Profiling human fallopian tubes with single-cell RNA sequencing

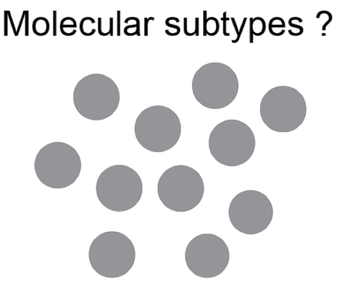
Zhiyuan Hu
MRC Weatherall Institute of Molecular Medicine

29 March 2019

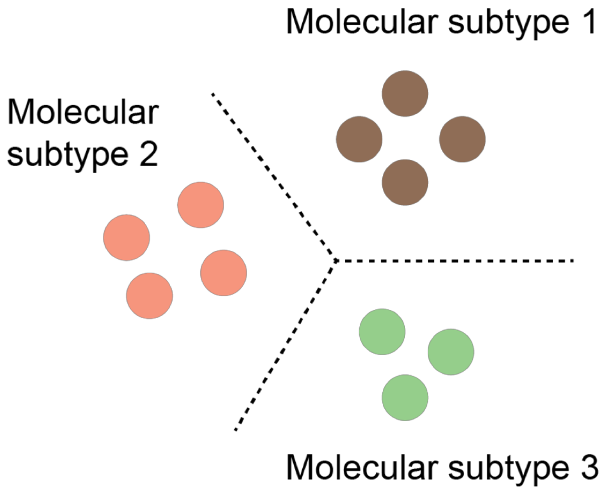
High-grade serous ovarian cancer (HGSOC) is the most aggressive subtype without robust molecular classification

High-grade serous ovarian cancer (HGSOC)

- ~80% ovarian cancer cases
- Five-year survival ~30%



Molecular subtypes ?



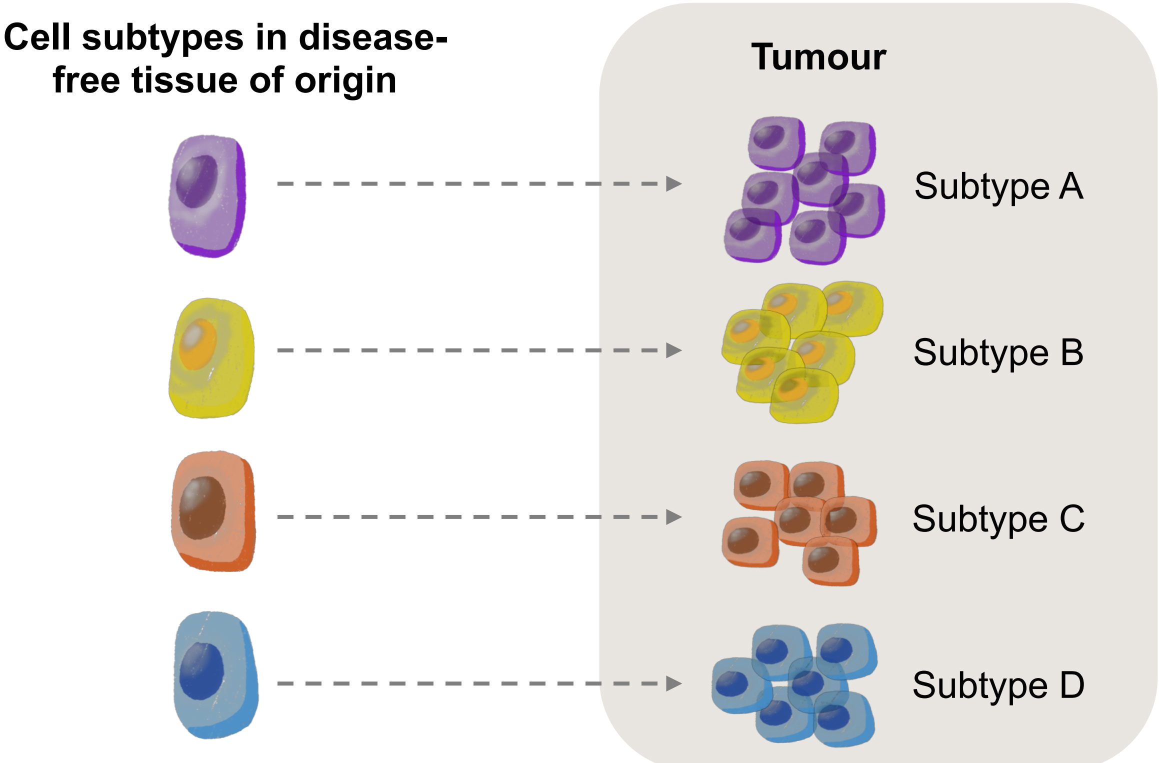
Molecular subtype 1

Molecular subtype 2

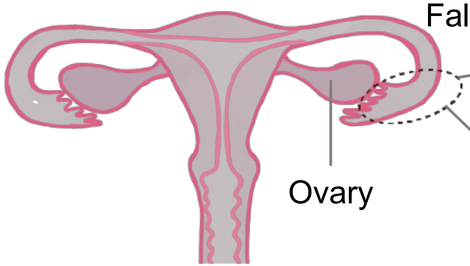
Molecular subtype 3

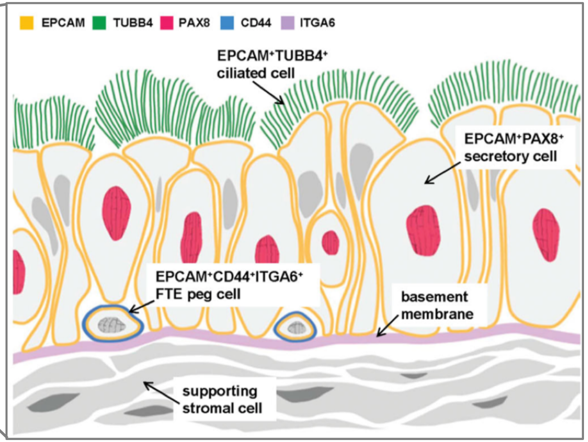
- Prognostic prediction
- Targeted therapy

Cell-of-origin may provide evidence for molecular subtyping



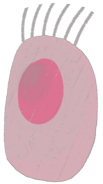
Knowledge of fallopian tube epithelium, the putative origin of HGSOC, was limited





Daniel Paik et al., *Stem Cells*, 2012

Secretory PAX8+



Ciliated TUBB4+

- Questions:
  - Are there additional cellular subtypes in the epithelium?
  - What is their connection to ovarian cancer?

Workflow: identification of novel subtypes in fallopian tube epithelium by single-cell RNA-seq

Fallopian tube samples from ten donors

Single-cell suspension

Flow cytometry to isolate single live cells

Single-cell RNA-seq (Smart-Seq2)

Data analysis

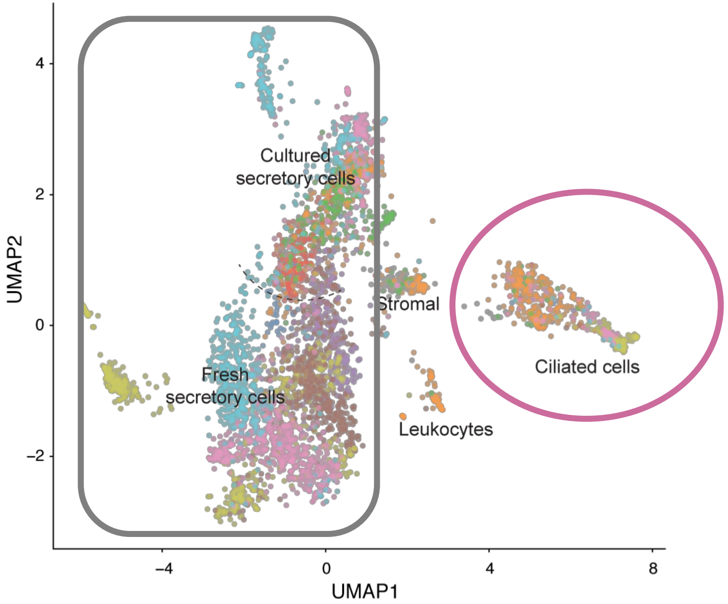
Validation by IF or IHC

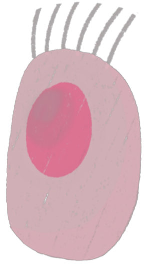
Normal tissues
Not cancer samples

Single-cell transcriptome recapitulates two major cell types in normal fallopian tube epithelium

~4K cells
0.5M reads/cell
3.5K genes/cell

Secretory
PAX8+



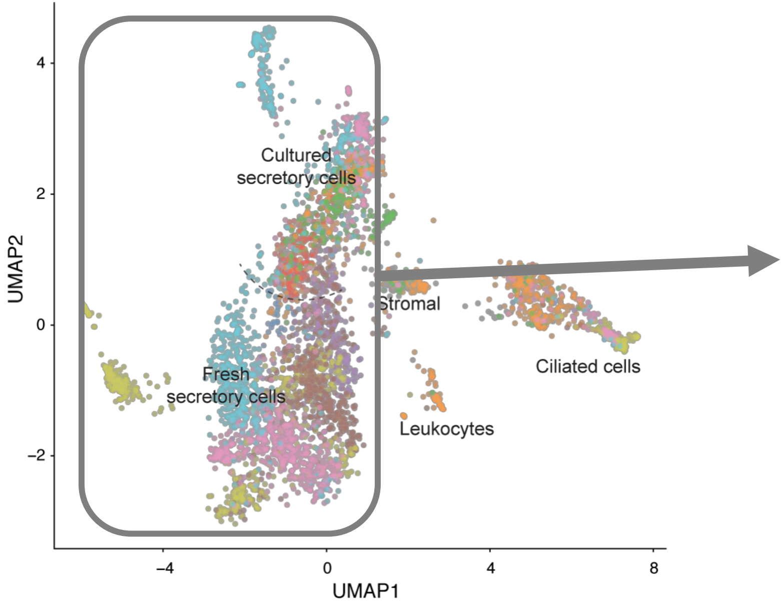


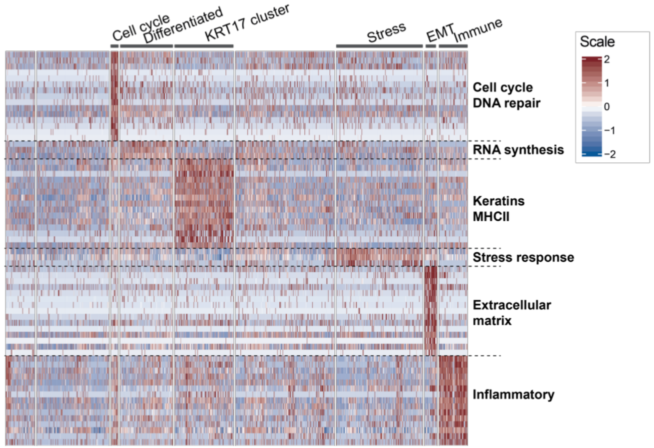
Ciliated cells
TUBB4+

UMAP: Dimensional reduction to project transcriptomic data onto a 2-dimensional plane

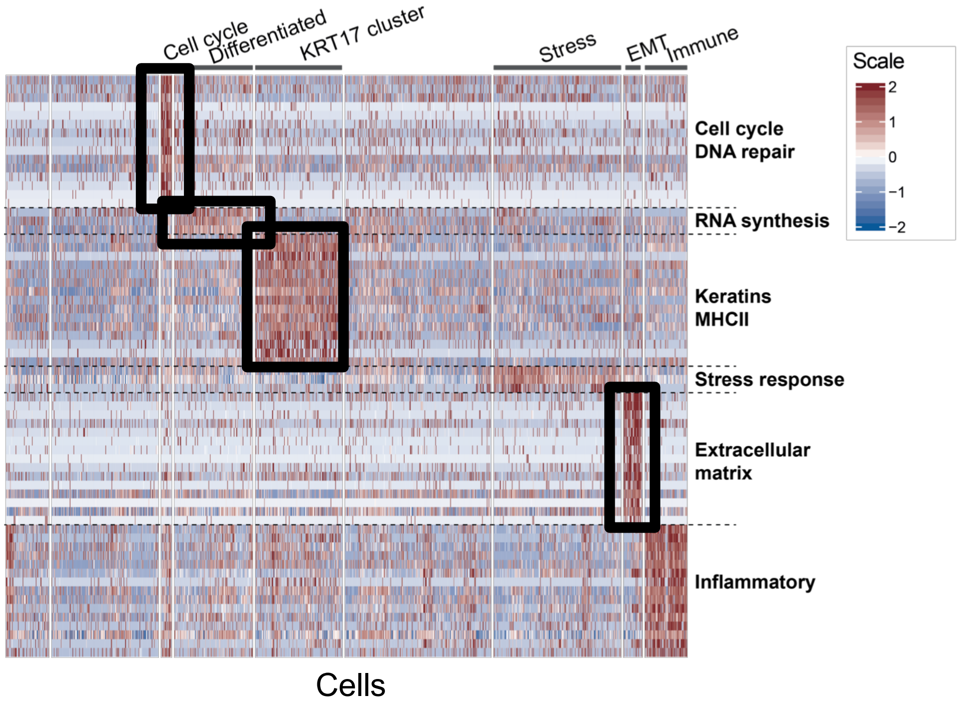
We identified four novel secretory subtypes in fallopian tube epithelium (normal tissue)

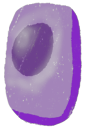
Secretory
PAX8+





We identified four novel secretory subtypes in fallopian tube epithelium (normal tissue)



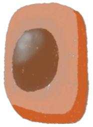


Differentiated cluster

RNA synthesis ↑

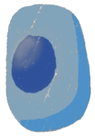
KRT17 cluster (KRT17+)

Keratins+ MHCII+ ALDH+



EMT cluster (RGS16+)

Extracellular matrix ↑

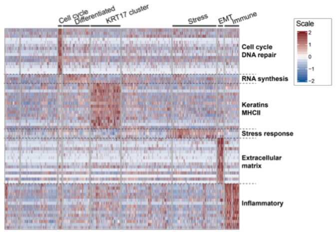


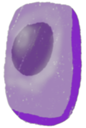
Cell cycle cluster (Stathmin+)

Cell cycle ↑ DNA repair ↑

Chromatin remodeling ↑

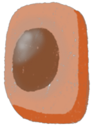
The *in silico* populations of secretory cells were validation at the protein level



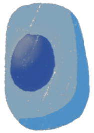


Differentiated cluster
RNA synthesis ↑

KRT17 cluster (KRT17+)
Keratins+ MHCII+ ALDH+

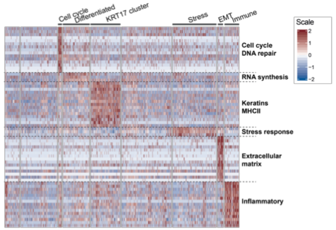


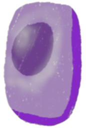
EMT cluster (RGS16+)
Extracellular matrix ↑



Cell cycle cluster (Stathmin+)
Cell cycle ↑ DNA repair ↑
Chromatin remodeling ↑

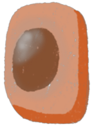
The *in silico* populations of secretory cells were validation at the protein level



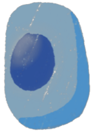


Differentiated cluster
RNA synthesis ↑

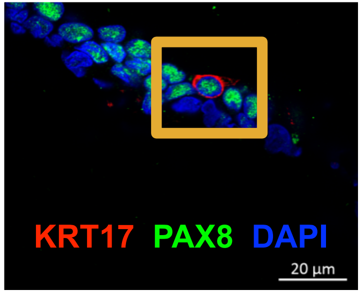
KRT17 cluster (KRT17+)
Keratins+ MHCII+ ALDH+



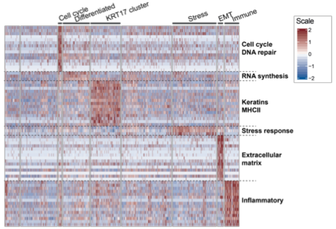
EMT cluster (RGS16+)
Extracellular matrix ↑



Cell cycle cluster (Stathmin+)
Cell cycle ↑ DNA repair ↑
Chromatin remodeling ↑



The *in silico* populations of secretory cells were validation at the protein level



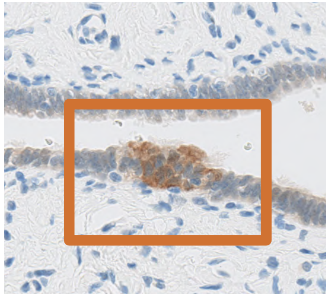
Differentiated cluster
RNA synthesis ↑

KRT17 cluster (KRT17+)
Keratins+ MHCII+ ALDH+

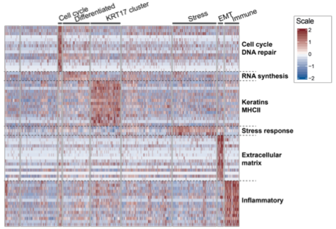
EMT cluster (RGS16+)
Extracellular matrix ↑

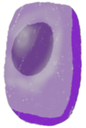
Cell cycle cluster (Stathmin+)
Cell cycle ↑ DNA repair ↑
Chromatin remodeling ↑

RGS16



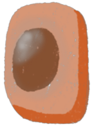
The *in silico* populations of secretory cells were validation at the protein level



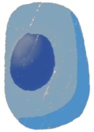


Differentiated cluster
RNA synthesis ↑

KRT17 cluster (KRT17+)
Keratins+ MHCII+ ALDH+

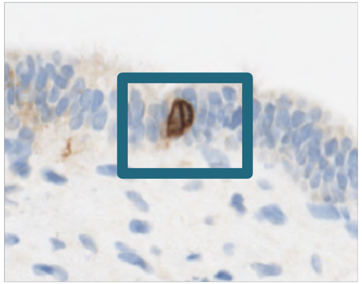


EMT cluster (RGS16+)
Extracellular matrix ↑



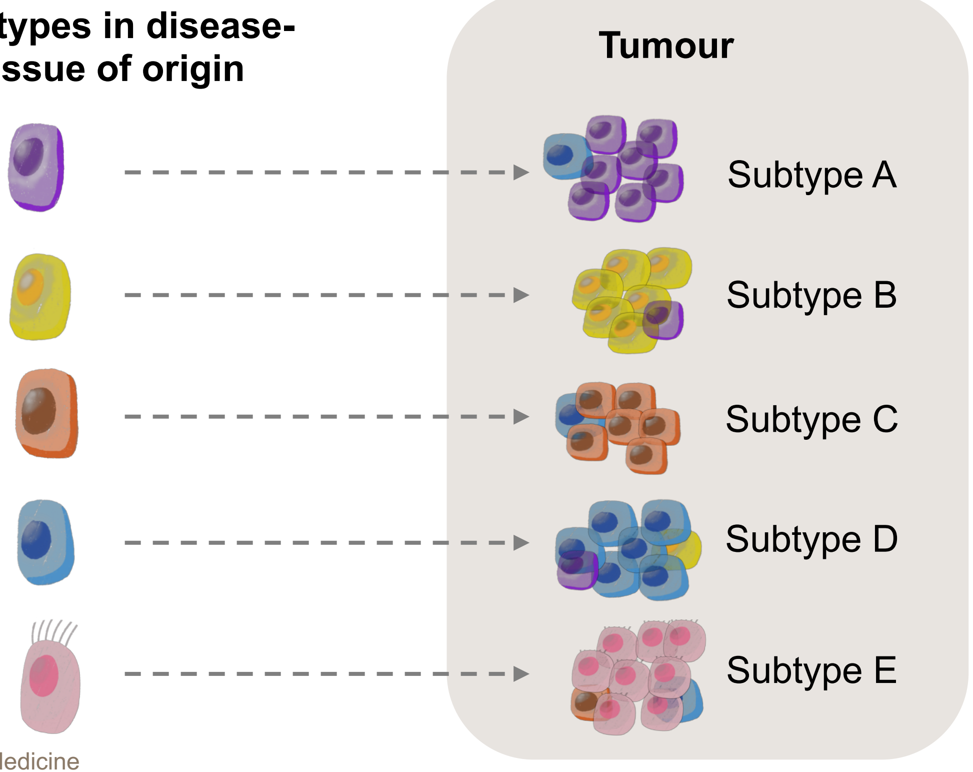
Cell cycle cluster (Stathmin+)
Cell cycle ↑ DNA repair ↑
Chromatin remodeling ↑

Stathmin/STMN1



Cell-of-origin may provide evidence for molecular subtyping

Profiled 4 secretory subtypes + 1 ciliated type in the normal tissue of origin



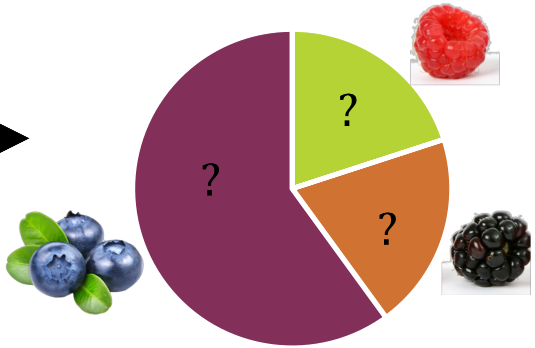
Next step: decomposing the bulk tumours into the proportions of individual cellular subtypes.

Deconvolution is decomposing bulk expression data with transcriptome of individual cellular subtypes

Cell subtypes profiled by single-cell RNA-seq

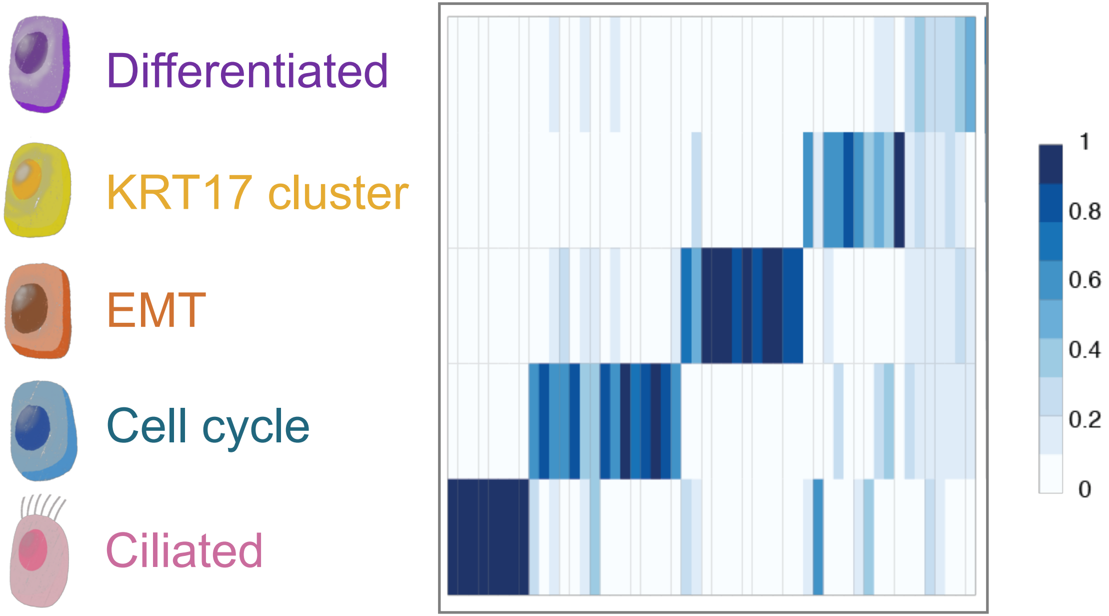
Bulk tumour sample

Deconvolution



Deconvolution step 1: constructing a reference matrix to represent cell subtypes profiled by single-cell RNA-seq

Cellular subtypes in normal fallopian tube epithelium


Differentiated

KRT17 cluster

EMT

Cell cycle

Ciliated

52 marker genes

Deconvolution step 2: estimating the proportion of each cell type in bulk tumour expression data

Estimating tumour composition

Tumour 1

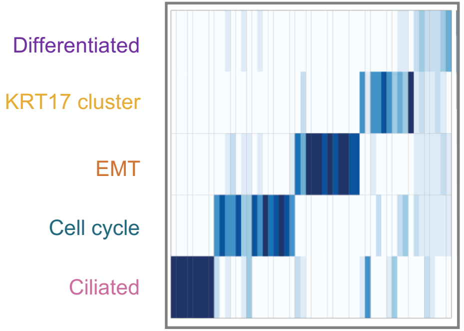
Tumour 2

Tumour 3

Deconvolution equation

Bulk Tumour = α + β + γ + δ +ε

Reference matrix of normal tissue-of-origin

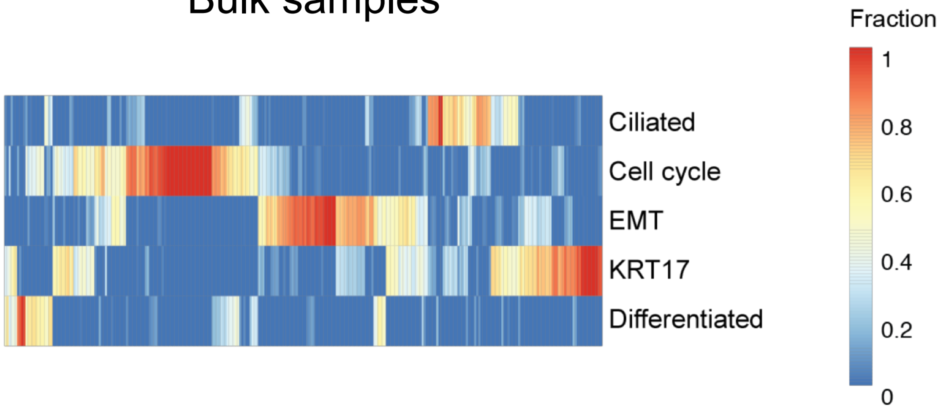


Deconvolution revealed linkage between cellular subtypes from normal tissue and HGSOC tumour subtypes

In normal tissue

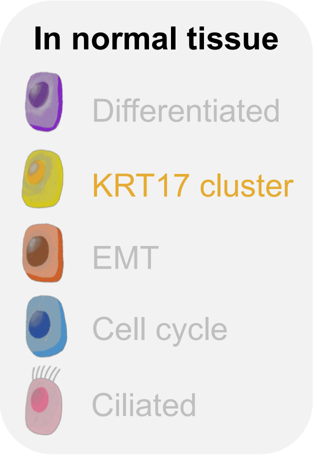
- Differentiated
- KRT17 cluster
- EMT
- Cell cycle
- Ciliated

High-grade serous ovarian cancer
Bulk samples



AOCS microarray dataset: Tothill et al., 2008

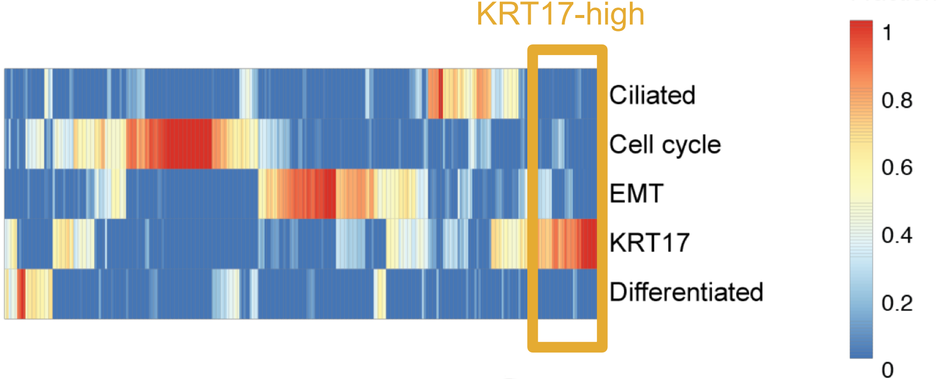
Deconvolution revealed linkage between cellular subtypes from normal tissue and HGSOC tumour subtypes

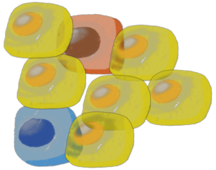


In normal tissue

- Differentiated
- KRT17 cluster
- EMT
- Cell cycle
- Ciliated

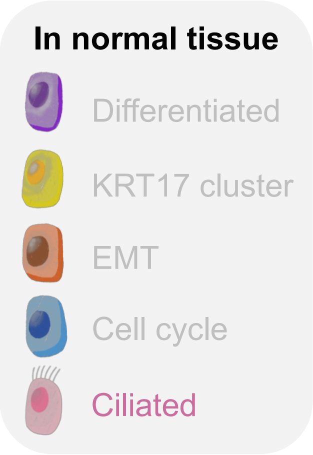
High-grade serous ovarian cancer
Bulk samples





AOCS microarray dataset: Tothill et al., 2008

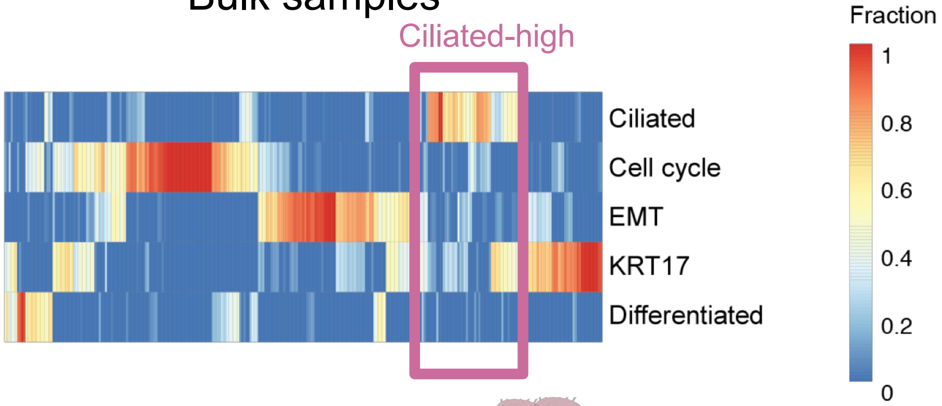
Deconvolution revealed linkage between cellular subtypes from normal tissue and HGSOC tumour subtypes

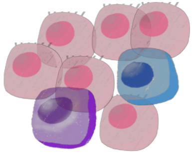


In normal tissue

- Differentiated
- KRT17 cluster
- EMT
- Cell cycle
- Ciliated

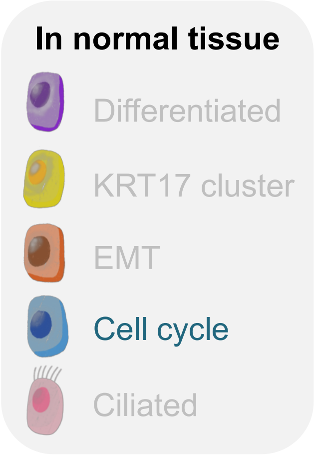
High-grade serous ovarian cancer
Bulk samples





AOCS microarray dataset: Tothill et al., 2008

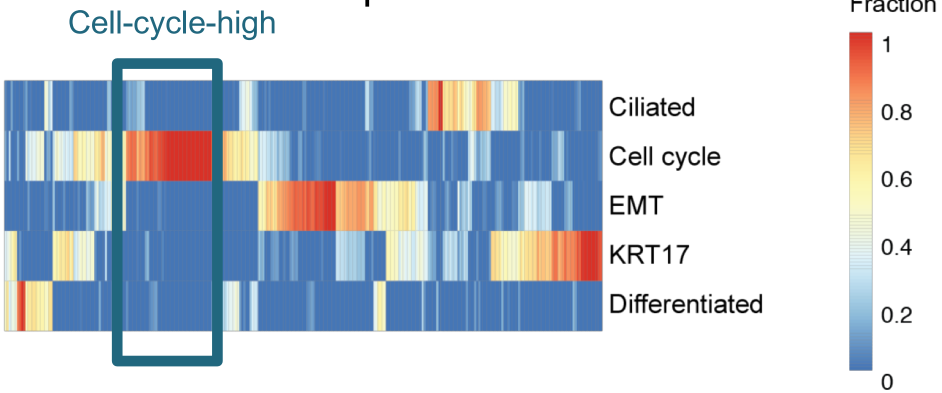
Deconvolution revealed linkage between cellular subtypes from normal tissue and HGSOC tumour subtypes

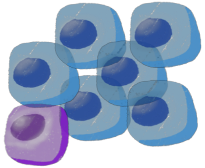


In normal tissue

- Differentiated
- KRT17 cluster
- EMT
- Cell cycle
- Ciliated

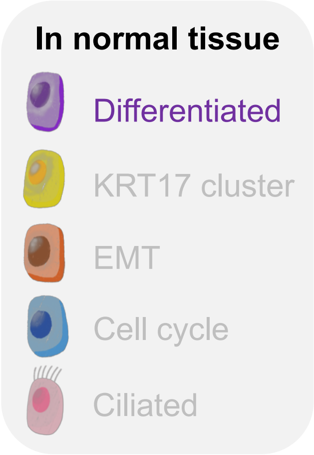
High-grade serous ovarian cancer
Bulk samples





AOCS microarray dataset: Tothill et al., 2008

Deconvolution revealed linkage between cellular subtypes from normal tissue and HGSOC tumour subtypes



In normal tissue

Differentiated

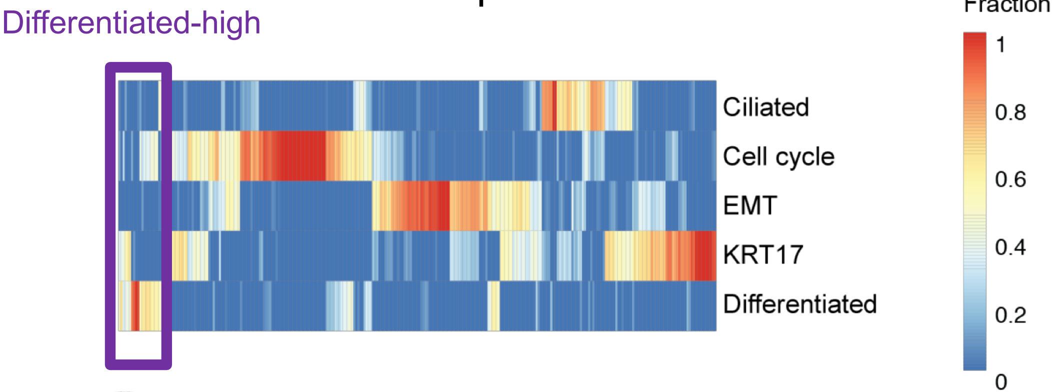
KRT17 cluster

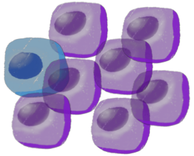
EMT

Cell cycle

Ciliated

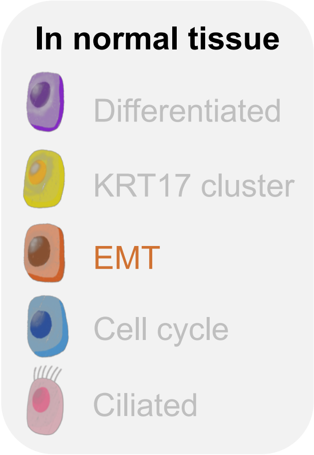
High-grade serous ovarian cancer
Bulk samples





AOCS microarray dataset: Tothill et al., 2008

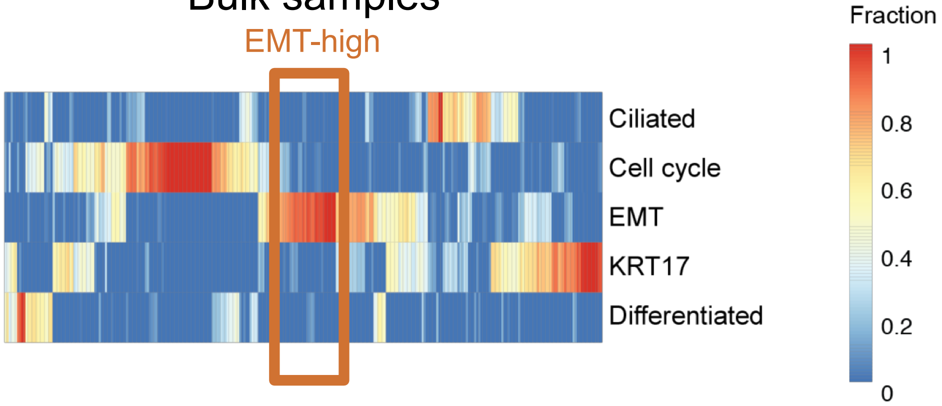
Deconvolution revealed linkage between cellular subtypes from normal tissue and HGSOC tumour subtypes

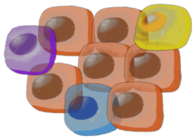


In normal tissue

- Differentiated
- KRT17 cluster
- EMT
- Cell cycle
- Ciliated

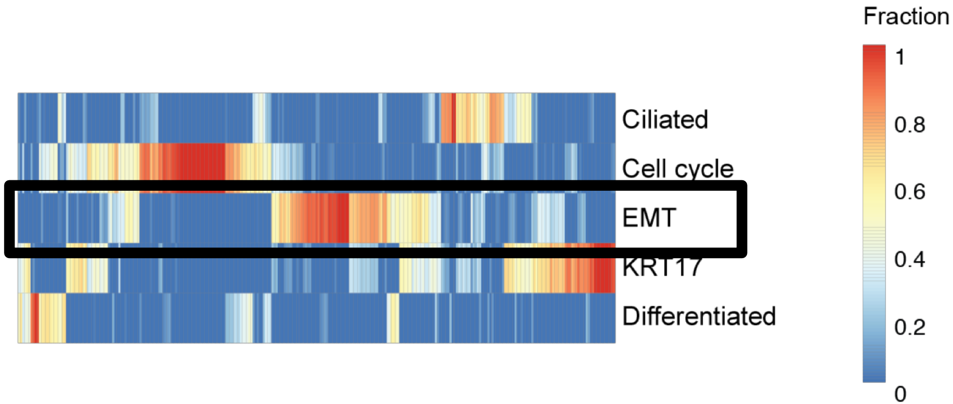
High-grade serous ovarian cancer
Bulk samples





AOCS microarray dataset: Tothill et al., 2008

What is the clinical significance of deconvolution results?



Test the relationship between EMT scores of tumours and overall survival of patients

AOCS microarray dataset:

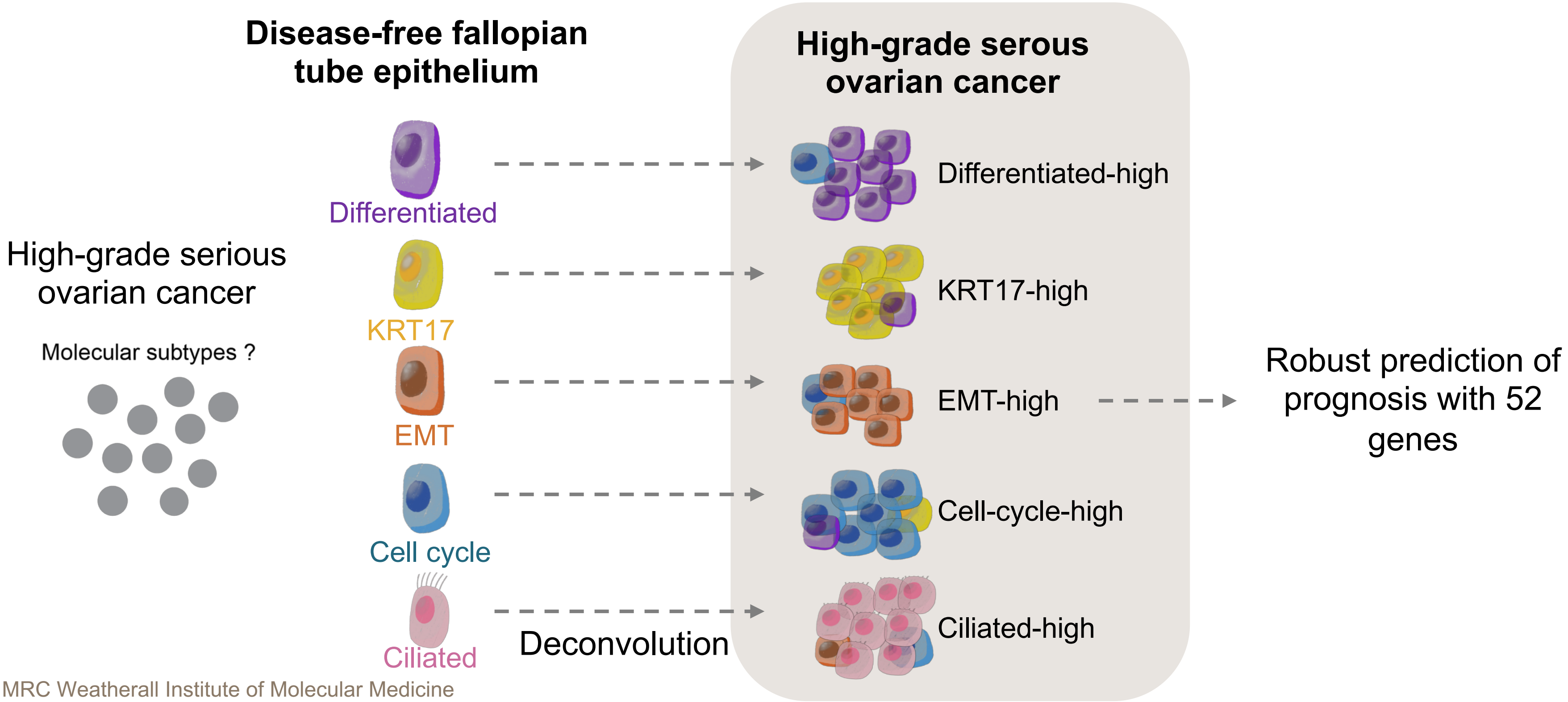
- Hazard ratio = 2.69
- P-value = 0.0004

EMT-high tumours are robustly correlated to poor prognosis of HGSOC

Correlation between EMT-high tumours and overall survival

Dataset (N > 100)	Hazard ratio	P-value	Citation
AOCS (Grade≥2)	2.69	0.0004	Tothill et al., 2008
GSE49997			Pils et al., 2012
GSE13876			Crijns et al., 2009
GSE26712			Bonome et al., 2008
GSE26193			Mateescu et al., 2011
GSE51088			Karlan et al., 2014
E.MTAB.386			Bentink et al., 2012
TCGA OvCa RNA-seq			Bell et al., 2011

Take-home messages



Acknowledgments

Thank you very much for listening!

Prof Ahmed Ahmed
Abdulkhaliq Alsaadi
Kay Chong
Laura Santana Gonzalez
Mara Artibani
Matteo Morotti
Mohammad Karaminejadranjbar
Nina Wietek
And all our colleagues

Dr Christopher Yau
Kaspar Märtens
Tammo Rukat
Yun Feng

Twitter @zhi_yuan_hu

WIMM Flow Cytometry Facility

WIMM Single Cell Facility

WIMM Computational Biology
Research Group (CBRG)

Nuffield Department of Women's
& Reproductive Health

St Cross College

Ovarian Cancer Action
CSC-NDM Studentship