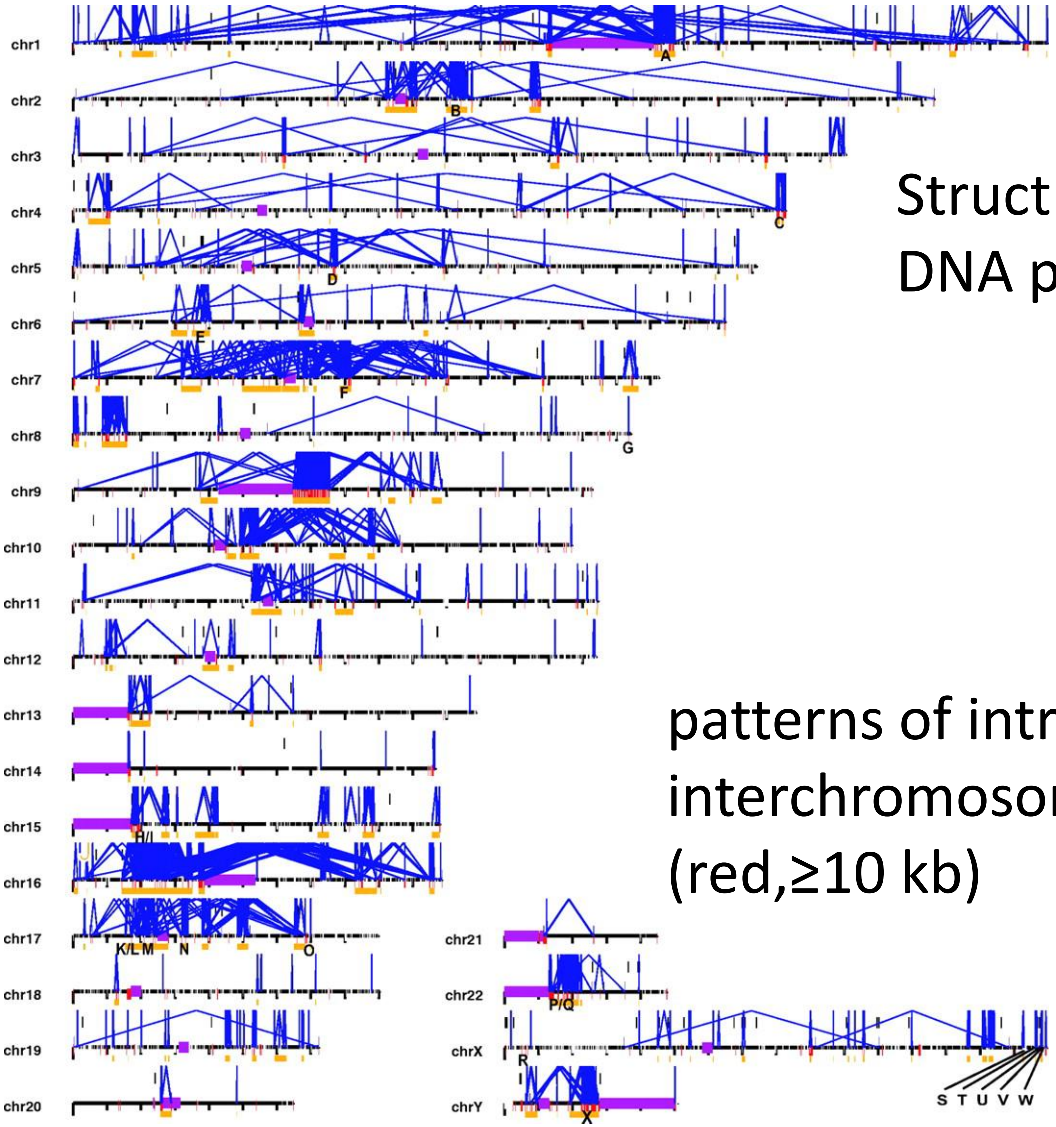
Functional Genomics and Single-Cell Genomics

Christine Queitsch
Department of Genome Sciences
queitsch@uw.edu

Outline

- Challenges
- Combining methods for better genome assembly
- Functional genomics – let's sequence it
- Principles and major findings
- Single cell genomics – why and how
- Applications and examples
- What's next?

Challenges: variants discovery and characterization

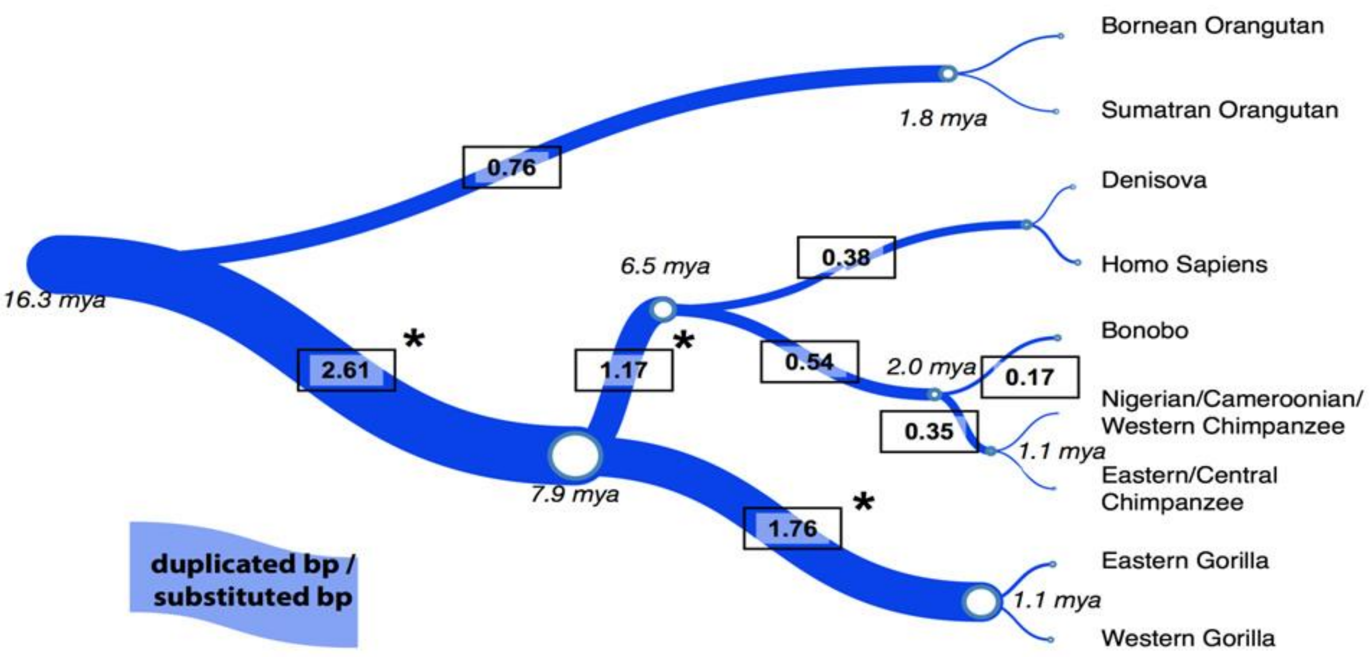


Structural variation and repetitive DNA pose unsolved challenges!

patterns of intrachromosomal (blue) and interchromosomal segmental duplication (red,≥10 kb)

Bailey et al Science 2002

Why do these regions matter?



- segmental duplications preceded divergence of humans
- segments encode human/great ape-specific gene families
- genes with functions in neurodevelopment, cell proliferation
- CNVs in segmental duplication regions are implicated in many human disease

How to assemble such regions?

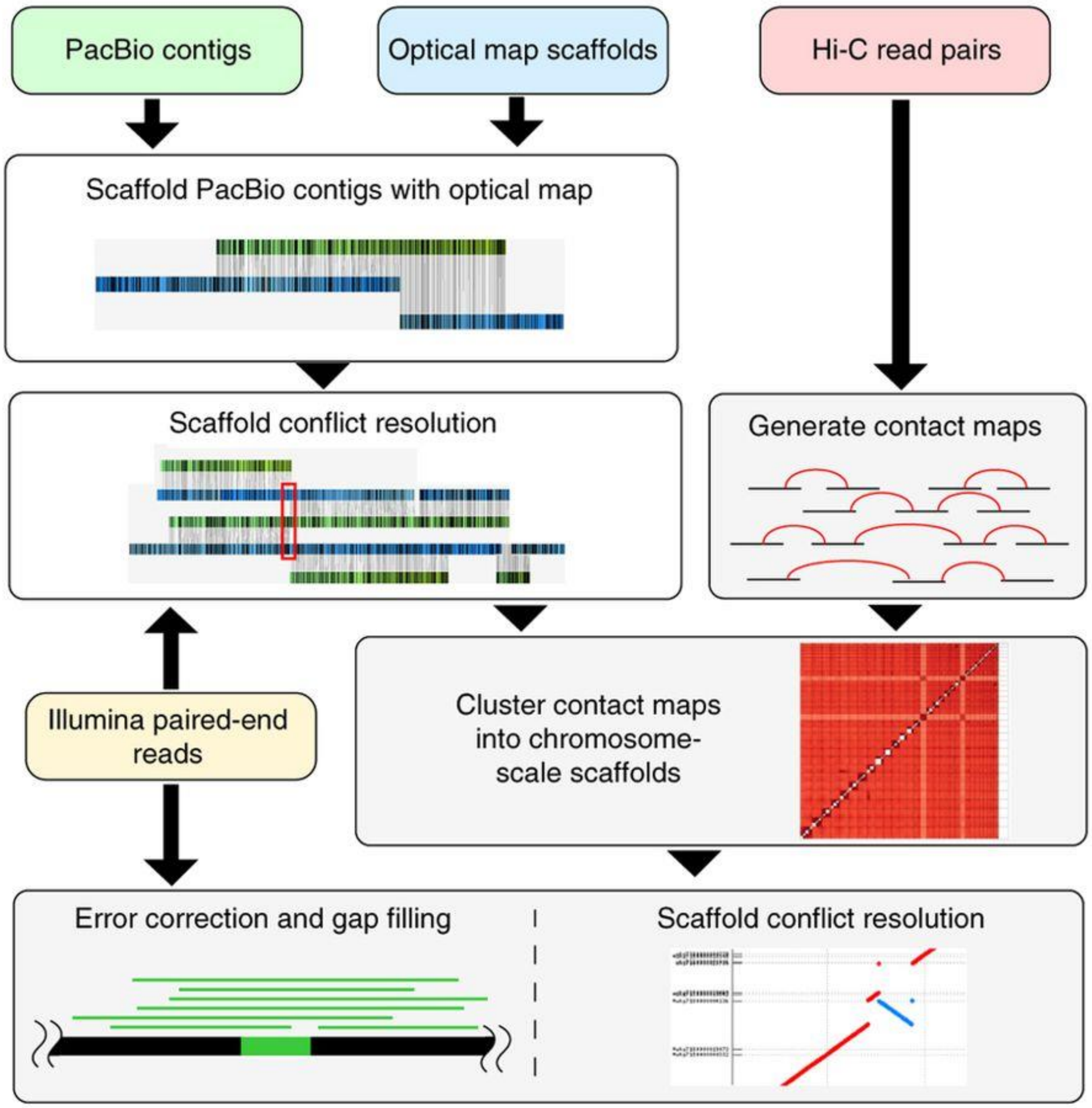
- reduce complexity by sequencing haploid genomes ->
  - hydatidiform moles
  - HAP1 and hapESC cells

- longer reads -> PacBio, nanopore and short-read sequencing for error corrections

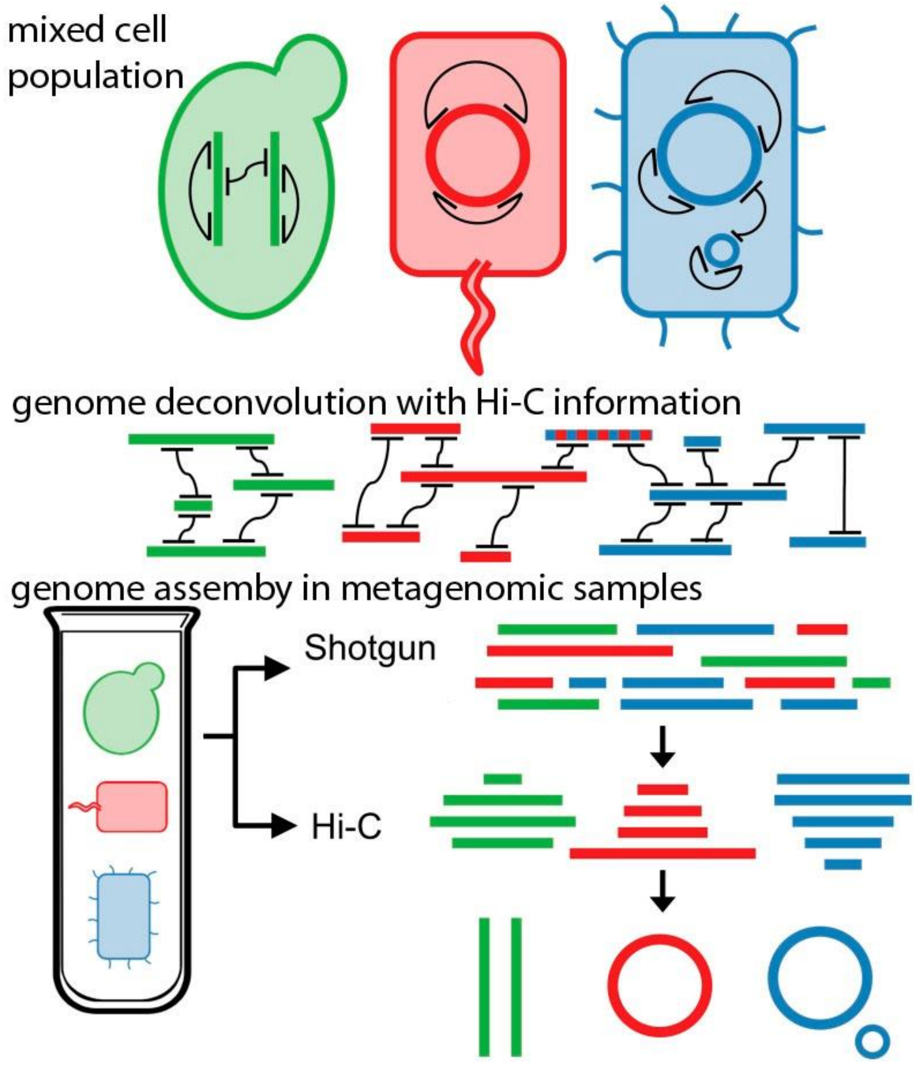
- adding proximity context –> Hi-C

Huddleston et al. 2017, Genome Research
Bickart, 2017, Nature Genetics
Burton, 2013, G3

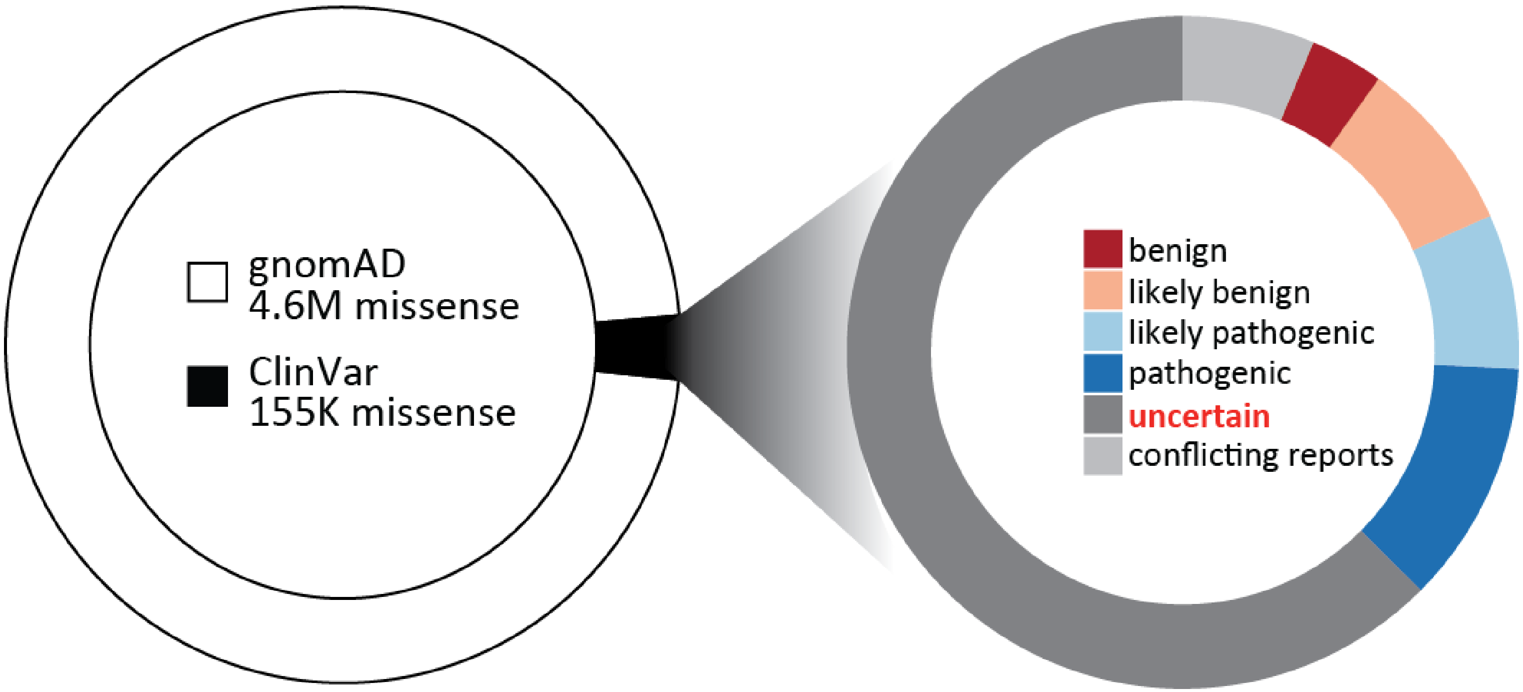
State-of-the-art – the most complete mammalian genome



Context allows deconvolution of complex samples



Variants are everywhere – what do they mean?

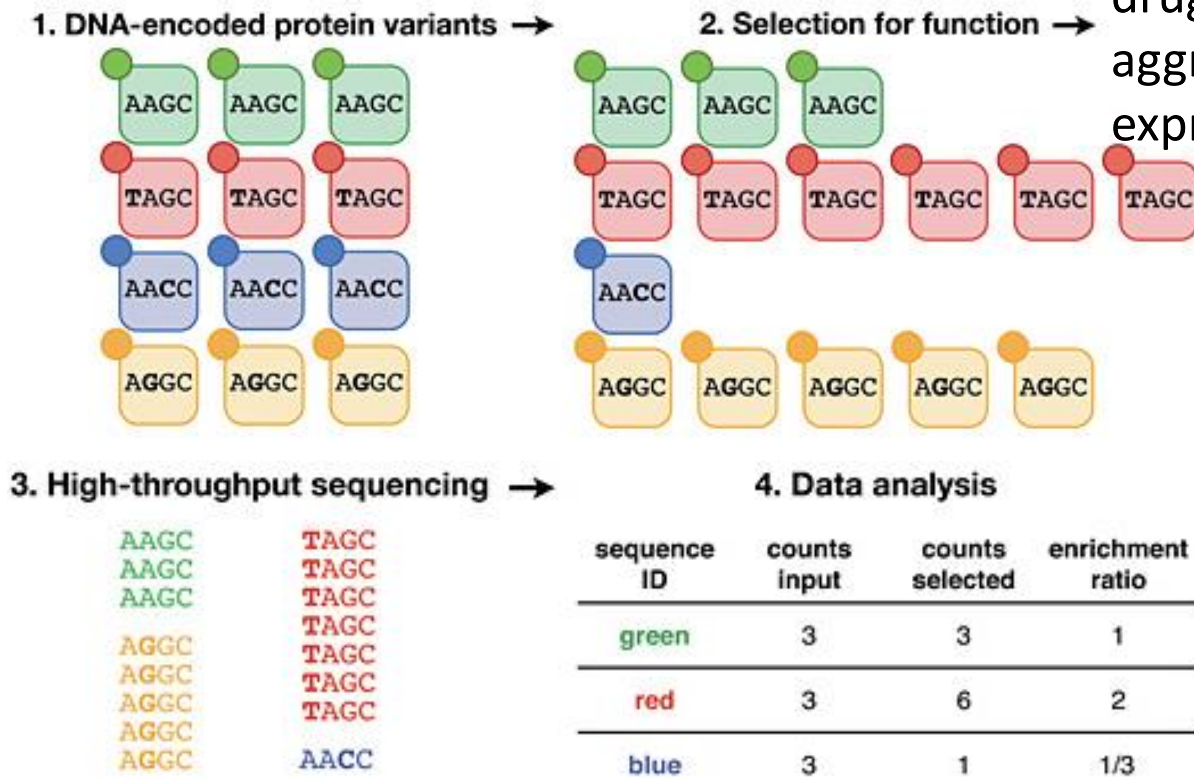


The challenge is enormous…this is just representing the coding portion of the genome (<1%).

Functional genomics – massively parallel assays

Saturation mutagenesis

cell growth
drug resistance
aggregation
expression of reporter



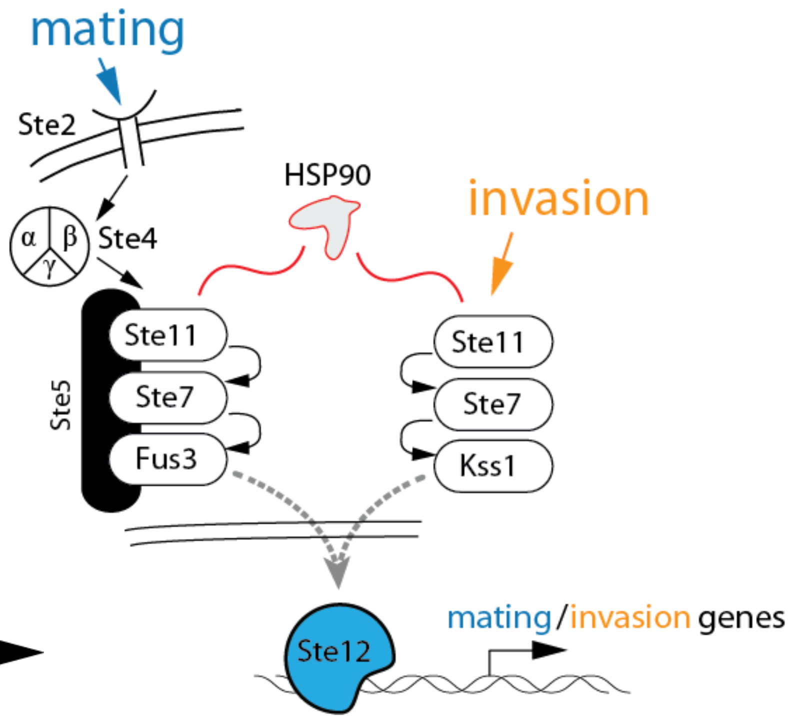
sequence ID	counts input	counts selected	enrichment ratio	
green	3	3	1	neutral
red	3	6	2	beneficial
blue	3	1	1/3	deleterious
yellow	3	5	5/3	possibly beneficial

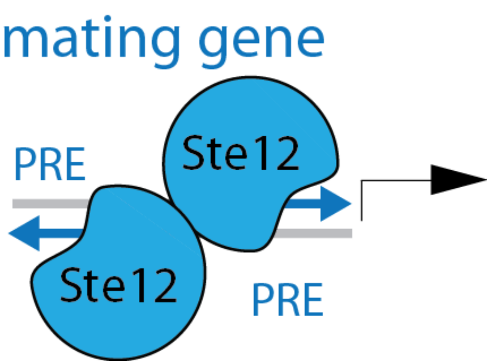
…any variant associated with a selectable function can be interrogated…

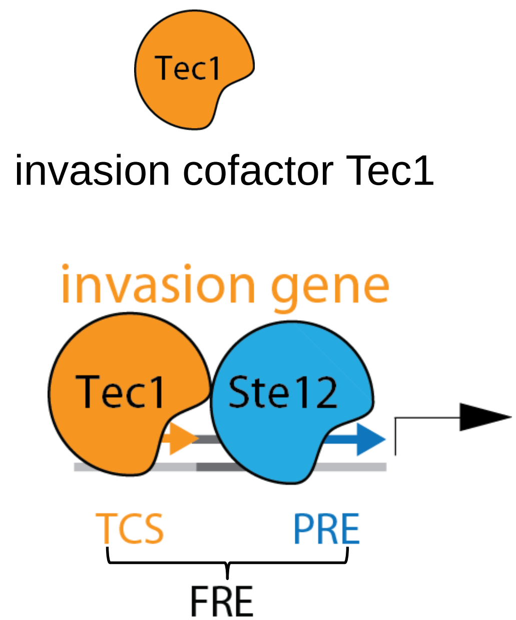
Fowler & Fields Nat. Methods 2010

Examples

Yeast mating – a canonical MAPK pathway with several phenotypic outputs



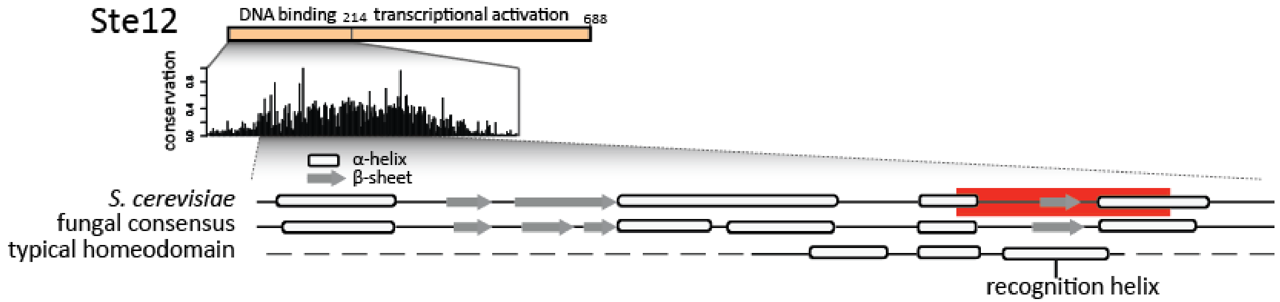




Mating and invasive growth converge on a single transcription factor: Ste12.

Dorrity et al., PNAS, 2018
Dolan, Kirkman & Fields, PNAS (1989)

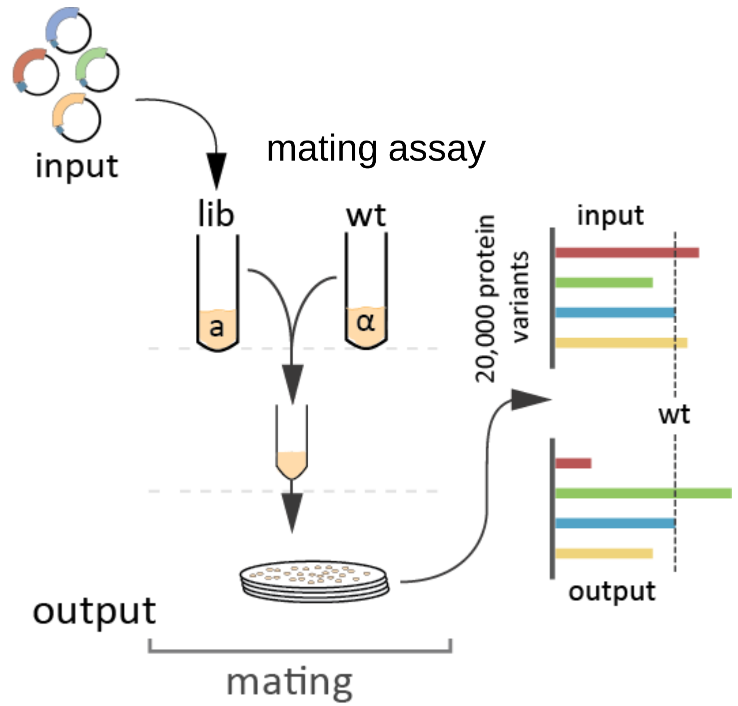
Ste12 variant performance in mating



Are there Ste12 DBD mutants that separate traits?

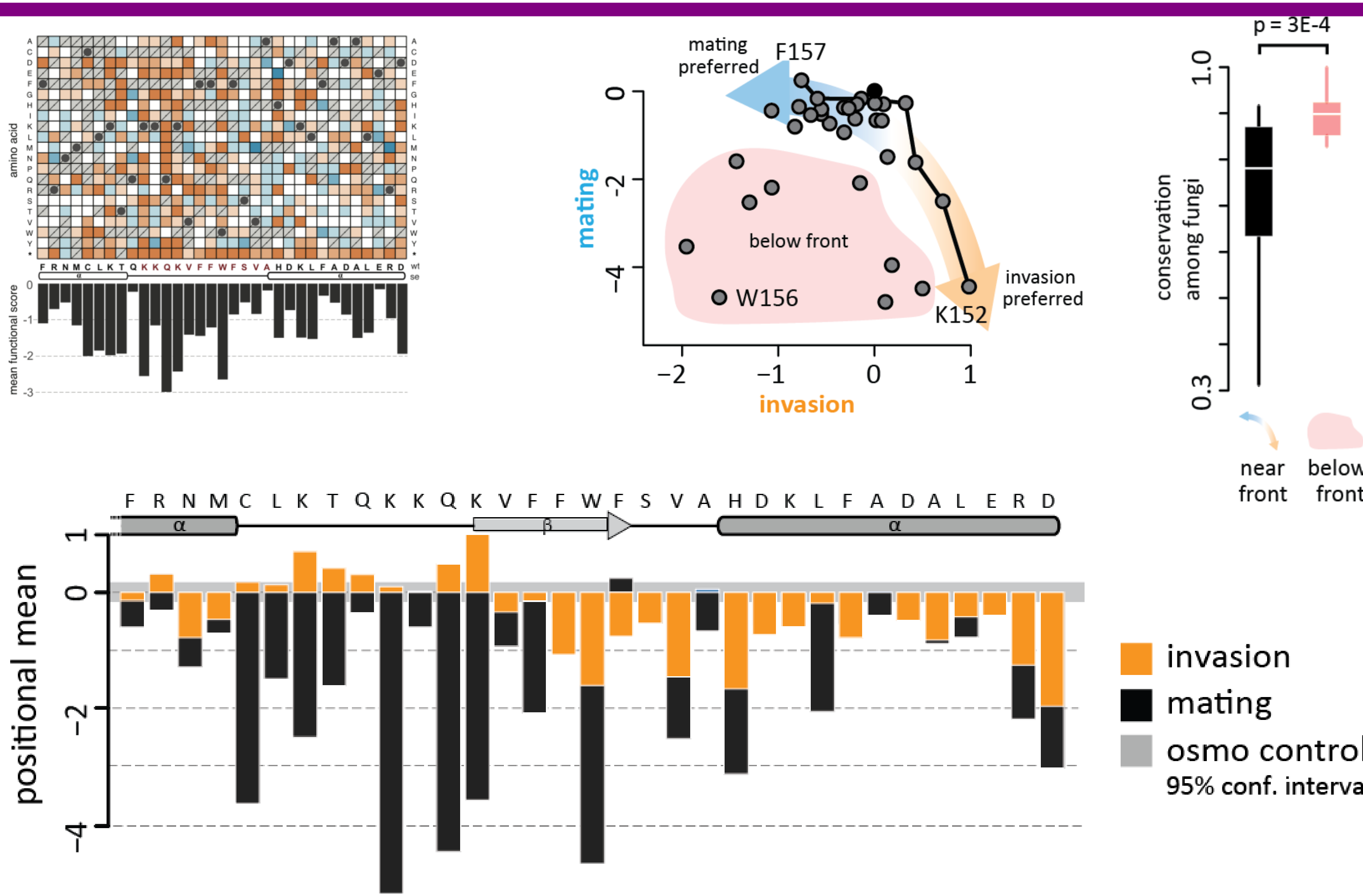
Phenotyping thousands of variants by sequencing

Input library of Ste12 variants



Deep mutational scanning allows assessing functional consequences of all (most) possible Ste12 DBD variants for mating and invasion under different conditions.

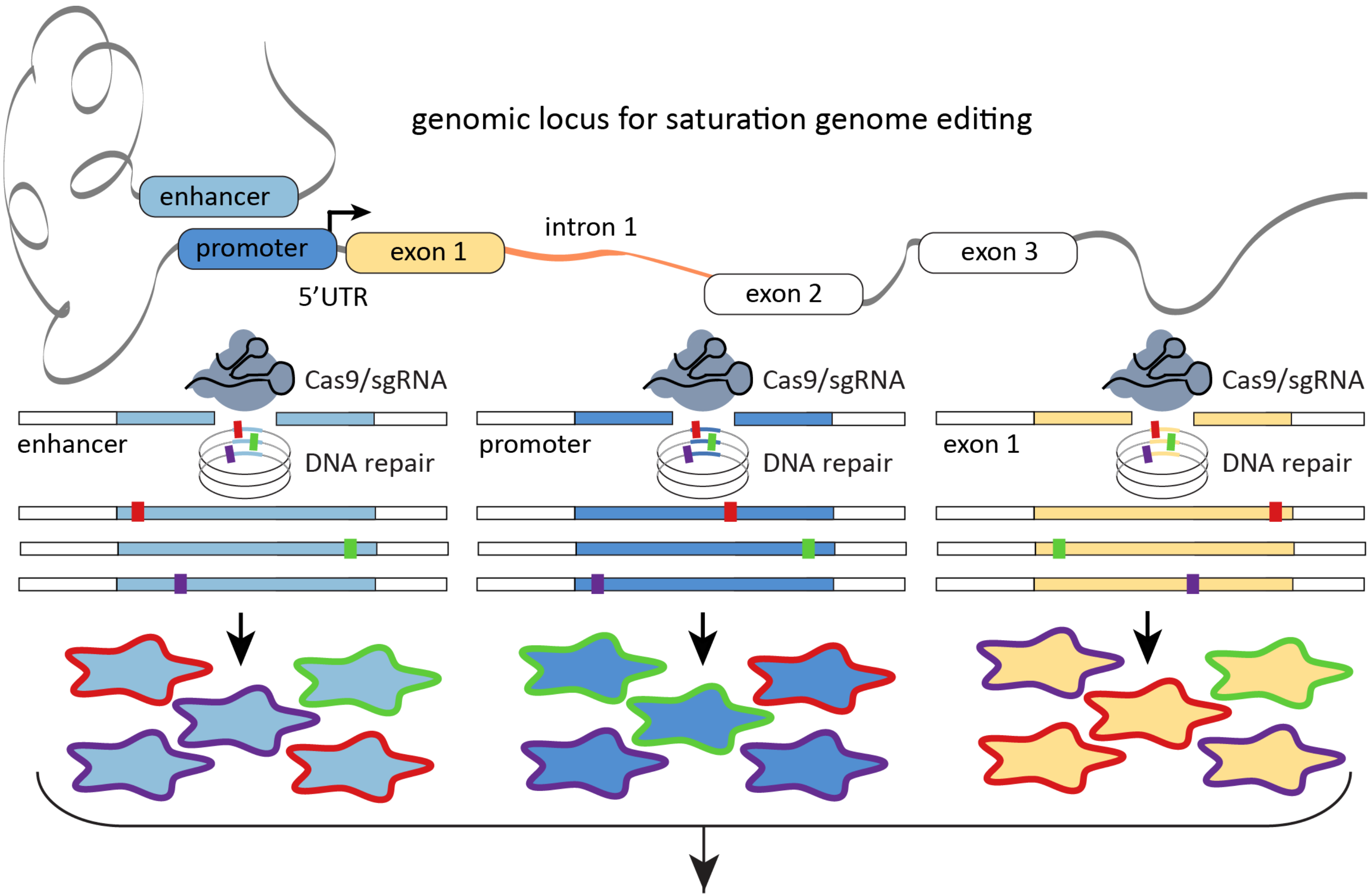
Single mutation suffices to shift between traits



Certain Ste12 variants confer hyperinvasion at the cost of mating.

Dorrity et al. PNAS 2018

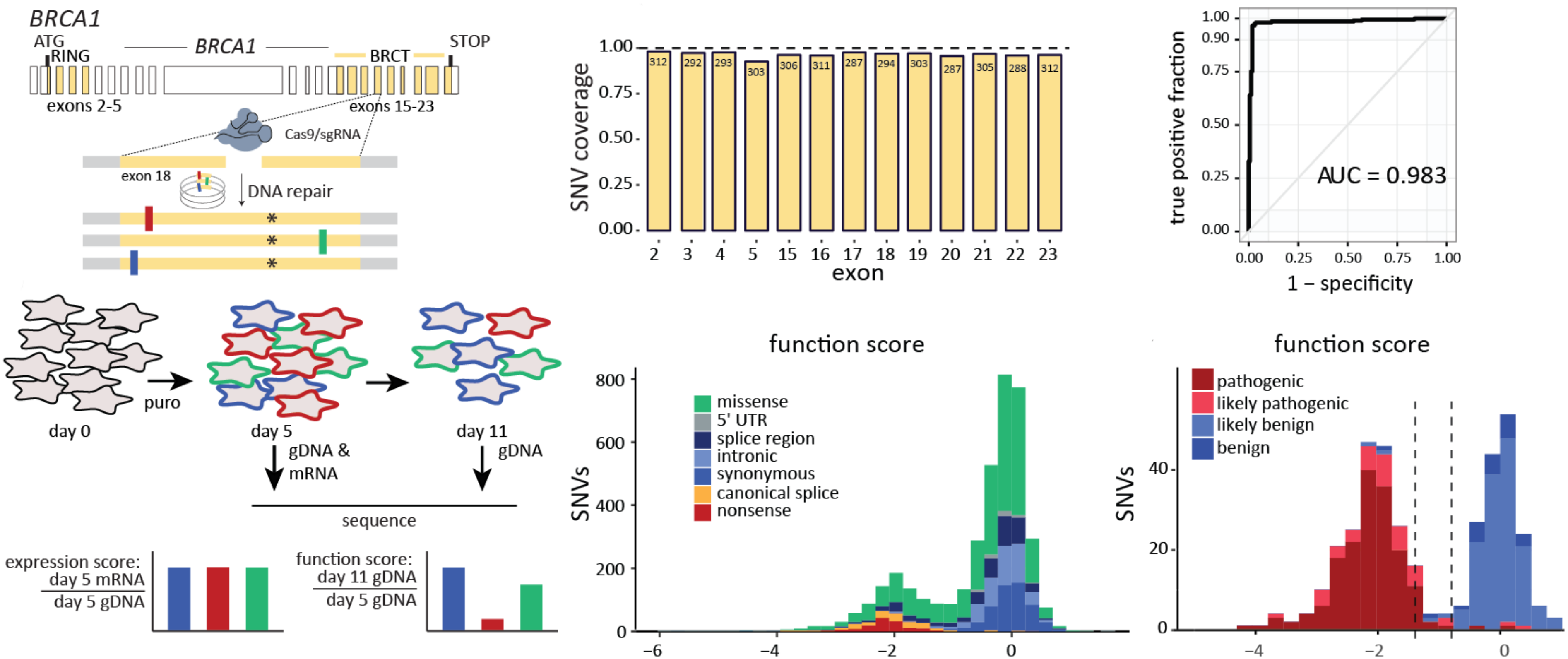
From yeast to human genes – within genomic context



high-content and high-throughput phenotyping:
molecular phenotypes
cellular phenotypes

BRCA1 – the breast cancer gene

BRCA1 saturation genome editing yields functional information on variants of previously unknown effects



Starita et al. Nature in press.
Starita et al. Genetics 2015

Challenge – how to scale up???

generating all possible variants:

- efficient insertion of variant libraries in genome OR efficient editing

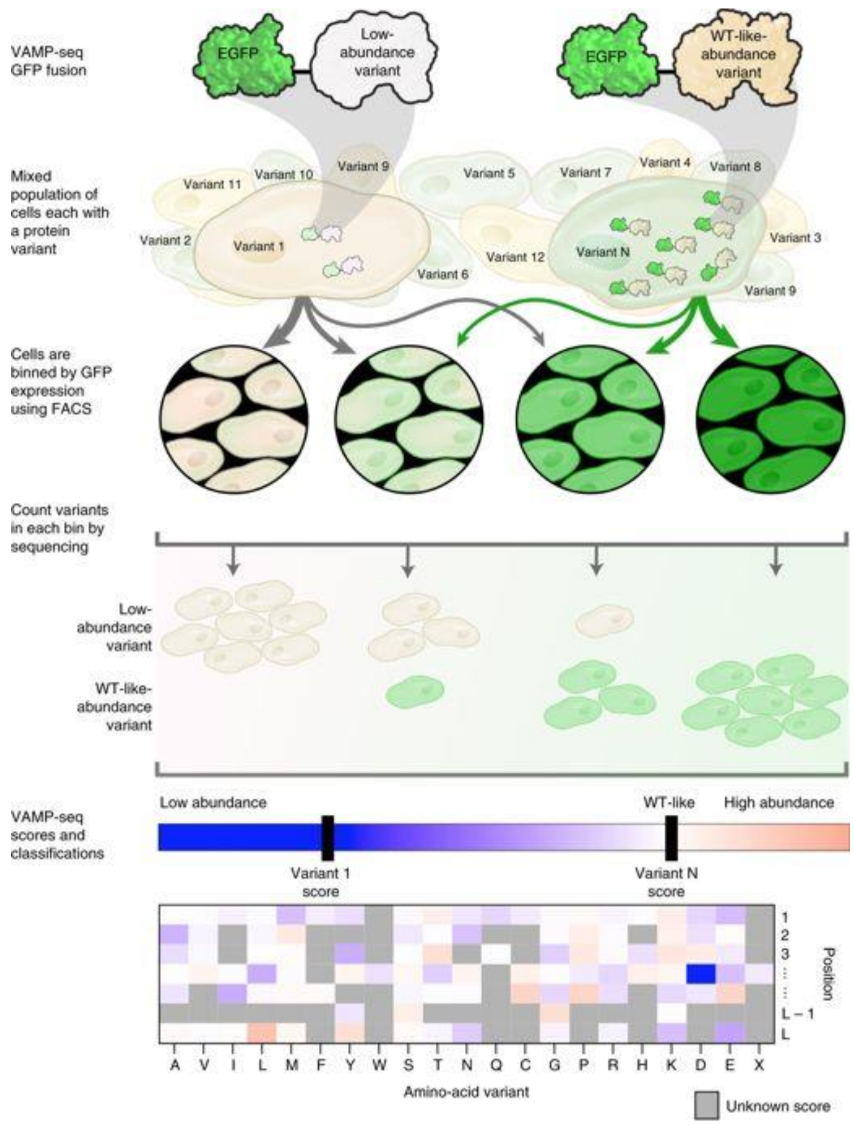
selection regimes and relevant phenotypes:

- protein aggregation holds functional information –> VAMP-seq
- cell shape -> advanced microscopy compatible with selection of cells
- nucleoli shape and size

NEXT Frontier: assessing function of proteins in similar fashion with mass spectrometry

Matreyek et al. 2018 Nature Genetics

VAMP-seq – deceptively simple…



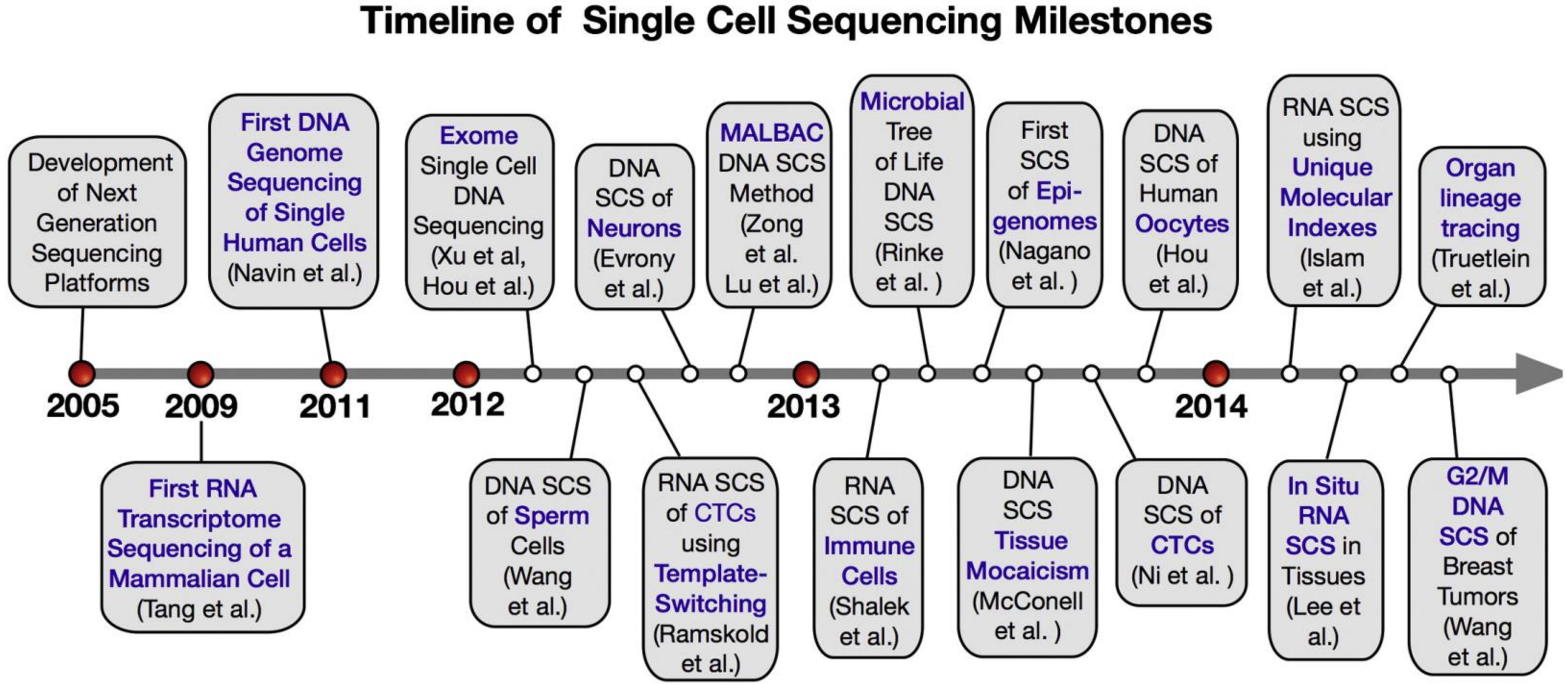
How does cellular gene expression differ for each cell carrying a particular variant?

Or accessibility?

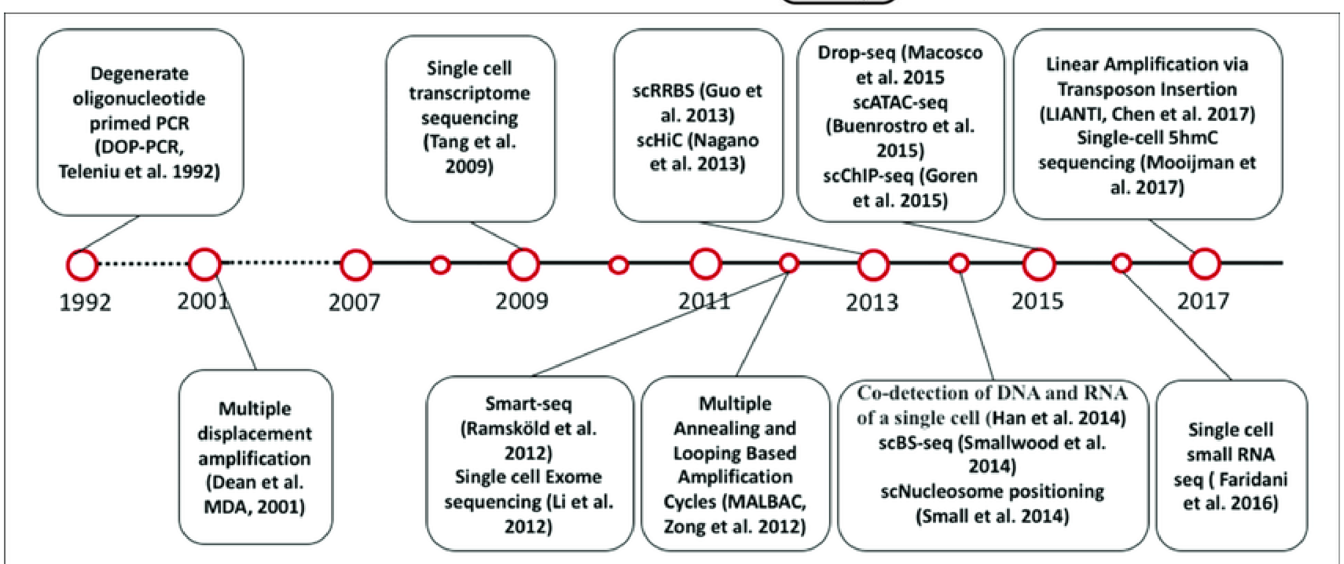
Or TF occupancy

Matreyek et al. 2018 Nature Genetics

Single-cell genomics – when?



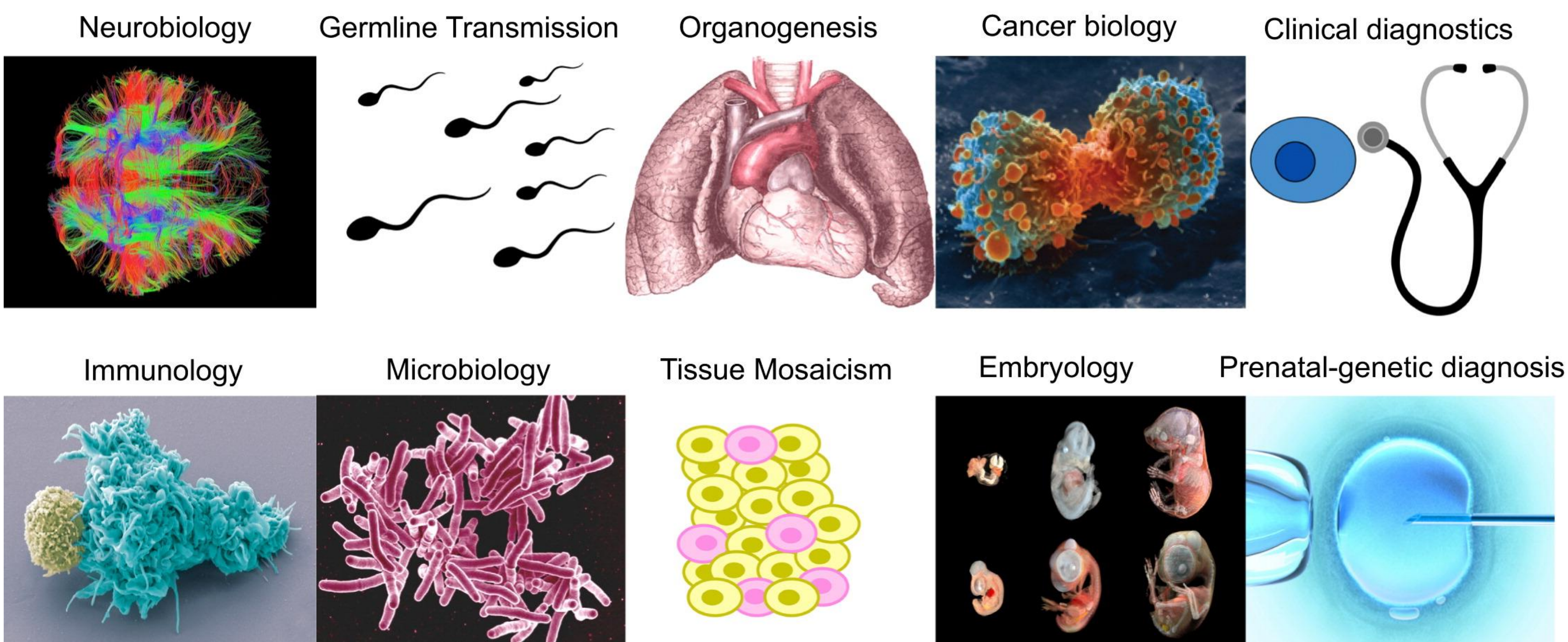
Wang & Navin 2015 Mol. Cell



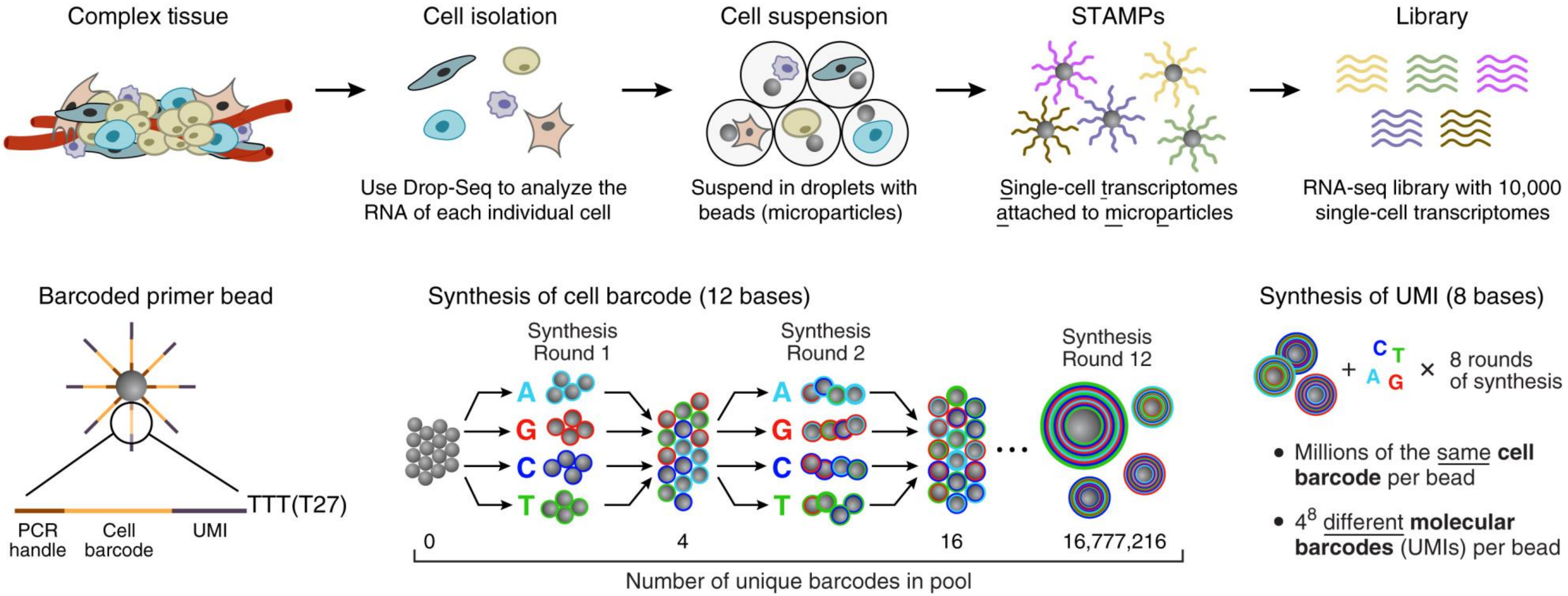
Hu et al. Frontiers in Cell and Developmental Biology 2018

Single-cell genomics – why?

- discover new cell types
- reveal developmental trajectories
- understand heterogeneity
- cancer
- neurodevelopment and memory



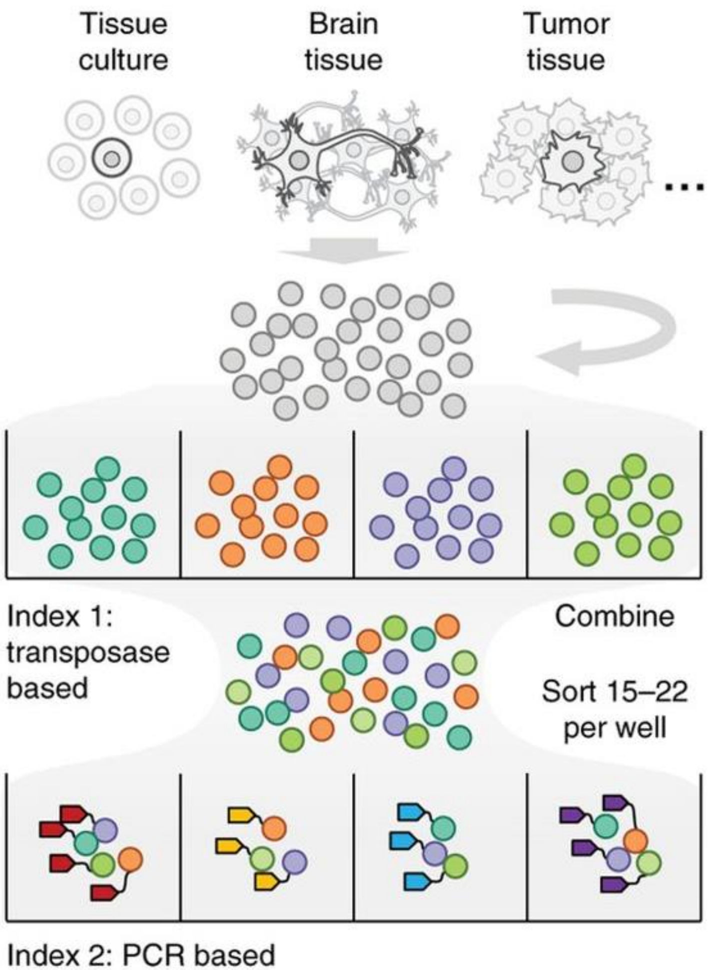
Single-cell genomics – technical principles



Macosko et al. 2015 Cell

Drop-seq

Sci-seq – combinatorial indexing



Vitak et al 2017 Nature Methods

REPORT

Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing

Darren A. Cusanovich[1], Riza Daza[1], Andrew Adey[2], Hannah A. Pliner[1], Lena Christiansen[3], Kevin L. Gunderson[3], Frank J. St...

+ See all authors and affiliations

Science 22 May 2015:
Vol. 348, Issue 6237, pp. 910-914
DOI: 10.1126/science.aab1601

Science 2015

RESEARCH ARTICLE

SINGLE-CELL GENOMICS

Comprehensive single-cell transcriptional profiling of a multicellular organism

Junyue Cao,[1,2]* Jonathan S. Packer,[1]* Vijay Ramani,[1]† Darren A. Cusanovich,[1]† Chau Huynh,[1] Riza Daza,[1] Xiaojie Qiu,[1,2] Choli Lee,[1] Scott N. Furlan,[3,4,5] Frank J. Steemers,[6] Andrew Adey,[7,8] Robert H. Waterston,[1]‡ Cole Trapnell,[1]‡ Jay Shendure[1,9]‡

Science 2017

Extreme heterogeneity of influenza virus infection in single cells

Alistair B. Russell[1], Cole Trapnell[2], Jesse D. Bloom[1,2]*

*For correspondence: jbloom@fredhutch.org

[1]Basic Sciences Division and Computational Biology Program, Fred Hutchinson Cancer Research Center, Seattle, United States; [2]Department of Genome Sciences, University of Washington, Seattle, United States

eLife 2018

LETTER

doi:10.1038/nature25981

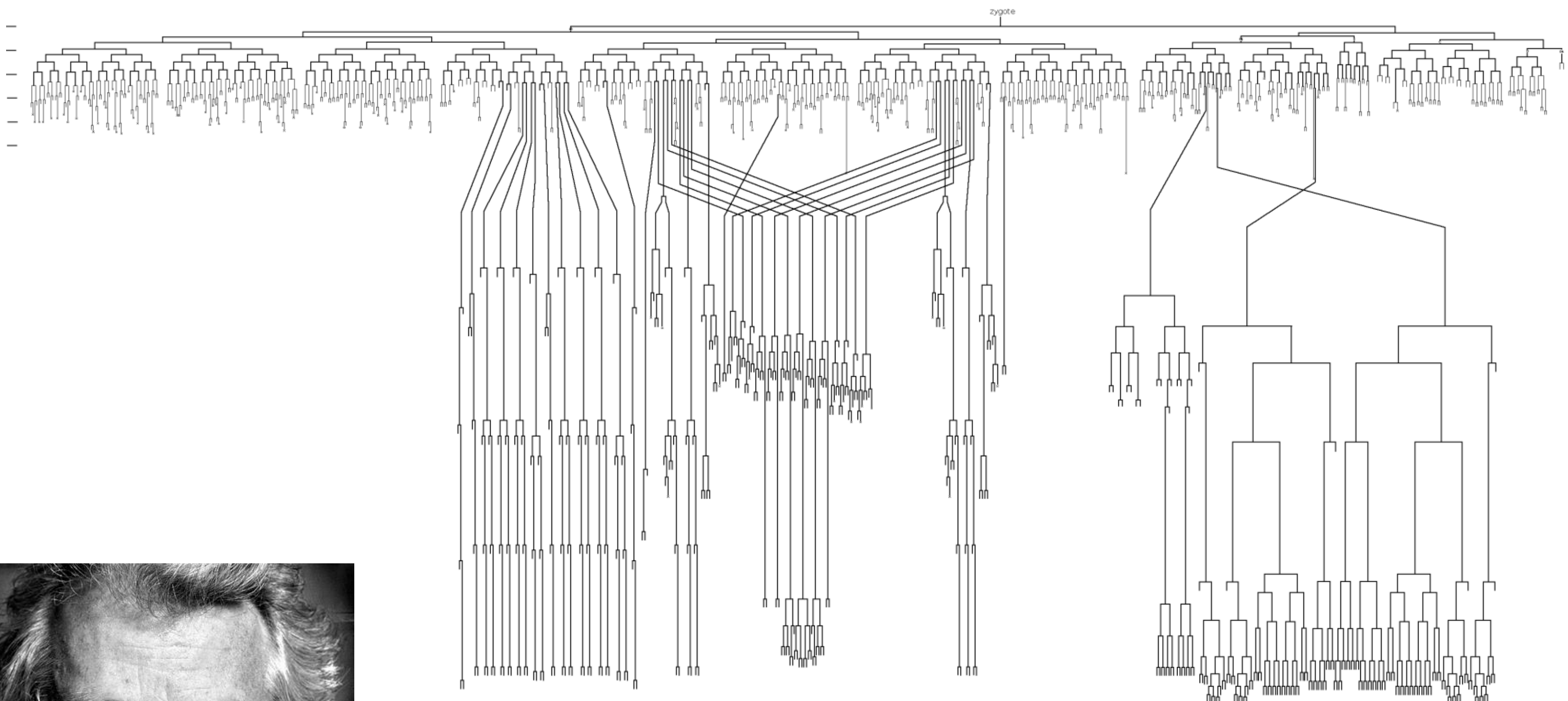
The *cis*-regulatory dynamics of embryonic development at single-cell resolution

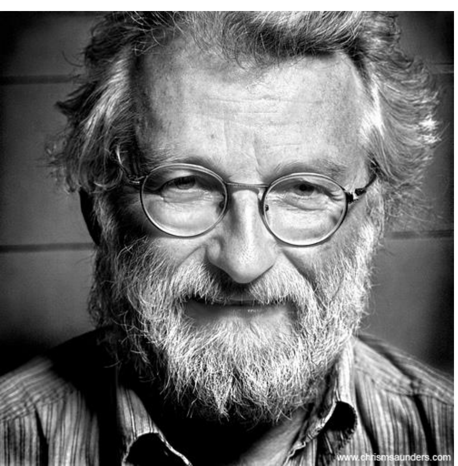
Darren A. Cusanovich[1]*, James P. Reddington[2]*, David A. Garfield[2]†*, Riza M. Daza[1], Delasa Aghamirzaie[1], Raquel Marco-Ferreres[2], Hannah A. Pliner[1], Lena Christiansen[3], Xiaojie Qiu[1], Frank J. Steemers[3], Cole Trapnell[1], Jay Shendure[1,4]§ & Eileen E. M. Furlong[2]§

Nature 2018

Exploring single biology in well-studied models

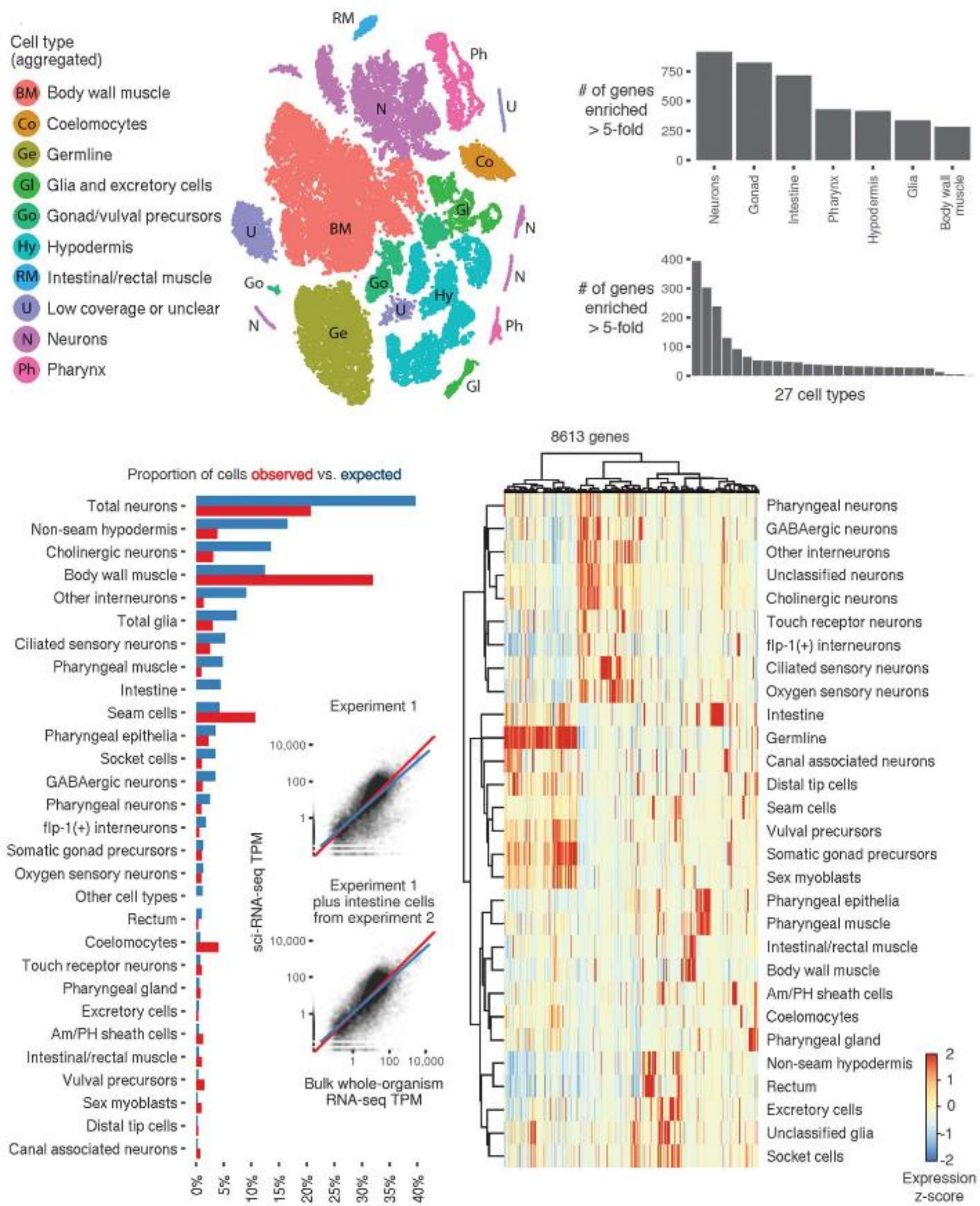
C. elegans - 959 cells





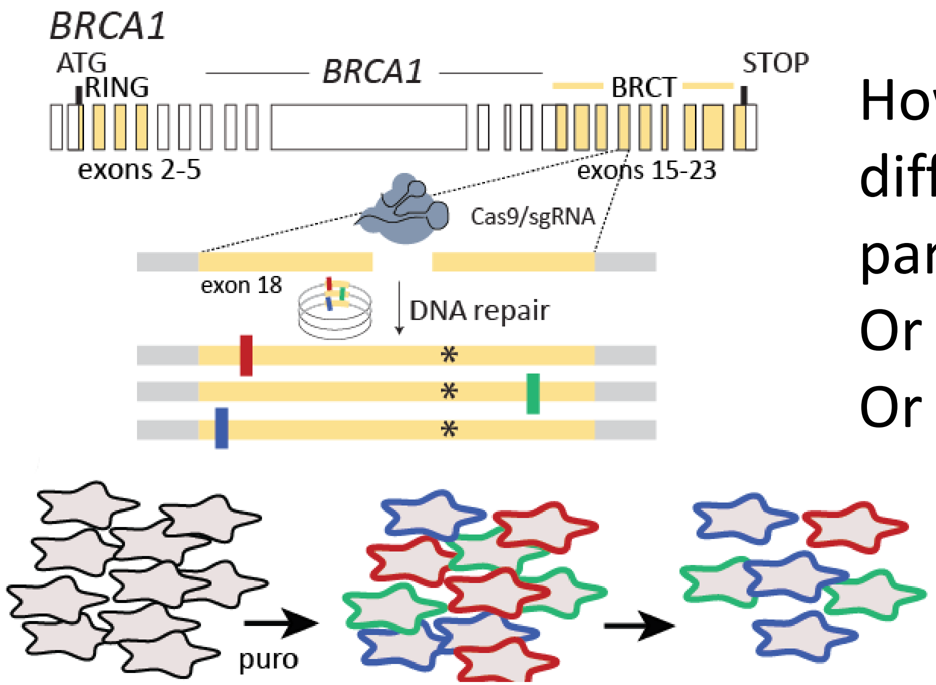
John Sulston - 1983 – embryonic development

What do we find – rare neuronal cells



What’s next?

Combine functional genomics with single-cell genomics



How does cellular gene expression differ for each cell carrying a particular enhancer or gene variant?
Or accessibility?
Or TF occupancy?
Or protein content?