

Dissecting cancer heterogeneity

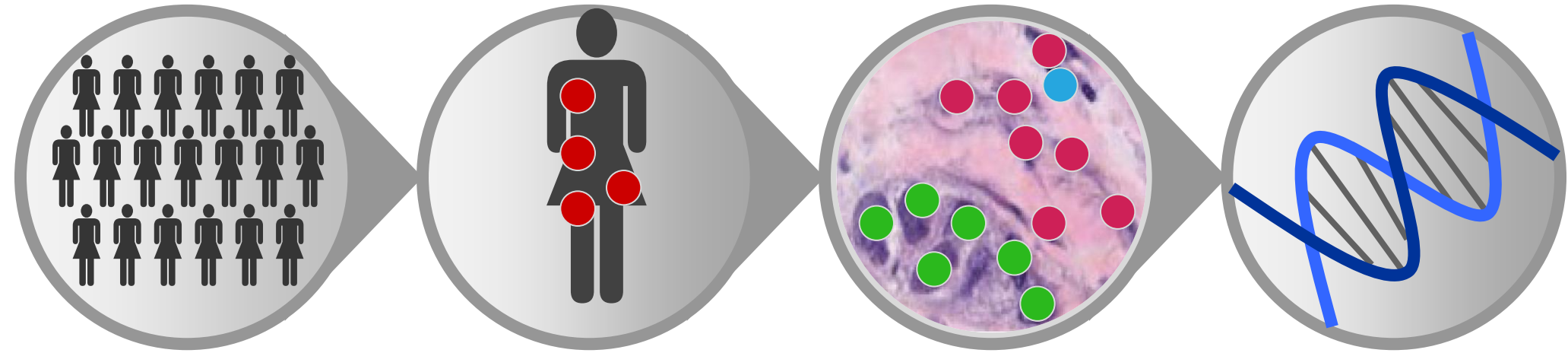
Population > patient > tissue > genome



Florian Markowitz
CRUK Cambridge Institute
www.markowitzlab.org



Heterogeneity in cancer



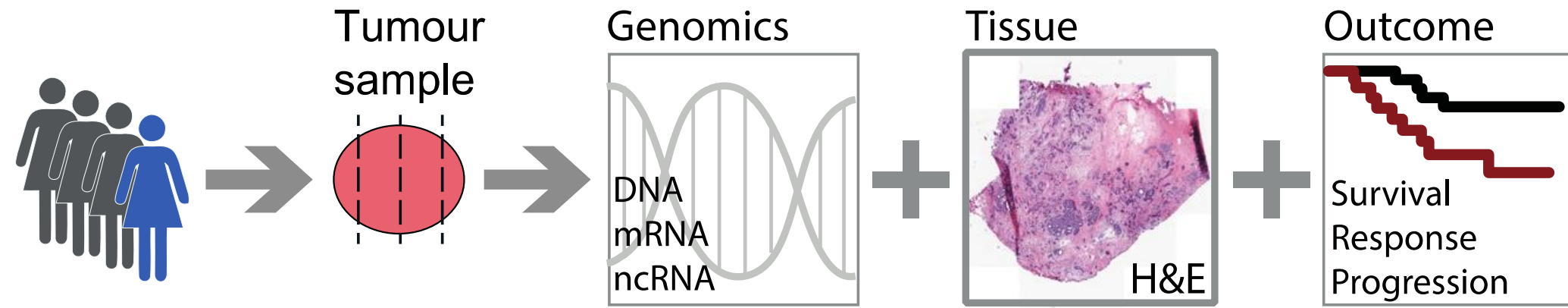
Inter-patient
population
subtypes

Intra-patient
spatial,
temporal

Intra-tumor
tissue

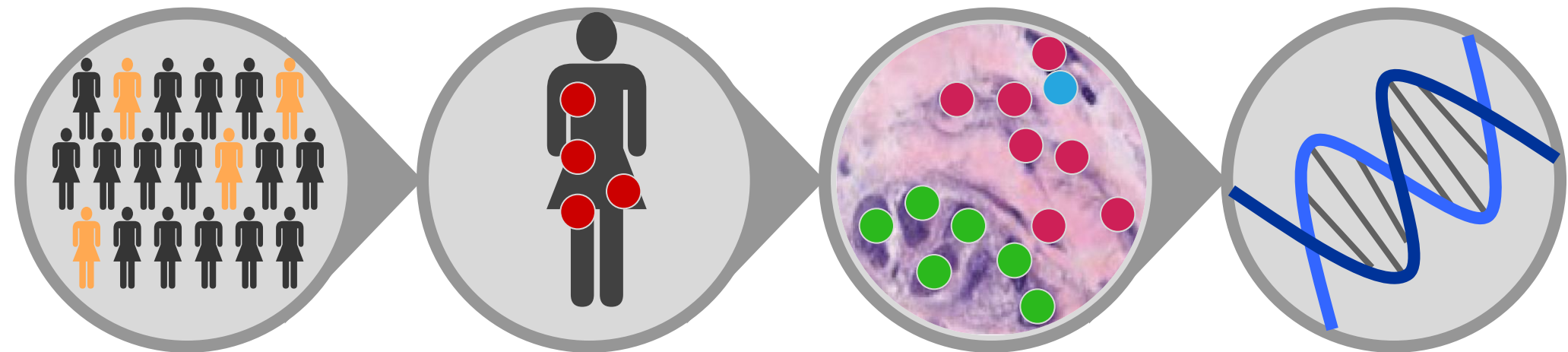
Intra-tumor
genetic

Systems Genetics of Cancer



- What are prognostic subtypes of cancer?
- Which genetic events drive tumour development?
- What are markers to predict disease progression?

Population heterogeneity



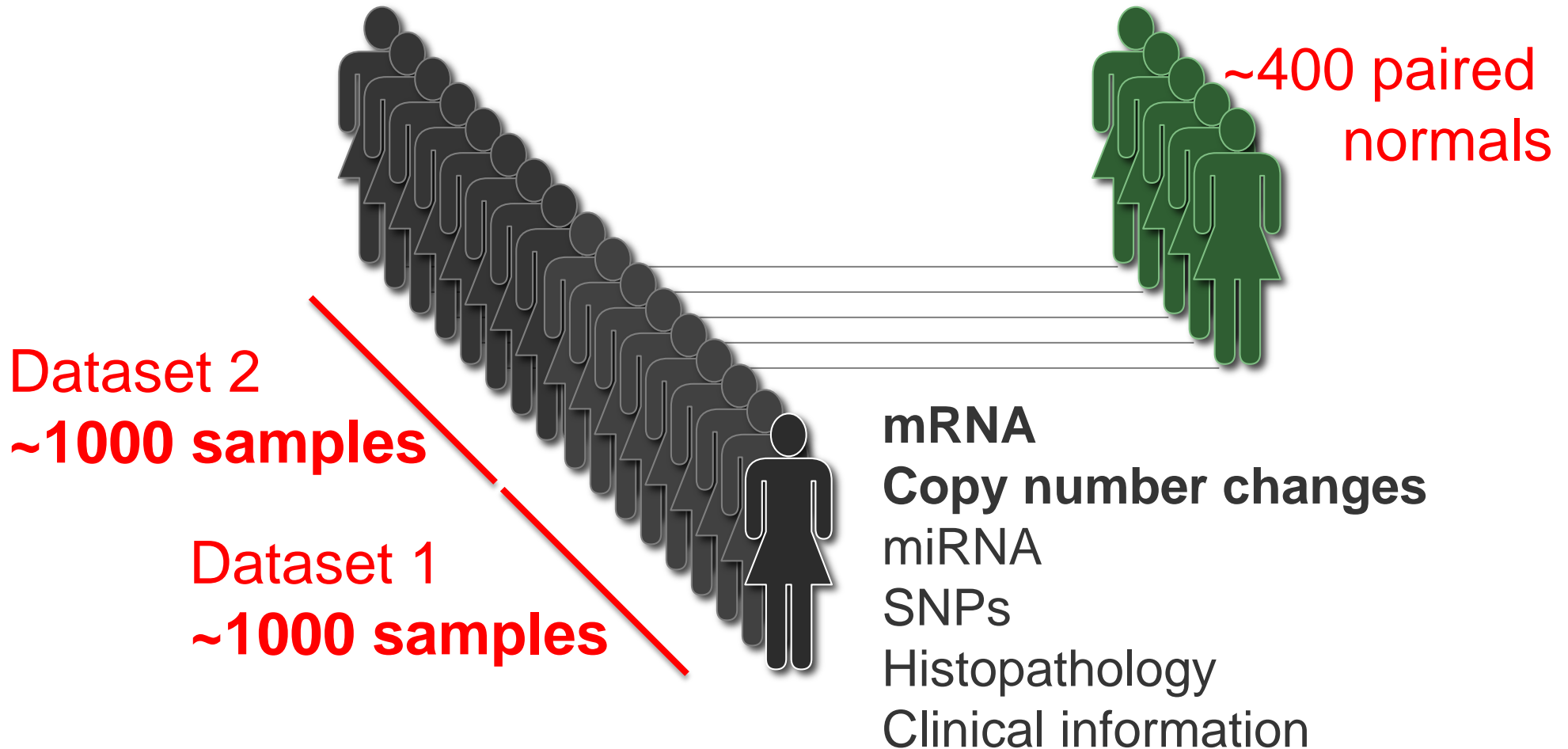
Curtis *et al*, Nature 2011

ARTICLE

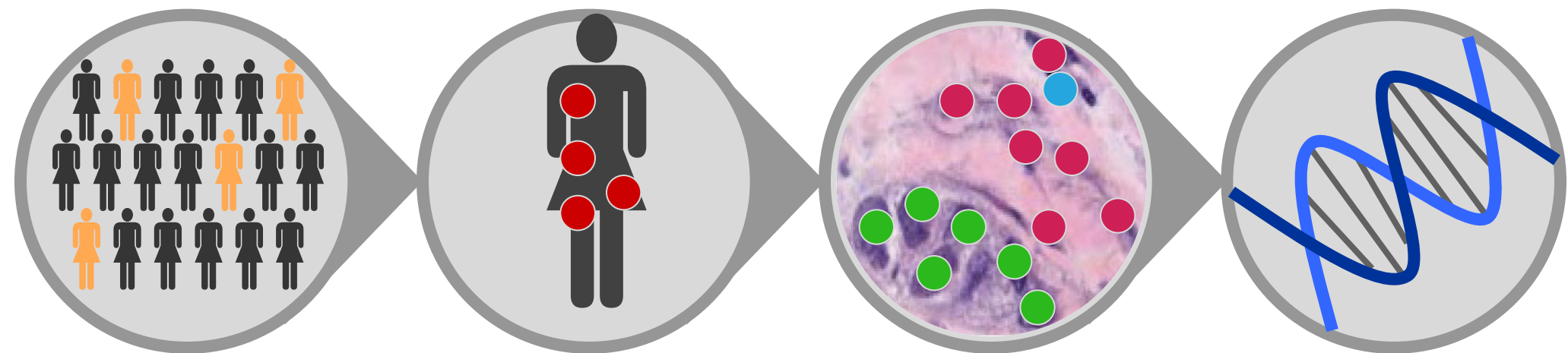
[doi:10.1038/nature10983](https://doi.org/10.1038/nature10983)

The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups

METABRIC – genomic and transcriptional landscape of breast cancer



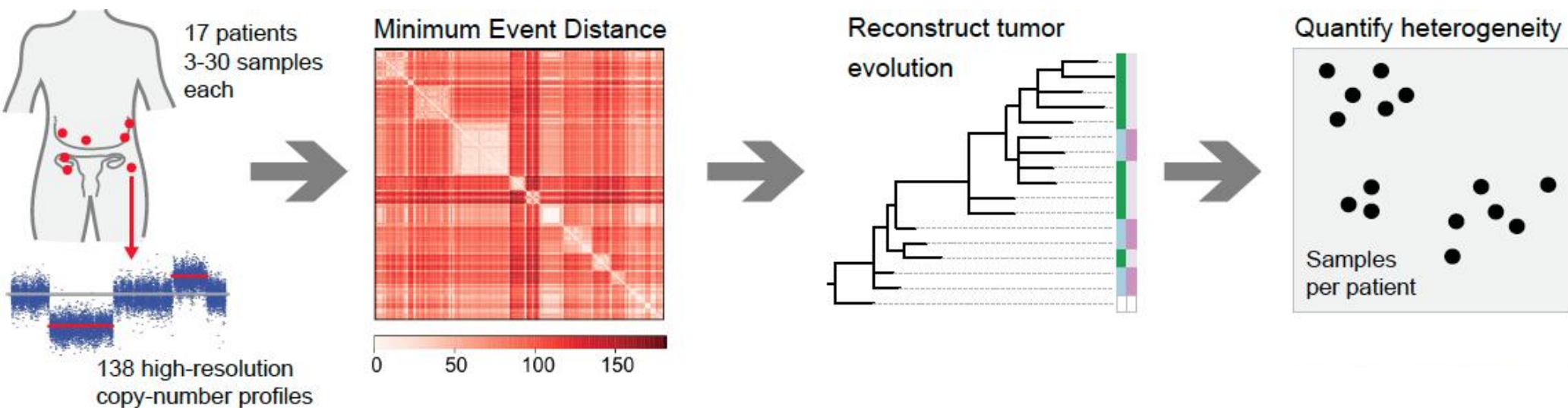
Intra-patient heterogeneity



Schwarz *et al*, submitted

Spatial and temporal heterogeneity
in ovarian cancer predicts survival

Intra-patient heterogeneity in HGSOC



HGSOC

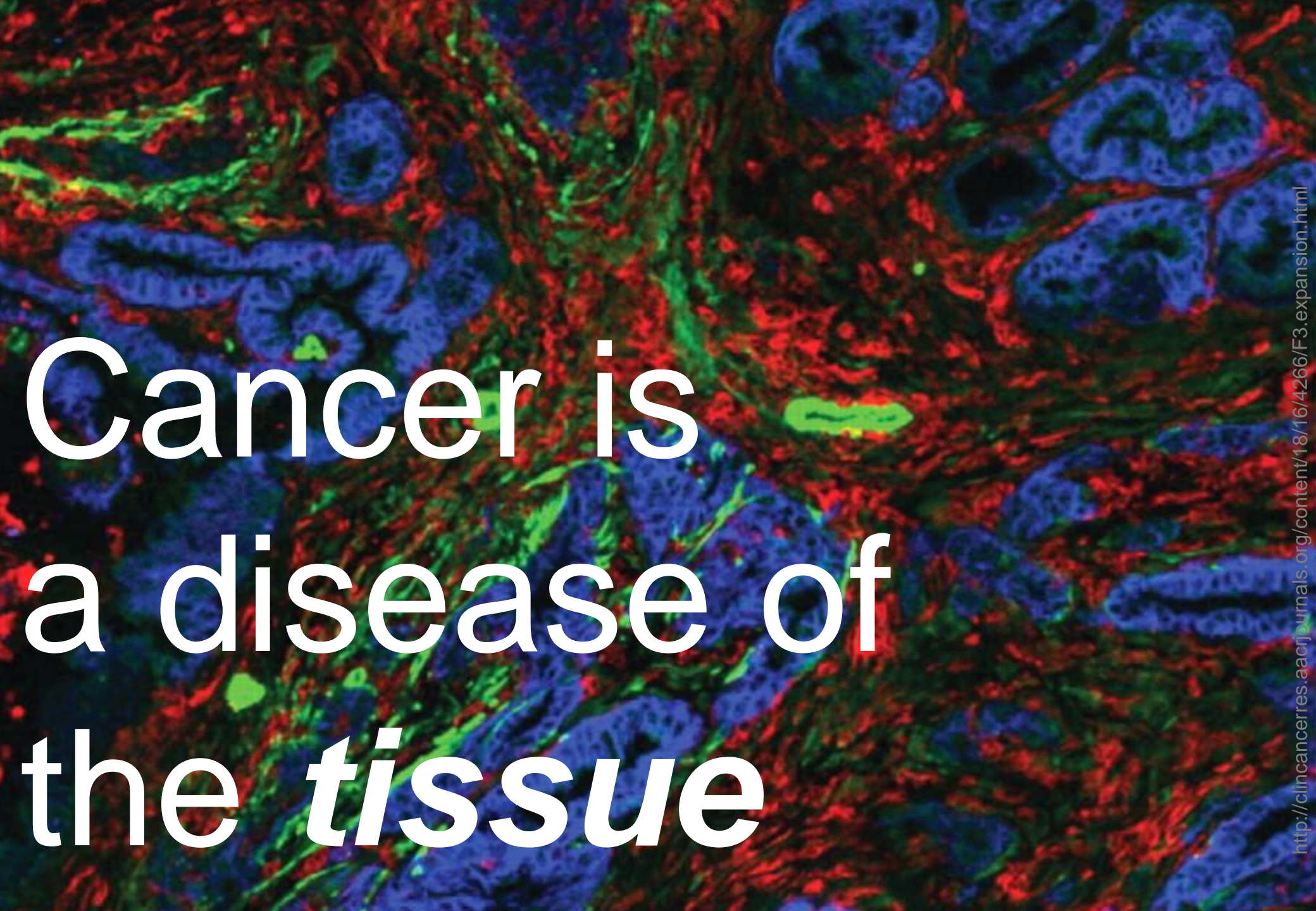
- Multiple metastases
- Good initial response
- Often resistant relapse
- Genomic rearrangements

OV03/04 study

- 17 patients
- 3-30 samples per patient
- Biopsy, surgery and relapse
- Pre- and post-chemotherapy

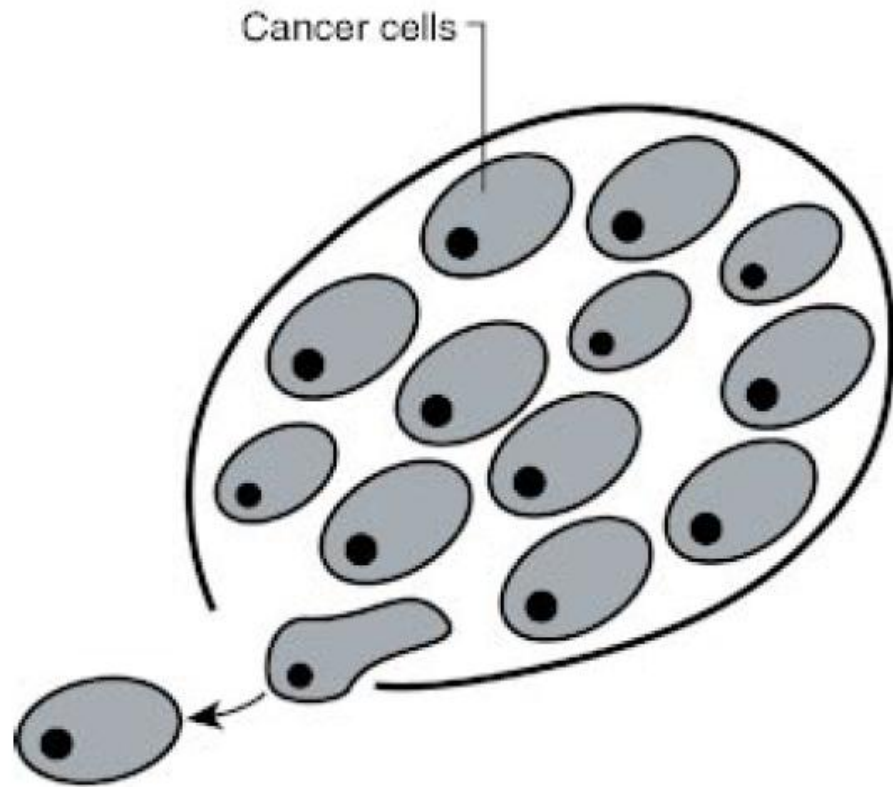


Cancer is
a disease of
the *genome*

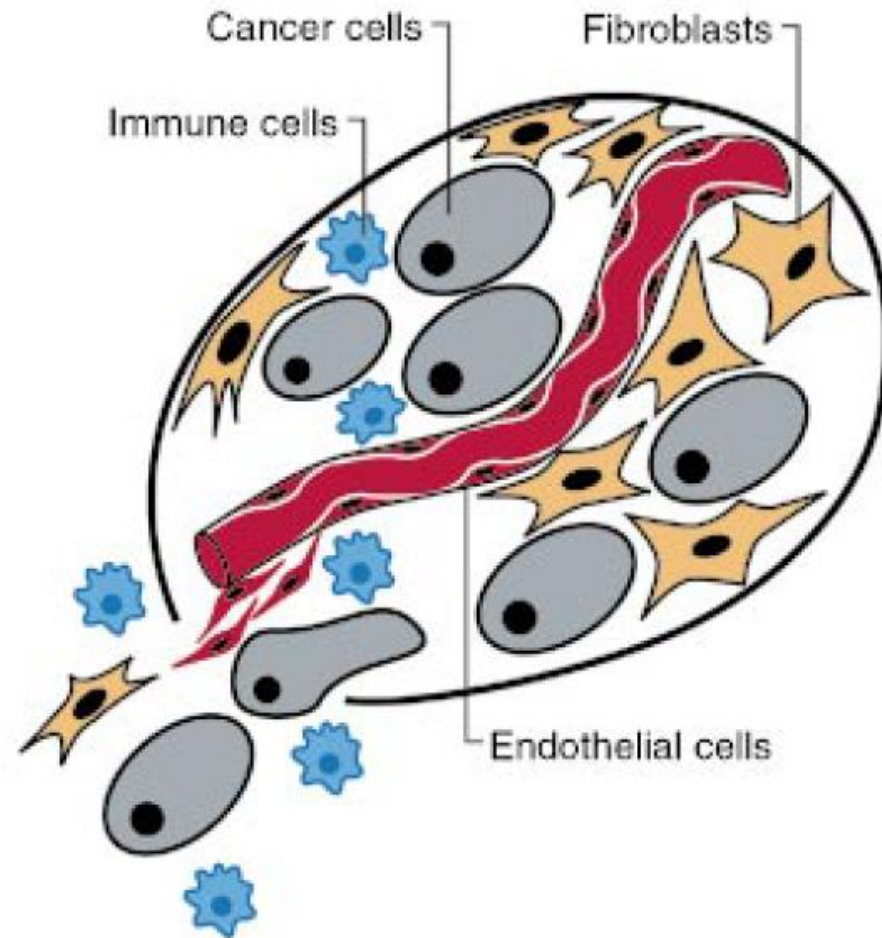
A fluorescence microscopy image of tissue, likely showing cellular structures. The image is composed of three color channels: blue, red, and green. The blue channel highlights nuclei, the red channel highlights a specific protein or marker, and the green channel highlights another. The overall appearance is a complex, textured pattern of these colors, representing the cellular architecture and the localization of the markers.

Cancer is
a disease of
the *tissue*

The Reductionist View

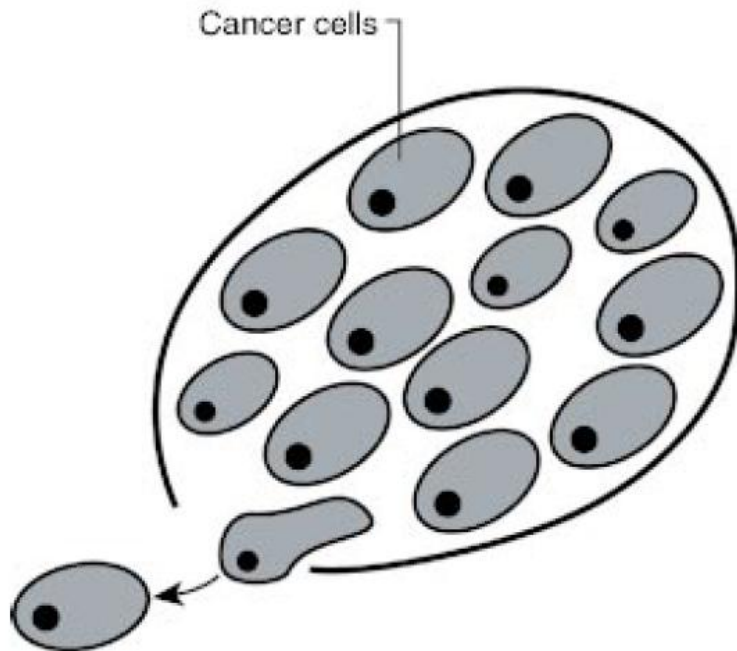


A Heterotypic Cell Biology

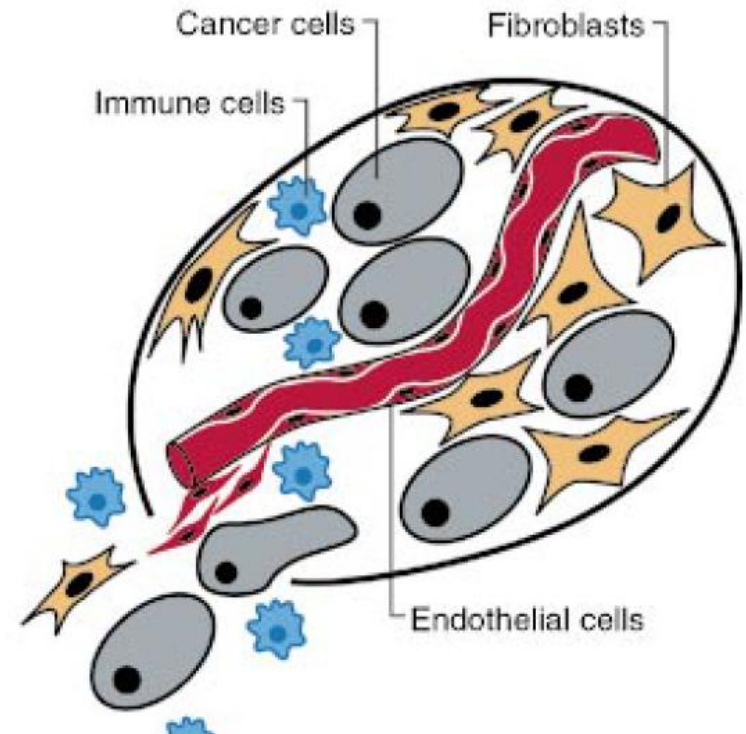


Comprehensive portraits of cancer

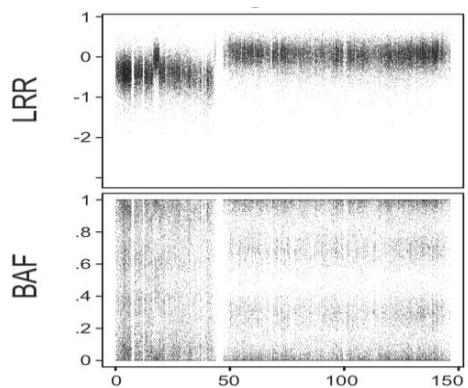
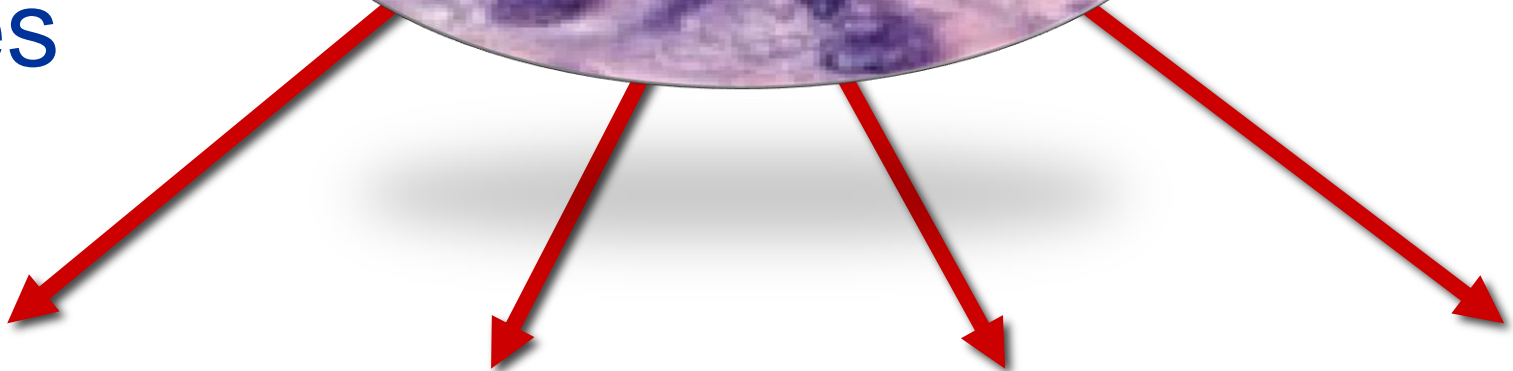
Genomics



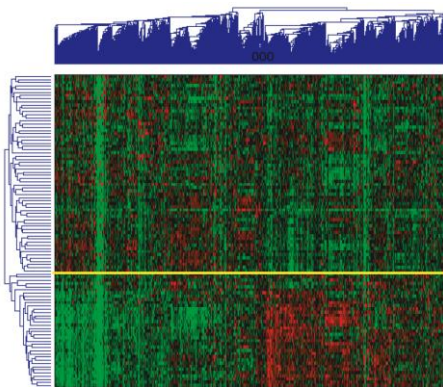
Tissue



Tumors are complex tissues

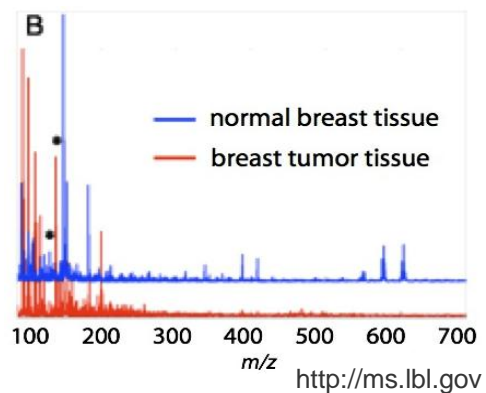


DNA

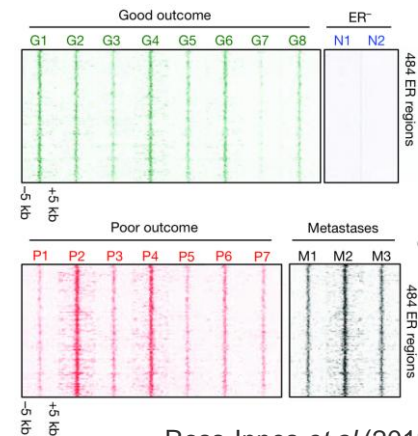


Van't Veer *et al* (2002)

RNA

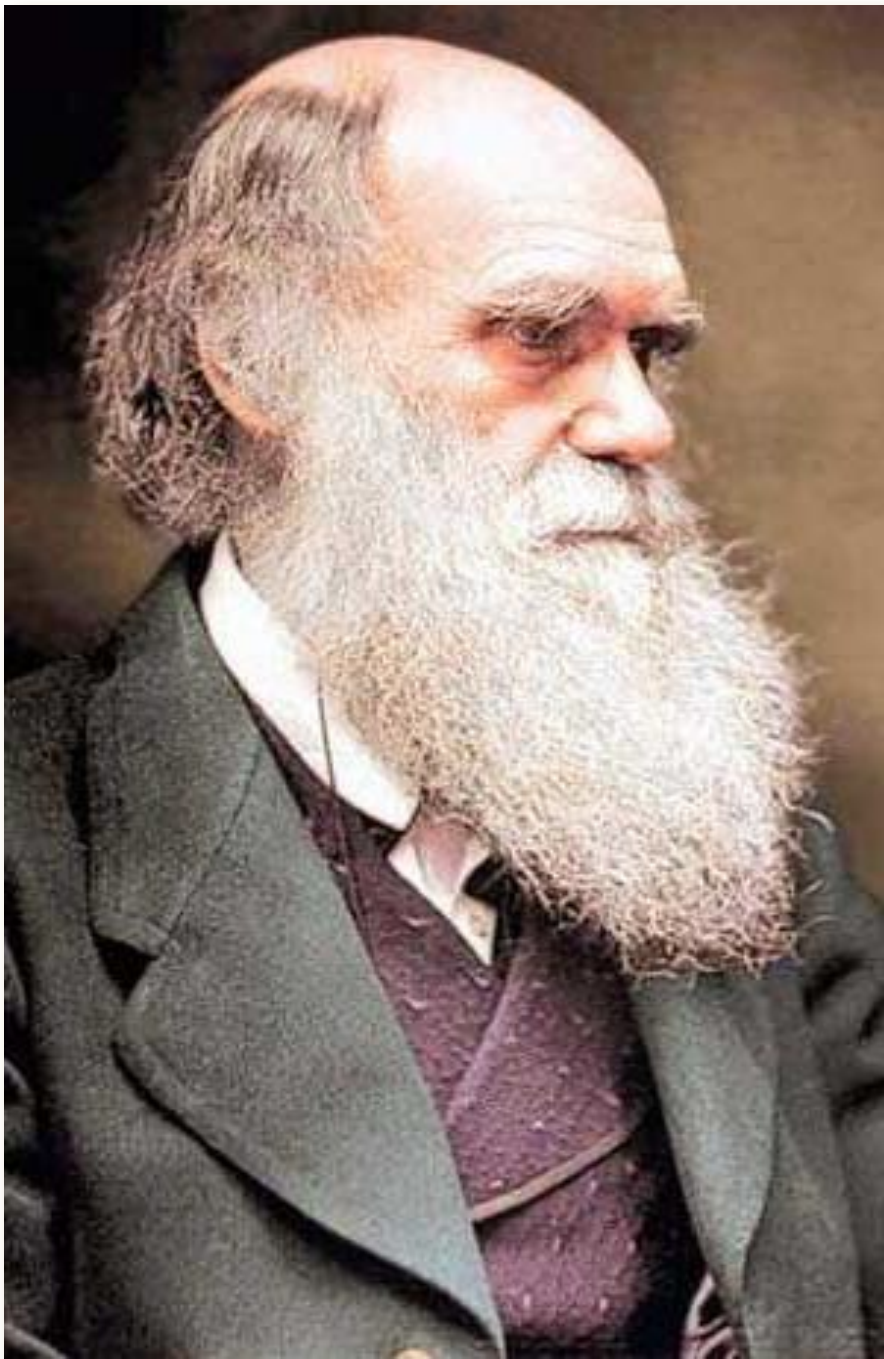


Protein

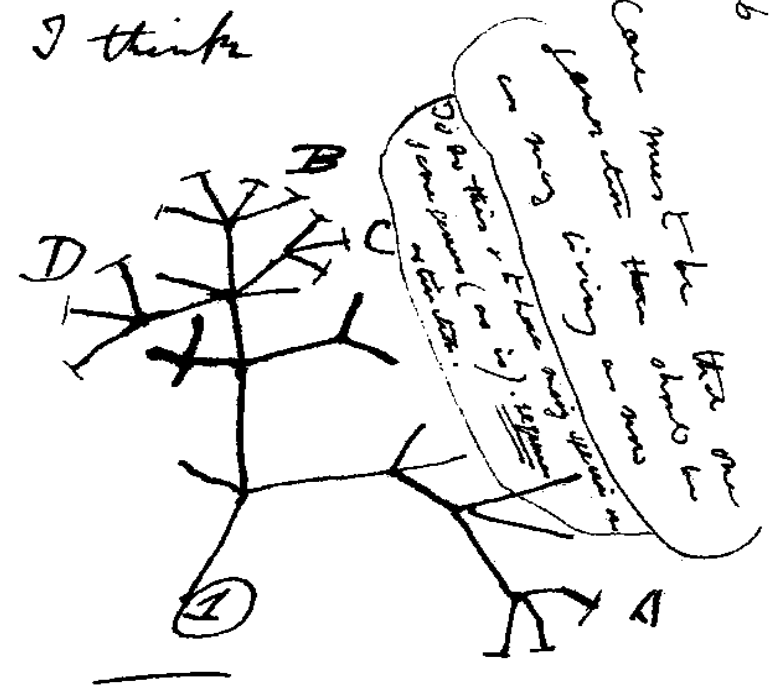


Ross-Innes *et al* (2012)

ChIP



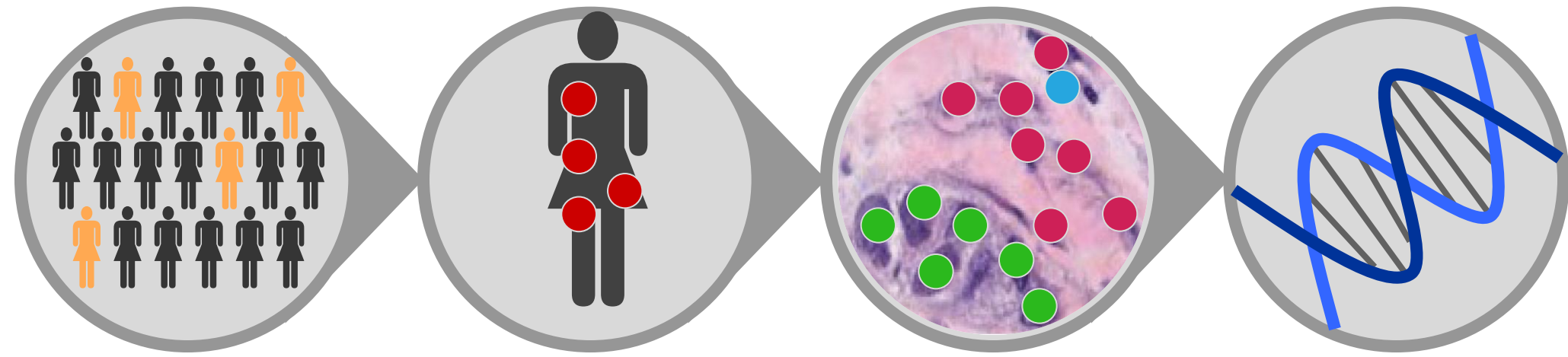
I think



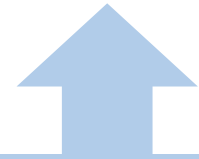
There between A & B. various
 sort of relation. C + B. The
 finest gradation, B + D
 rather greater distinction
 than former would be
 formed. - binary relation



Intra-tumor heterogeneity



Yuan *et al*, Science Trans Med 2012

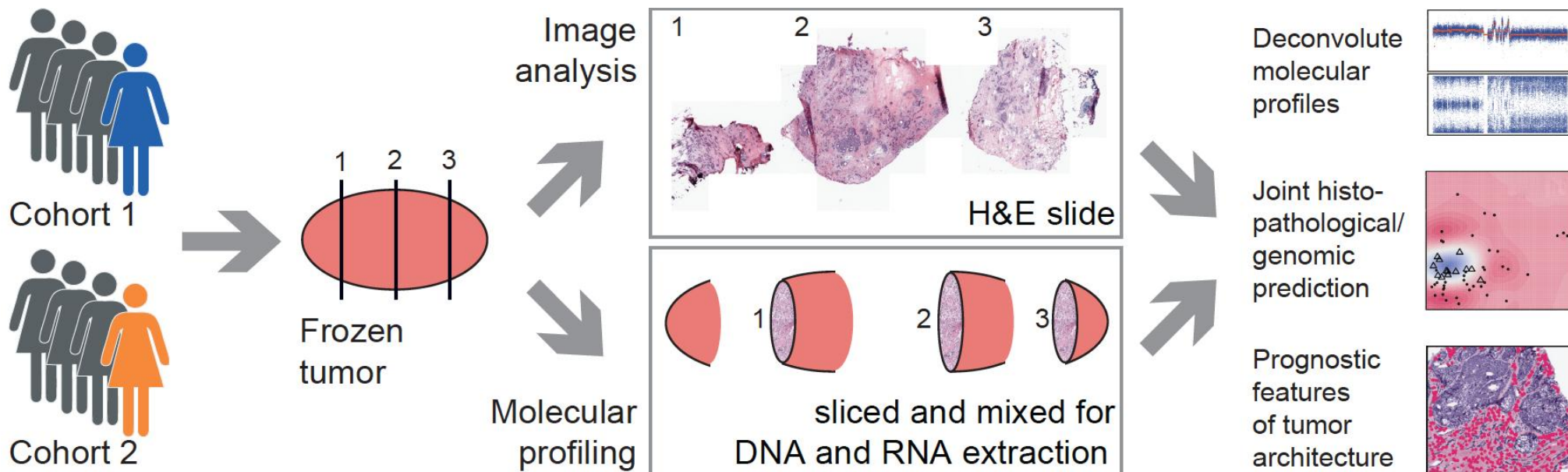


Quantitative image analysis of cellular heterogeneity complements genomics

CANCER

Quantitative Image Analysis of Cellular Heterogeneity in Breast Tumors Complements Genomic Profiling

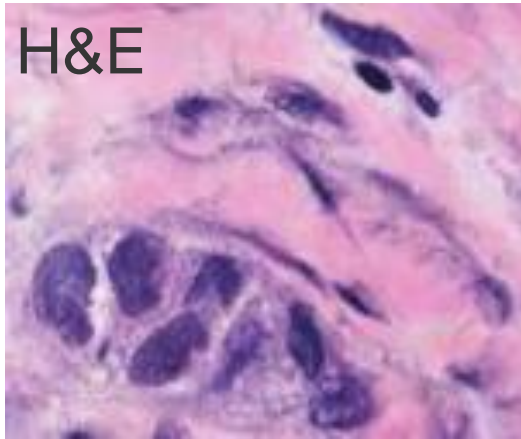
Yinyin Yuan,^{1,2,*†} Henrik Failmezger,^{3,4‡} Oscar M. Rueda,^{1,2‡} H. Raza Ali,^{1,2‡} Stefan Gräf,^{1,2§} Suet-Feung Chin,^{1,2} Roland F. Schwarz,^{1,2} Christina Curtis,⁵ Mark J. Dunning,¹ Helen Bardwell,¹ Nicola Johnson,⁶ Sarah Doyle,⁶ Gulisa Turashvili,^{7,8} Elena Provenzano,⁹ Sam Aparicio,^{7,8} Carlos Caldas,^{1,2,9,10} Florian Markowetz^{1,2,*}



Automated image analysis



Yinyin Yuan

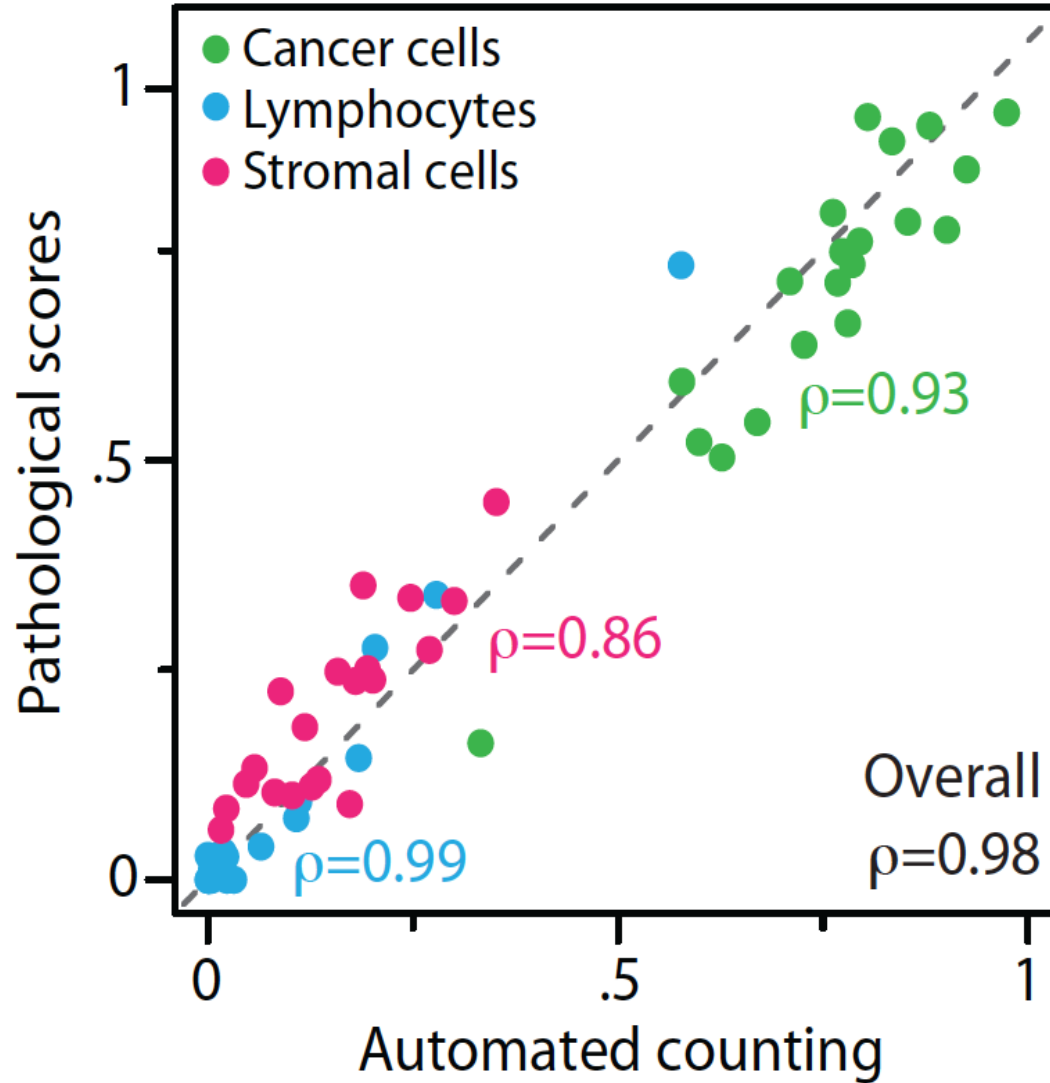




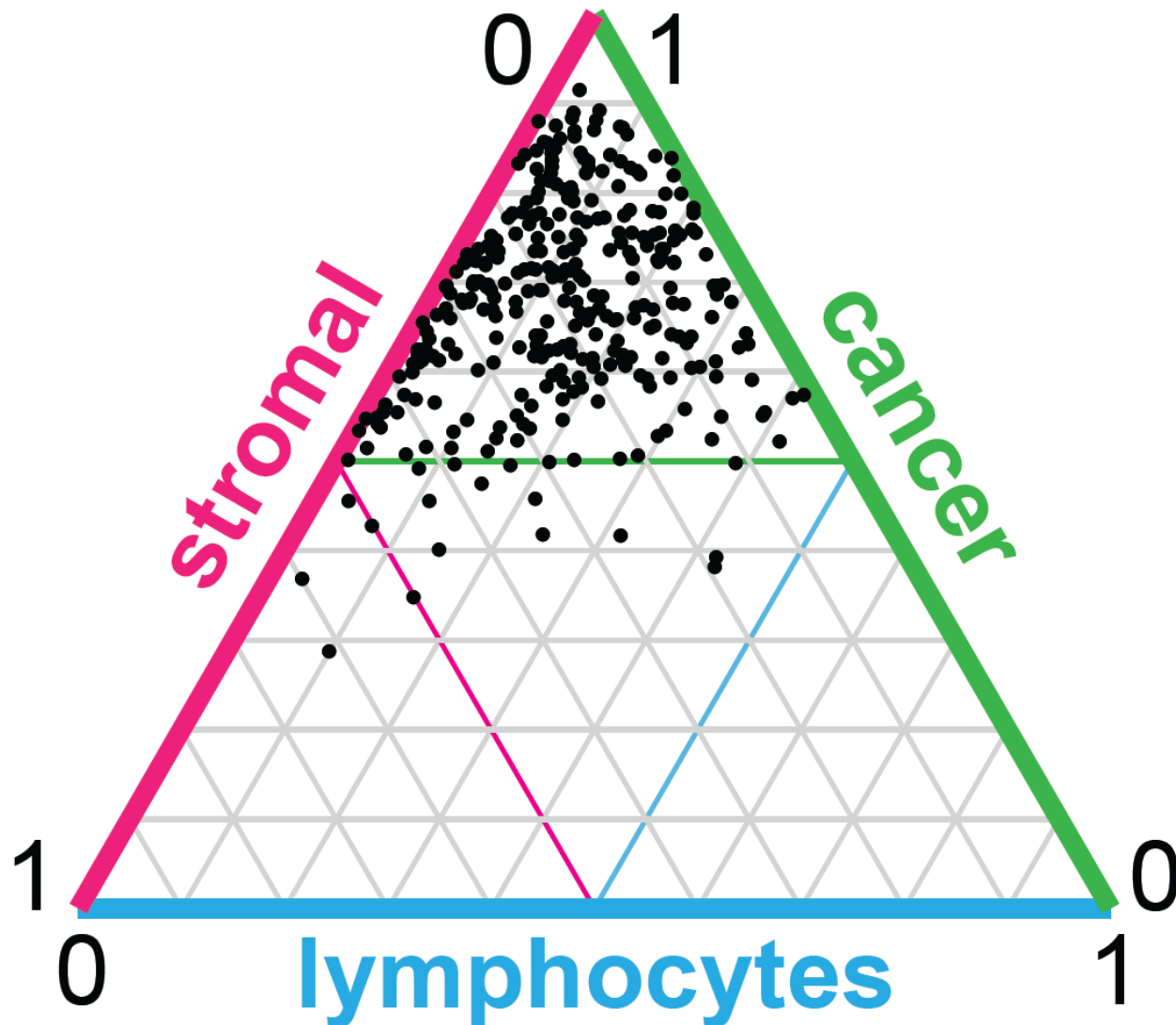
CRImage

The image displays a histological section of tissue, likely stained with hematoxylin and eosin (H&E). The nuclei are stained dark purple, while the cytoplasm and extracellular matrix are pink. The nuclei are segmented and numbered from 1 to 275, indicating the output of a cell detection algorithm. The numbers are placed near the corresponding nuclei, and the segmentation is shown as thin green outlines around the nuclei.

Man vs Machine

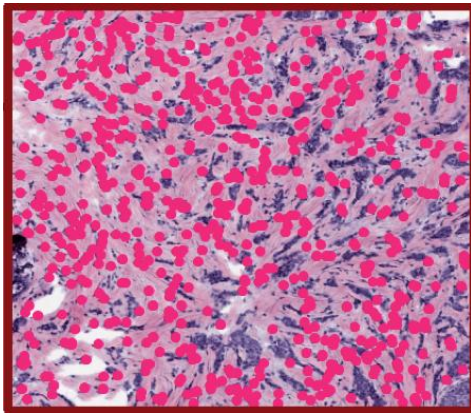


Quantitative analysis of tumour composition

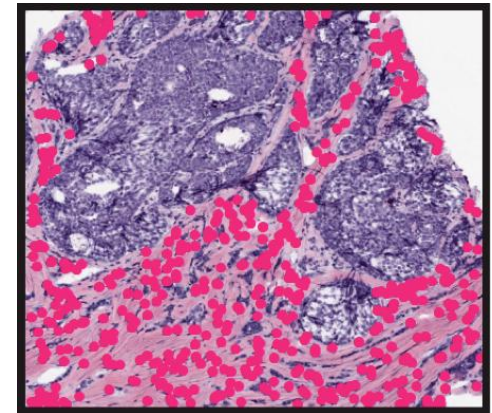


Spatial features of tissue organisation

Uniform

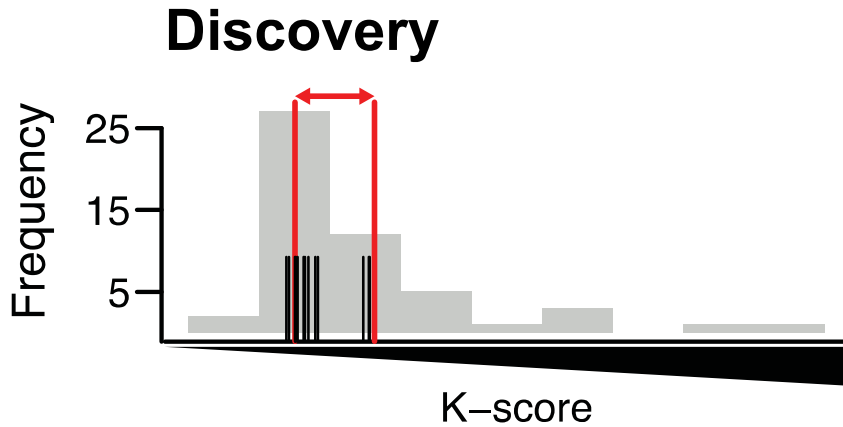


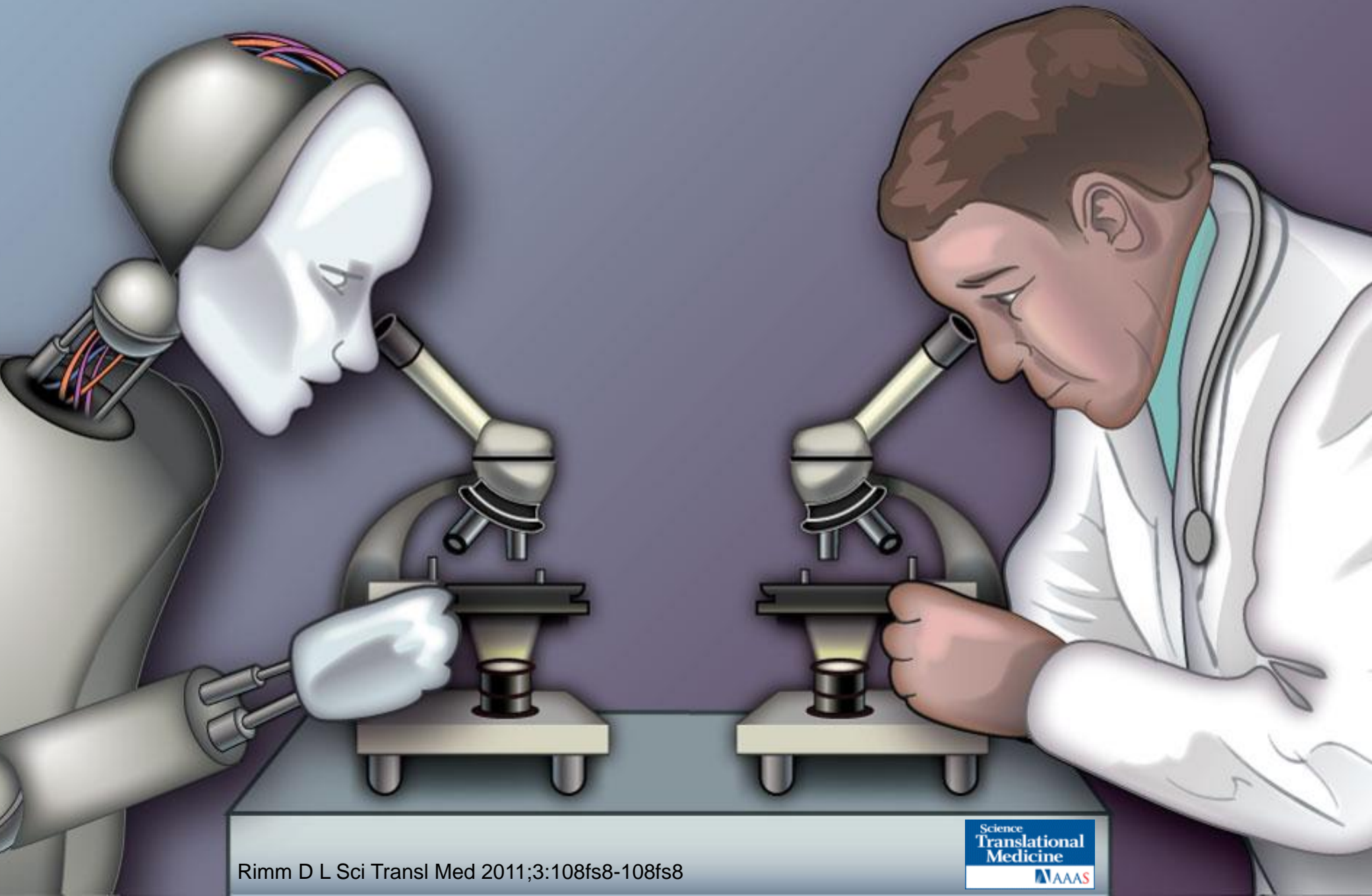
Clustered



Spatial statistics
(K-score)

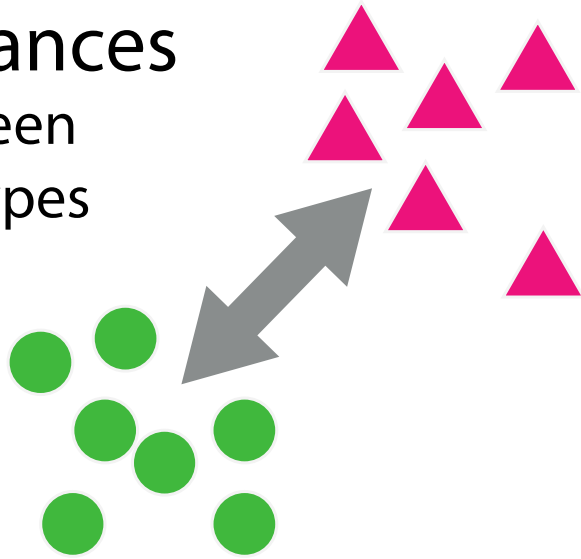
Spatial features of tissue organisation





Spatial features of tumour tissue

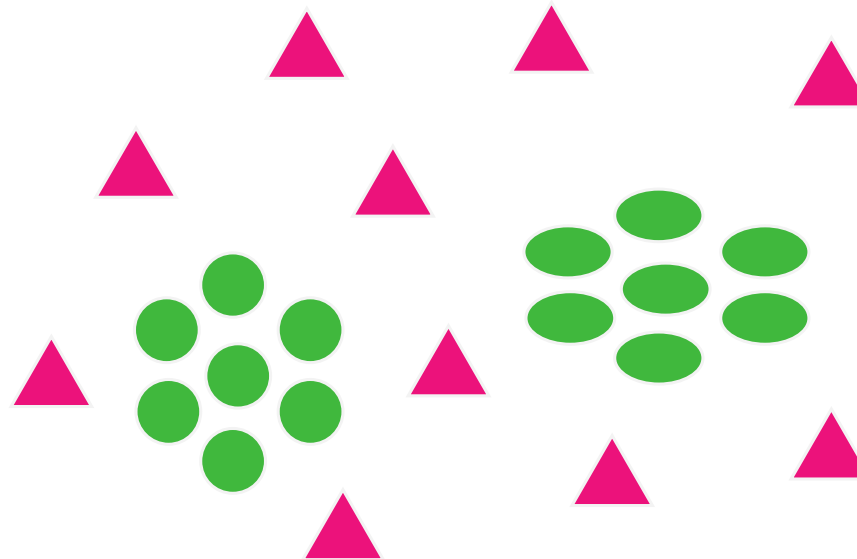
Distances
between
cell types



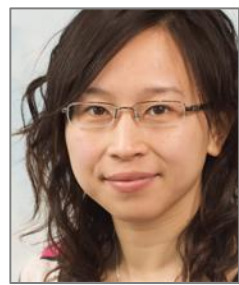
Co-location
e.g. are tumour
cells surrounded
by immune cells



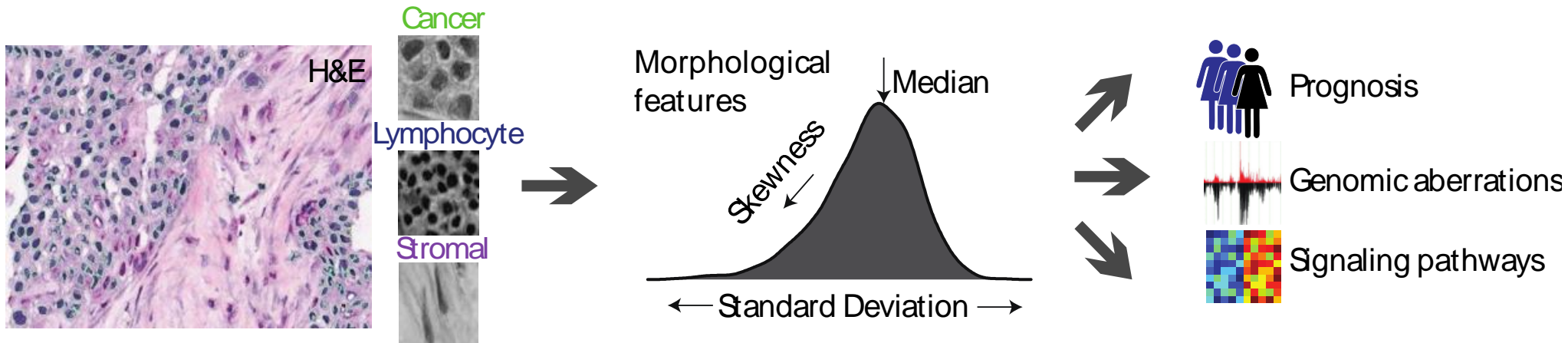
Distribution
of cells and
morphologies
across tumour



Morphological heterogeneity



Yinyin Yuan



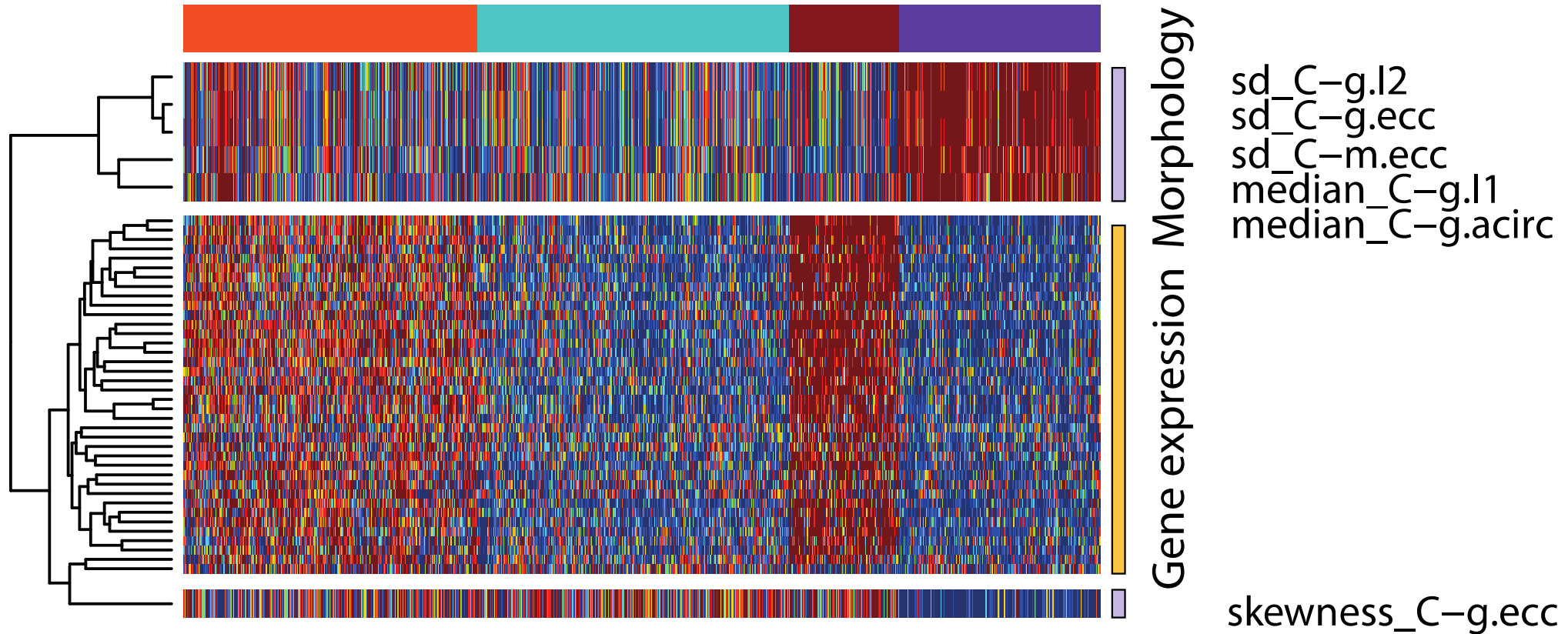
Morphological features

1. Fraction of pixels outside of the circle with radius r
2. Shape factor,
3. 1st Hus translation/scale/rotation invariant moment
4. Eccentricity calculated based on geometric information
5. Eccentricity calculated based on image moments

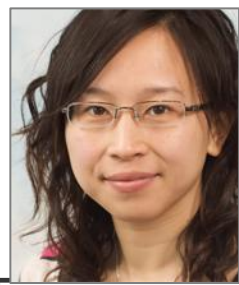
Morpho-genomic subtypes



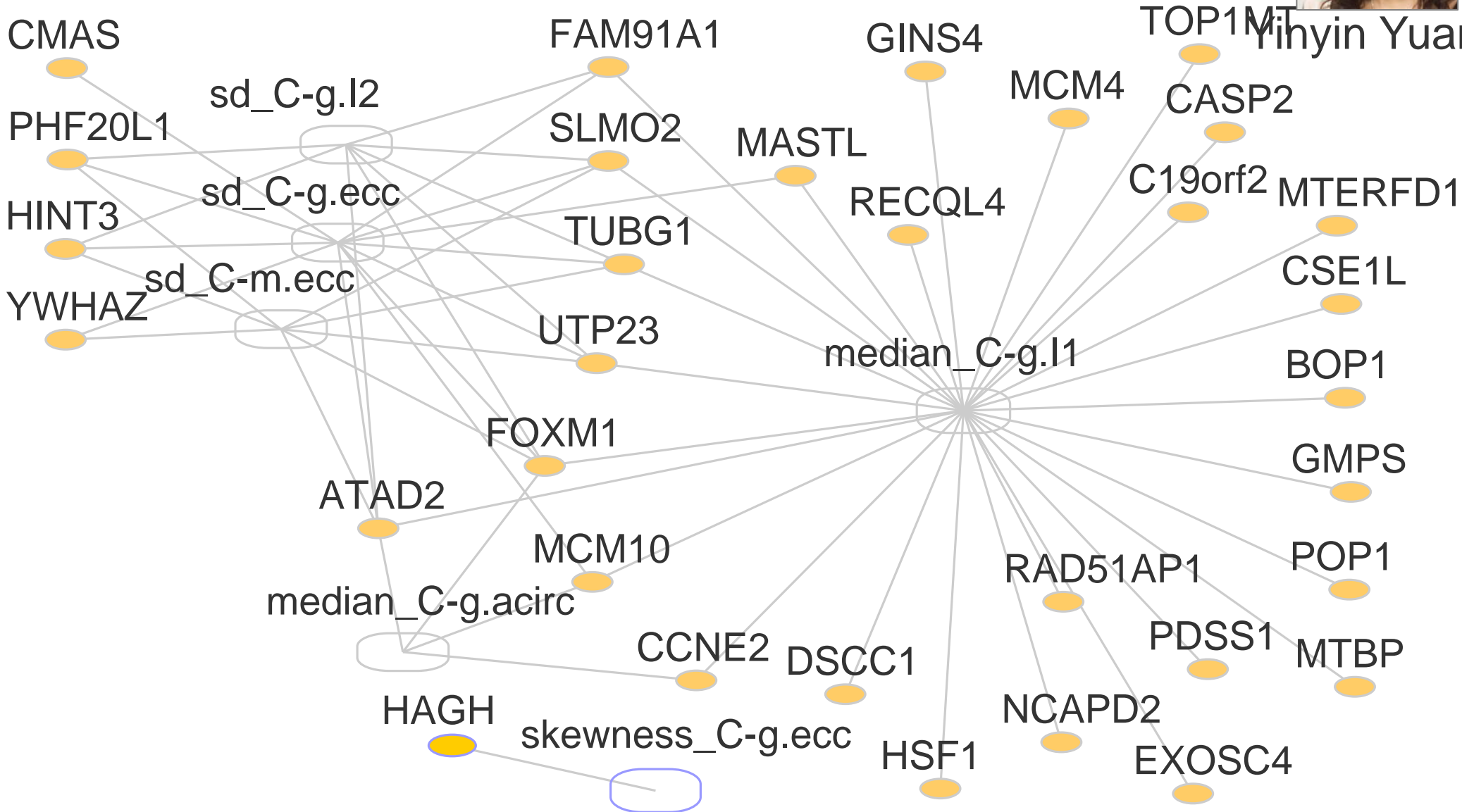
Yinyin Yuan



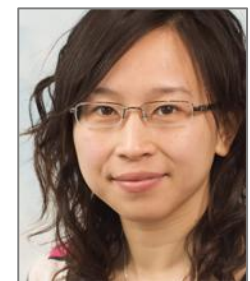
Morphology <-> Gene expression



Yinyin Yuan



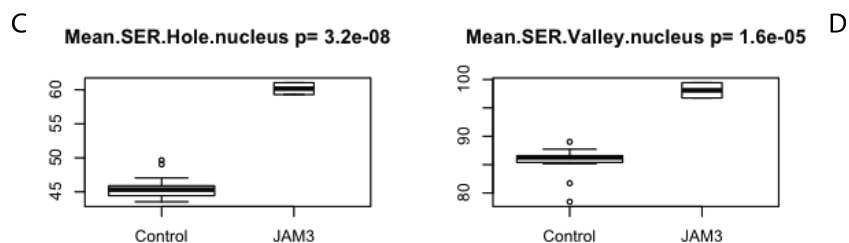
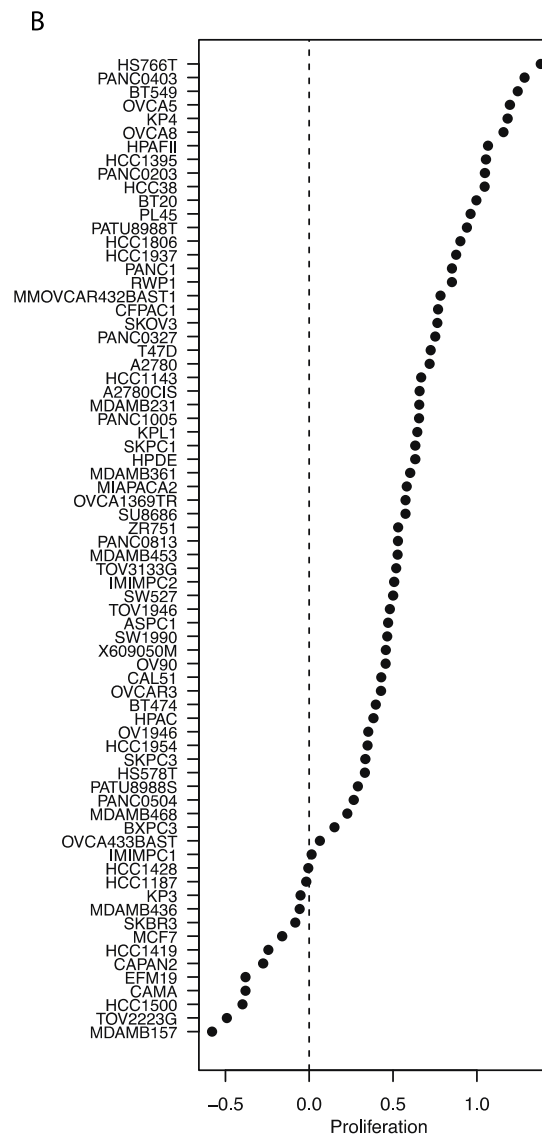
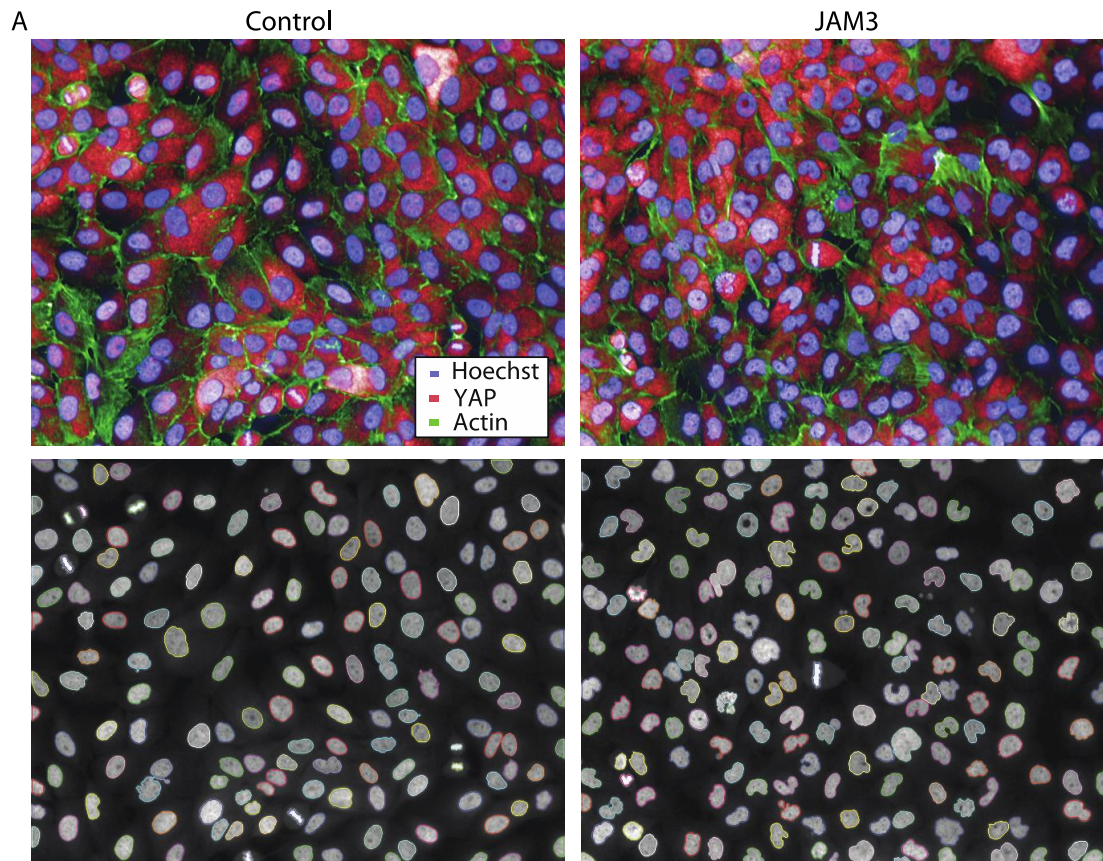
JAM3 – driver of cell morphology



Yinyin Yuan



Chris Bakal





Xin Wang

Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions

Felipe De Sousa E Melo^{1,7}, Xin Wang^{2,7}, Marnix Jansen³, Evelyn Fessler¹, Anne Trinh², Laura P M H de Rooij¹, Joan H de Jong¹, Onno J de Boer³, Ronald van Leersum¹, Maarten F Bijlsma¹, Hans Rodermond¹, Maartje van der Heijden^{1,4}, Carel J M van Noesel³, Jurriaan B Tuynman⁵, Evelien Dekker⁶, Florian Markowitz², Jan Paul Medema^{1,7} & Louis Vermeulen^{1,4,7}

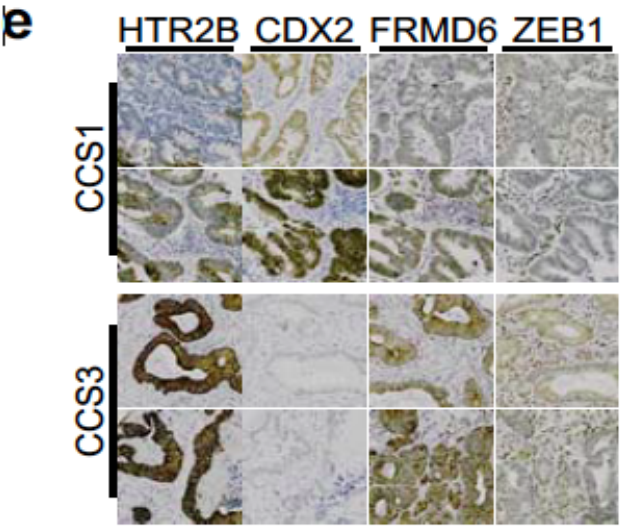
Colon cancer is a clinically diverse disease. This heterogeneity makes it difficult to determine which patients will benefit most from adjuvant therapy and impedes the development of new targeted agents¹. More insight into the biological diversity of colon cancers, especially in relation to clinical features, is therefore needed. We demonstrate, using an unsupervised classification strategy involving over 1,100 individuals with colon cancer, that three main molecularly distinct subtypes can be recognized. Two subtypes have been previously identified and are well characterized (chromosomal-unstable

24 patients with CCS3 (49, 24, and 27%, respectively) (Fig. 1b). We validated this classifier in six independent data sets and found comparable proportions of patients being assigned to each subtype (Fig. 1c and Supplementary Table 3). We could also classify colorectal cancer cell lines into the three subtypes (Fig. 1d)⁴. Moreover, the subtypes were generally maintained upon xenografting of cell lines and primary tumors (Supplementary Fig. 3d–g)^{5,6}, suggesting they reflect persistent genetic or epigenetic features of tumor cells rather than differences in stroma or immune infiltrates, which have been used previously to stratify patients^{7,8}. To further characterize the subtypes, we determined the

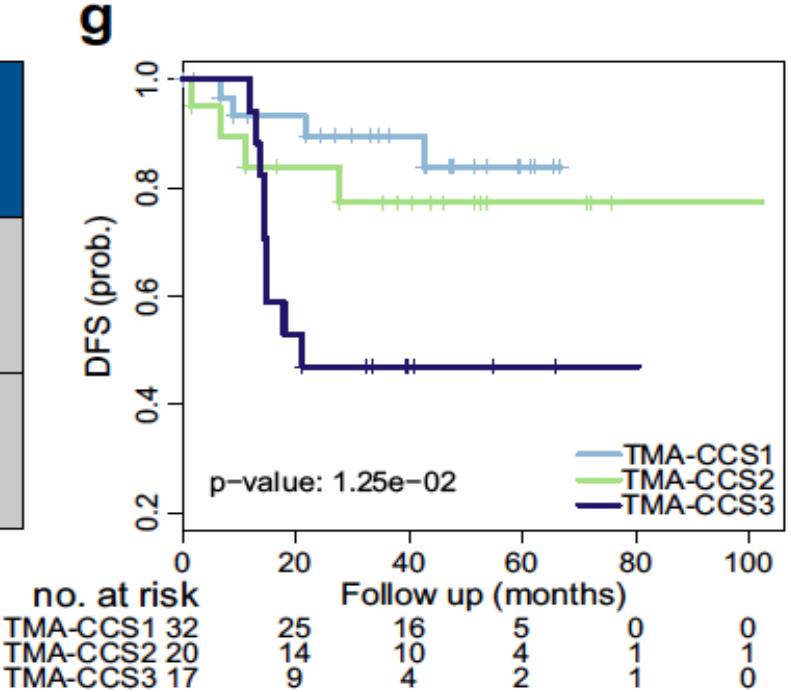
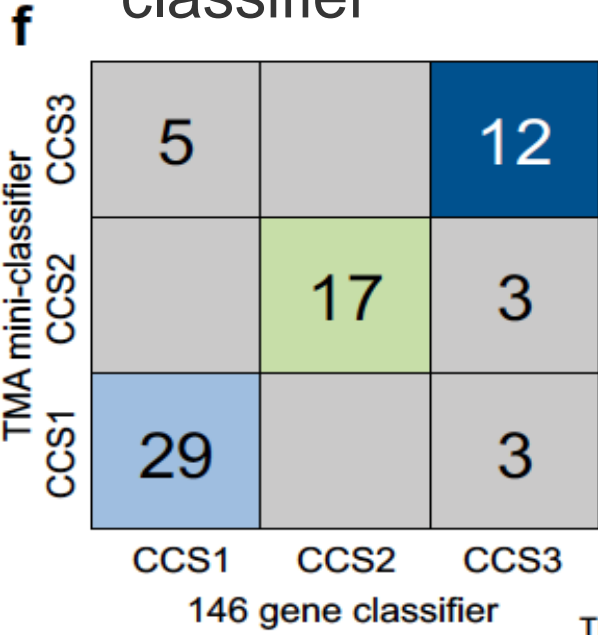


Anne Trinh

Stainings in tissue microarrays



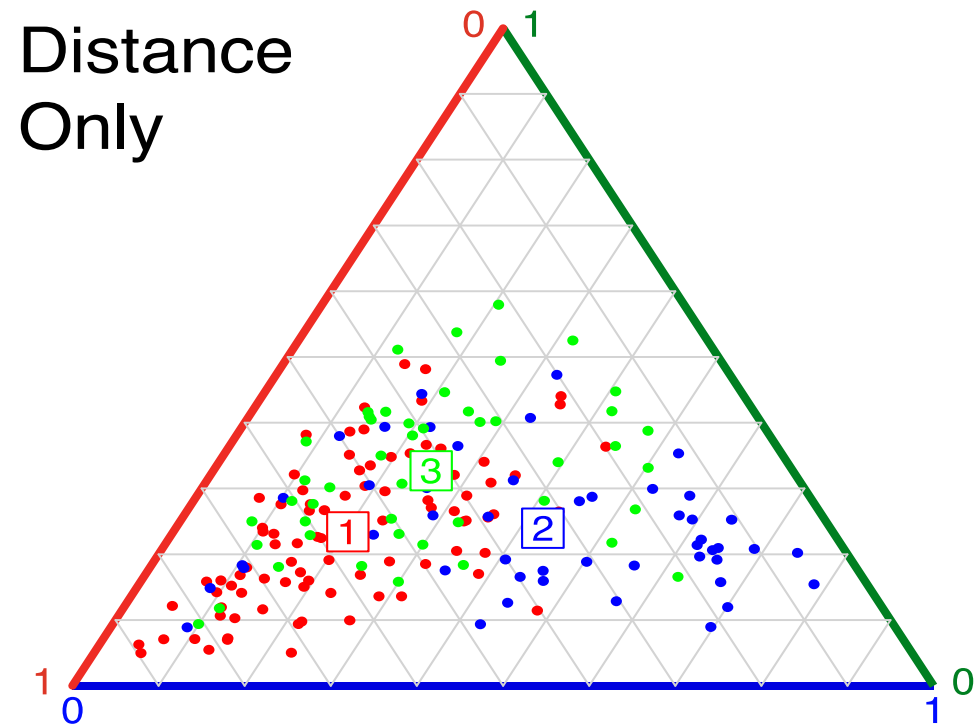
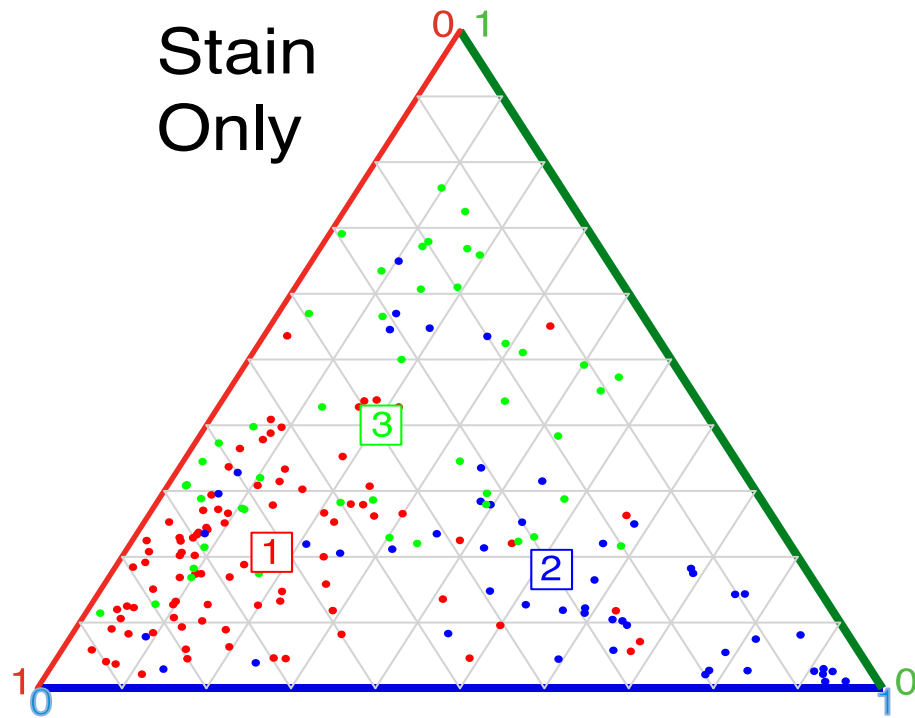
Comparison to gene expression classifier



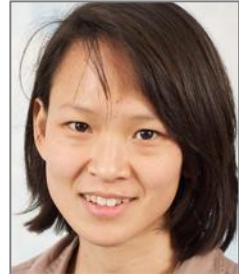
Spatial features are predictive



Anne Trinh



ASUMT: A Still Unnamed MATLAB Toolbox



Anne Trinh

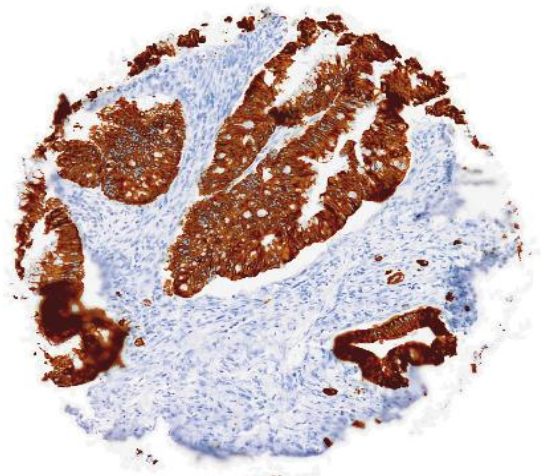
Image Properties

Image Properties:		Regional Information:		Cellular Information:	
Image Size	1543 1543	TMA Area	986000	Stained Cells	306572
Percentage Image Area	41	Epithelial Area	NA	Unstained Cells	NA
Percentage Brown	31	Stromal Area	NA	Epithelial Density	NA
Average Brown Intensity	7	Background Area	NA	Stromal Density	NA
Cell Stain Thresh (Br)	NA	Ep/Str Ratio	NA	Total Cell Count	NA

View Results

Brown Area [Dropdown] [Zoom On] [Clear All]

Plotting Area



Brown Stained Regions determined by Otsu Thresholding

Cellularity Counter. Cool Logo goes here

Feedback Box:
* Instructions
* Errors

Load & Preprocess:
* Load TMA
* Detect Outline
* Find Brown Area

Regional Segmentation:
* Label Dataset (file or directly)
* KMeans-MRF (grayscale & RGB)

Cell Count:
* Background Thresholding
* H-minima Watershed
* SVM Cell Classification

Save:
* mat file
* csv & images
* classifiers

Reg_Segmenter
Cellularity Counter. Cool Logo goes here

Image Properties

Image Properties:		Regional Information:		Cellular Information:	
Image Size	1543 1543	TMA Area	986000	Stained Cells	306572
Percentage Image Area	41	Epithelial Area	NA	Unstained Cells	NA
Percentage Brown	31	Stromal Area	NA	Epithelial Density	NA
Average Brown Intensity	7	Background Area	NA	Stromal Density	NA
Cell Stain Thresh (Br)	NA	Ep/Str Ratio	NA	Total Cell Count	NA

Step 1: Load & Preprocess
Insert image from MATLAB workspace or file:

Select Thresholds:
Entropy: Size: [Preprocess] [Remove Region]

Step 2: Label Regions
 Labelled Image Label Data

Brush Diameter:
or Label Image:

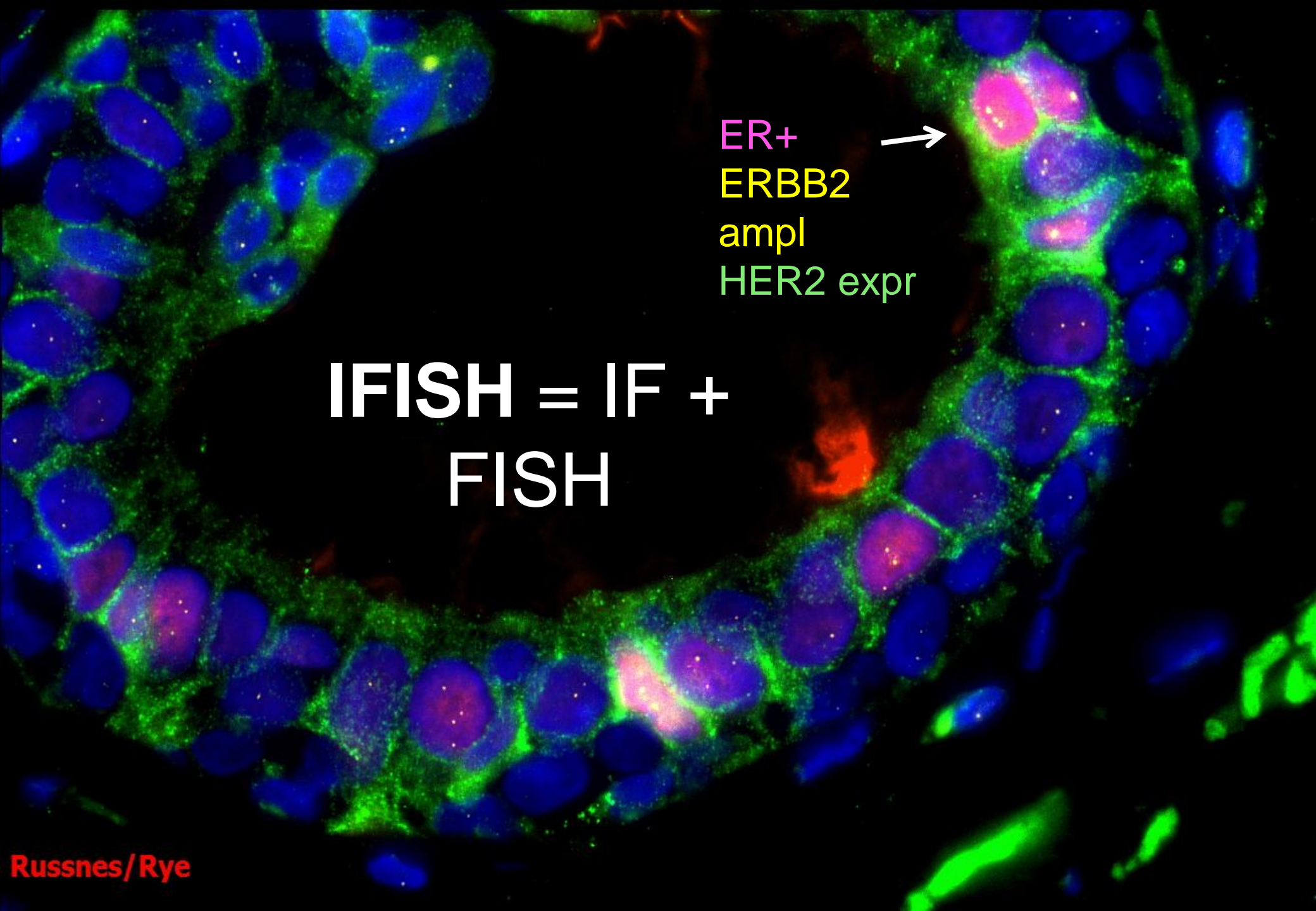
Step 3: Region Detection
Grid Size:
Beta:
Number of Iterations:
Convergence Value:
 Shift - merge
 RGB independent [Run KMeans MRF]

Step 4: Cellular Detection
Set Background Threshold:
Optimal cell constraints:
Min Size: Max Size:
Solidity: MinInt:
Eccentricity: ArDev:
 nuclear cytoplasmic [Start Watershed]

Step 5: Cell Classification
Insert Classifier:
or find example cells:
Insert Label: [Update]
SVC parameters:
Cost:
Gamma: [Start Classification]

Step 6: Save to File

 mat file images & csv [Save!]
Regional Classifier: Cell Classifier:

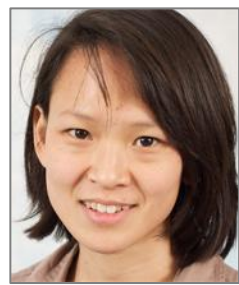


ER+
ERBB2
ampl
HER2 expr



**IFISH = IF +
FISH**

Go IFISH: a toolbox for semi-automated detection of nuclei, membrane and spots

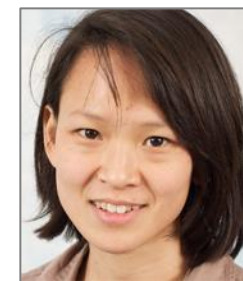


Anne Trinh

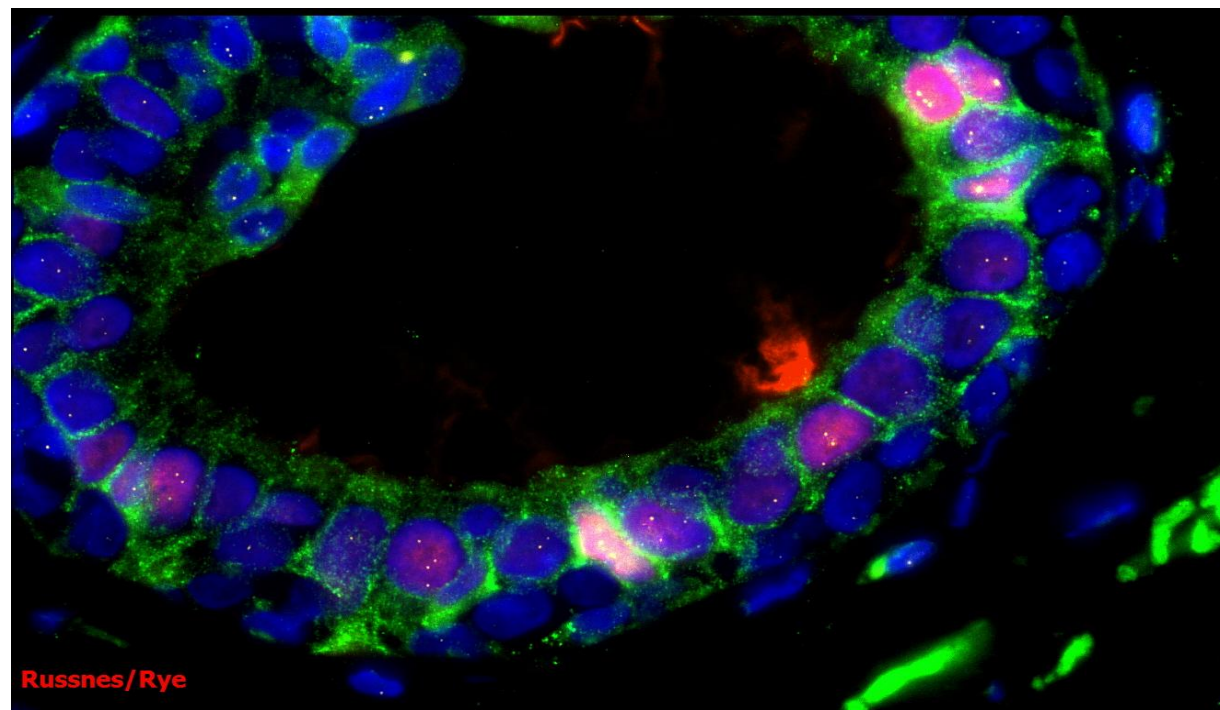
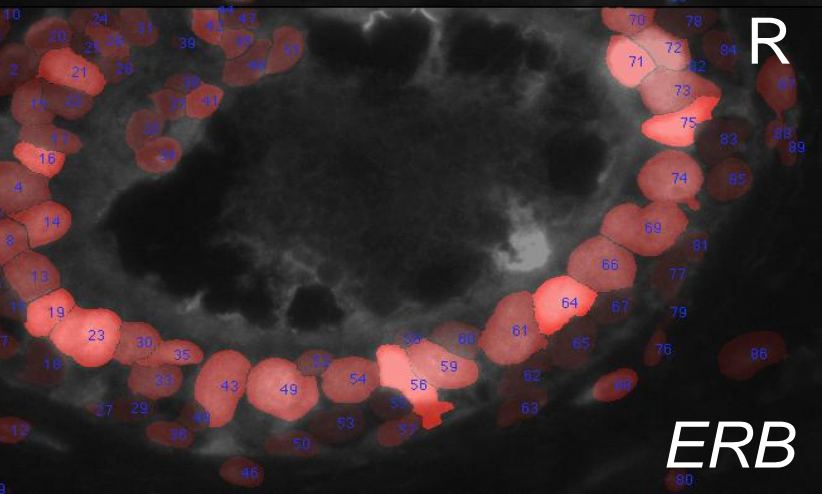
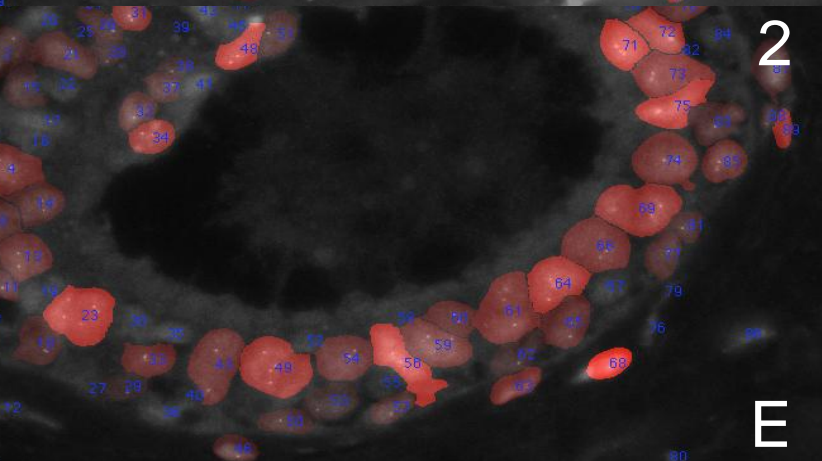
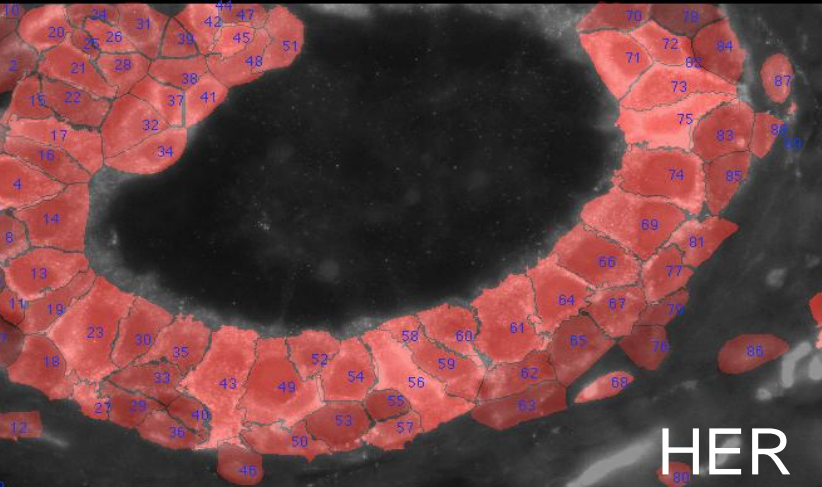
The screenshot displays the Go IFISH software interface, titled "CellSegmenter". The main window is divided into several sections:

- Top Left:** "Go IFISH" title, "User Edited Segmentation" dropdown, "Zoom Off" button, and "Edge View" (checked) and "Number Cells" (unchecked) checkboxes.
- Top Center:** Five buttons labeled "Stain1" through "Stain5".
- Main View:** A grayscale microscopy image of cells with red outlines indicating segmentation. A blue line is drawn on one of the cells.
- Right Panel (Step 1: Load & Preprocess):** Includes "Insert image from MATLAB workspace or file:" with a folder icon and "PrePro" button. Below are dropdowns for "DAPI", "Membr...", "FISH", "Nuclear", and "FISH". A "Resolution (x Magnification)" field is set to 40.
- Right Panel (Step 2: Manual Fixing Tools):** Contains icons for various tools: scissors, a bottle, a trash can, a brush, a lasso, a selection tool, a circular selection tool, and two arrow icons. A "10" field is present.
- Right Panel (Step 3: DAPI Segmentation):** Includes "Select DAPI background thresh" with a slider, "TLow, THigh" fields set to 60, a "Run" button, "Select watershed method" dropdown set to "-- Select WatershedMethod --", and a "Clean Up Missegmentations" checkbox. Below are icons for different watershed methods and a "Classify" button.
- Right Panel (Step 4: Membrane Segmentation):** Includes parameters: μ (65), λ_1 (1), ϵ (1.5), NIt (100), v (0), λ_2 (1), σ (1.5), and dt (2). A "-Active Contour Smooth..." dropdown and a "Run" button are also present.
- Right Panel (Step 5: FISH Detection):** Includes "Brush" (5), "Intensity" (0.10), "Gamma" (1.15), and a "Run" button.
- Right Panel (Step 4: Save to File):** Includes a "Generate Summary HeatMaps and Labels" button, checkboxes for "Progress File" and "images & csv", and a save icon.

Single cell analysis of stain intensities



Anne Trinh

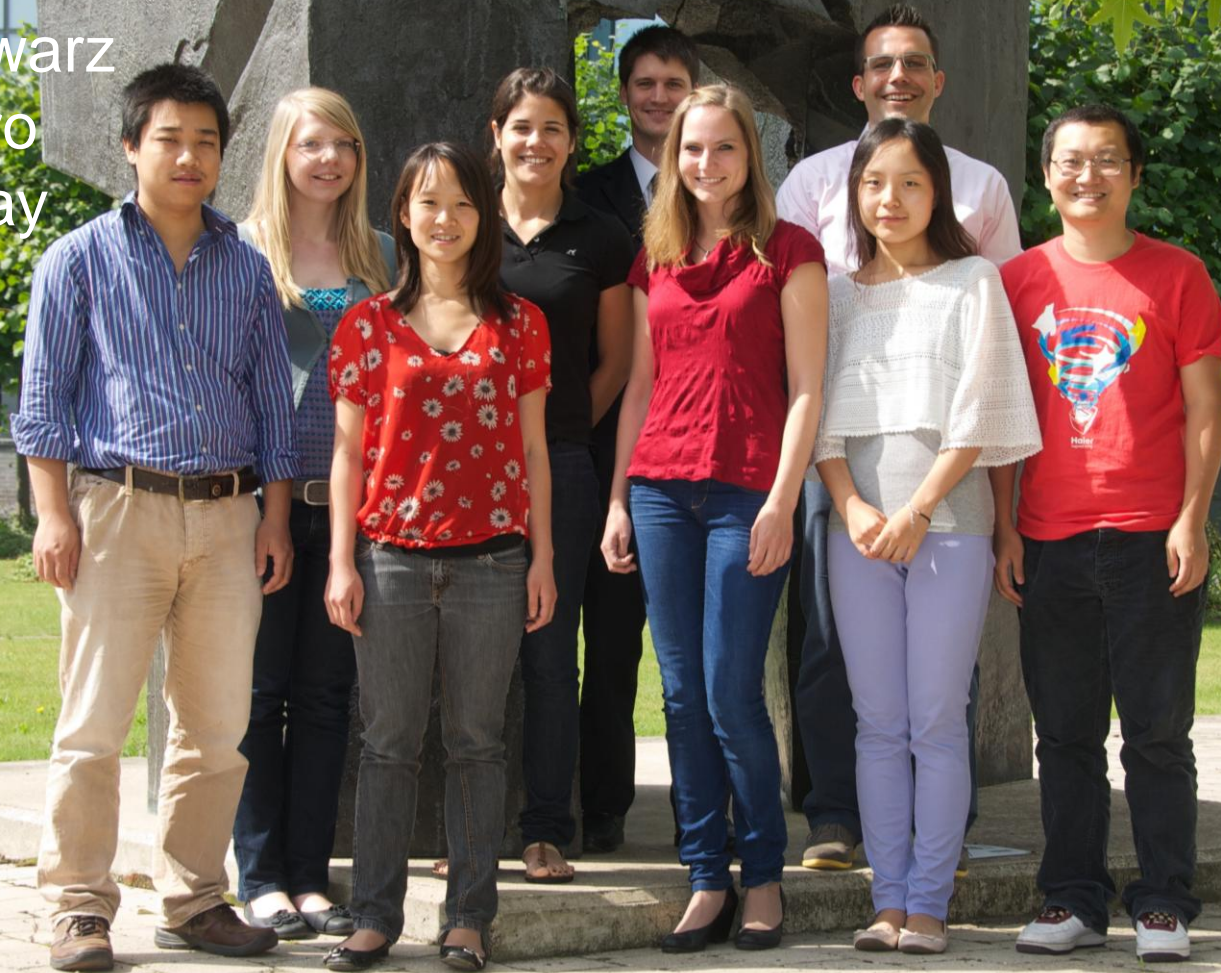


Key collaboration partners

- **Carlos Caldas, Raza Ali, Suet-Feung Chin, Oscar Rueda, Stefan Gräf** @ [University of Cambridge](#)
- **Yinyin Yuan + lab** @ [Institute for Cancer Research](#)
- **Chris Bakal + lab** @ [Institute for Cancer Research](#)
- **JP Medema, Louis Vermeulen** @ [Amsterdam Medical Center](#)
- **Anne-Lise Børresen-Dale, Hege Russnes, Inga** @ [Oslo University](#)

Alumni:
Xin Wang
Yinyin Yuan
Roland Schwarz
Mauro Castro
Gökmen Altay

the team



UNIVERSITY OF
CAMBRIDGE



wellcome trust
Hutchison Whampoa



CANCER
RESEARCH
UK

Paul Pharoah

*Strangeways
Laboratories,
Cambridge*

- Genetic epidemiology



FMlab



Carlos Caldas

- Breast Cancer Functional Genomics
- Cambridge Breast Cancer Research Unit



Jason Carroll

- ER biology
- ChIP-seq in tumors



Doug Fearon

- Tumor immunology
- Tumor microenvironment

Stephen Friend

Sage Bionetworks



Dissecting cancer heterogeneity

Thank you

!



Florian Markowitz
CRUK Cambridge Institute
www.markowitzlab.org



Systems Genetics = genome × phenotypes × conditions

