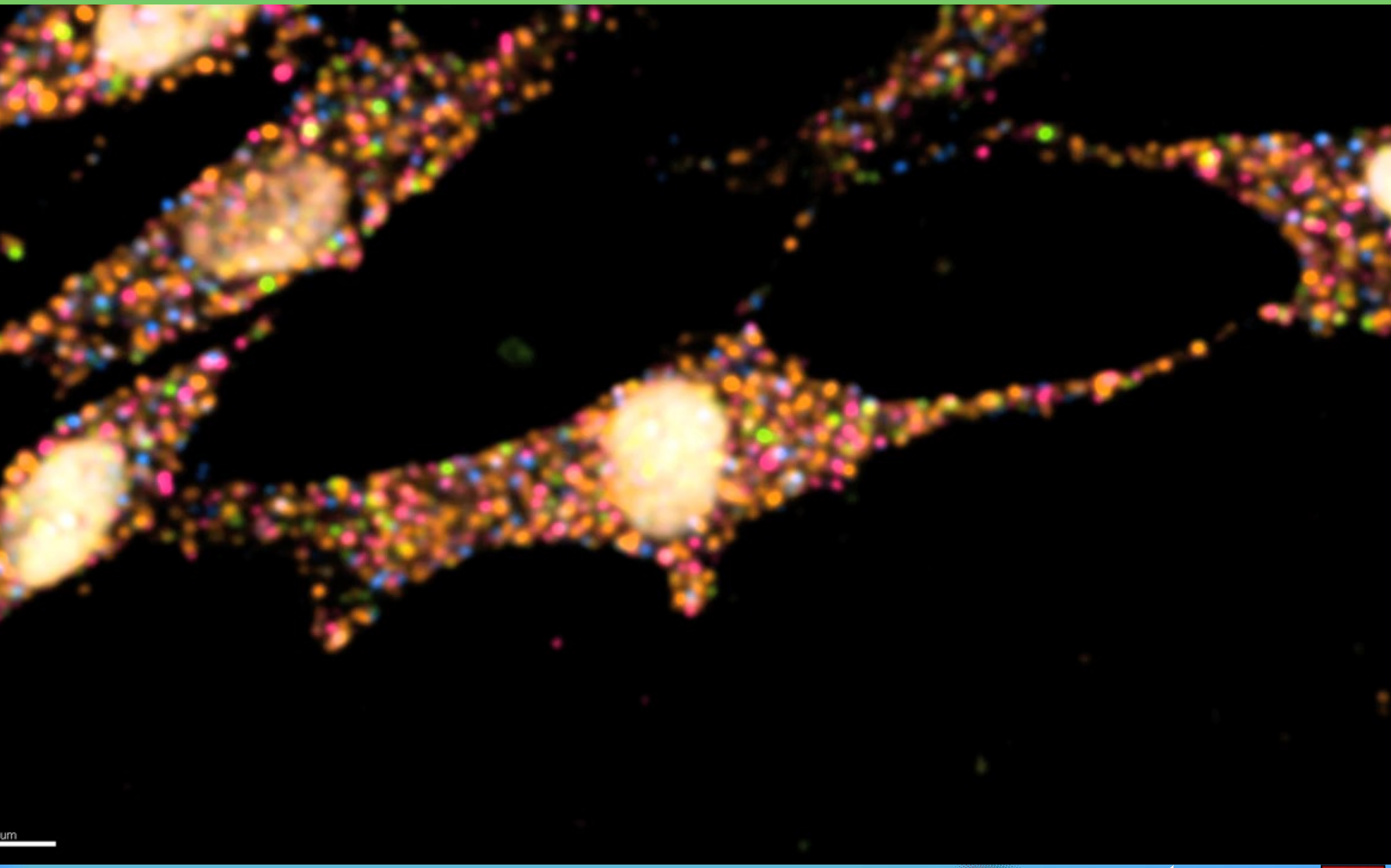
RNA FISSEQ Data



RNA FISSEQ Data



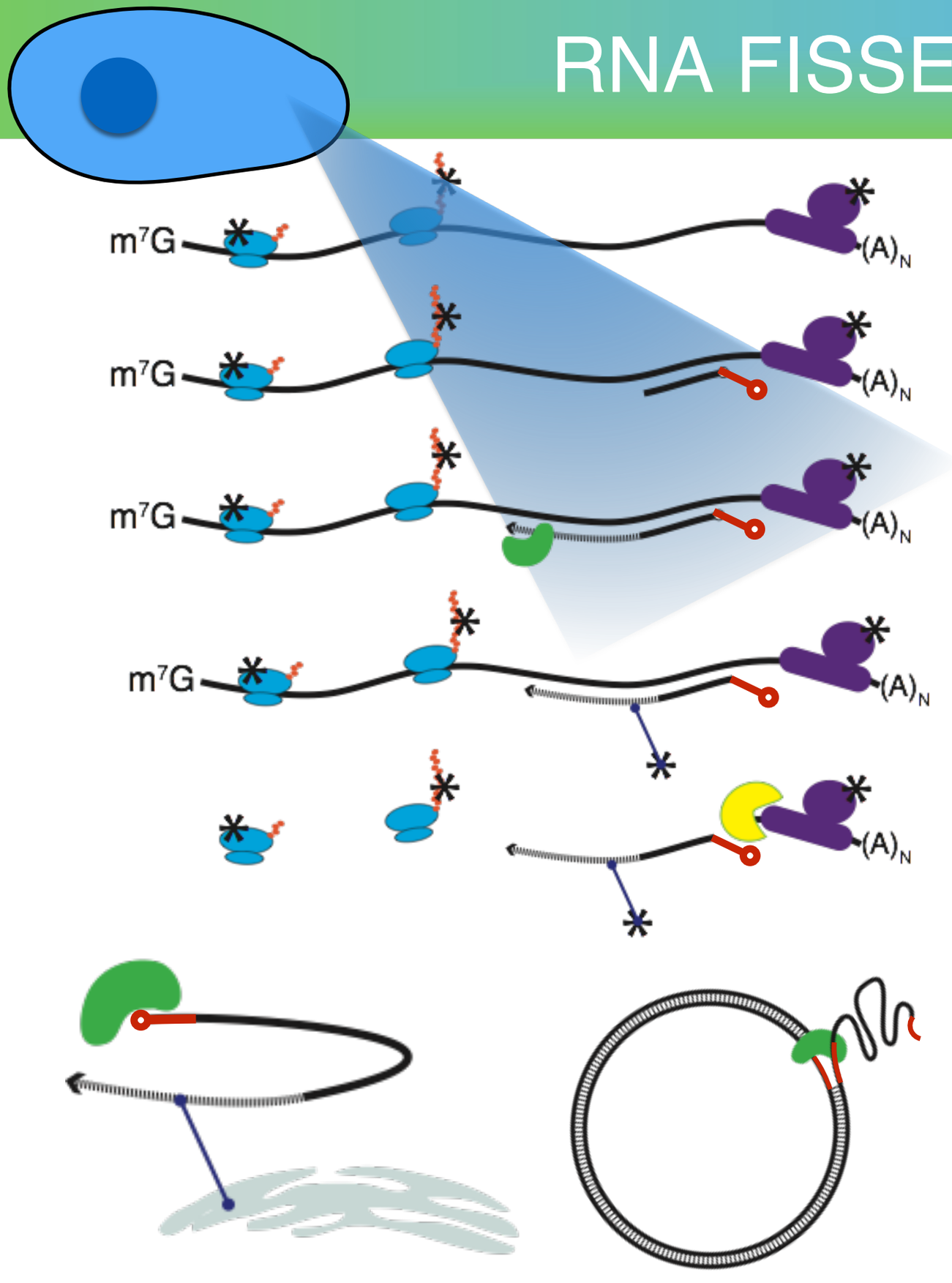
Molecular detection across spatial scales



Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

RNA FISSEQ Protocol



1. Fix RNA in place

2. Add RT primer (random hex)

3. Reverse transcription incorporating aminoallyl-dUTP

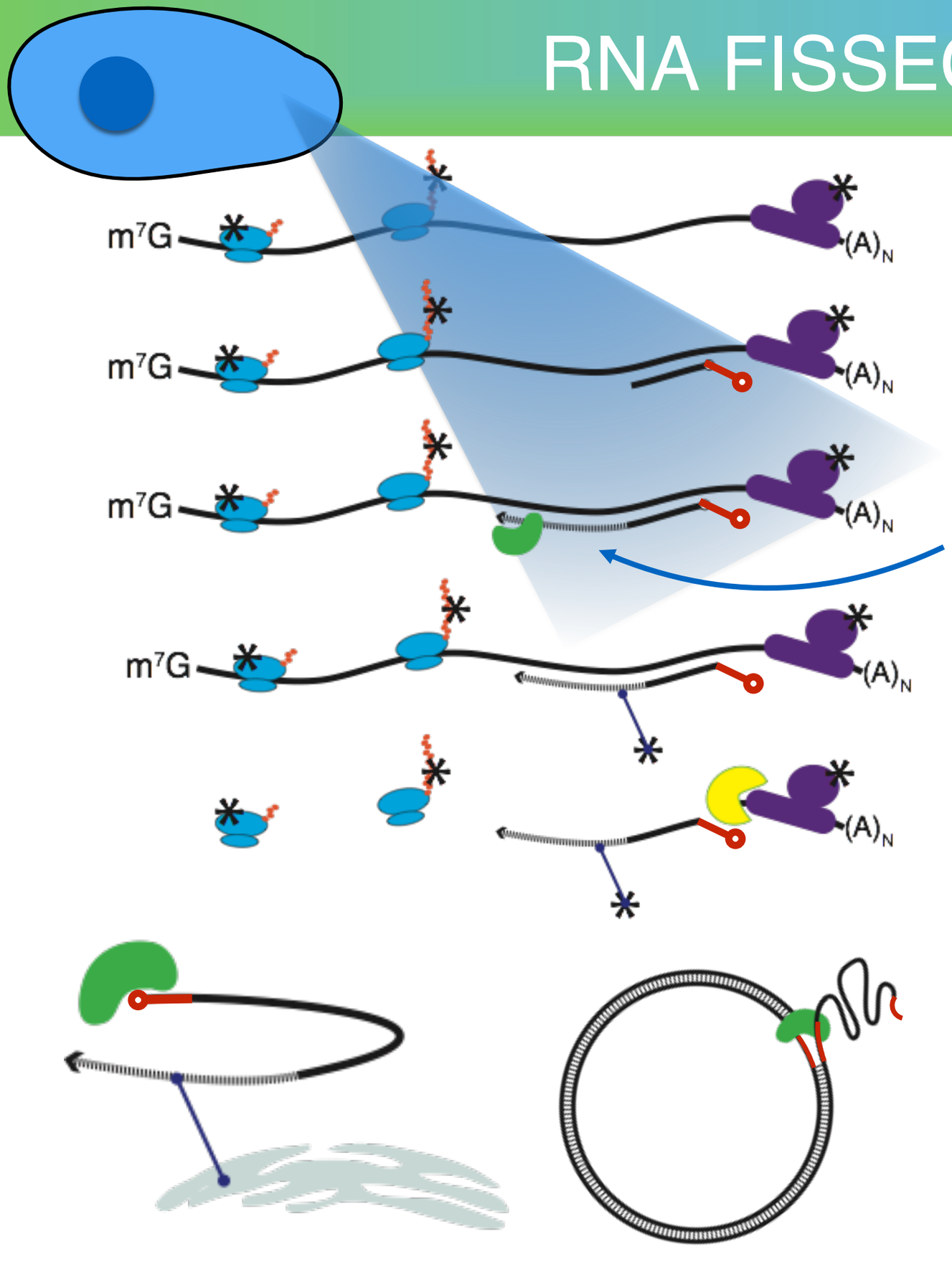
4. Cross-link cDNA using BS(PEG)9

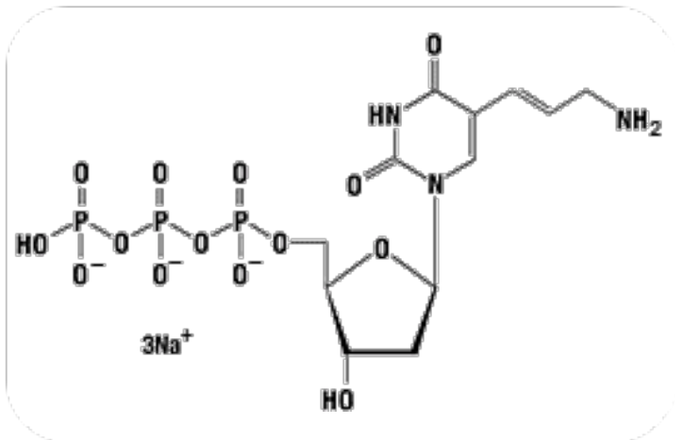
5. RNase to free cDNA ends

6. Circularize cDNA

7. Rolling circle amplification to generate sequencing amplicon “rollony”

RNA FISSEQ Protocol





1. Fix RNA in place

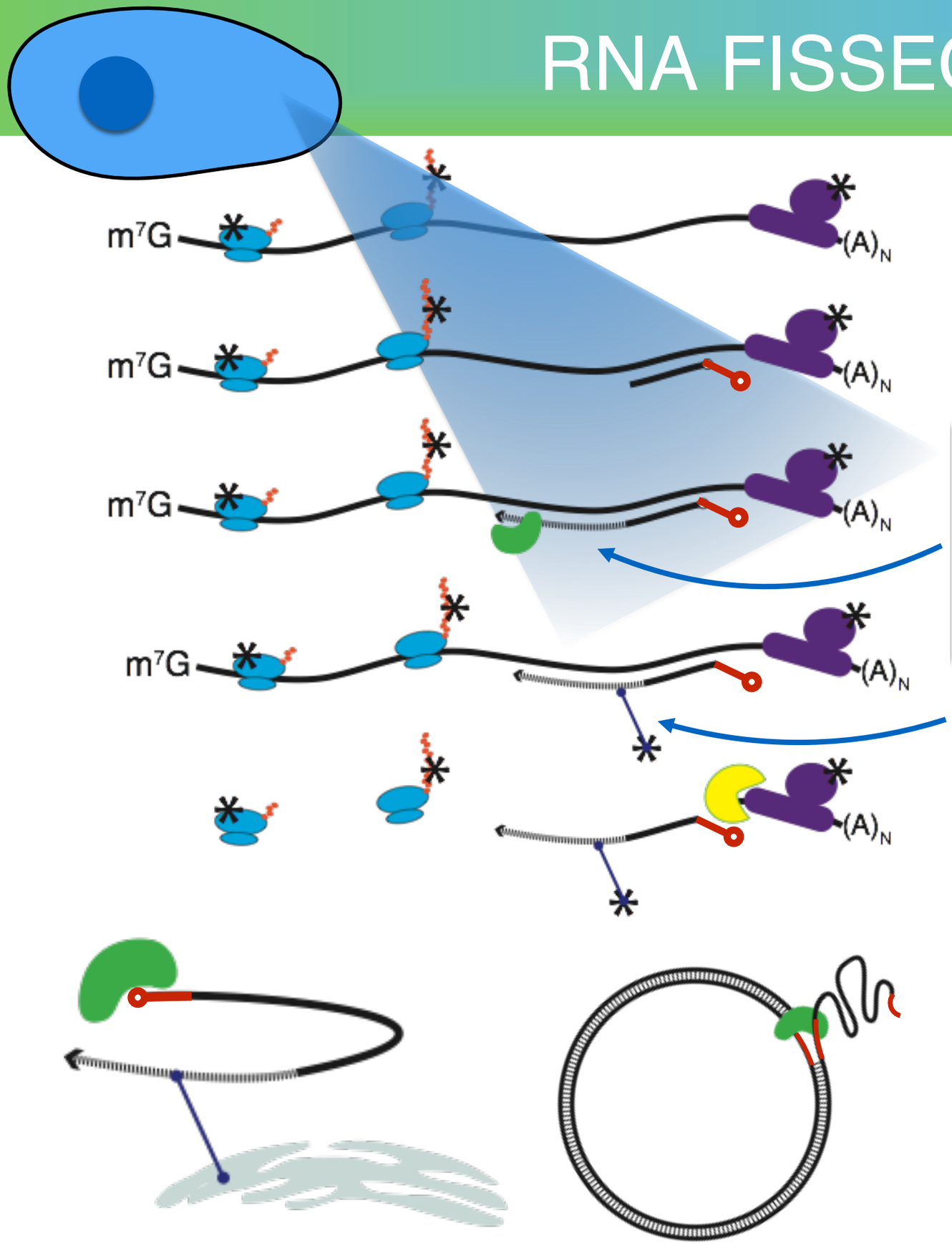
2. Add RT primer (random hex)

5. RNase to free cDNA ends

6. Circularize cDNA

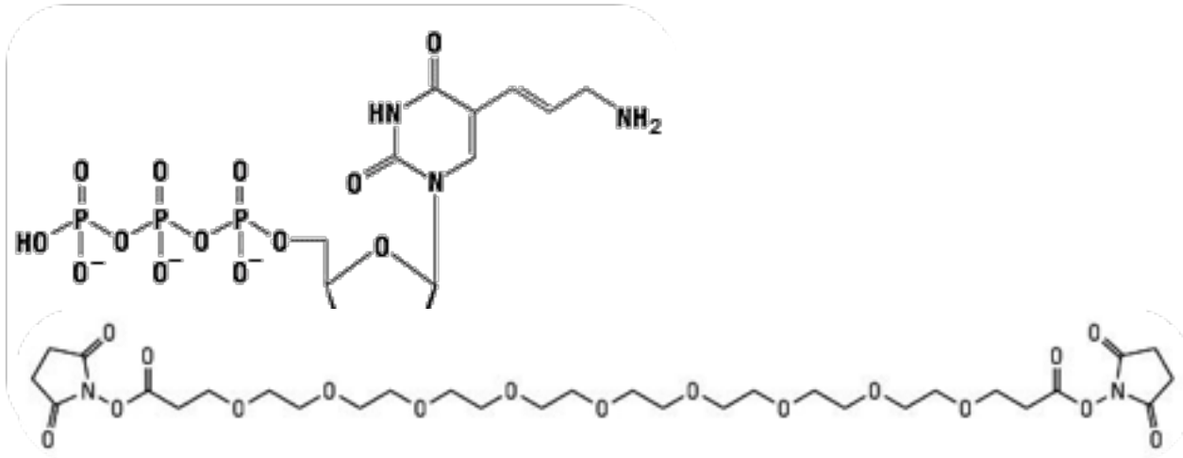
7. Rolling circle amplification to generate sequencing amplicon “rollony”

RNA FISSEQ Protocol



1. Fix RNA in place

2. Add RT primer (random hex)



5. RNase to free cDNA ends

6. Circularize cDNA

7. Rolling circle amplification to generate sequencing amplicon “rollony”

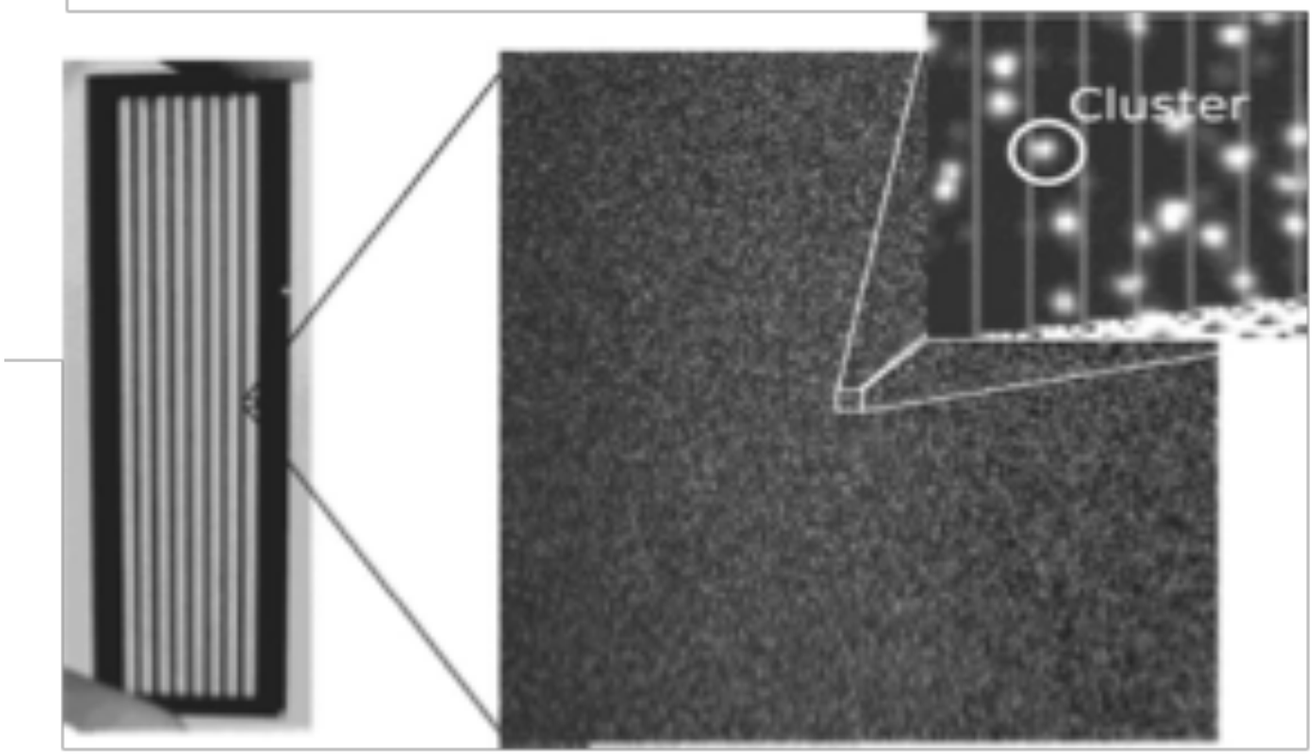
The *in situ* sequencing library

single RNA molecule capture

↓

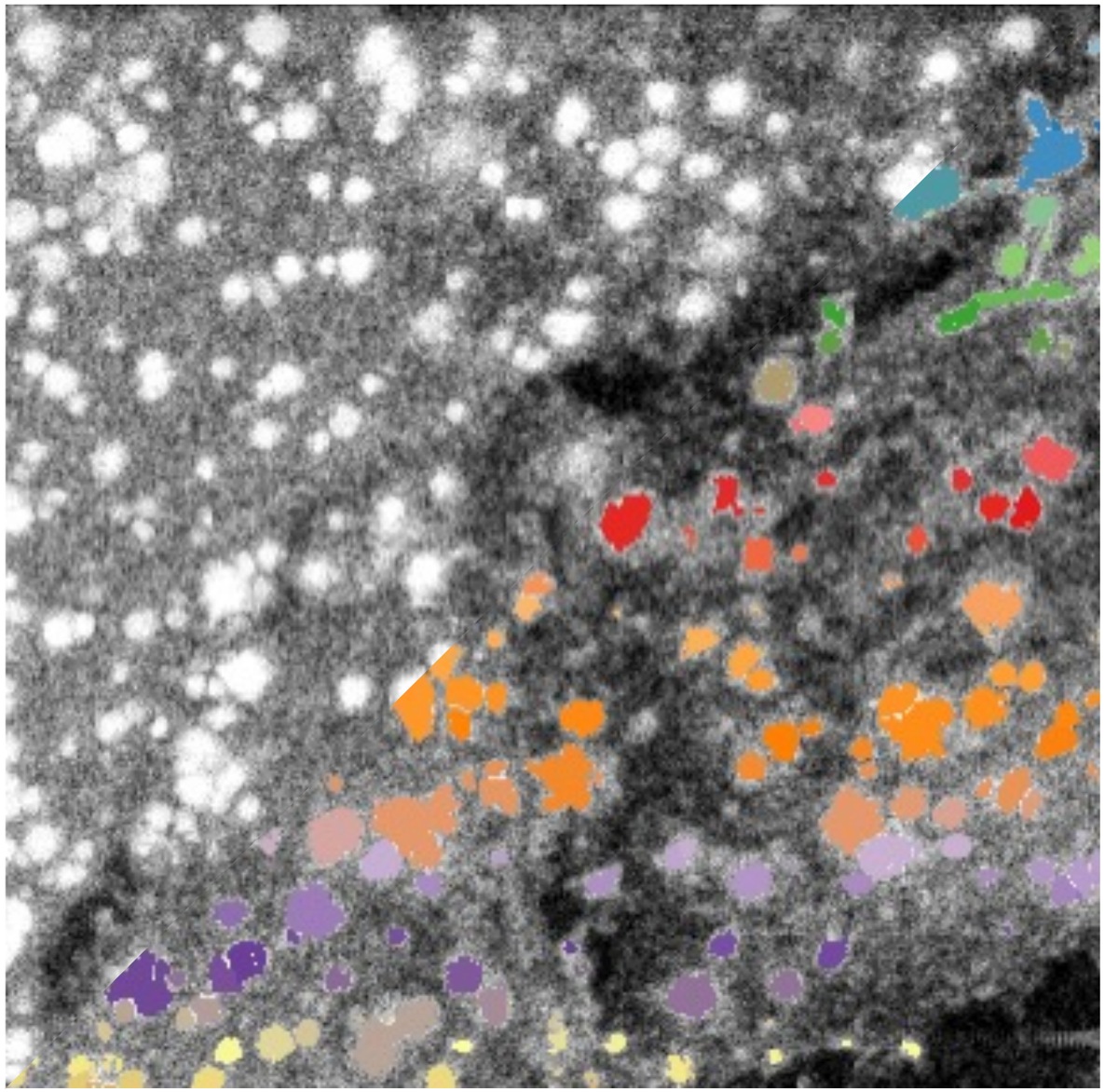
amplified sequencing template

traditional NGS



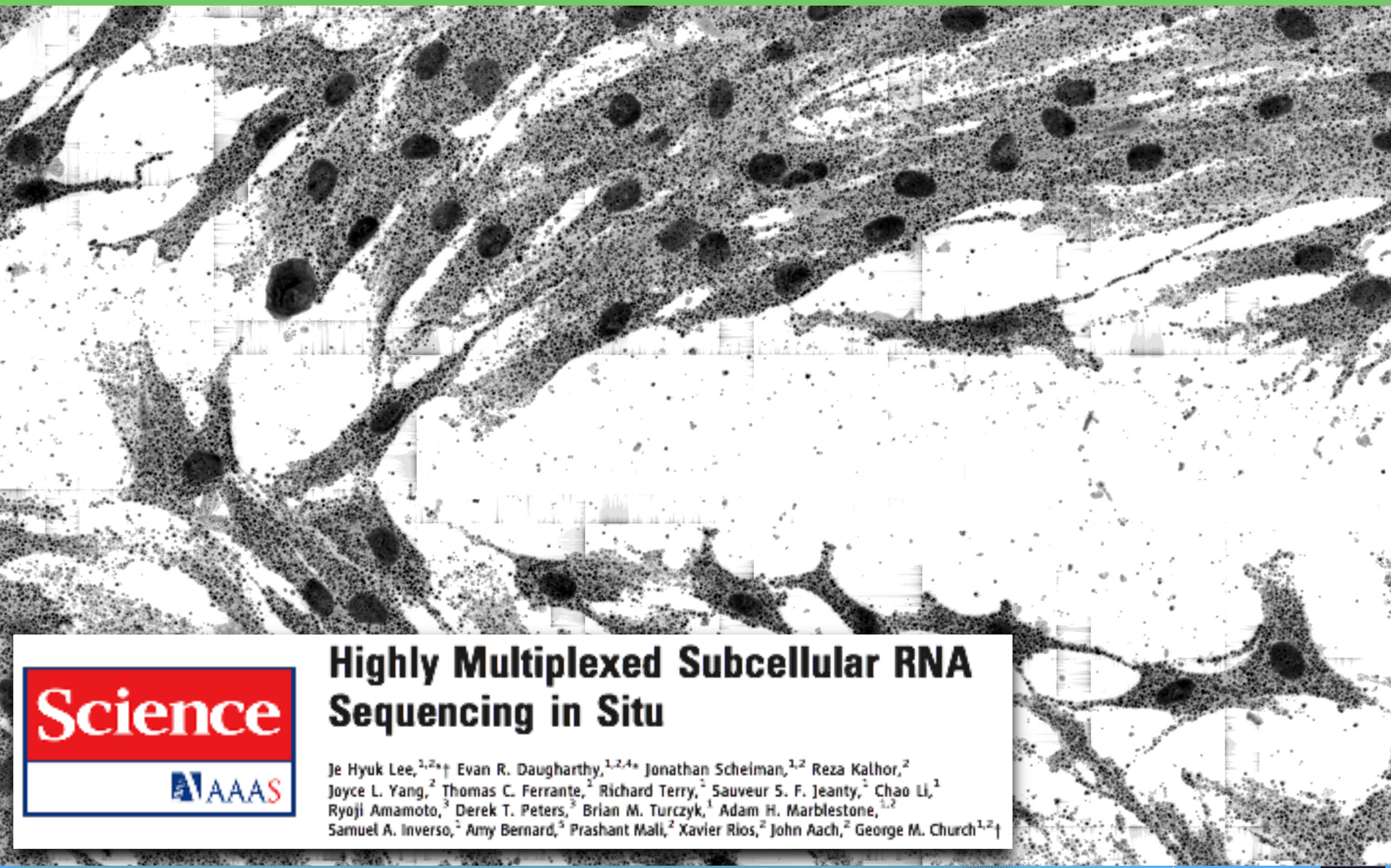
template = sequence

FISSEQ



sequence & position

RNA-FISSEQ of primary fibroblast wound healing



Science

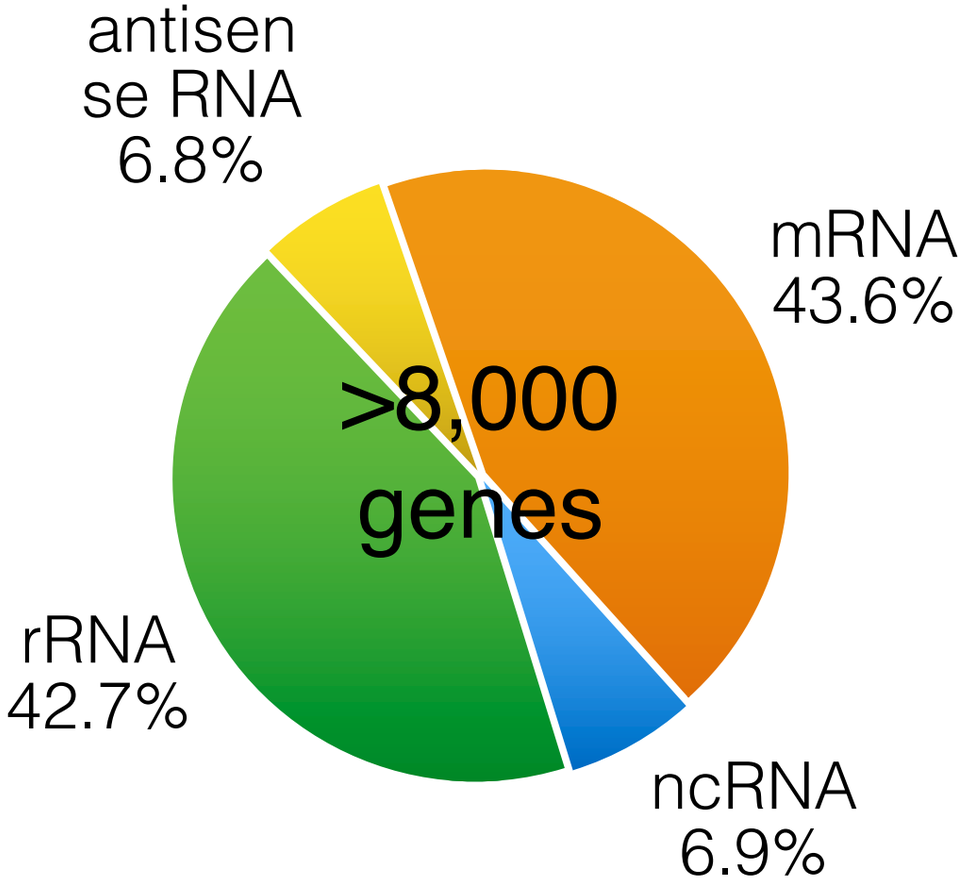
AAAS

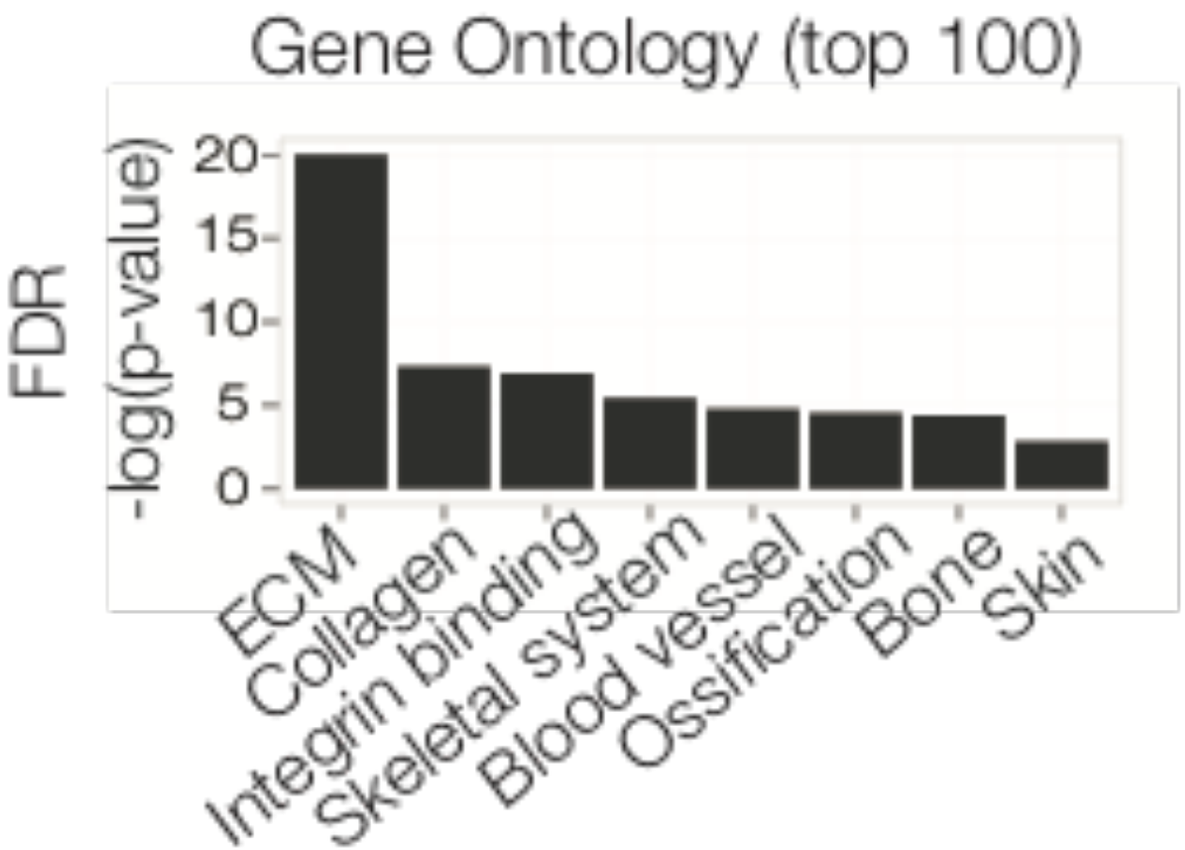
Highly Multiplexed Subcellular RNA Sequencing in Situ

Je Hyuk Lee,[1,2]* † Evan R. Daugharthy,[1,2,4]* Jonathan Scheiman,[1,2] Reza Kalhor,[2] Joyce L. Yang,[2] Thomas C. Ferrante,[1] Richard Terry,[1] Sauveur S. F. Jeanty,[1] Chao Li,[1] Ryoji Amamoto,[3] Derek T. Peters,[3] Brian M. Turczyk,[1] Adam H. Marblestone,[1,2] Samuel A. Inverso,[1] Amy Bernard,[5] Prashant Mali,[2] Xavier Rios,[2] John Aach,[2] George M. Church[1,2] †

RNA-FISSEQ data is RNA-seq data

Random hexamer reverse transcription captures from the whole transcriptome

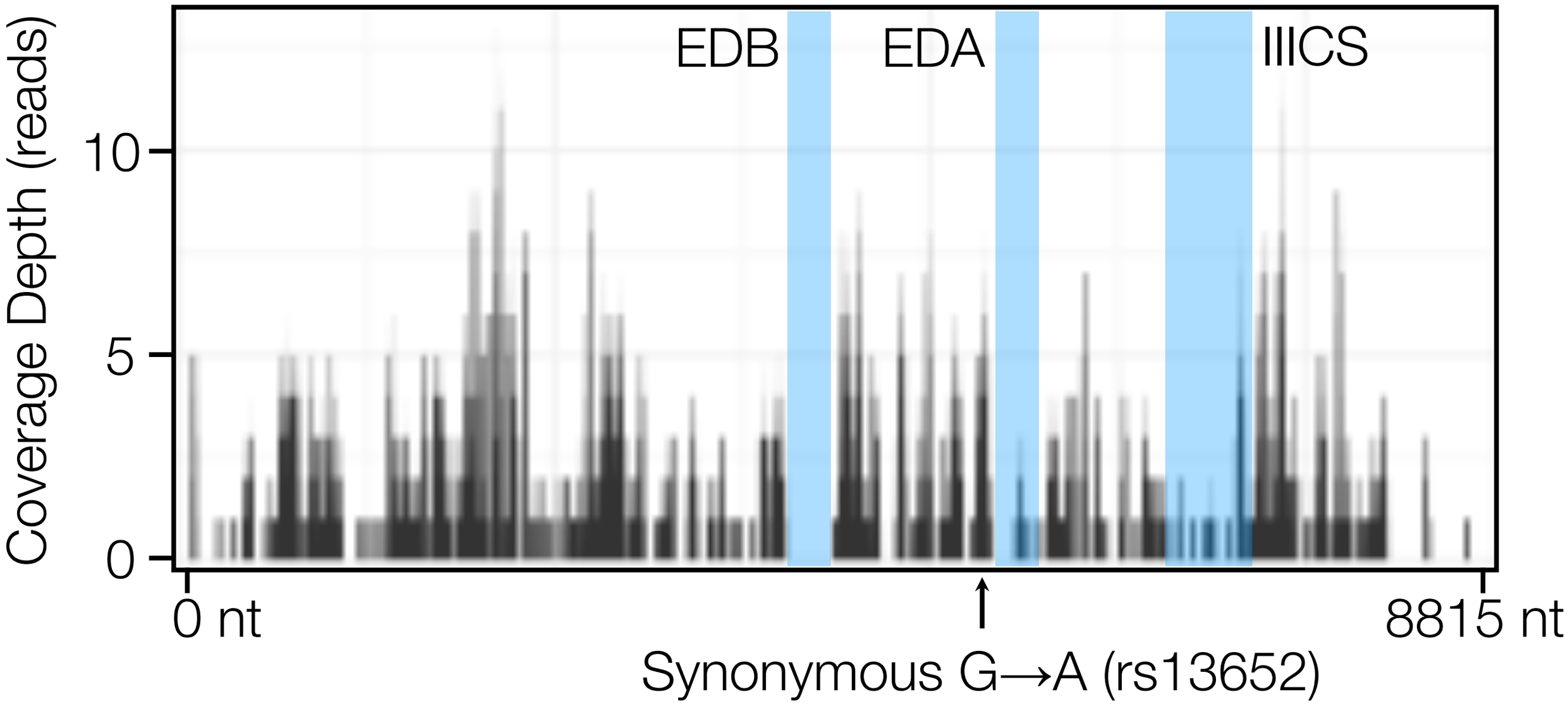




human primary fibroblast data

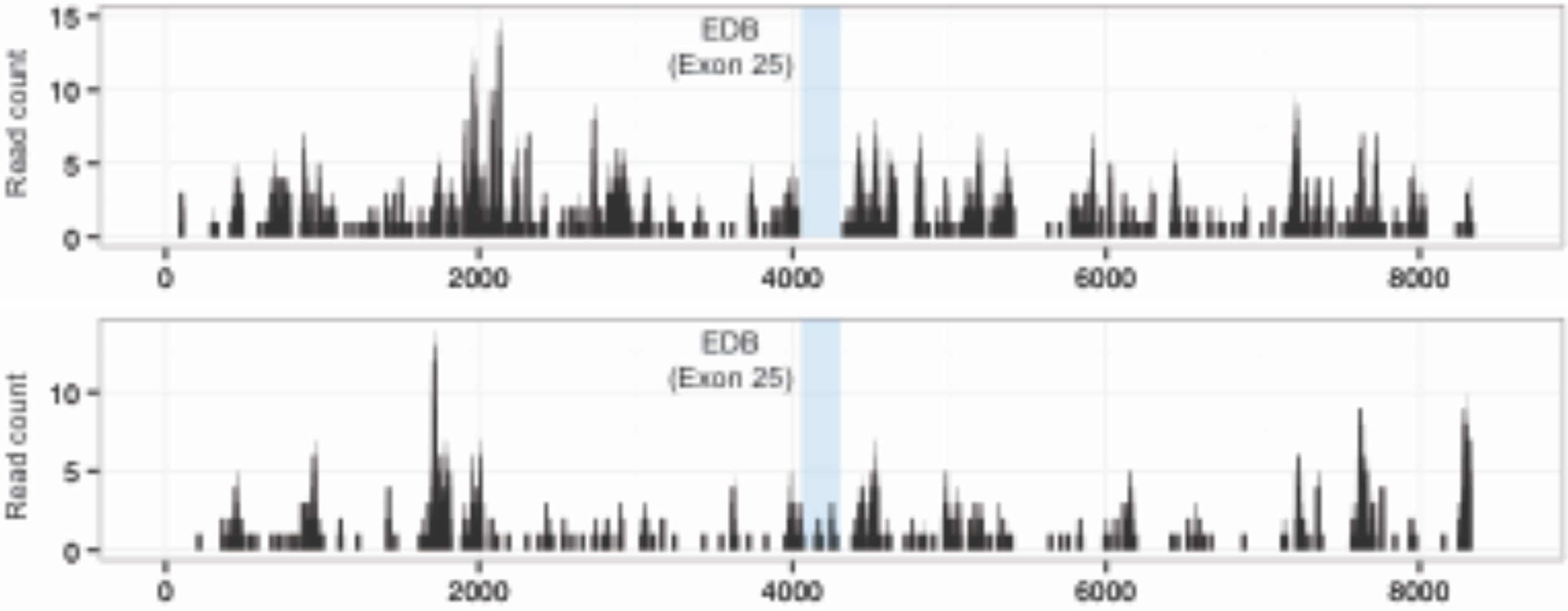
RNA-FISSEQ data is RNA-seq data

FN1 average per-base coverage 1.6×
(527 reads / 8.9 kb)



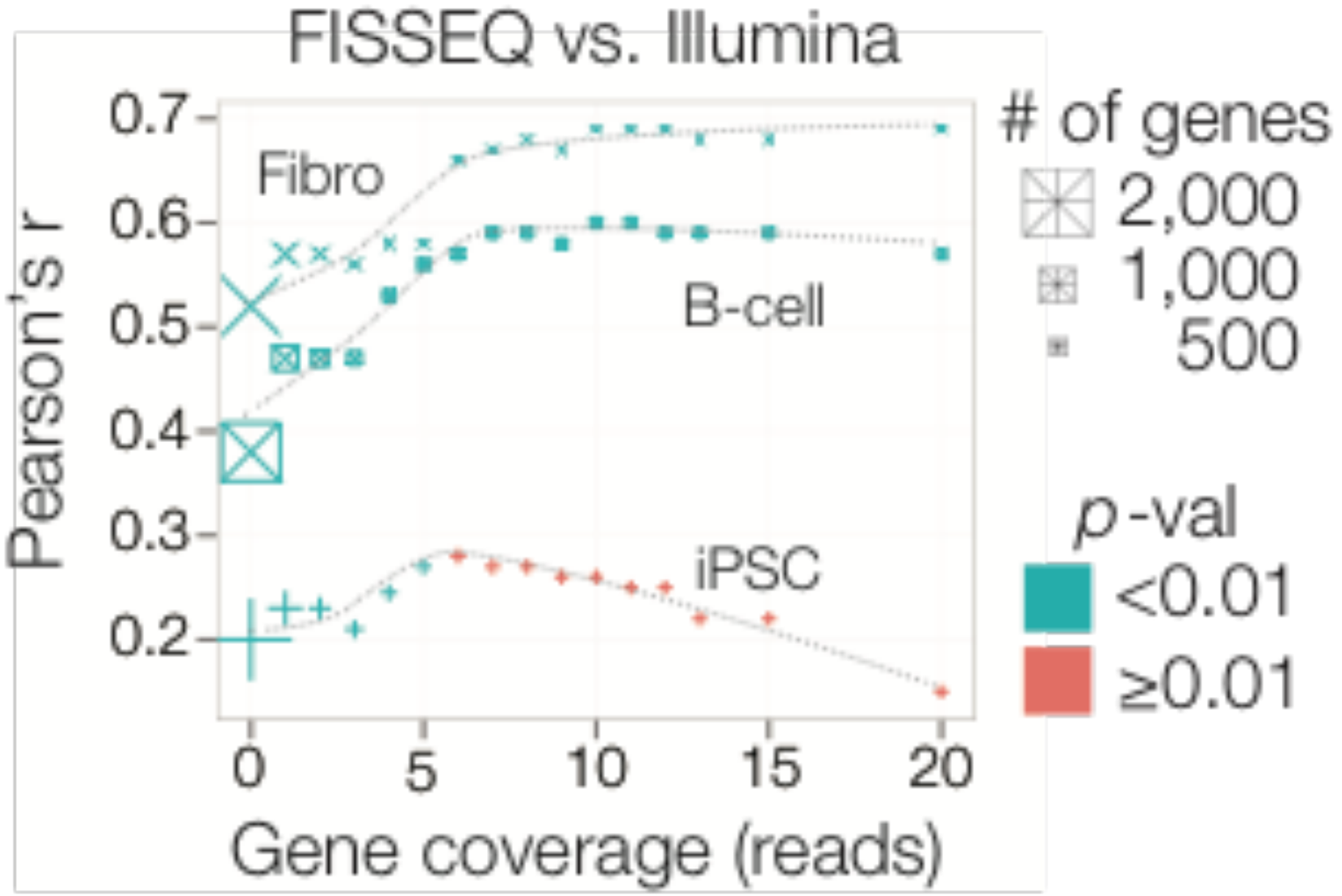
RNA-FISSEQ data is RNA-seq data

Media-dependent splicing of fibronectin in human primary fibroblasts reflects mesenchymal-epithelial transition



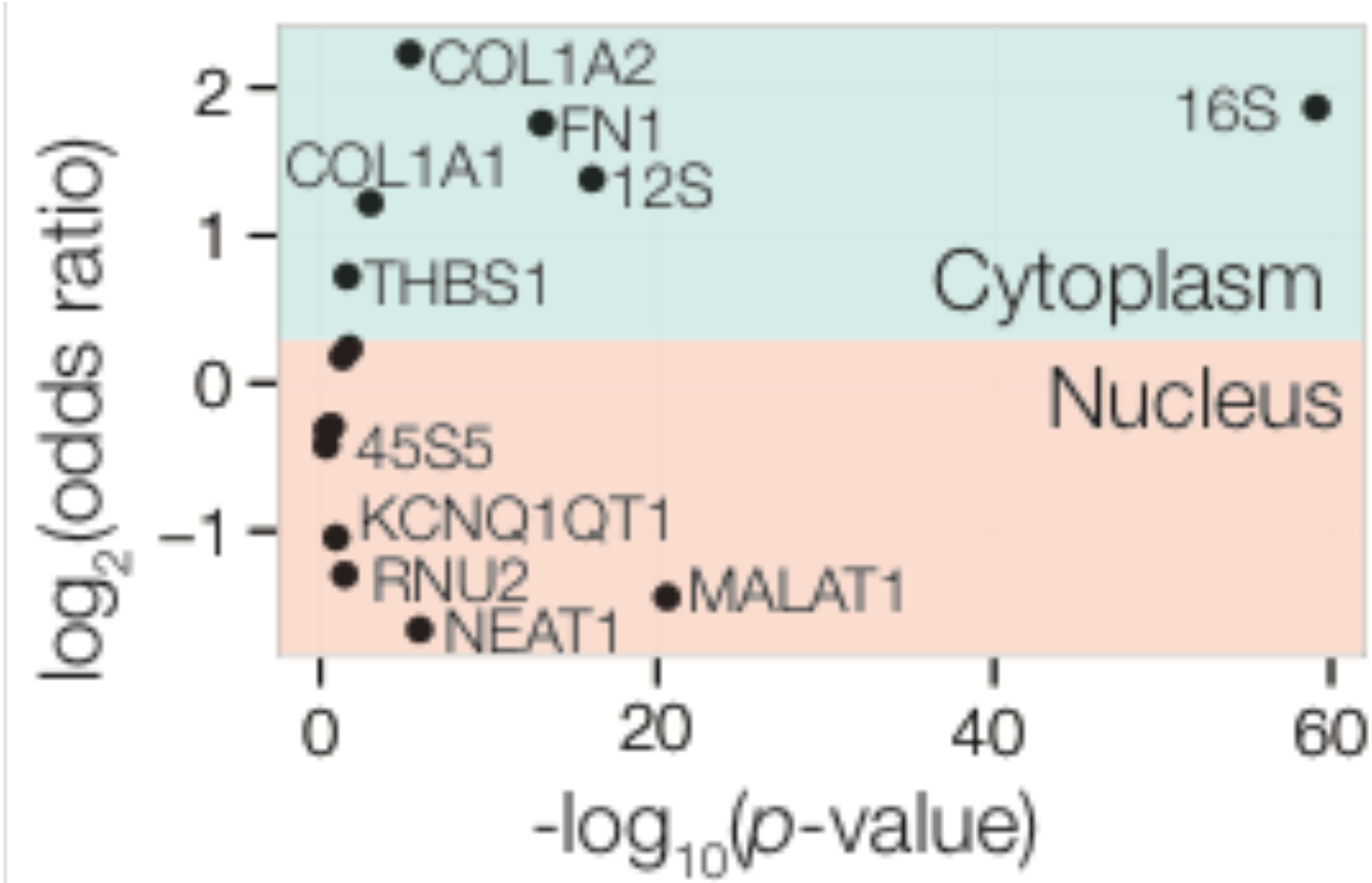
EDB Expression p<1E-16 in FBS vs. EGF Media

RNA-FISSEQ data is quantitative



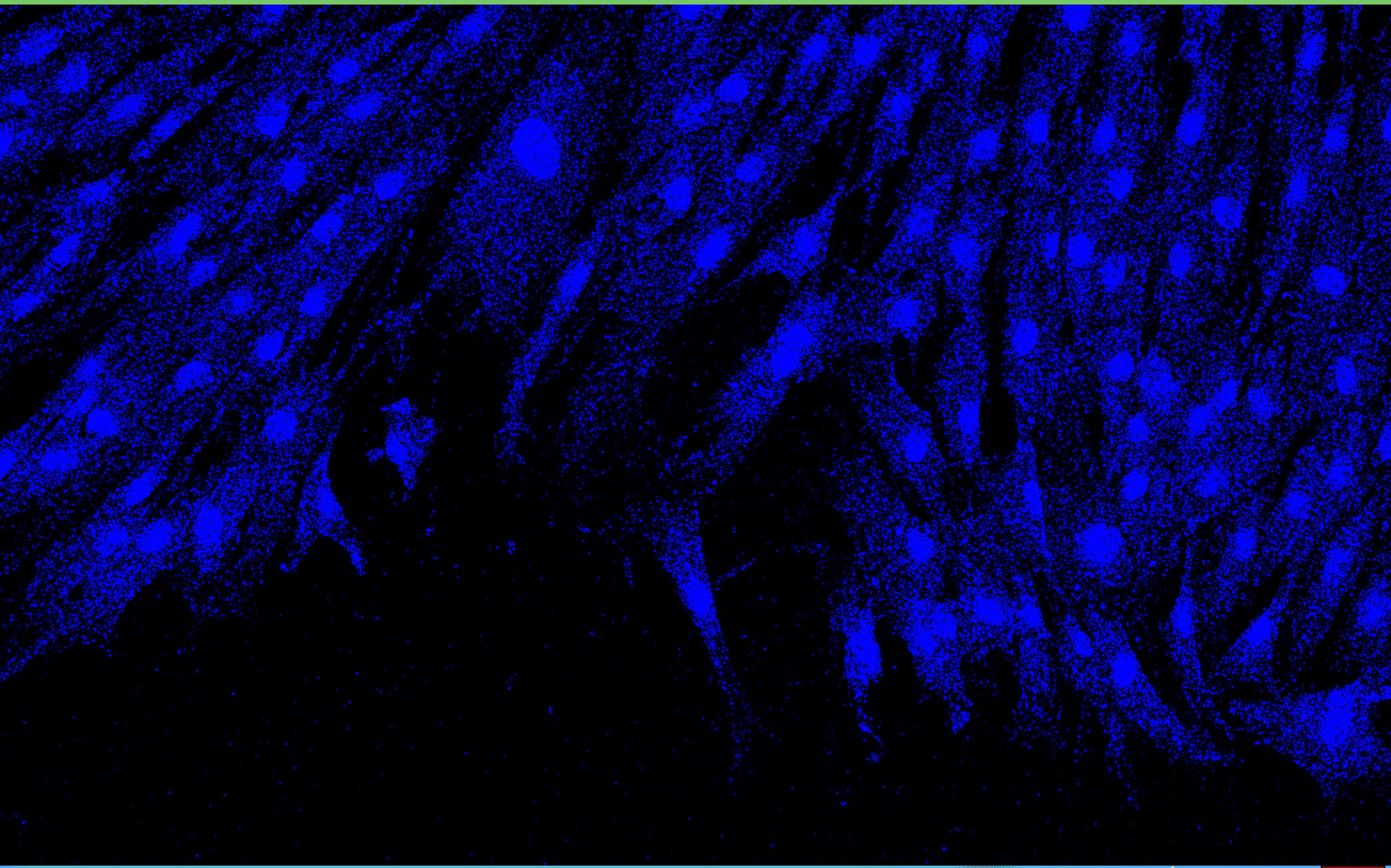
human primary fibroblast data

RNA-FISSEQ data is spatially resolved

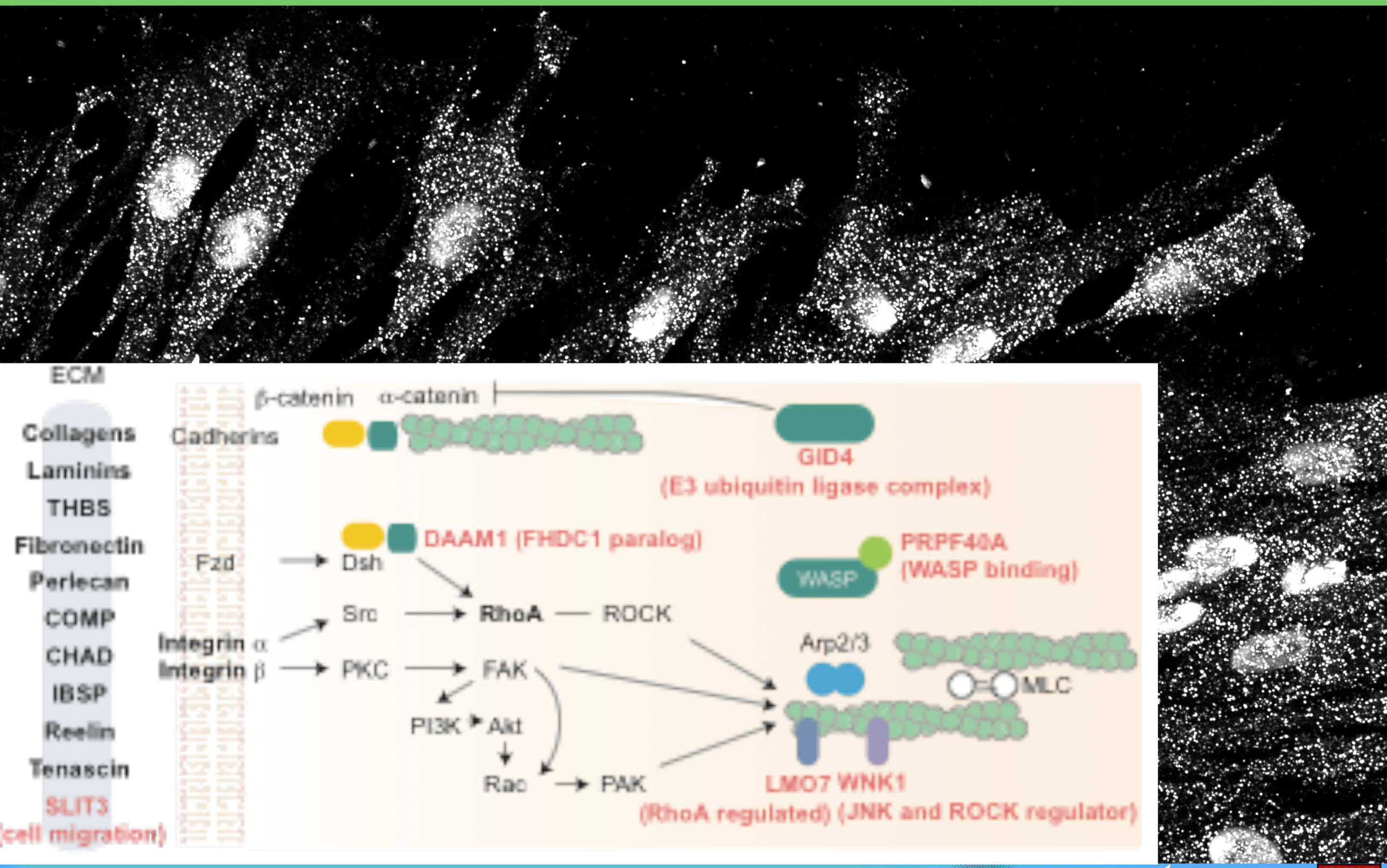


human primary fibroblast data

Wound healing model



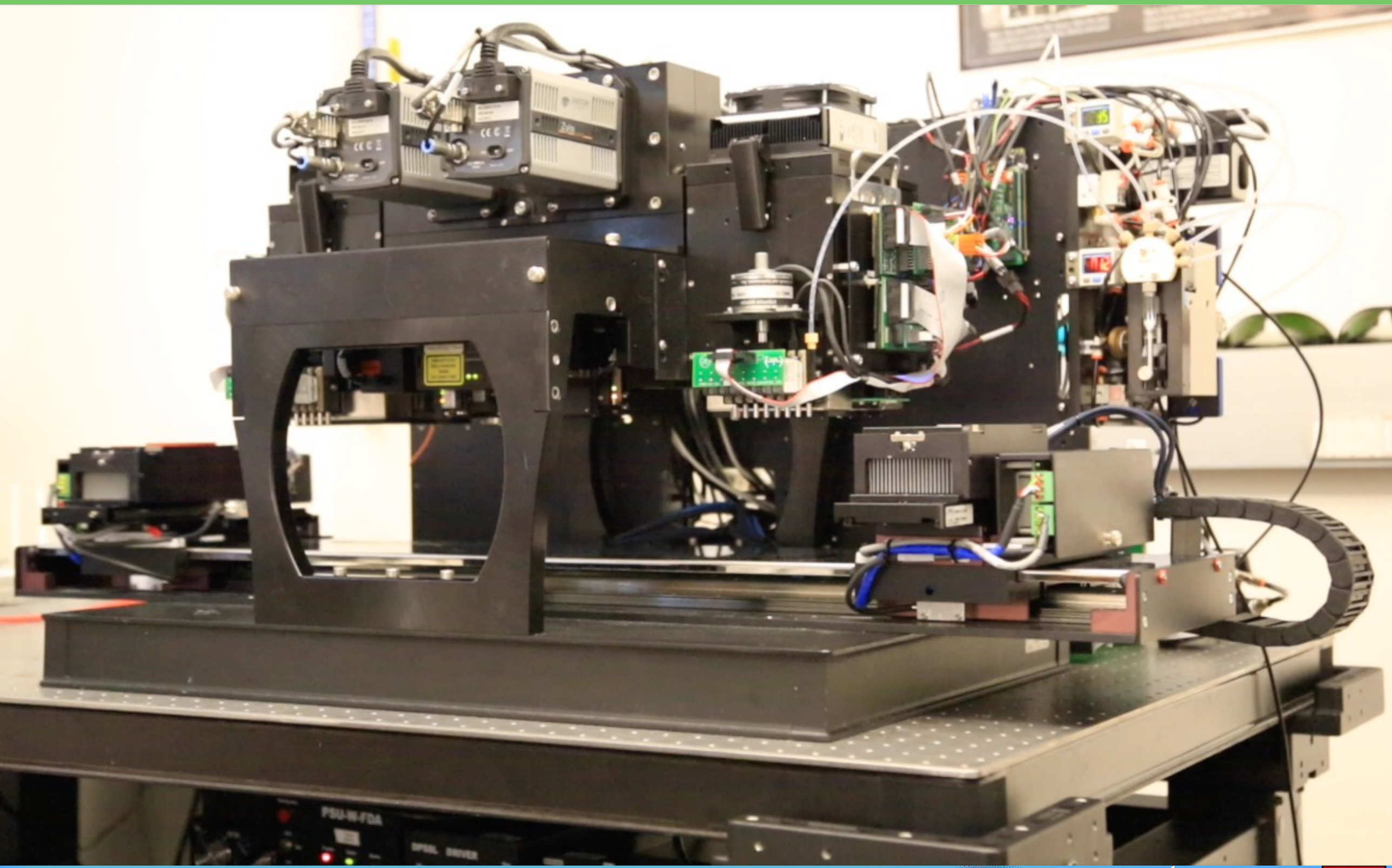
Molecular phenotype of wound sensing & response



Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Polonator H12



Other Related Technologies

Technology	Encoding	Transmission/ Decoder	Used For	Exemplary Companies
Nanopore Sequencing	None (direct measurement) Or synthesis of complementary strand	Charge/impedance changes as molecule passes through a pore (electrical signals)	DNA, RNA, protein Any polymers	Oxford Nanopore Genia (Roche)
Single-Cell RNA/DNA Sequencing	DNA/RNA enzymatic processing adds barcode (information tag) to each cell in a micro-droplet before mixing for NGS	Fluorescence / Optical (NGS)	DNA, RNA	10X Genomics Bio-Rad
Single-Cell/Spatial Protein	DNA barcoding	Fluorescence / Optical (NGS)	Proteins	CITE-Seq (10X) Akoya
Imaging Mass Spec	None (direct measurement) - ionizes the molecules on the surface of the sample and collects a mass spectrum at each pixel	Mass spectrometry (electrical signal)	Proteins	Bruker Thermo

Evan Daugharthy
evan.daugharthy@gmail.com

Other Related Technologies

Technology	Encoding	Transmission/ Decoder	Used For	Exemplary Companies
smFISH	Hybridization Reaction	Fluorescence / Optical (spectral multiplexing)	RNA/DNA	ACD RNAscope ACD Base Scope
Multiplex FISH	Hybridization Reaction(s)	Fluorescence / Optical (temporal multiplexing / barcode detection)	RNA/DNA	ACD High Multiplex MERFISH OligoFISSEQ
Spatial Transcriptomics	Spatial DNA tagging (DNA/RNA processing biochemistry)	Fluorescence / Optical (NGS)	RNA	10X Visium
FISSEQ	DNA/RNA processing biochemistry	Fluorescence / Optical (NGS)	RNA / DNA	ReadCoor / 10X Genomics
ISS (in situ sequencing) & Targeted FISSEQ	Hybridization Reaction(s) & DNA/RNA biochemistry	Fluorescence / Optical (temporal multiplexing / barcode detection)	RNA / DNA	Cartana / 10X Genomics

Evan Daugharthy
evan.daugharthy@gmail.com

Pause

End of technology – any questions/discussion

Application discussion to follow

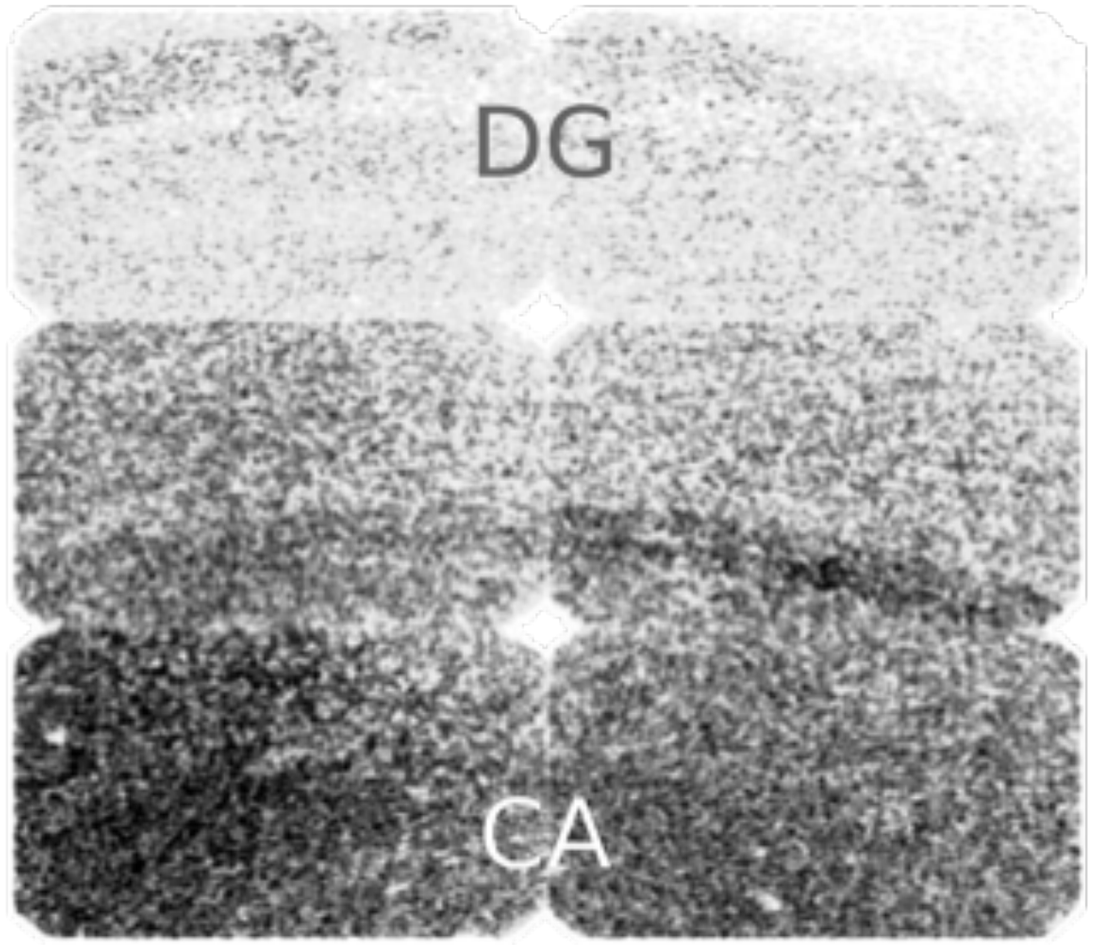
Evan Daugharthy
evan.daugharthy@gmail.com

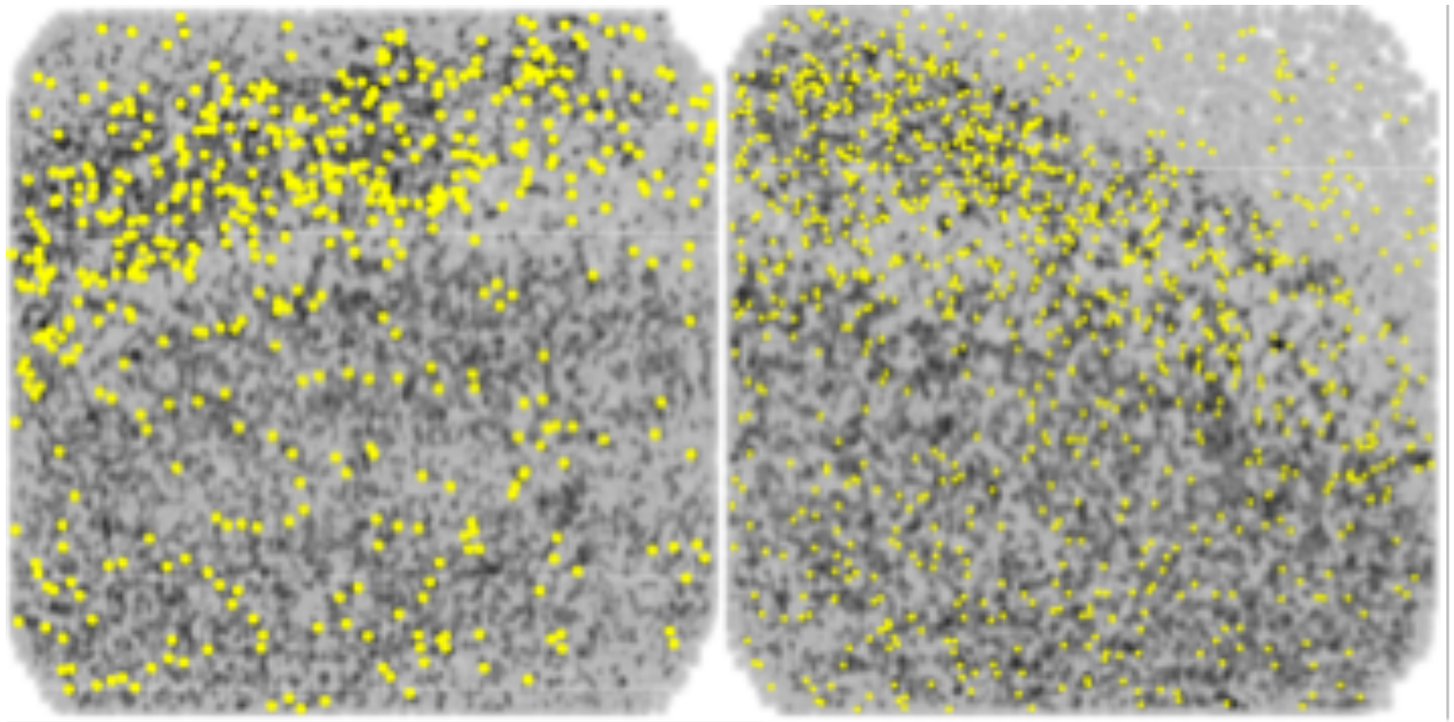
Brain FISSEQ

1989	2006	2015
210 cell types	145 neuron types	53 single neurons sequenced
1 neuron		
Alberts, Molecular Biology of the Cell	Vikaryous, Human cell type…	Dueck, Deep sequencing…

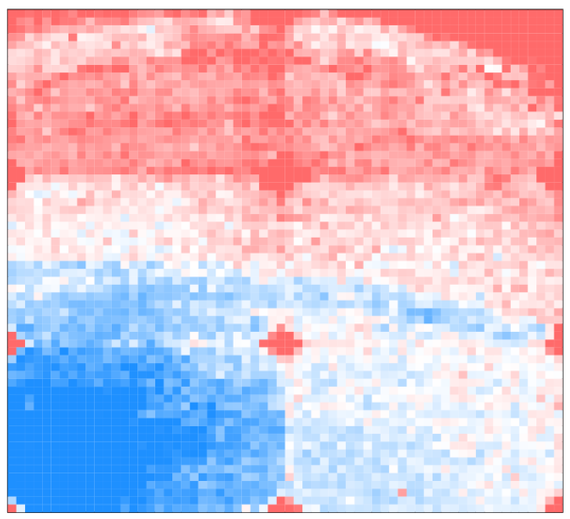
mouse hippocampus

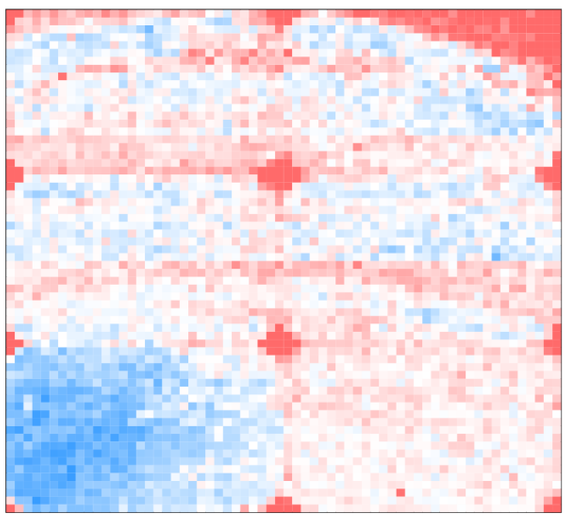
Brain FISSEQ

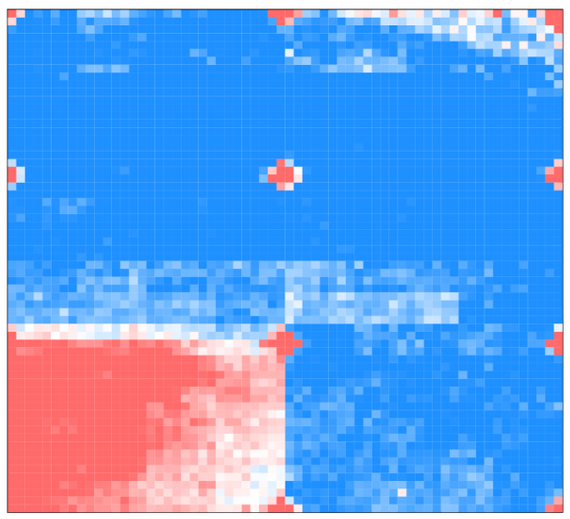


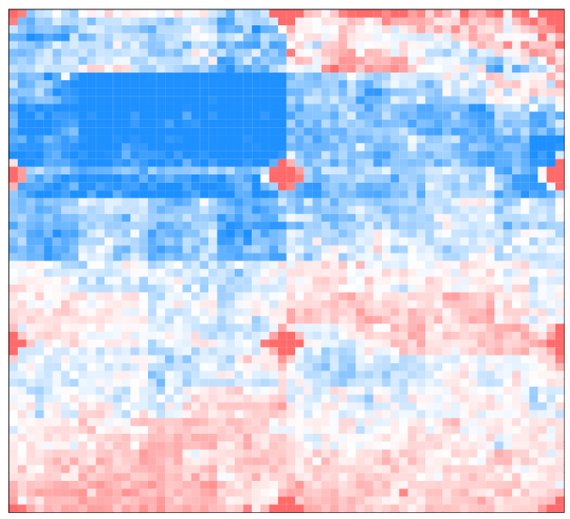


1 exocytosis/secretion/transport
2 cytoskeletal processes
3 homeostasis and macromolecule modification
4 epithelial cell migration &c.



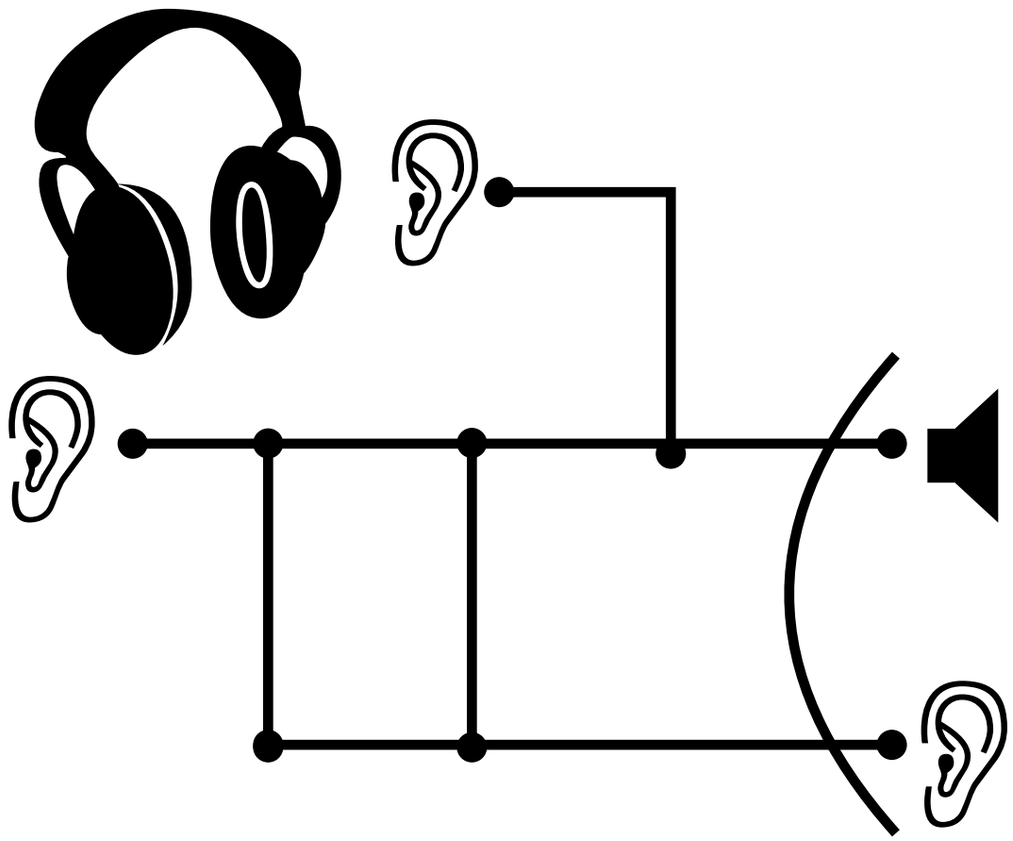


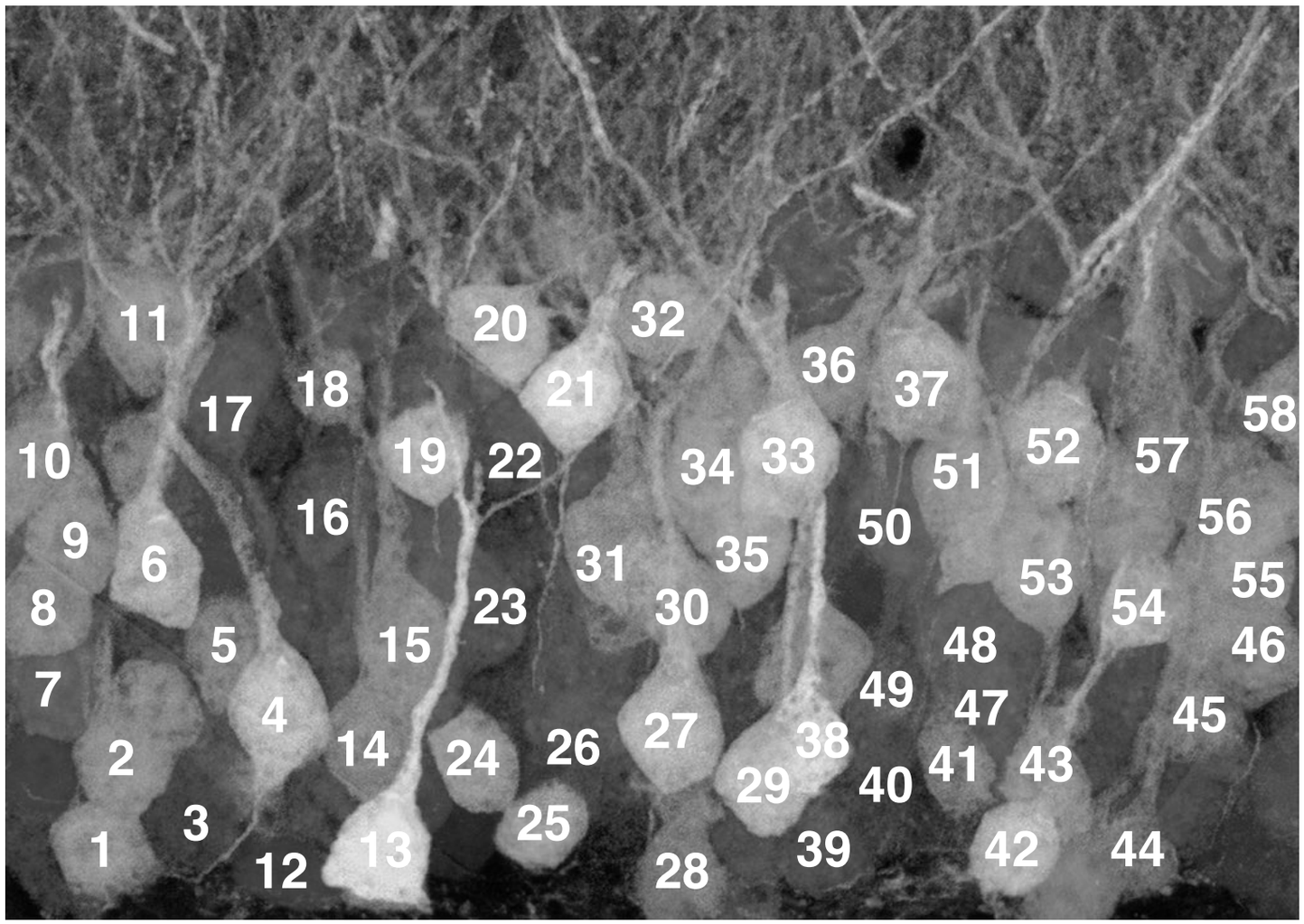




Rosetta Brain

Connectomic reconstruction is possible by barcoding neuronal connections

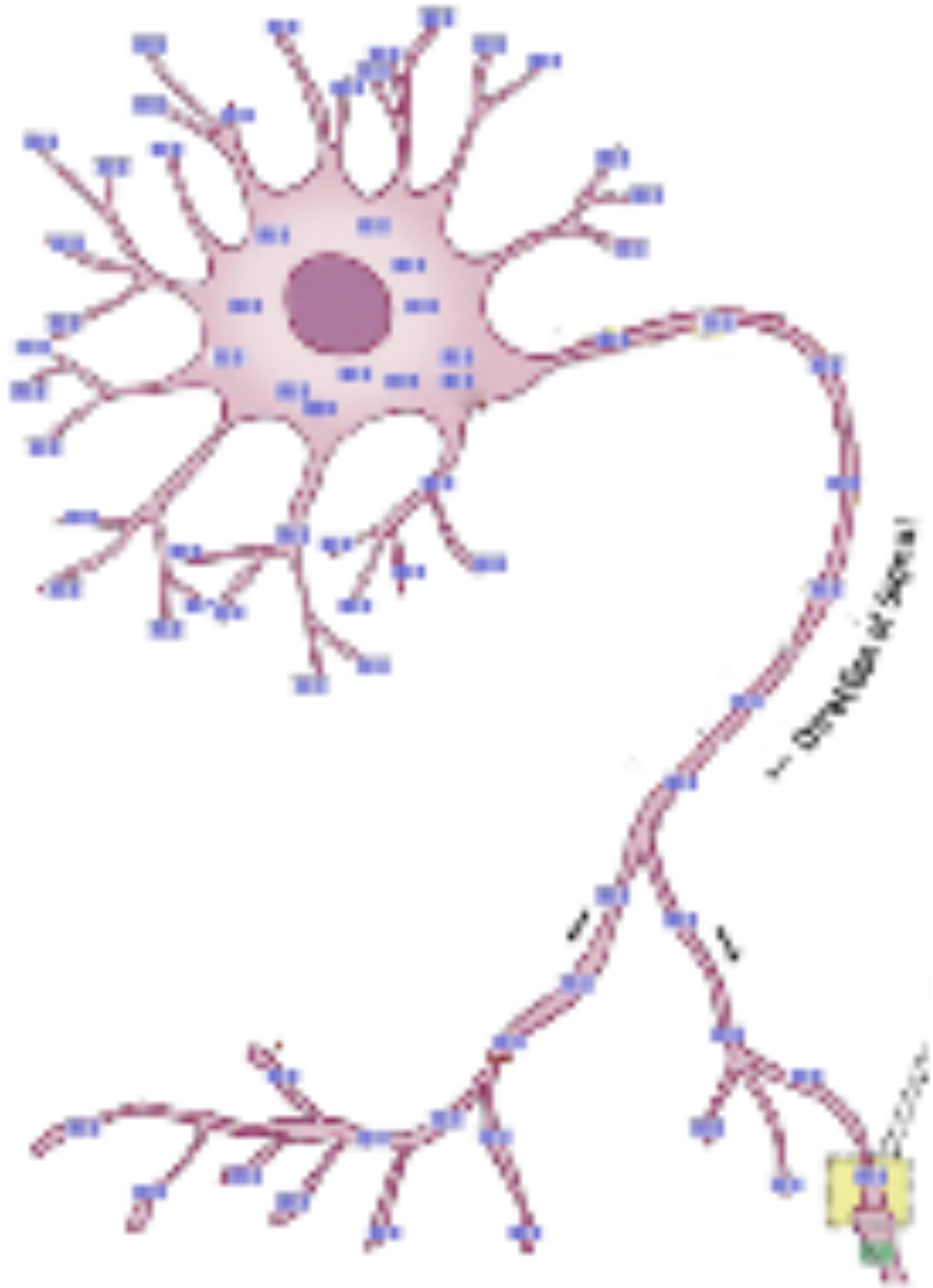




Marblestone, Daugharthy et al. ***arXiv*** (2014) arXiv:1404.5103

Marblestone, Daugharthy et al. ***arXiv*** (2014) doi:10.1101/001214

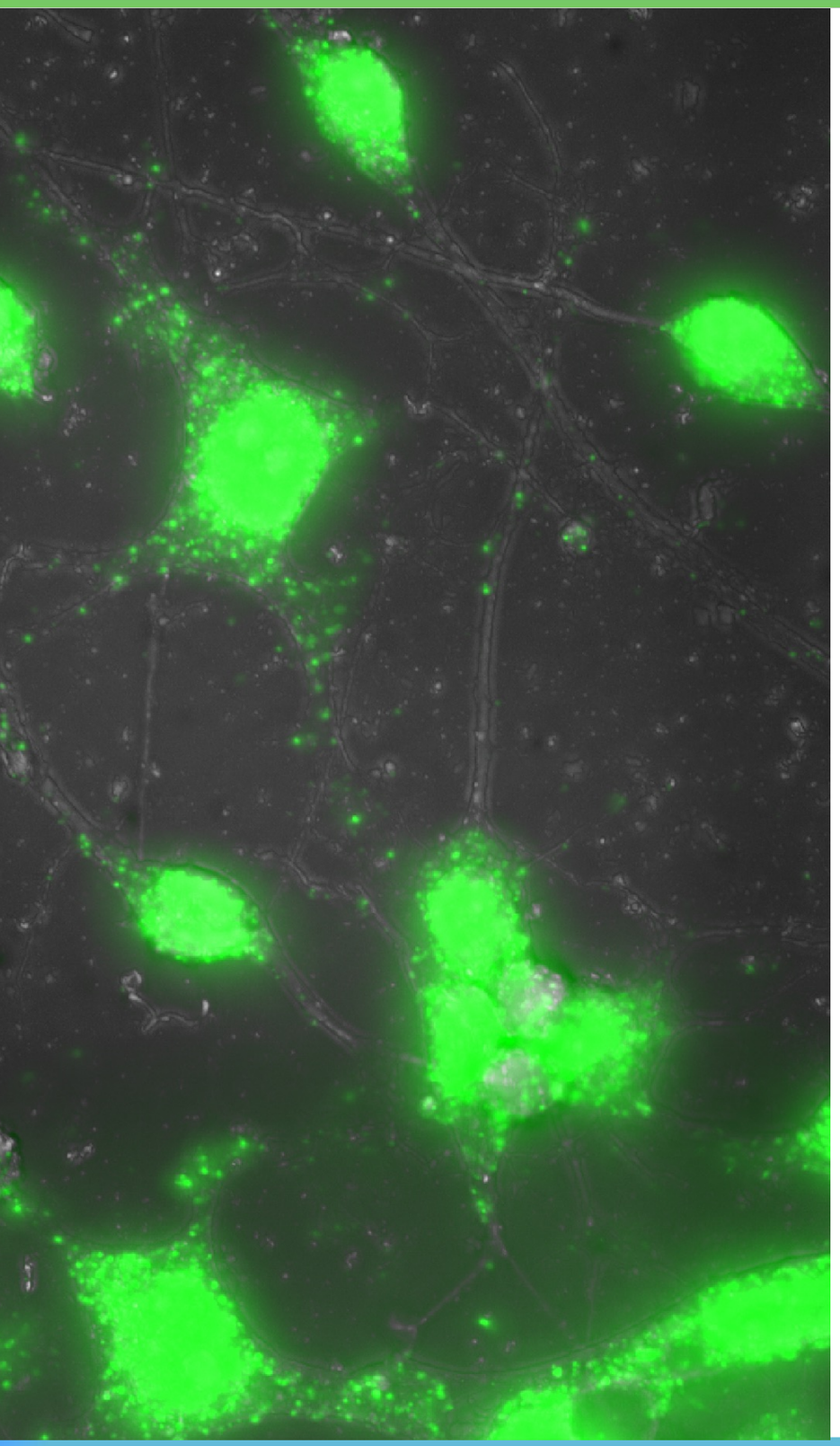
Rosetta Brain



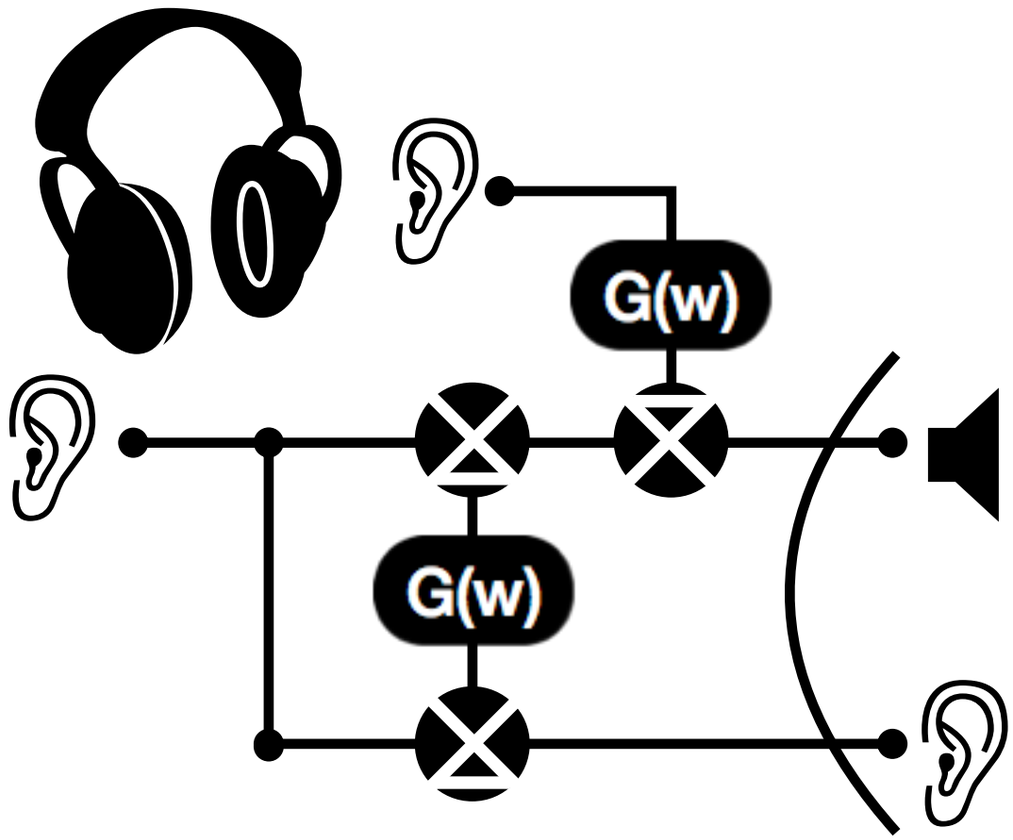
Marblestone, Daugharthy et al. *arXiv* (2014) arXiv:1404.5103

Marblestone, Daugharthy et al. *arXiv* (2014) doi:10.1101/001214

Rosetta Brain

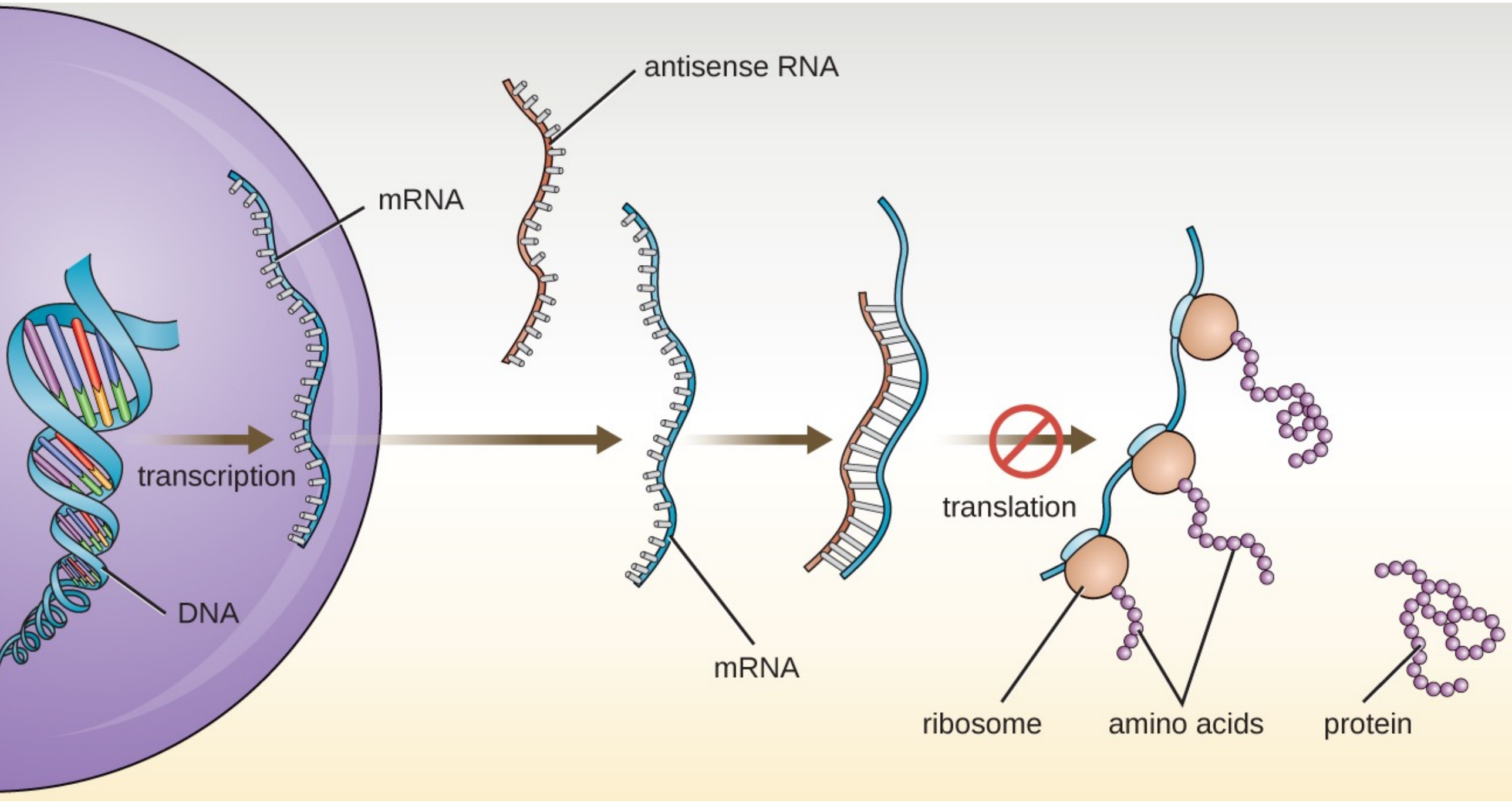


Simultaneous RNA-FISSEQ reveals the cellular identity of each neuron



Marblestone, Daugharthy et al. ***arXiv*** (2014) arXiv:1404.5103
Marblestone, Daugharthy et al. ***arXiv*** (2014) doi:10.1101/001214

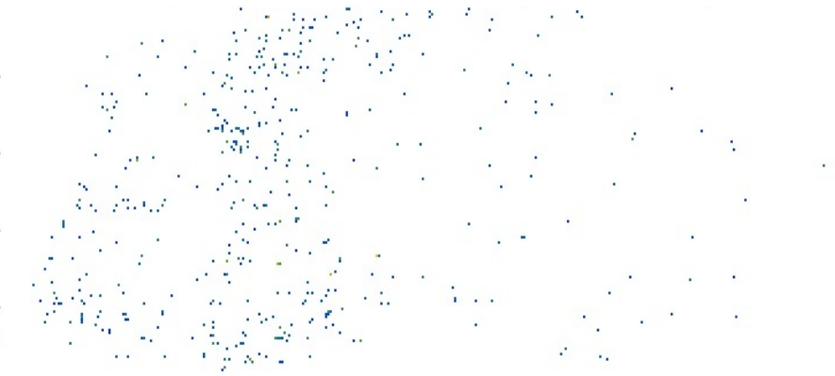
Application: Gene therapy



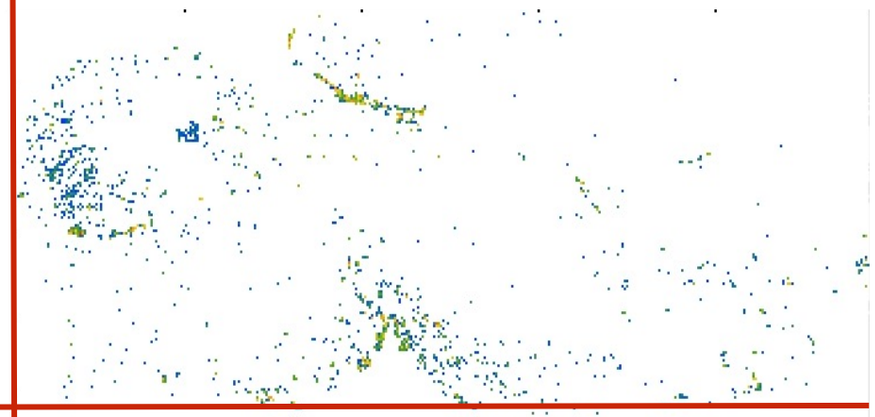
Whole brain therapeutic detection & MALAT1 KD

- Quantification of functional therapeutic A and corresponding knockdown of target MALAT1 expression level, compared with negative control therapeutic B in whole mouse brain

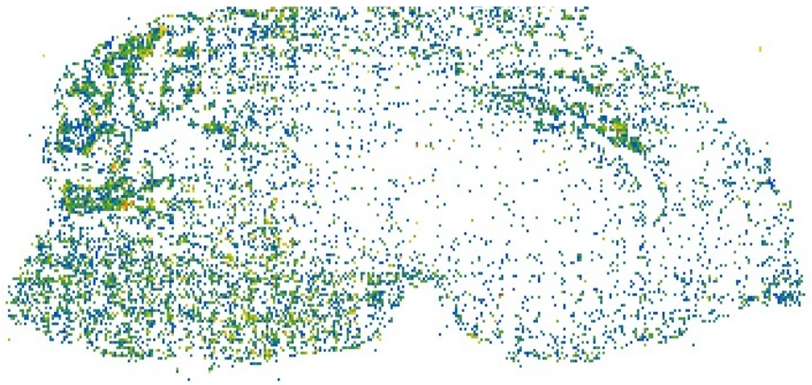
Dosed with control therapeutic B
Detect functional therapeutic A



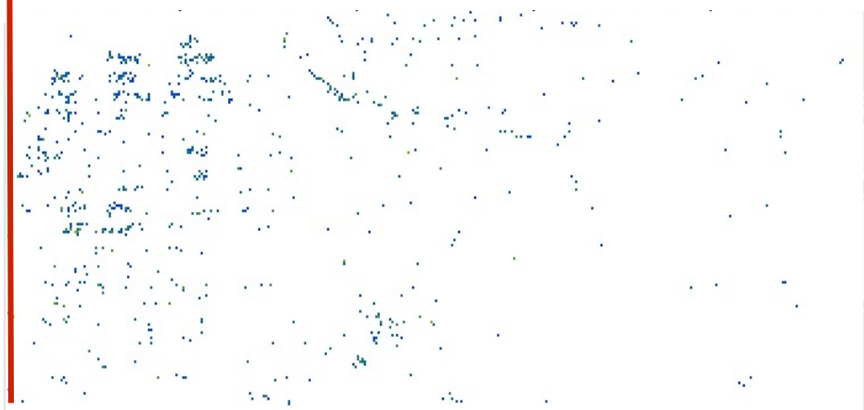
Dosed with functional therapeutic A
Detect therapeutic A

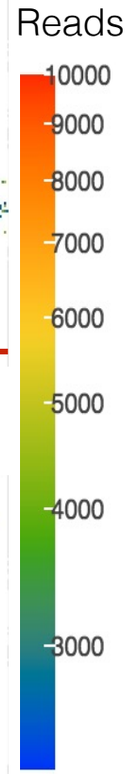


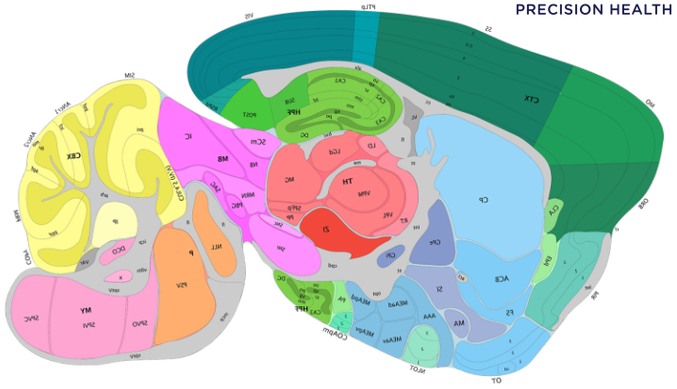
Dosed with control therapeutic B
Detect MALAT1



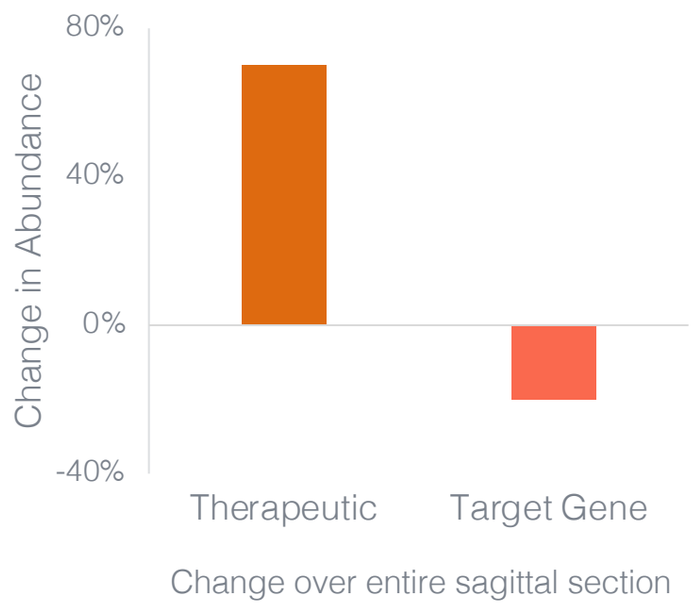
Dosed with functional therapeutic A
Detect MALAT1







ASO 3174 vs Control Treatment

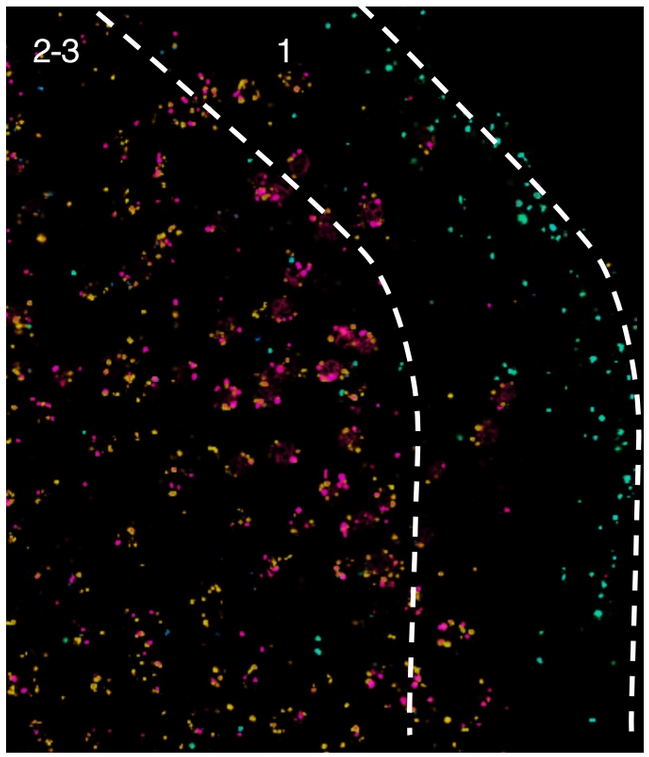


Change over entire sagittal section

Evan Daugharthy
daugharthy@fas.harvard.edu

Spatial uptake of gene therapy

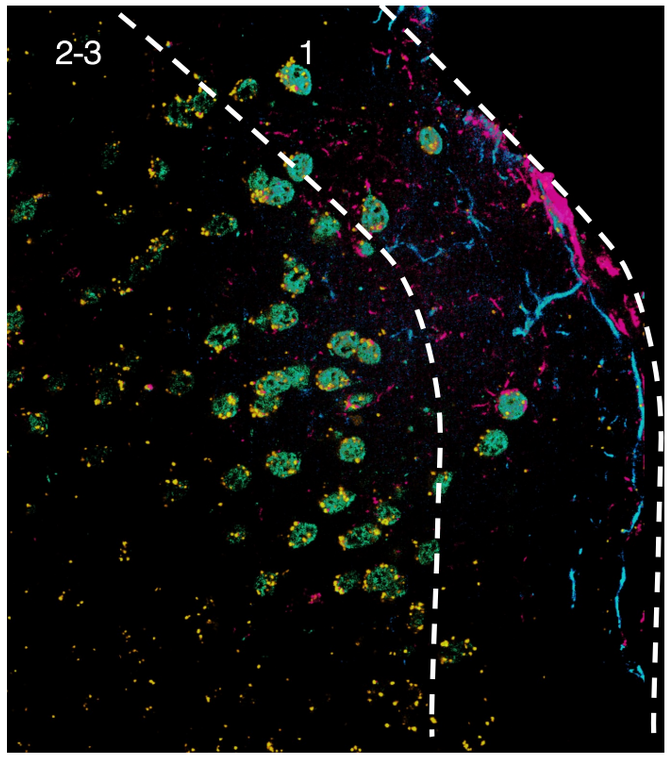
Single-cell Phenotyping by RNA Expression



Astrocyte
APOE, GFAP

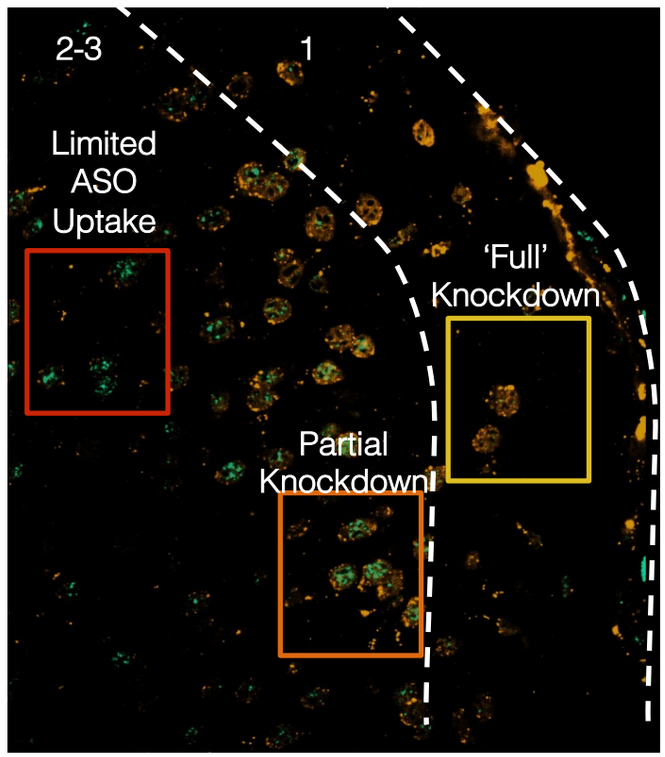
Neuron
SNAP25, SYN1

Cell & Tissue Morphology by Antibody



Astrocyte
Neuron
Microglia
Registration Marker

Therapeutic uptake and knockdown response are spatially localized



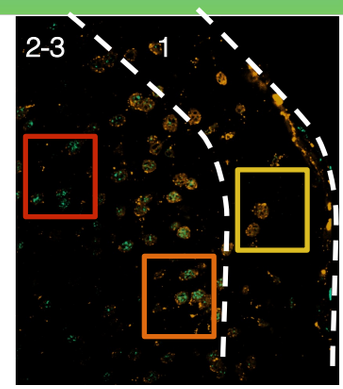
Therapeutic
Target Gene

Data from mouse visual cortex

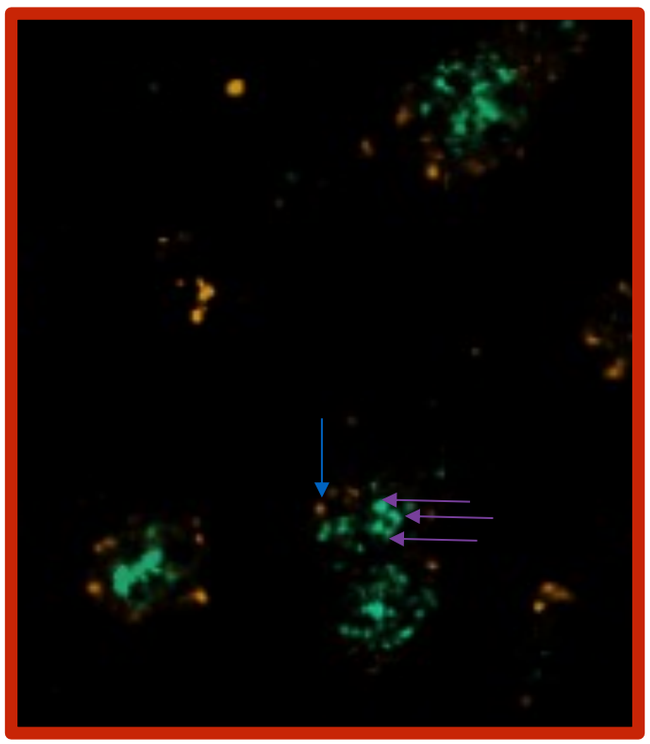
AGBT™

Spatial uptake of gene therapy

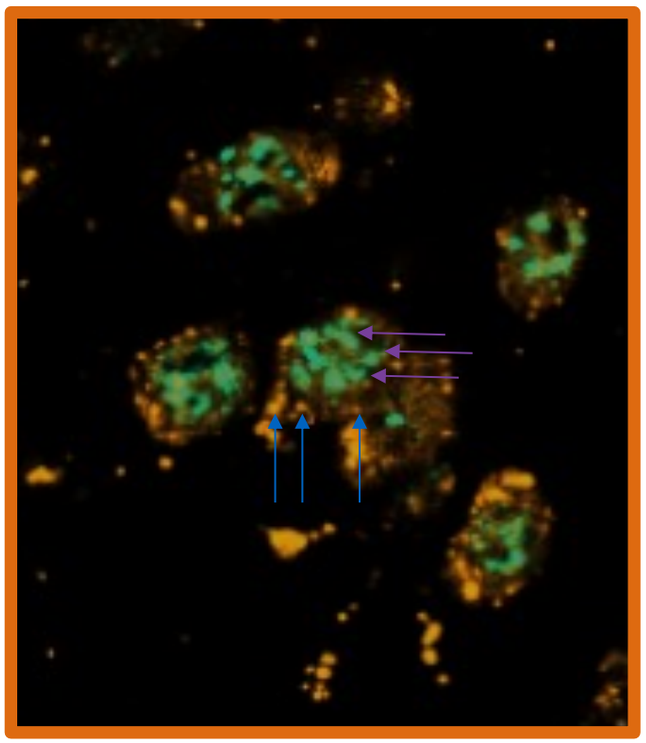
- Uptake and knockdown are cortical layer and cell type dependent



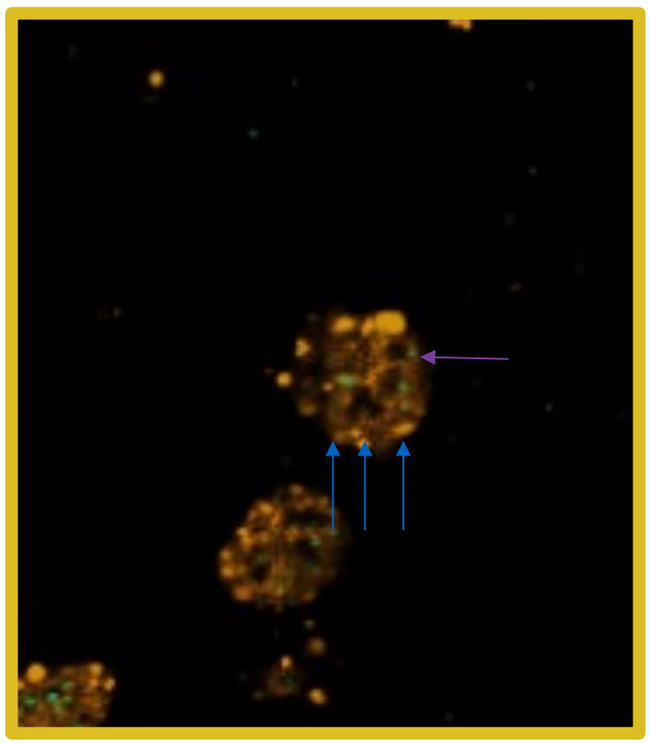
Limited ASO uptake
Limited MALAT1 knockdown



ASO Fills Cell
Partial MALAT1 Knockdown

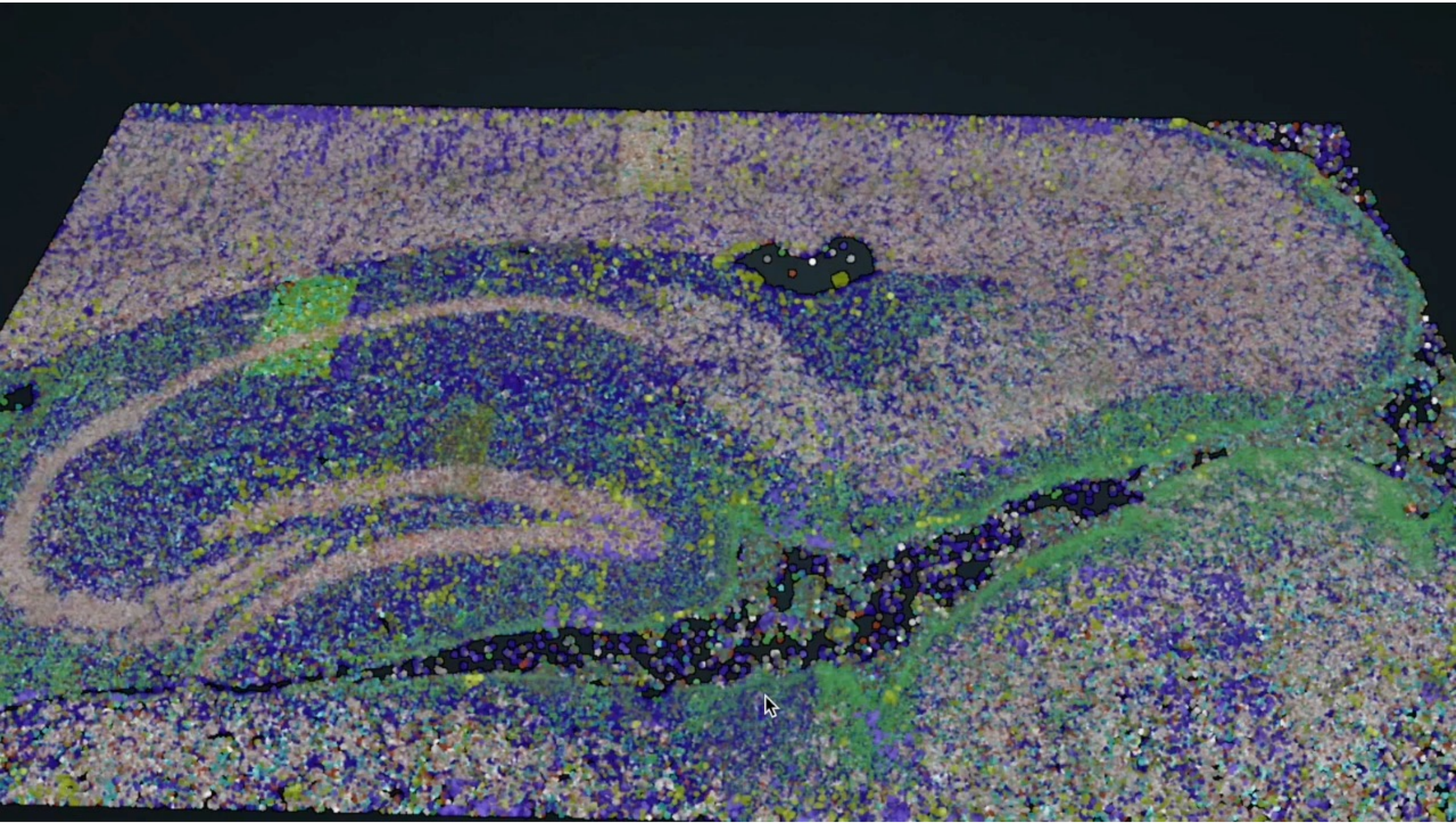


ASO Fills Cell
Nearly Full MALAT1 Knockdown



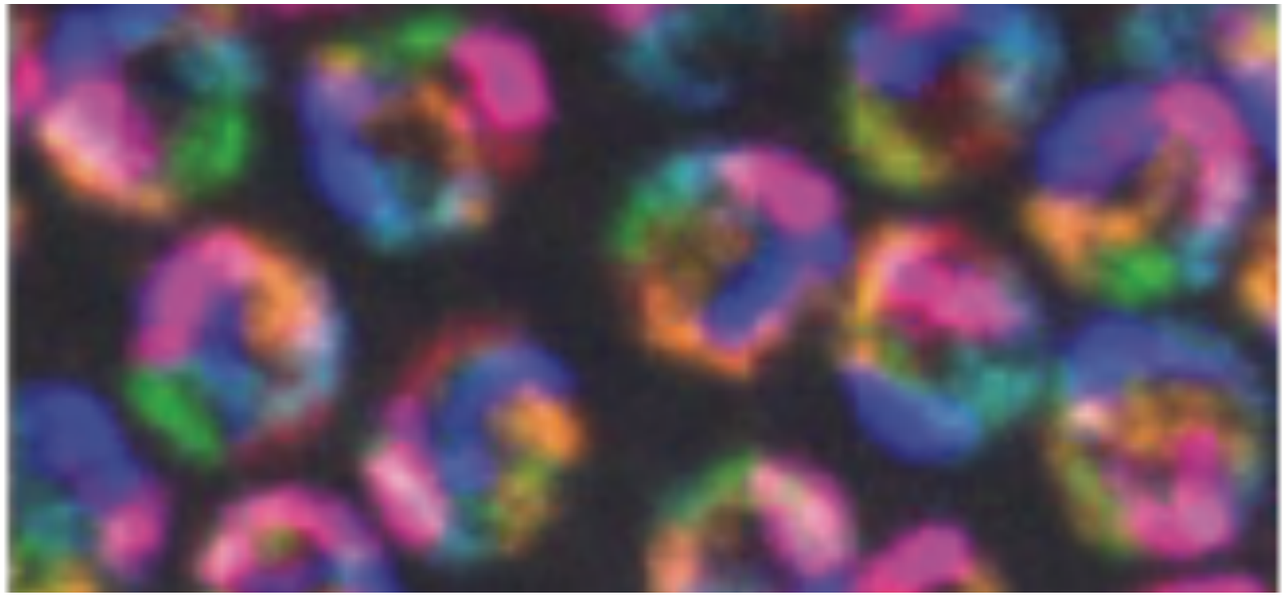
Data from mouse visual cortex

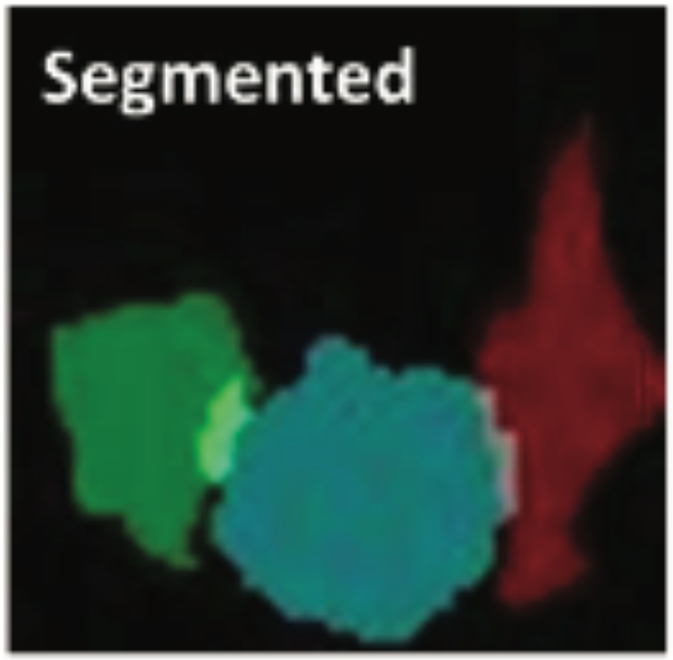
AGBT™

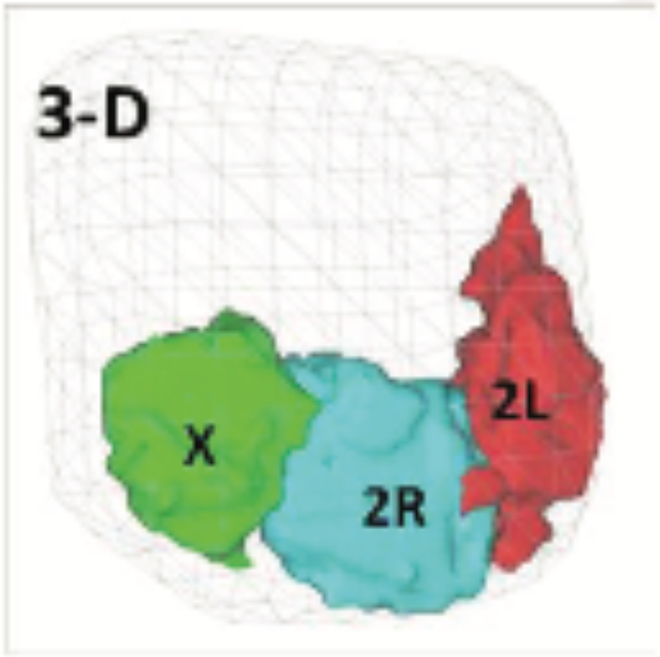


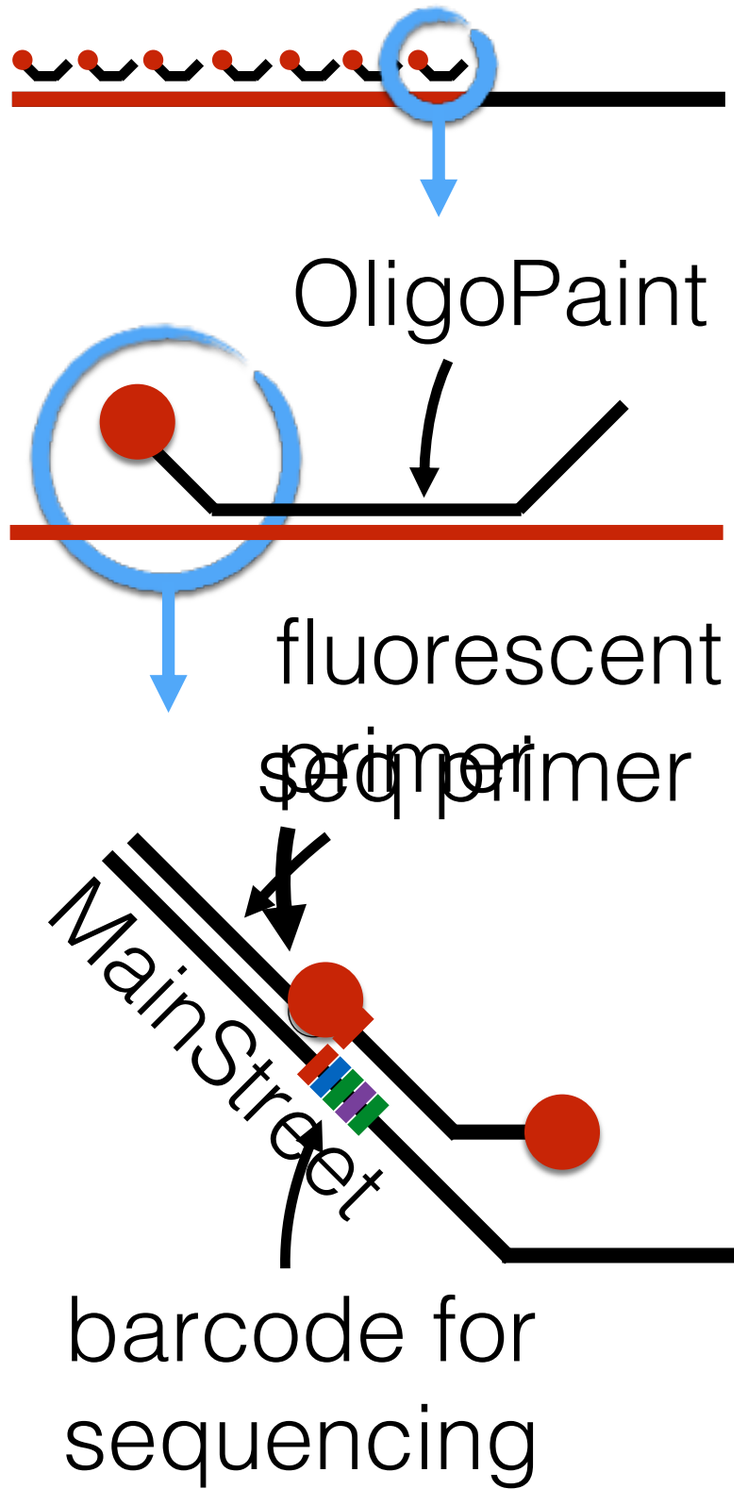
Scaling OligoPaints to whole genome

C. elegans whole genome OligoPaints



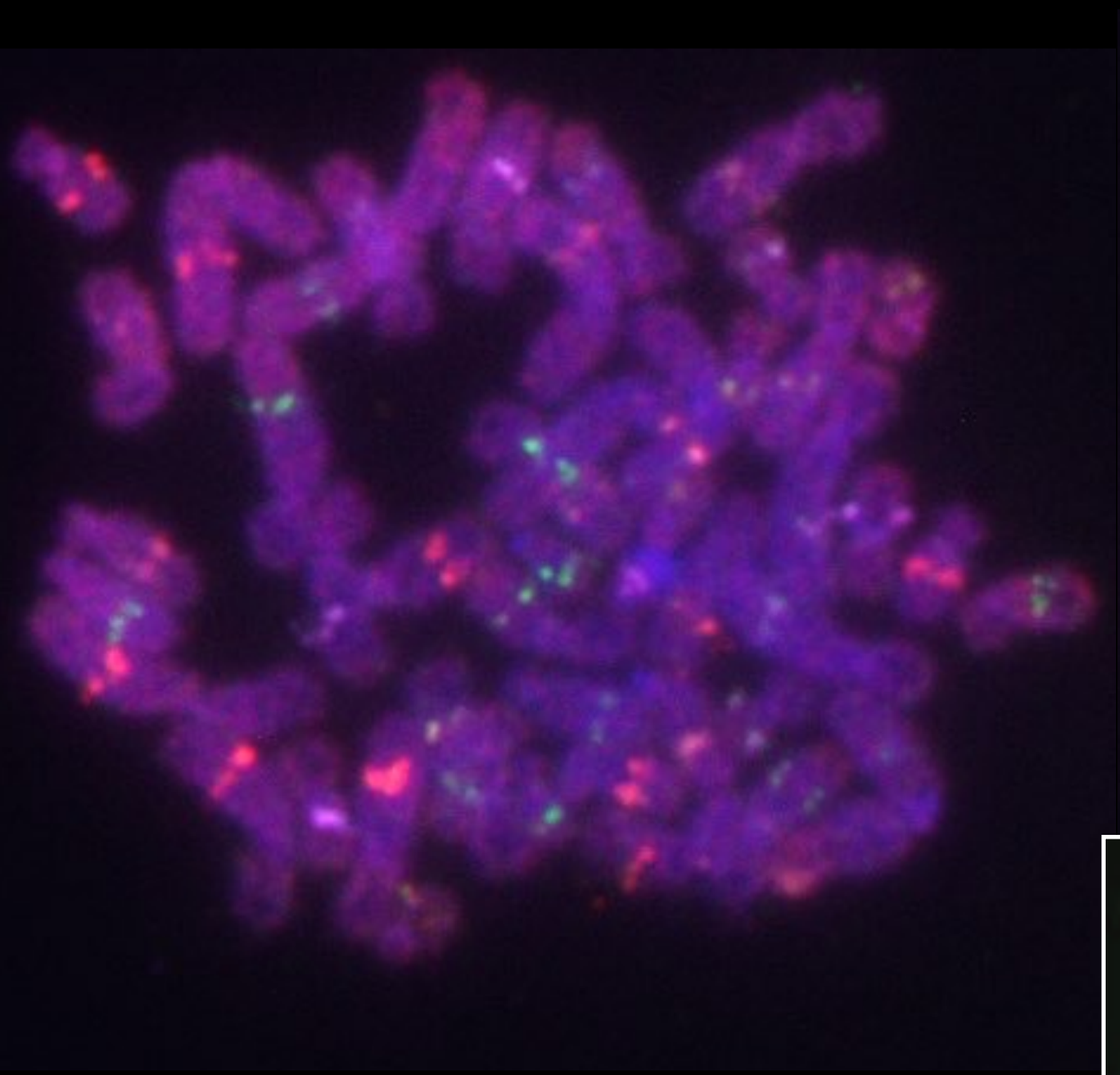


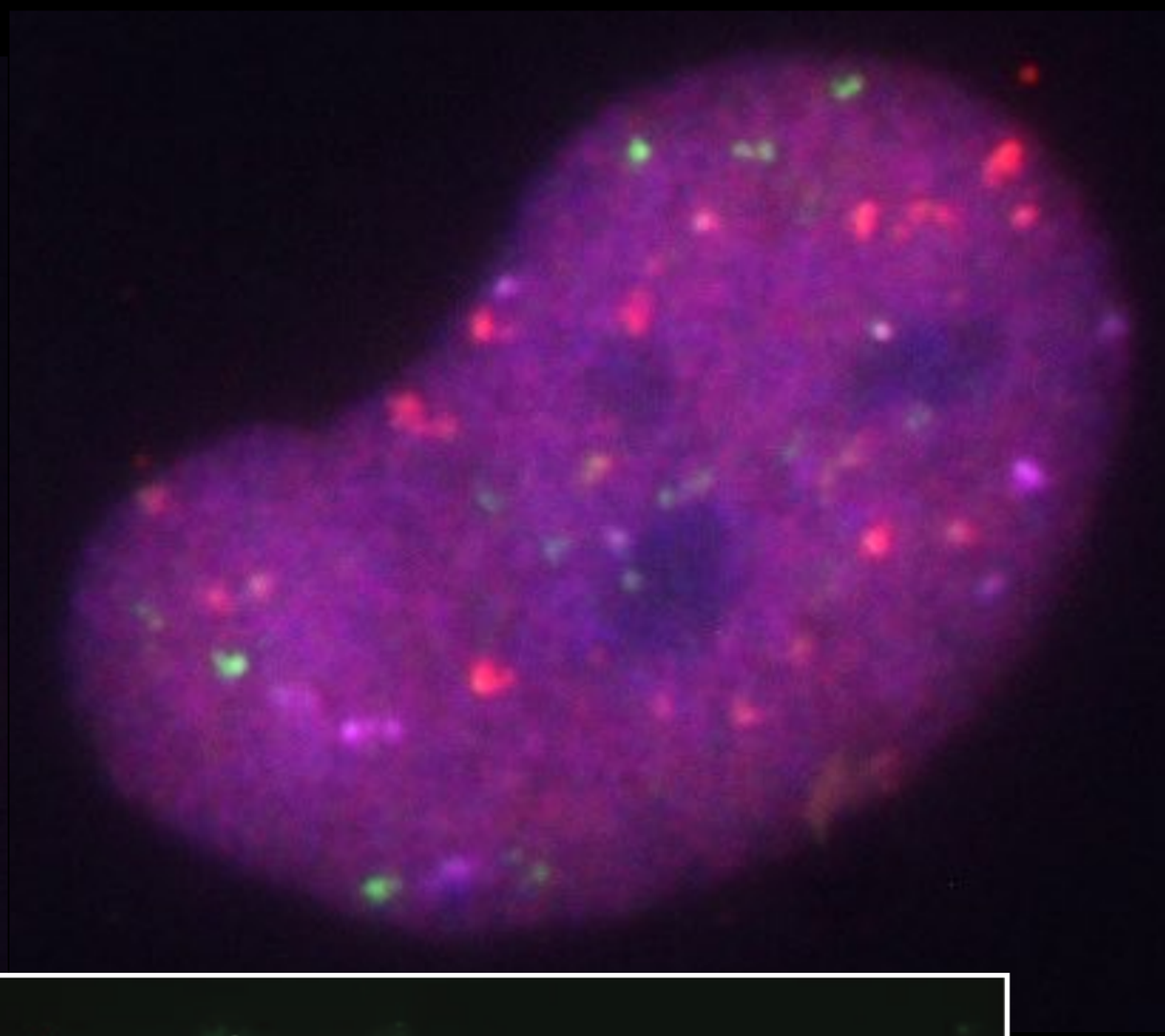


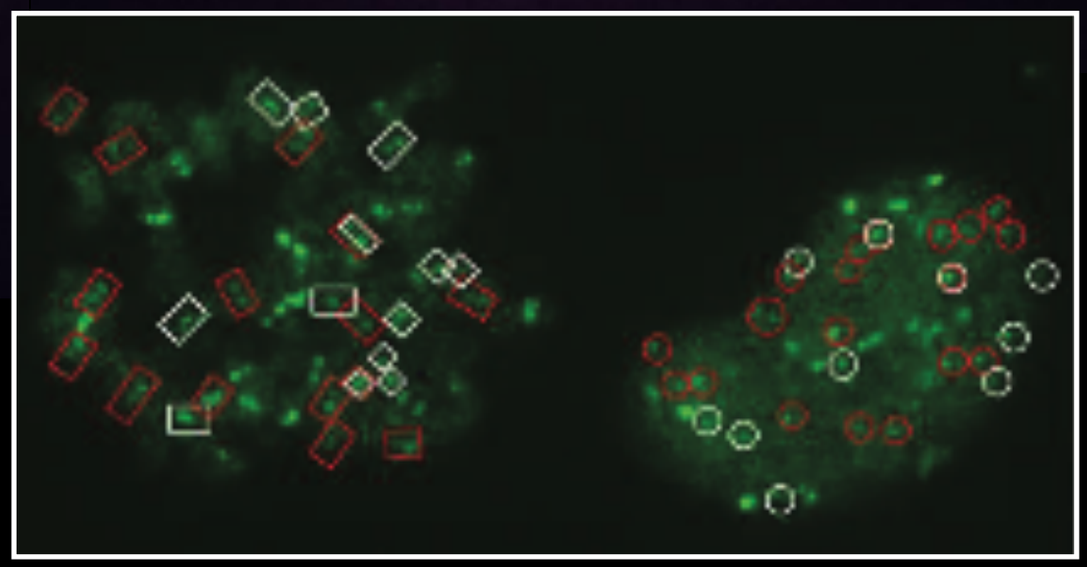


Courtesy of Brandon Fields, Scott Kennedy, Son Nguyen & Ting Wu

Preliminary data from genome OligoFISSEQ

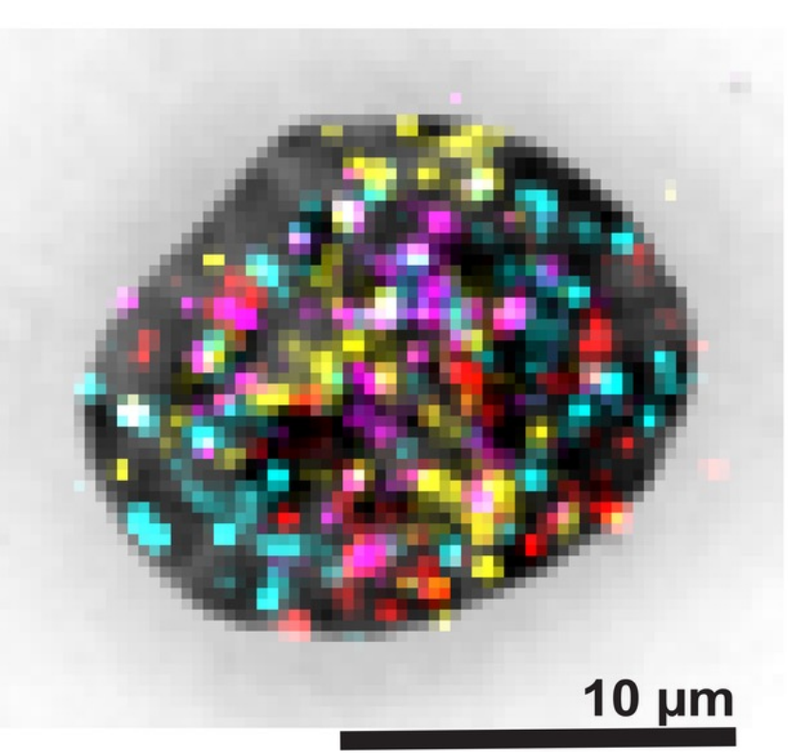


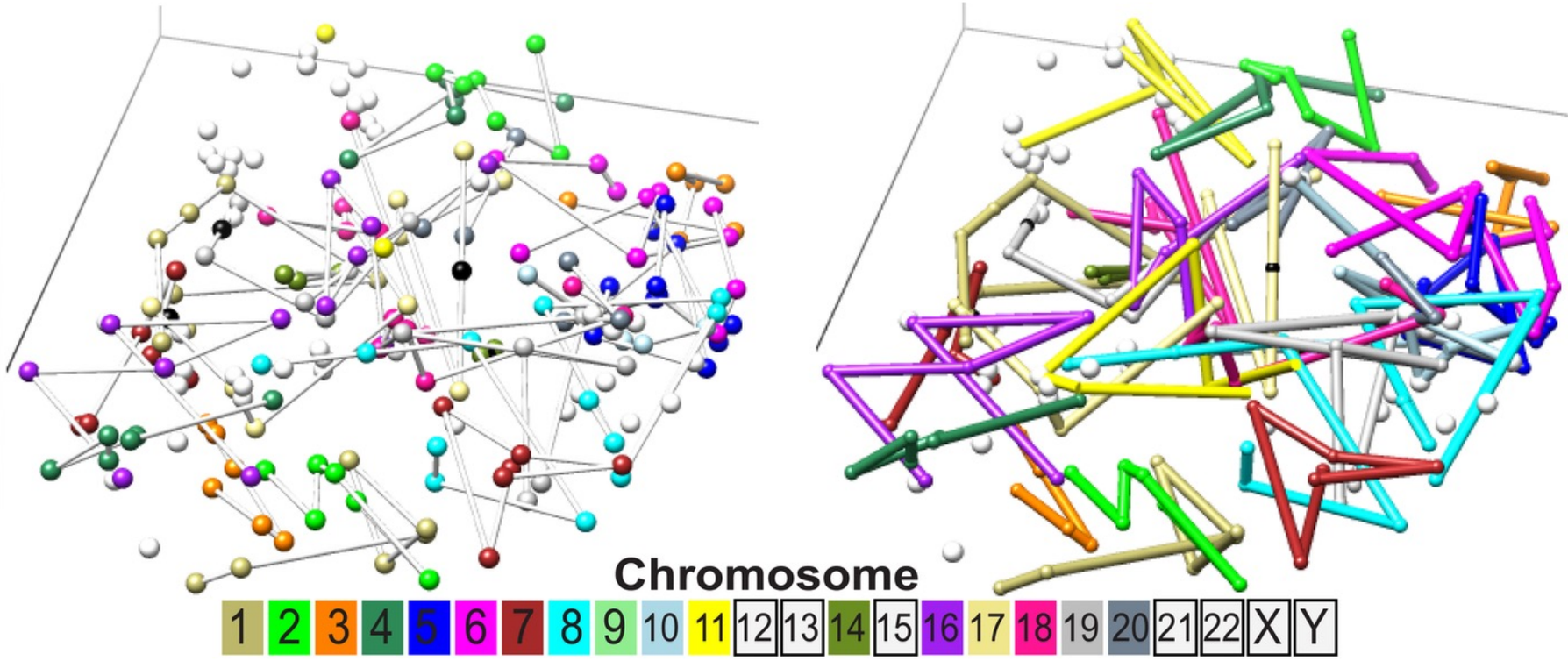




1 sequence/chromosome

Publication



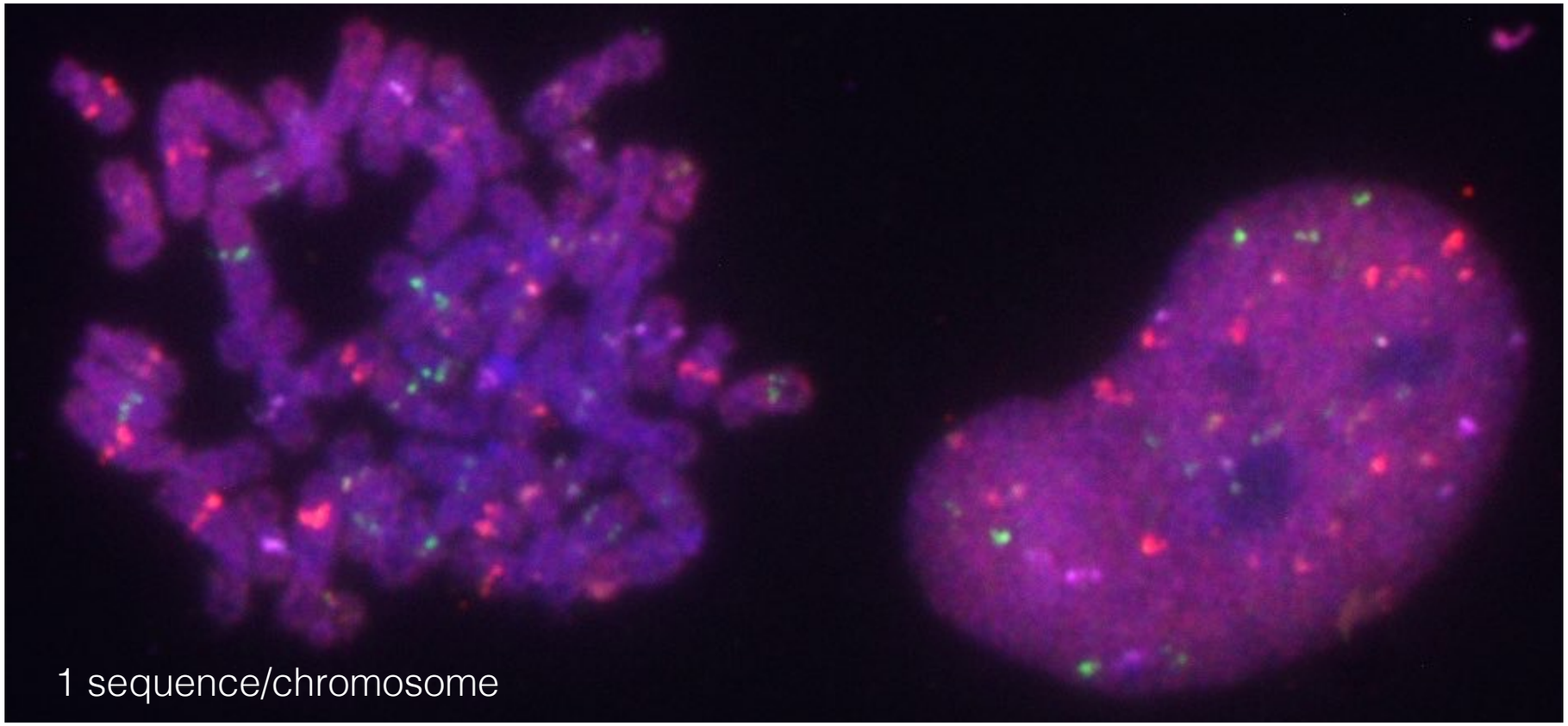


Nguyen, Huy Q., et al. "3D mapping and accelerated super-resolution imaging of the human genome using in situ sequencing." *Nature Methods* 17.8 (2020): 822-832.

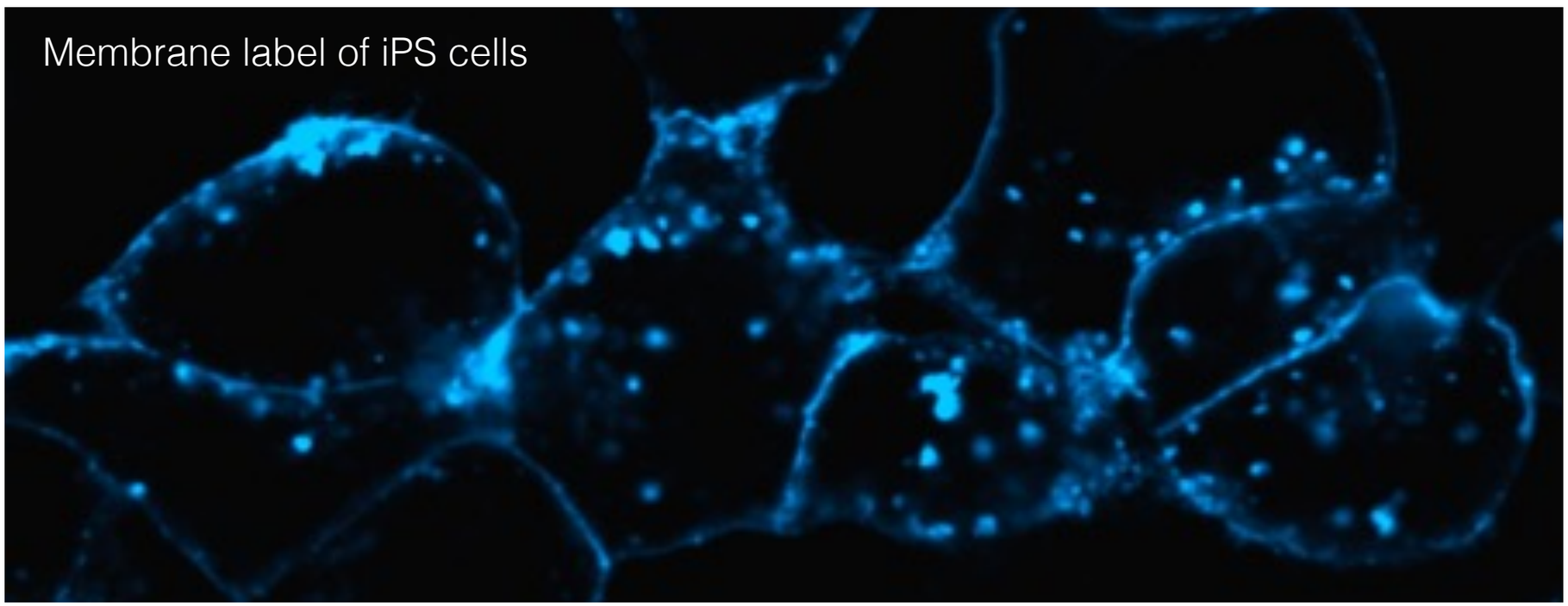
Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Preliminary data from genome & proteome

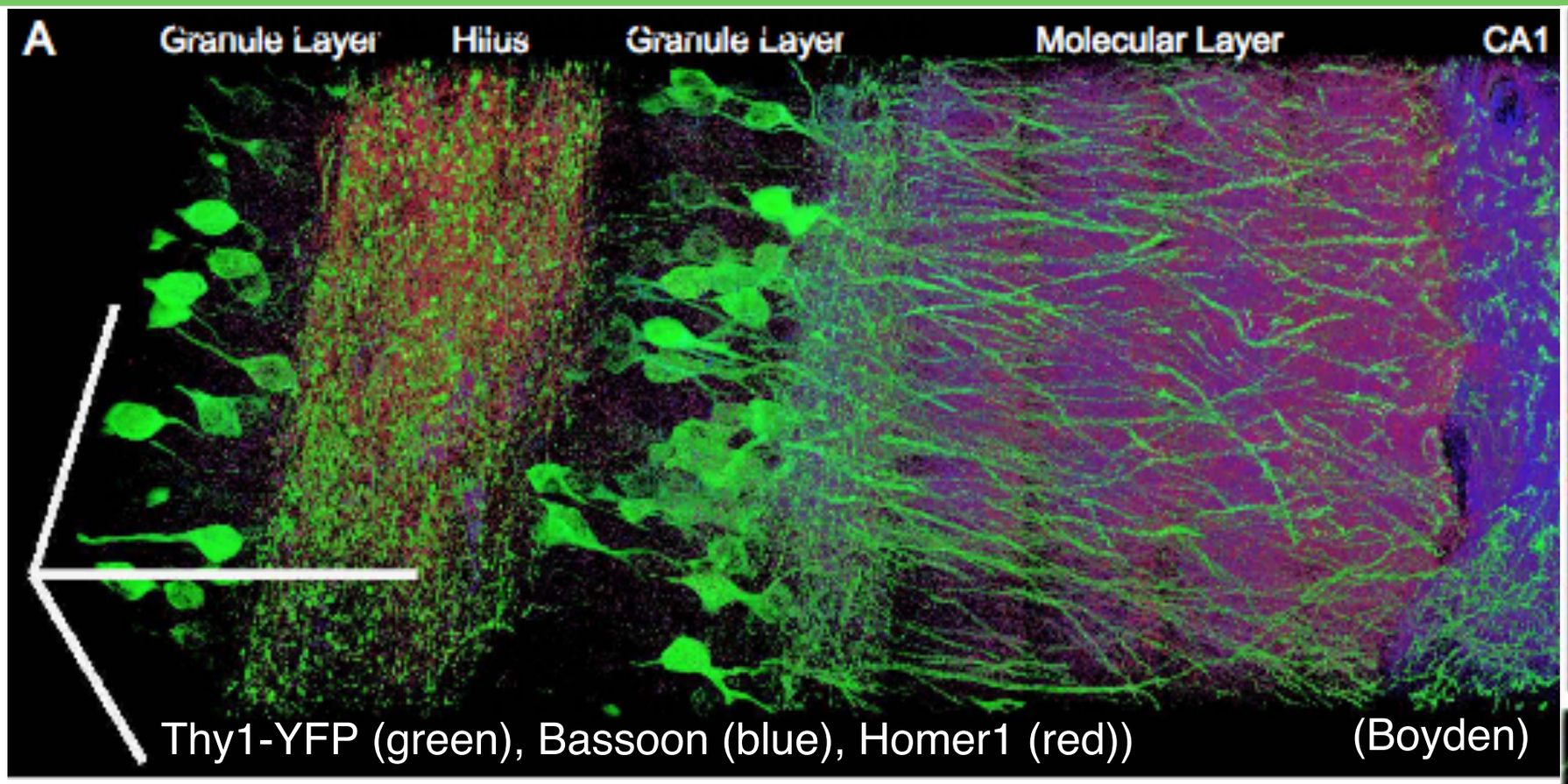


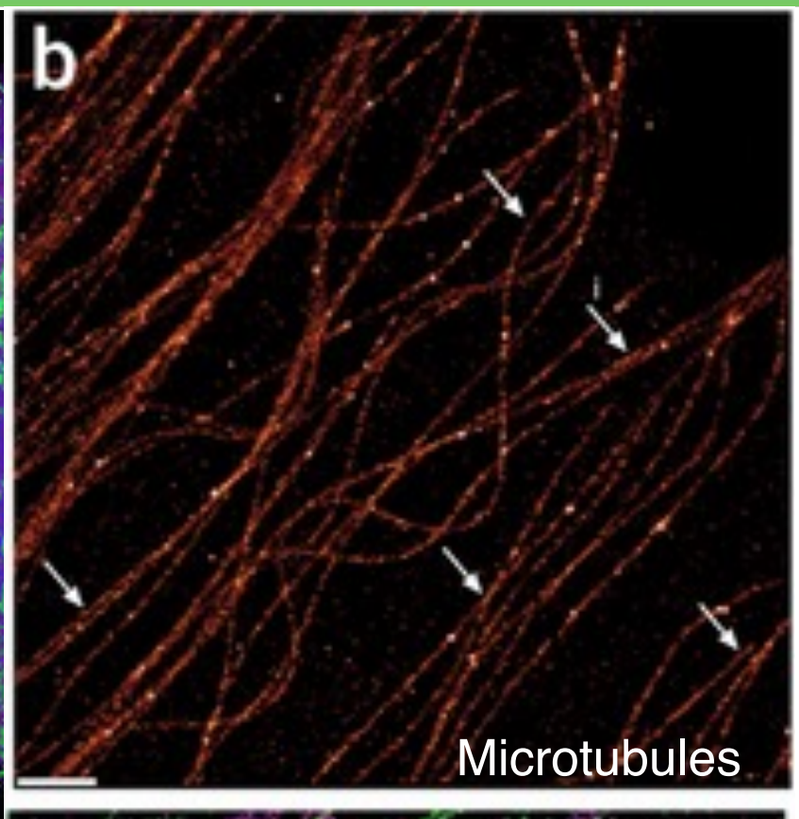
Direct genomic sequencing (unpublished)

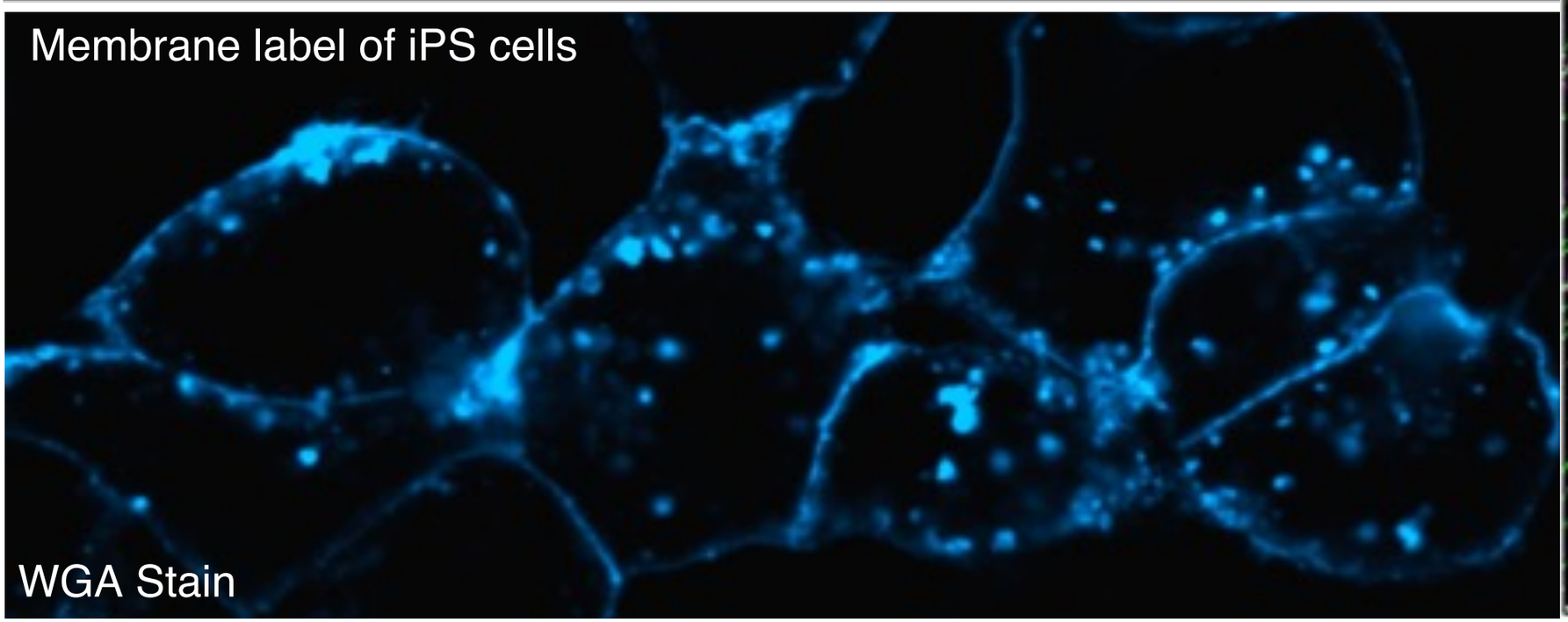


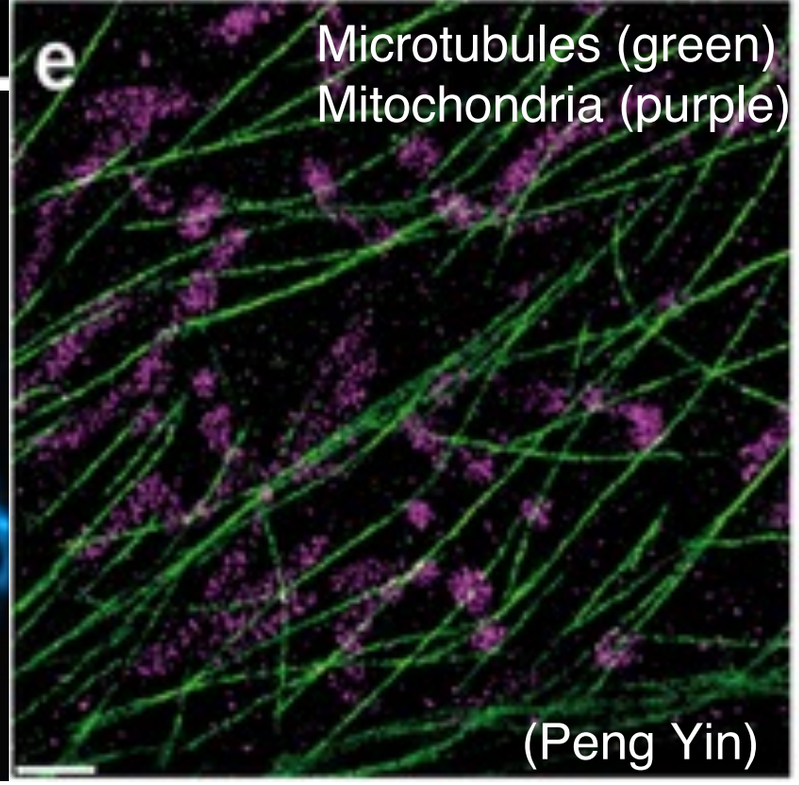
DNA-barcoded antibody labeling

Protein FISSEQ









Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Toward perfect detection

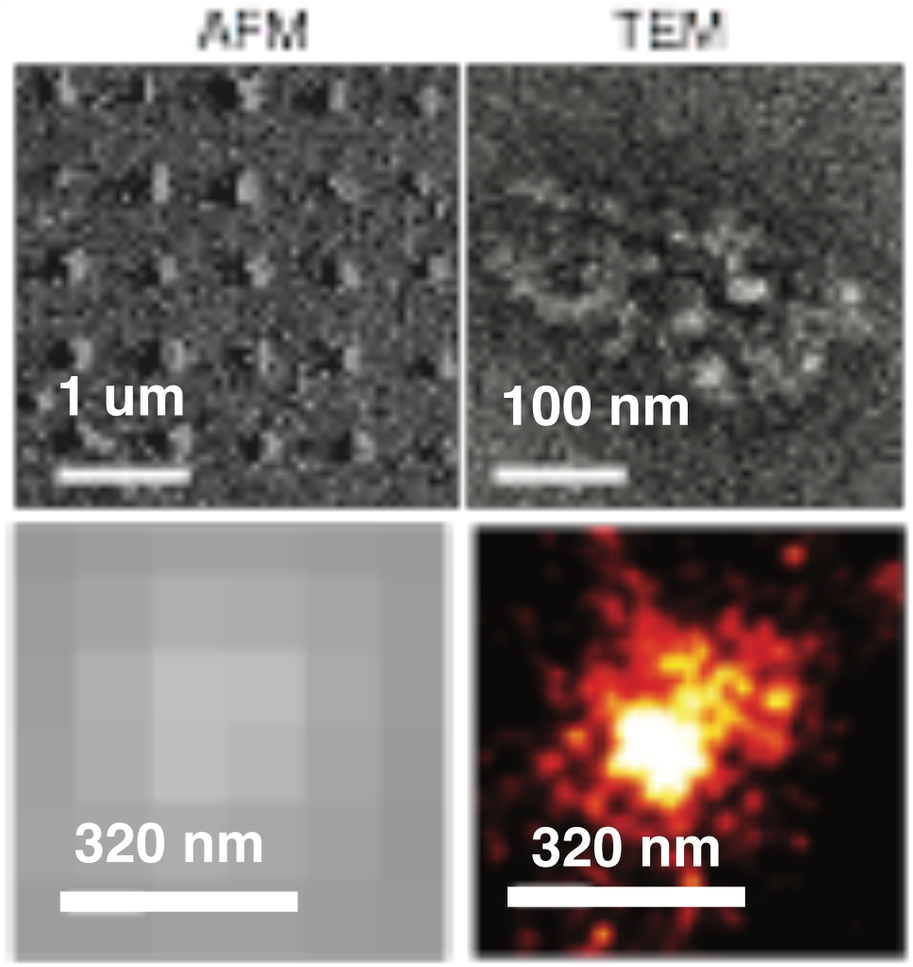
We can detect up to [cell volume] / 0.04 μm^3 RNA's per cell

Cell type	Cell volume	Upper Limit
erythrocyte	100 μm^3	2,500
neutrophil	300 μm^3	7,500
beta cell	1,000 μm^3	25,000
enterocyte	1,400 μm^3	35,000
fibroblast	2,000 μm^3	50,000
HeLa	3,000 μm^3	75,000
hair cell	4,000 μm^3	100,000
osteoblast	4,000 μm^3	100,000
macrophage	5,000 μm^3	125,000
cardiomyocyte	15,000 μm^3	375,000
megakaryocyte	30,000 μm^3	750,000
fat cell	600,000 μm^3	15,000,000
oocyte	4,000,000 μm^3	100,000,000

mRNA: 10k-[50k-300k]-1m / cell

Genome: 11m bp / 0.04 μm^3

Protein: 80,000 / 0.04 μm^3



Evan Daugharthy
evan.daugharthy@gmail.com

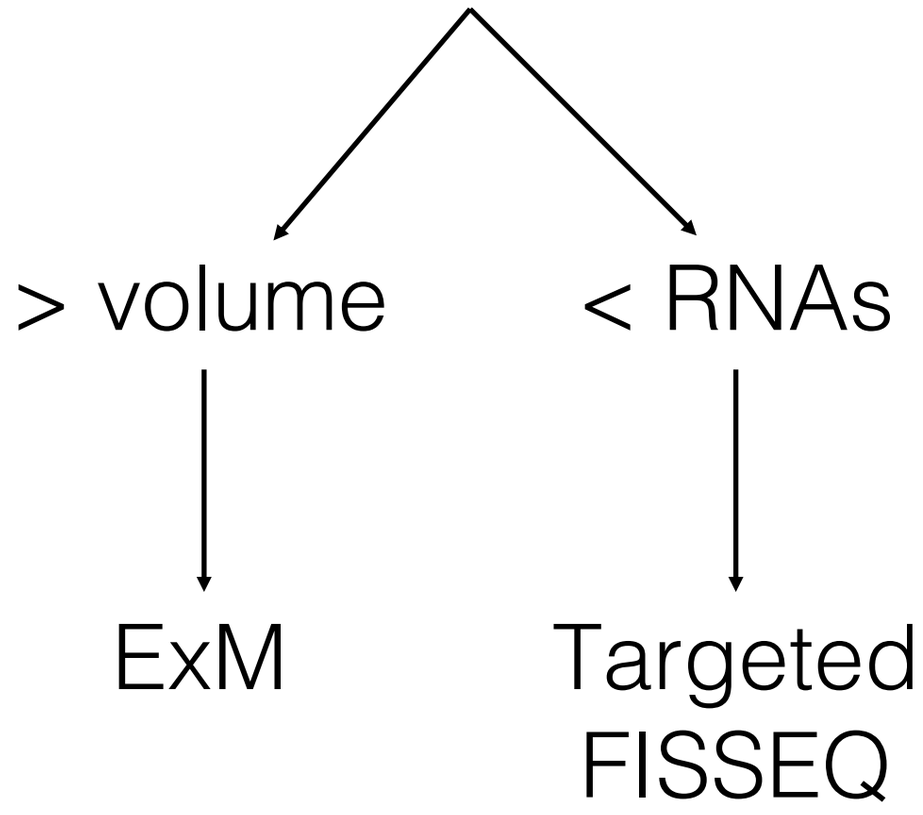
READCOOR WYSS INSTITUTE

Increased sensitivity

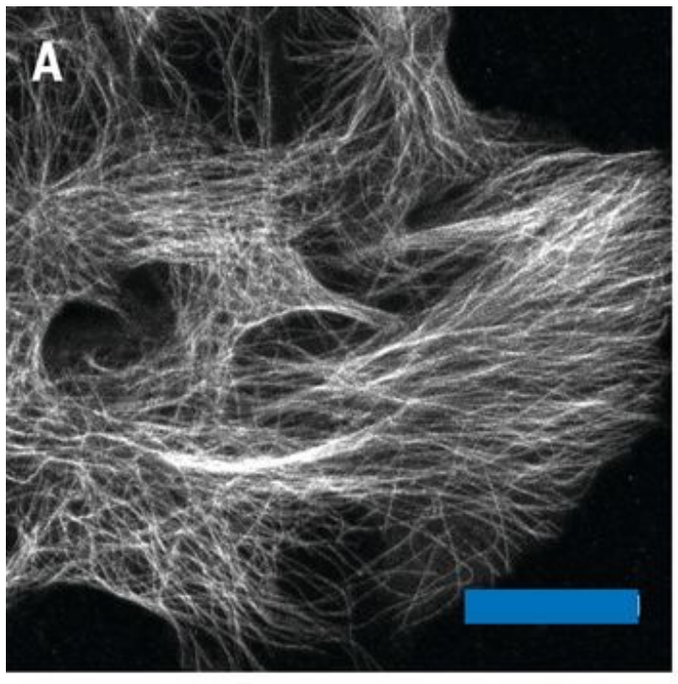
We can detect up to [cell volume] / 0.04 μm^3 RNA's per cell

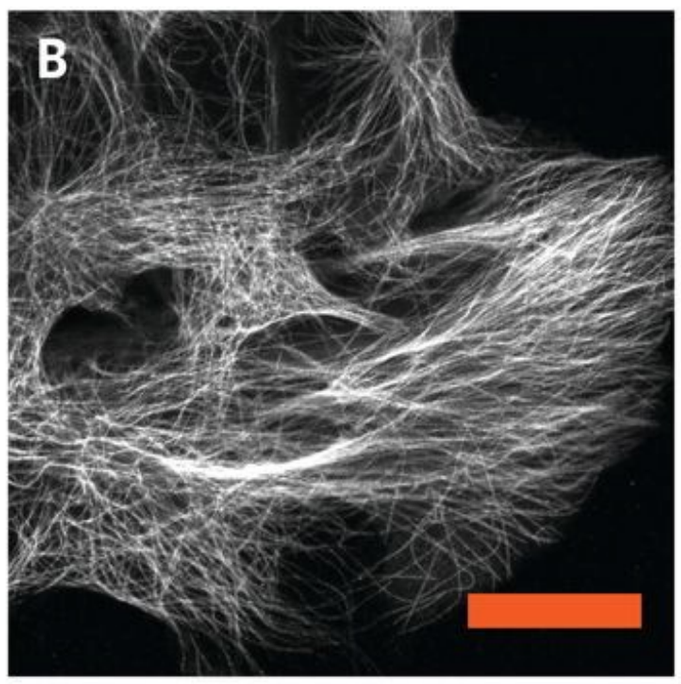
Cell type	Cell volume	Upper Limit
erythrocyte	100 μm^3	2,500
neutrophil	300 μm^3	7,500
beta cell	1,000 μm^3	25,000
enterocyte	1,400 μm^3	35,000
fibroblast	2,000 μm^3	50,000
HeLa	3,000 μm^3	75,000
hair cell	4,000 μm^3	100,000
osteoblast	4,000 μm^3	100,000
macrophage	5,000 μm^3	125,000
cardiomyocyte	15,000 μm^3	375,000
megakaryocyte	30,000 μm^3	750,000
fat cell	600,000 μm^3	15,000,000
oocyte	4,000,000 μm^3	100,000,000

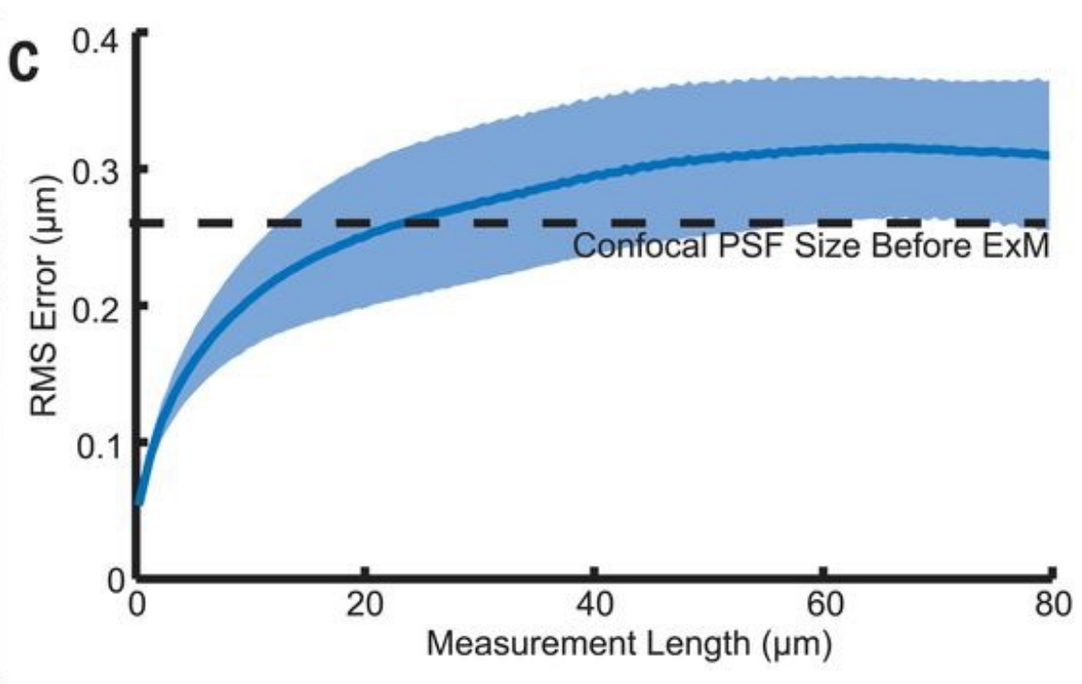
How to get more sensitivity per RNA molecule?

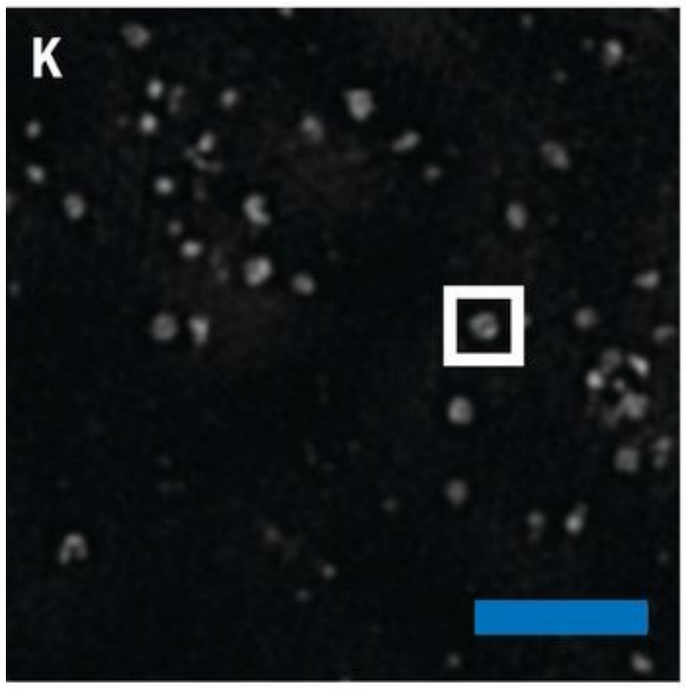


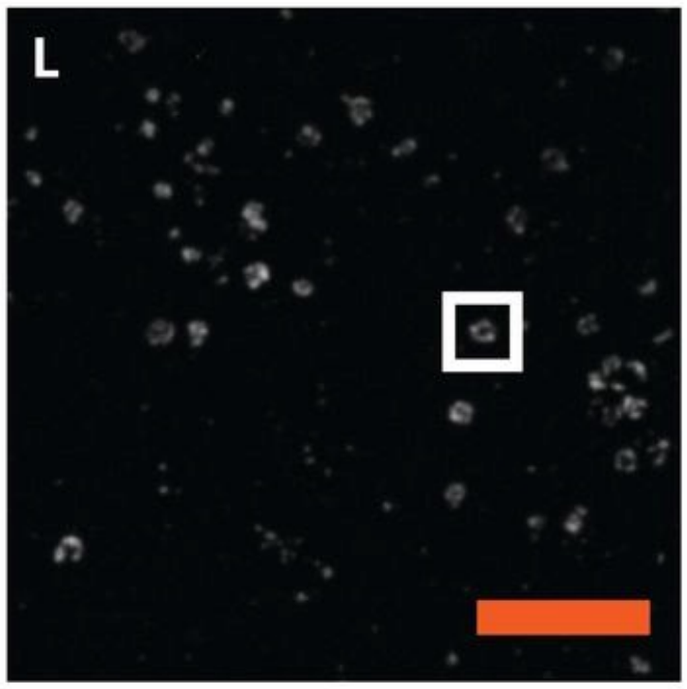
FISSEQ with expansion microscopy

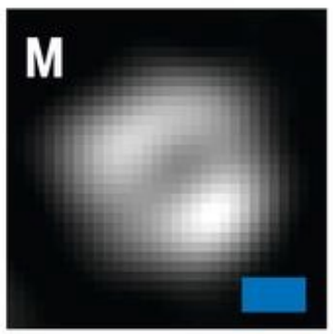


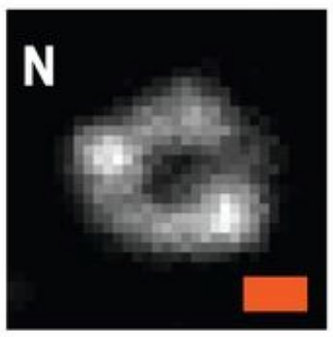


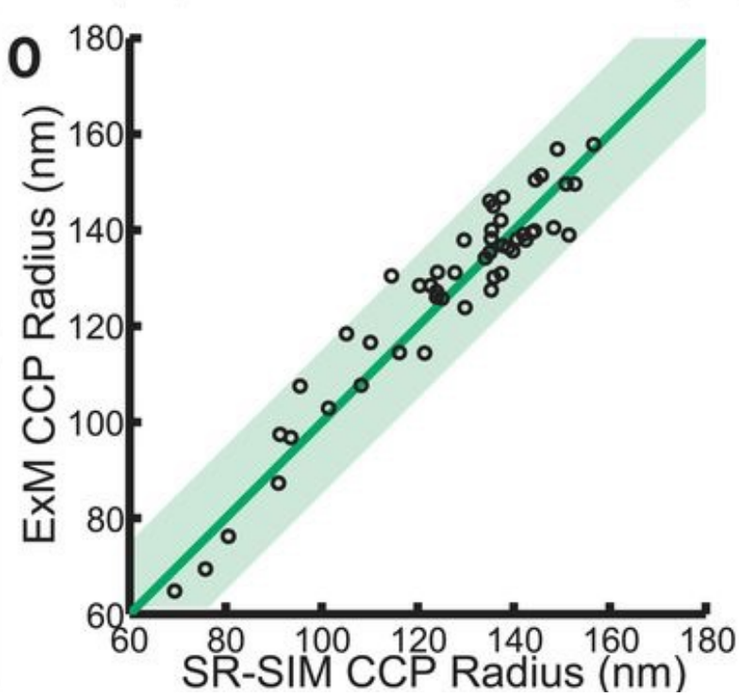






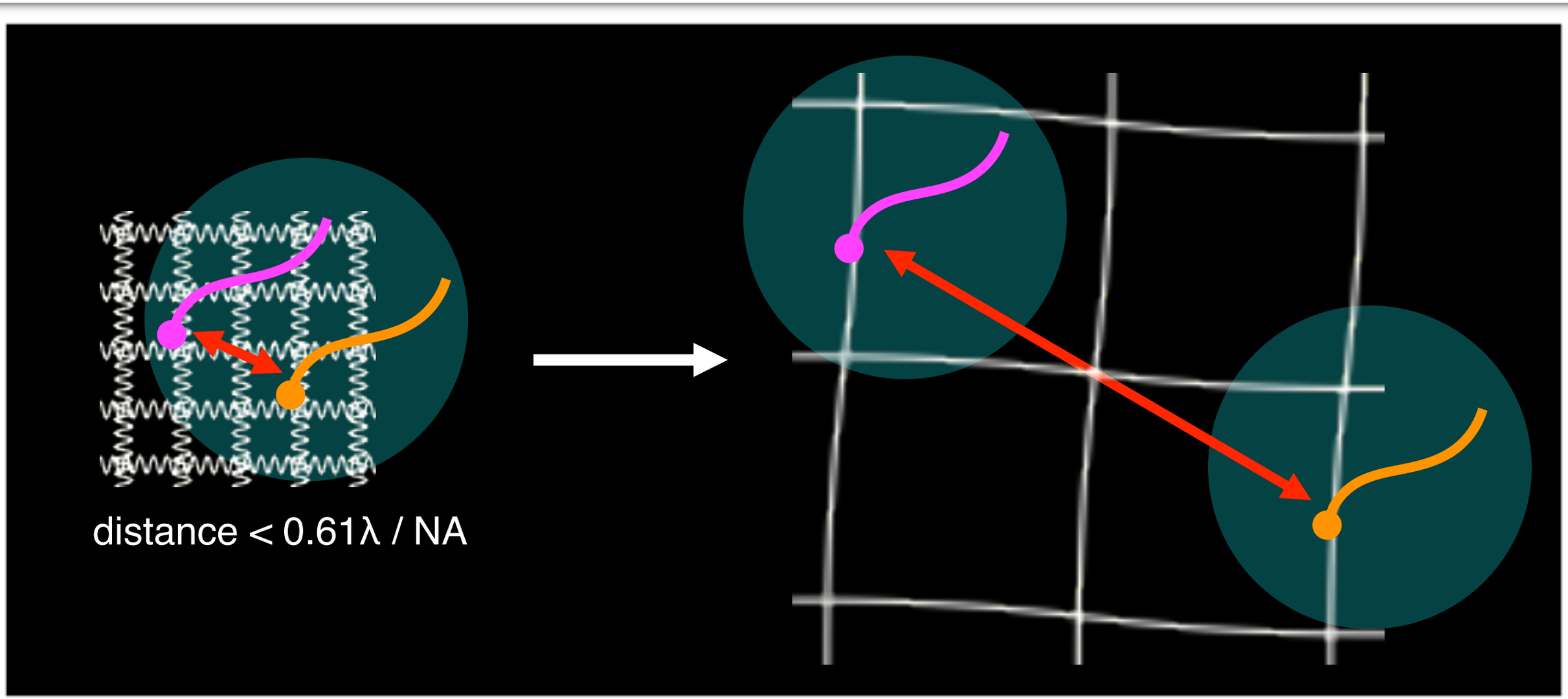






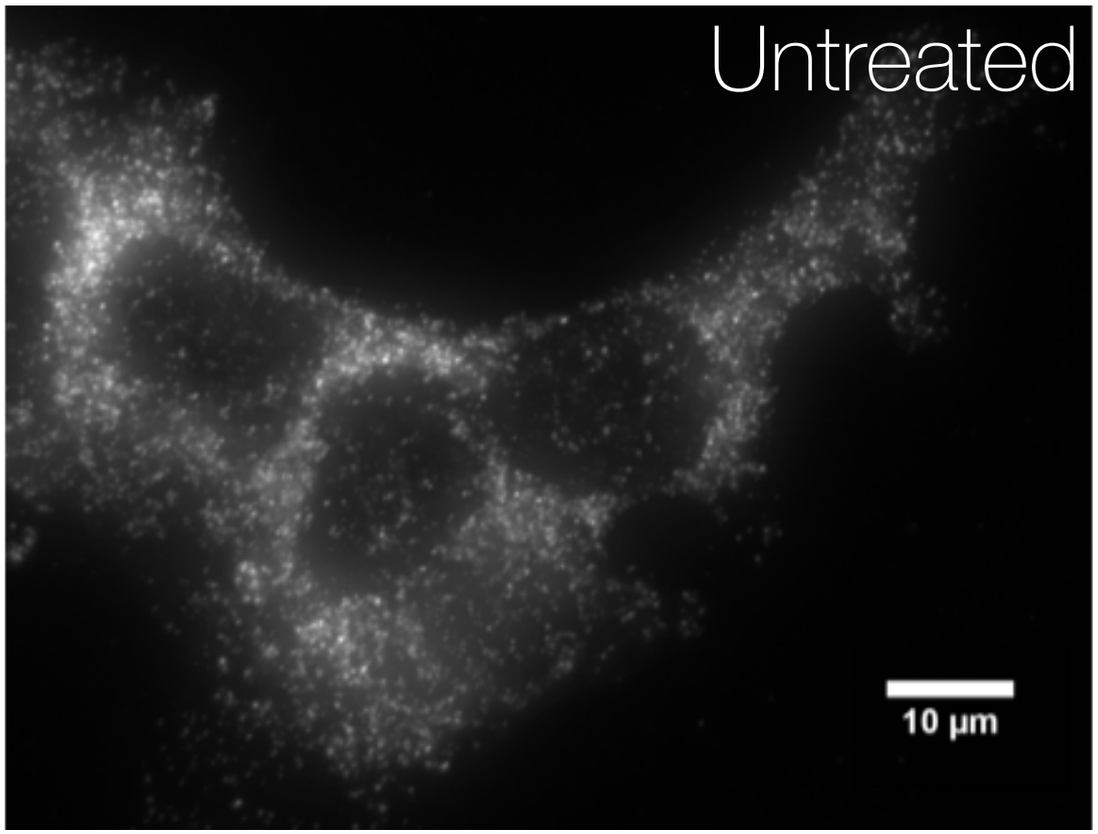
Preliminary data on FISSEQ with expansion

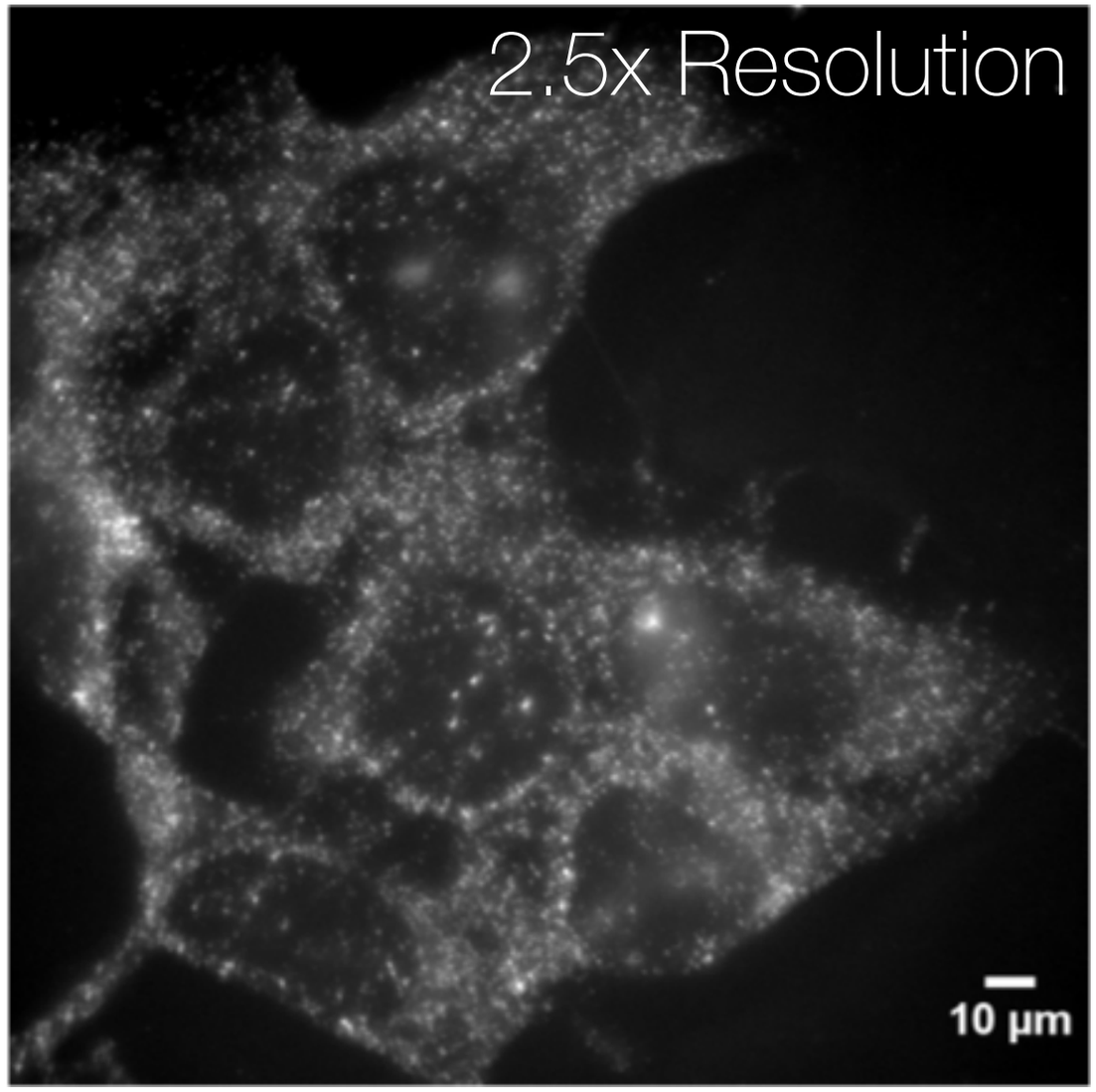
tether RNA & expand isotropically



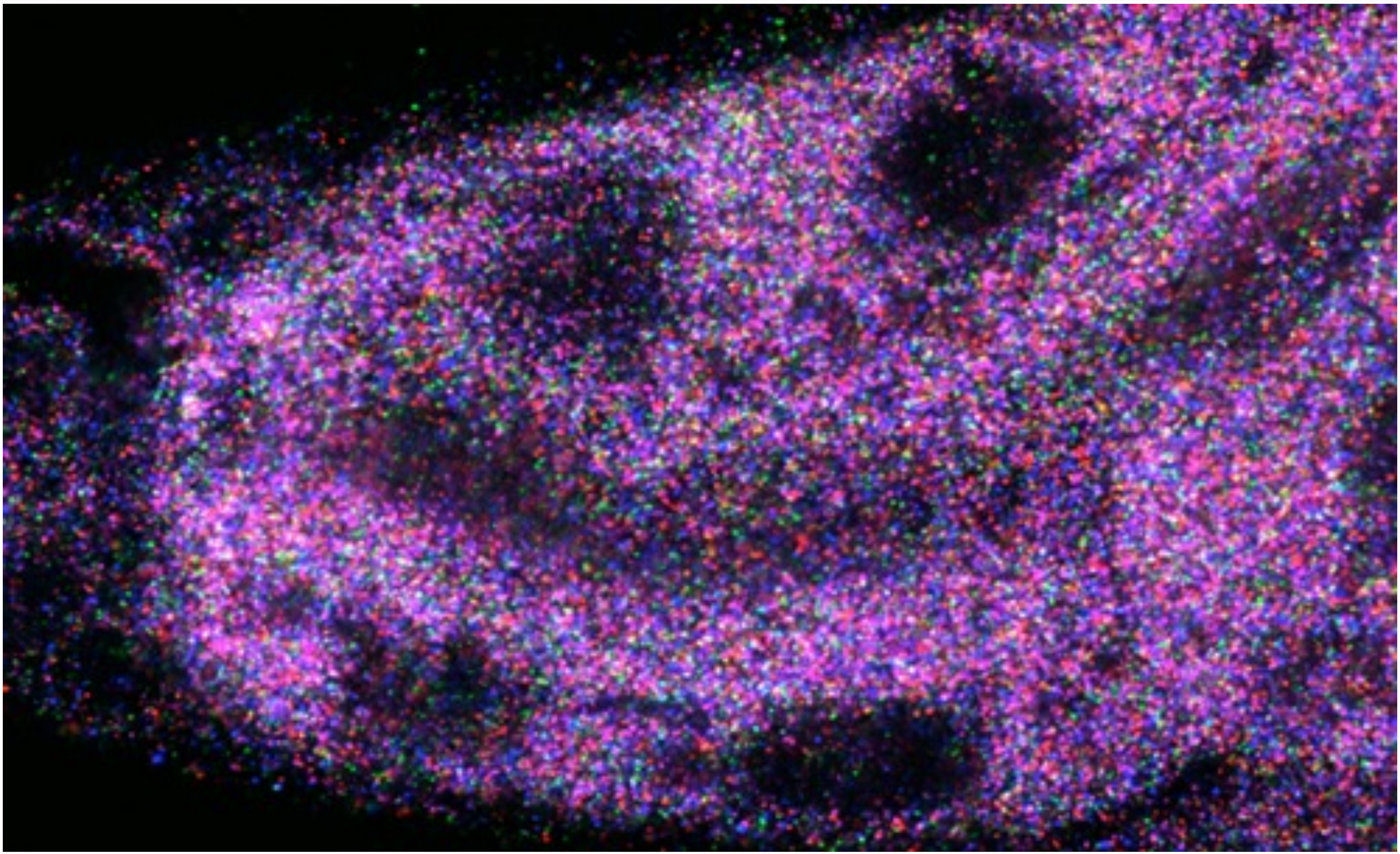
Preliminary data on FISSEQ with expansion

RNA-FISH of GAPDH demonstrates efficient RNA expa

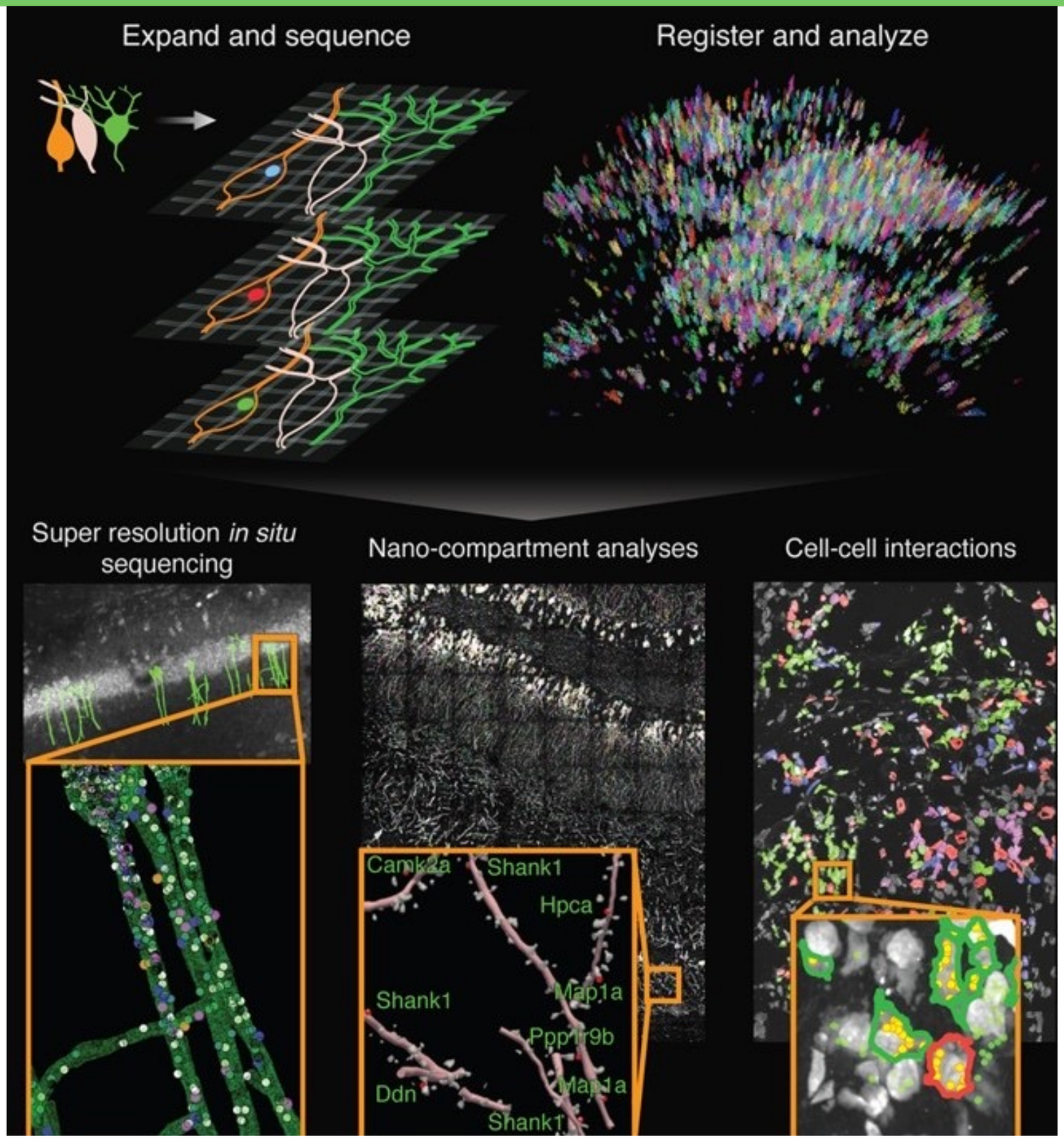


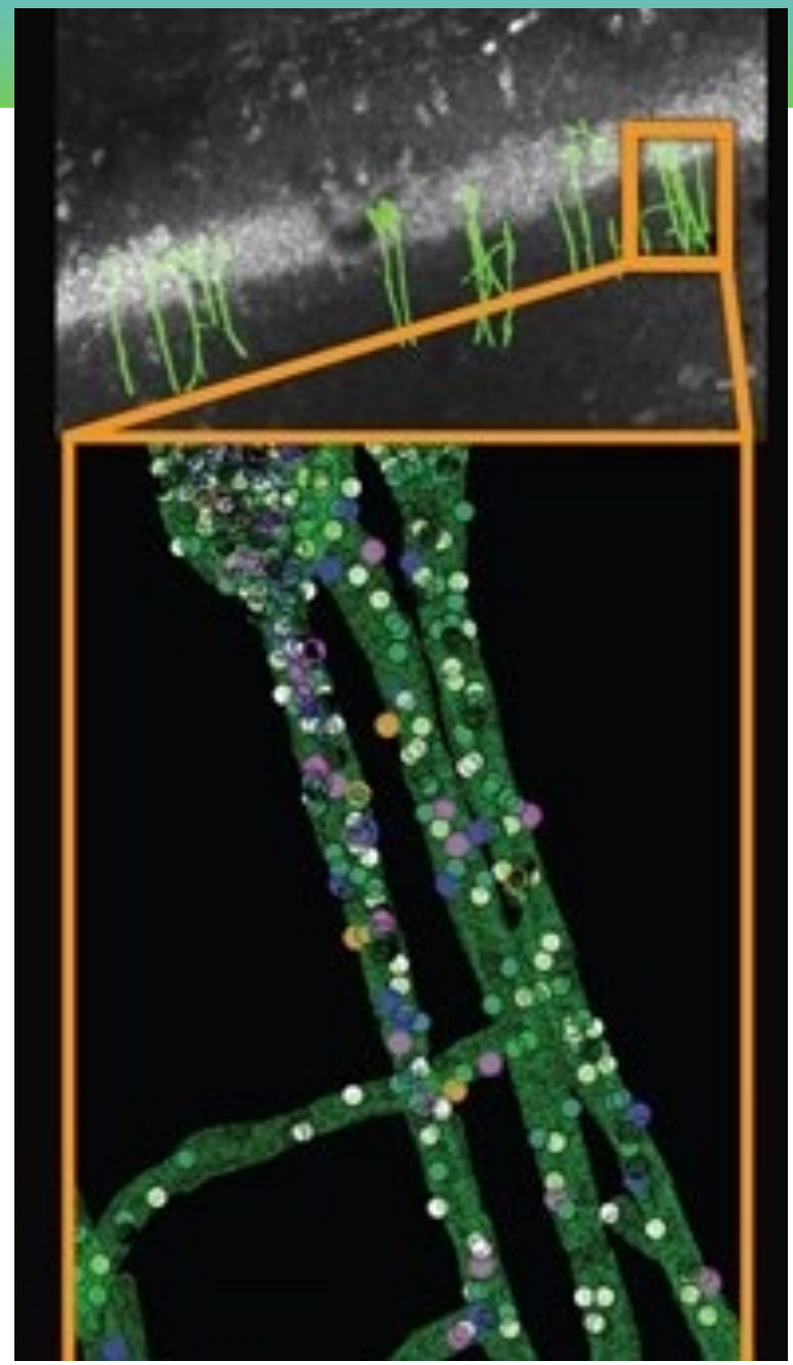


C. elegans



ExSeq





Alon, Shahar, et al. "Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems." *Science*371.6528 (2021).

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Perfect resolution

Genome	“B” form DNA helix 2 nm diameter and 3.4 nm long per 10 base pairs
Transcriptome	1.5-3 nm per 10 nt
Proteome	~nm scale

Not easy, but within reach

Expansion Microscopy (ExM)	300 nm / 150× expansion = 2 nm
SIM + ExM	150 nm / 75× expansion = 2 nm
DNA PAINT + ExM	10 nm / 5× expansion = 2 nm

Thanks

Wyss Institute
Jonathan Braff
Nicholas Conway
Jessica Duda
Kevin Esvelt
Thomas Ferrante
Sam Inverso
David Kalish
Seth Kroll
Kathleen Leeper
Daniel Levner
Chao Li
Allison Martin
Steven Perrault
Ben Pruitt
Michael Sismour
Richard Terry
Brian Turczyk
Frederick Vigneault
Daniel Wiegand

DAC
Peng Yin
Peter Sorger
John Quackenbush

Boyden Lab
Ed Boyden
Shahar Alon
Fei Chen
Paul Tillberg
Asmamaw Wassie

Church Lab
John Aach
Volker Busskamp
Sven Dietz
Nancy Feng
Dan Goodman
Kettner Griswold
Jeremy Huang
Eswar Iyer
Reza Kalhor
Gleb Kuznetsov
Jay Lee
Nathan Lewis
Prashant Mali
Adam Marblestone
Kalim Mir
Pierce Ogden
Srivatsan Raman
Paul Reginato
Jonathan Scheiman
Yu Wang

VisiTech
Steve Coleman

Kharchenko Lab
Peter Kharchenko
Joseph Herman
Fan Jean

Zador Lab
Tony Zador
Ian Peikon

Systems Biology
Tim Mitchison
Andrew Murray
Emily Runey
Sam Reed
Hattie Chung
Siting Gan
Antonina Hafner
Stephanie Hays
Adrian Jinich
Jose Reyes
Cameron Myhrvold
Mashaal Sohail
Eric Solis
Matthieu Landon
Alex Ng

Yin Lab
Peng Yin
Maier Avendano
Mingjie Dai
Ralf Jungmann
Cameron Myhrvold
Luvena Ong
Florian Schuederr
Johannes Woehrstein

Yanai Lab
Itai Yanai
Maayan Baron

Students
Brian Ahern
Tiffany Chen
Vivek Dasari
Joshua Lehrer

Broad Institute
Xian Adiconis
Martin Aryee
Joshua Levin
Alex Shalek
Aviv Regev

Allen Brain Institute
Amy Bernard
Allan Jones
Bosiljka Tasic

Megason Lab
Sean Megason
Kishore Mosaliganti

Depace Lab
Angela Depace
Meghan Bragdon
Tara Martin

Kennedy Lab
Scott Kennedy
Brandon Fields

Colaiacovo Lab
Monica Colaiacovo
Jinmin Gao

Beck Lab
Andrew Beck
Octavian Bucur
Jong Cheol
Humayun Irshad
Alex Lancaster

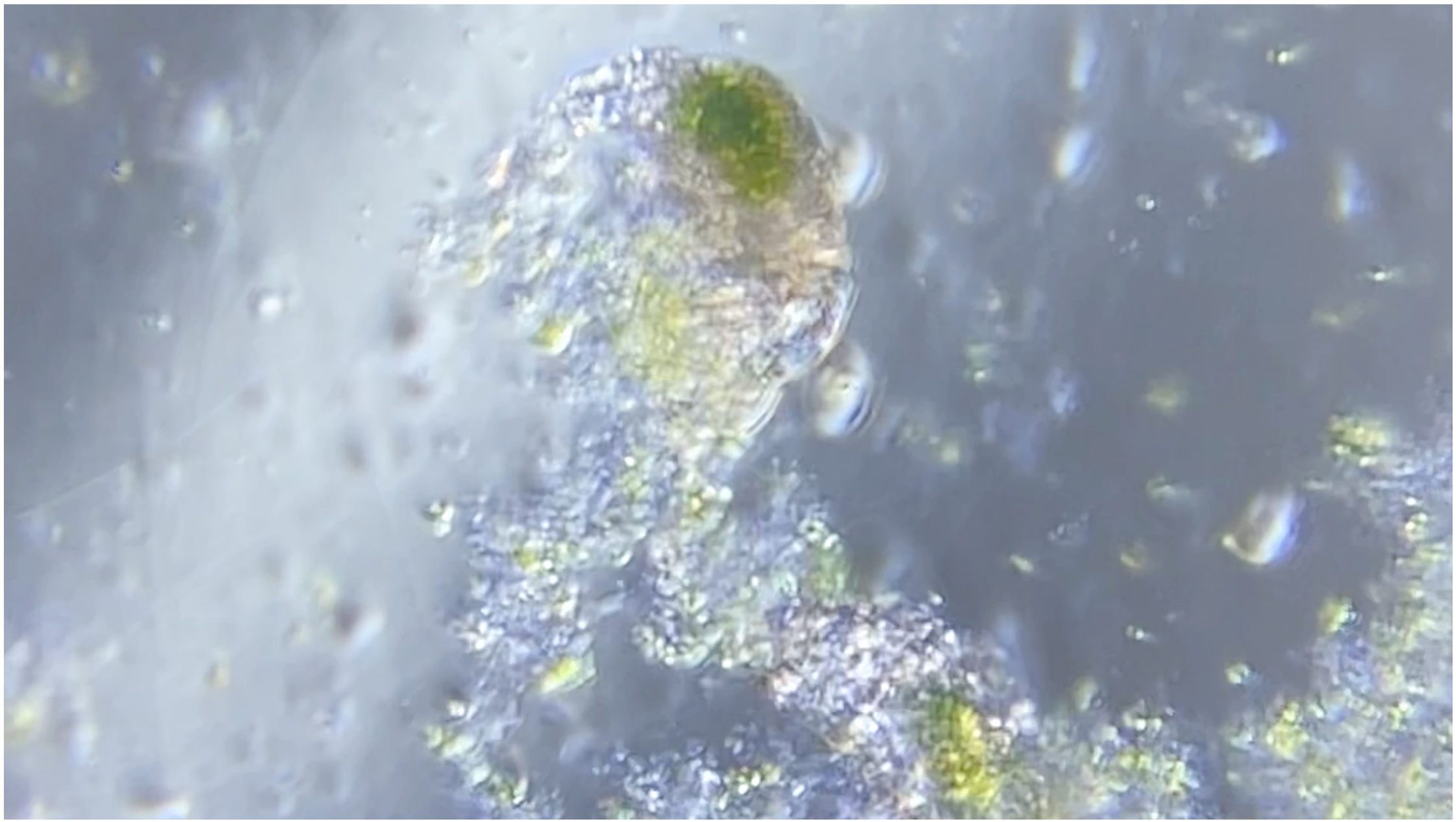
Wu Lab
Ting Wu
Huy Nguyen
Son Nguyen

Funding
NSF Fellowship DGE1144152
CEGS P50 HG005550
NIH 1R01MH103910-01
NIMH MH098977
NHBLI RC2HL102815
Allen Institute for Brain Science

Evan Daugharthy
evan.daugharthy@gmail.com

Homework Preview

Part 1: FoldScope



Samples from Carolina.com

Evan Daugharthy
evan.daugharthy@gmail.com

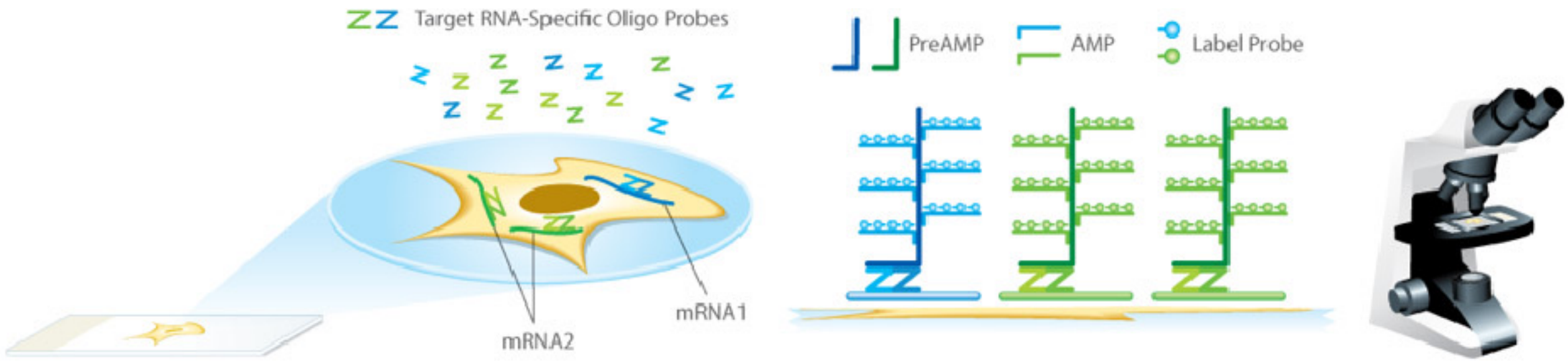
READCOOR

WYSS INSTITUTE

Homework Preview

Part 2: FISH Probe Design

Use bioinformatics & Python tools to screen potential FISH probes (encoders)



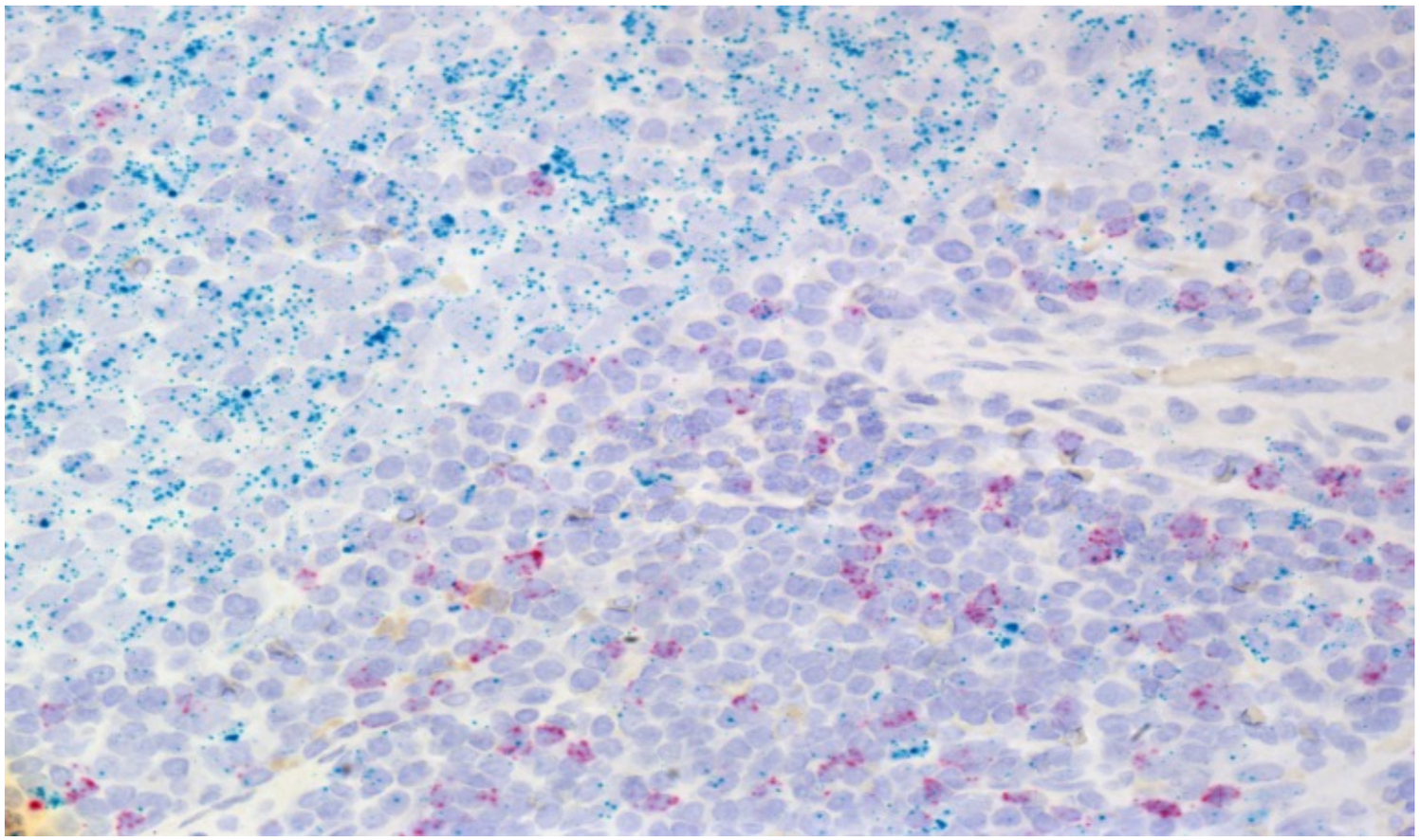
Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE

Homework Preview

Part 3: FIJI Image analysis (smFISH)

To be supplied by recitation – experiment ongoing today!



Advanced Cell Diagnostics

Thanks!

Evan Daugharthy
evan.daugharthy@gmail.com

READCOOR WYSS INSTITUTE