Phylogenies Derived from Matched Transcriptome in Breast Cancer Brain Metastases

Yifeng Tao¹,², Haoyun Lei¹,², Adrian V. Lee³, Jian Ma¹, Russell Schwartz¹,⁴

¹Computational Biology Department, School of Computer Science, Carnegie Mellon University
²Joint Carnegie Mellon-University of Pittsburgh Ph.D. Program in Computational Biology
³Department of Pharmacology and Chemical Biology, UPMC Hillman Cancer Center, Magee-Womens Research Institute, University of Pittsburgh
⁴Department of Biological Sciences, Carnegie Mellon University
Background: Cancer Progression

- Cancer: mainly caused by accumulated somatic alterations
- Tumor cells: heterogeneous populations/clones
- Tumor phylogeny: tumor cells follow a clonal evolution
- Metastasis: transfer from primary site to other sites
- Cell communities vs. cell clones

[Uchi, R. et al., PLOS Genetics. 2016]
Breast cancer: 2nd common cause of death from cancer in women

Metastatic breast cancer
  - Causes majority of those deaths
  - Limited viable treatment options
  - Early detection is important

Mechanism of tumor progression/evolution during metastasis?
Tumor Evolution Derived from Match Bulk Transcriptome

- Given matched primary and metastatic bulk transcriptome:
  - Q1: How to cope with high-dim, noisy, and uninformative transcriptome?
  - Q2: What model and solver to unmix/deconvolve clones?
  - Q3: How to infer evolutionary trajectory and perturbed pathways/functions?

- Yes! We proposed a three-step pipeline.

[Vareslija, D. et al., Journal of the National Cancer Institute. 2018]
Step 1: Mapping to Gene Modules and Cancer Pathways

Q1: How to cope with high-dim, noisy, and uninformative RNA?
   - Gene modules
     - Compress high dimensional and noisy data → accurate deconvolution
   - Cancer pathways
     - Markers/probes → interpretation purpose
Step 2: Deconvolution of Cell Communities

Q2: What model to unmix/deconvolve clones?

- Matrix factorization
- C: expression profiles of communities
- F: fractions of communities in samples

However, it is non-convex and not trivial to solve…
Step 2: Deconvolution of Cell Communities

Gradient descent by backpropagation

\[
\min_{C,F} \| B - CF \|^2_{Fr}, \\
\text{s.t.} \quad F_{lj} \geq 0, \quad \sum_{l=1}^{k} F_{lj} = 1
\]

- Q2: What model and solver to unmix clones?
  - Neural network deconvolution (NND)
  - # components: trade-off of model complexity vs. sample size
    - Mask matrix for cross-validation in NND

Yifeng Tao et al. @ ISMCO’19
Phylogenies of Breast Cancer Brain Metastases
Step 3: Inference of Phylogeny

Q3: How to infer evolutionary trajectory and perturbed pathways?
- Phylogeny skeleton built using neighbor-joining algorithm
- Pathway of Steiner nodes inferred by minimizing the elastic potential energy:

\[
\begin{align*}
\min_x & \quad U(x, y; \mathcal{W}) = \sum_{(u,v) \in \mathcal{E}} \frac{1}{2} w_{uv} (x_u - x_v)^2 + \sum_{v \leq k-2} \frac{1}{2} w_{uv} (x_u - y_v)^2 \\
& \quad \uparrow \\
\min_x & \quad \frac{1}{2} x^T P(\mathcal{W}) x + q(\mathcal{W}, y)^T x
\end{align*}
\]

\[G = (\mathcal{V}, \mathcal{E})\]

PC1: recurrent feature between primary and metastatic samples
PC2+PC3: variability between patients
Effective in separating primary tumors from metastatic tumors

Gene expression
\[ z_{\text{MSD}} = -2.6 \]

Pathway
\[ z_{\text{MSD}} = -13.4 \]

[Park, Y. et al., Transactions on Computational Biology and Bioinformatics. 2009]
Differentially Expressed Cancer Pathways

- Neurotransmitter and calcium homeostasis
- ErbB2/HER2 pathway
- Immune activity
- Dysregulation promotes tumor growth

Cancer pathway (FDR)
- cAMP signaling pathway (6.88e-03)
- ErbB signaling pathway (2.09e-02)
- Calcium signaling pathway (4.39e-02)
- Cytokine-cytokine receptor interaction (4.37e-06)
- Apoptosis (8.53e-04)
- JAK-STAT signaling pathway (8.53e-04)
- Wnt signaling pathway (3.97e-03)
- Hedgehog signaling pathway (4.50e-03)
- PI3K-Akt signaling pathway (1.35e-02)
- TGF-beta signaling pathway (4.56e-02)
- Notch signaling pathway (4.56e-02)
The deconvolution provides more fine-grained landscape of tumor cell communities
Common temporal order of perturbed pathways during metastasis
- Gained ErbB caused by early events
- Expansion of minor clonal populations with lost PI3K-Akt and gained RET
Conclusion and Future Work

- **Conclusion**
  - Pipeline to infer tumor evolution using matched bulk transcriptome
  - Common temporal order of perturbed pathways in breast cancer brain metastases

- **Open source code, data and supp:**
  - [https://github.com/CMUSchwartzLab/BrM-Phylo](https://github.com/CMUSchwartzLab/BrM-Phylo)

- **Further exploration:** multiple metastatic sites

---

Yifeng Tao et al. @ ISMCO’19

Phylogenies of Breast Cancer Brain Metastases
Acknowledgment

○ Authors

Prof. Russell Schwartz  Prof. Jian Ma  Prof. Adrian V. Lee  Haoyun Lei

○ Fundings

National Institutes of Health

Pennsylvania Department of Health

BREAST CANCER ALLIANCE

Pittsburgh Health Data Alliance

Center for Machine Learning and Health
Carnegie Mellon University

Yifeng Tao et al. @ ISMCO’19  Phylogenies of Breast Cancer Brain Metastases