肿瘤信息学

徐鹰

佐治亚大学生化系、吉林大学计算机学院
Lecture V

Biology of cancer metastasis and metastatic cancers
Understanding Cancer at the Tissue Level

- It is well established that a cancer tumor goes through repeated hypoxia and reoxygenation cycles as the cancer progresses.
Understanding Cancer at the Tissue Level

**Questions:**

- How do cancer cells deal with the two extreme oxygen conditions: highly hypoxic and oxygen rich due to angiogenesis?
- What cellular characteristics do repeated hypoxia-reoxygenation cycles tend to select in the survived cancer cell population(s)?
Identification of Subpopulations

• We have examined if each tissue sample may consist of multiple large subpopulations across three cancer types
  • Breast carcinoma: 994 samples
  • Colon adenocarcinoma: 233 samples
  • Lung adenocarcinoma: 470 samples

• Specifically we used 256 genes related to hypoxia, angiogenesis, cell proliferation and a few other functional groups, to examine if there are cell subpopulations showing distinct expression patterns within each tissue, which are consistent across all tissues of each cancer type
Identification of Subpopulations

• This problem can formulated as a matrix factorization problem of the gene expression dataset $X_{m \times n}$, with $n$ samples and $m$ genes.

• Specifically to find $k$ independent vectors $S_{m \times k}$ and a proportion matrix $P_{k \times n}$ so that

$$\max_i \{ \|E_i\| \} \; \text{is minimized}$$

with $E_i$ being the $i^{th}$ column of $E = X_{m \times n} - S_{m \times k} \cdot P_{k \times n}$, $k \ll \min\{m, n\}$

under the constraint that $\sum_i P_{i \times j}$ is approximately 1.0 for each $1 \leq j \leq n$. 
Identification of Subpopulations

• We have examined different k’s and found k = 3 gives substantial better results than k = 2 but roughly similar results for k = 4

• Hence we conclude that each tissue has three subpopulations each having the same characteristic across all samples of each cancer type
Identification of Subpopulations

• The three characteristic vectors, termed signatures, have the following features:
  • The first one shows highly expressed hypoxia marker genes
  • The second one shows highly expressed reoxygenation marker genes
  • The third one shows low expressions in both the above marker genes (which seems to represent a transition population)

• We focus our analysis on the first two subpopulations, termed as hypoxia and reoxygenation subpopulations
Characteristics of Subpopulations

- The mutation rates in the hypoxia subpopulation is substantially higher than those of the corresponding reoxygenation subpopulation across all three cancer types.
Characteristics of Subpopulations

• The hypoxia subpopulation shows over-expression of the following pathways: **cell proliferation, glycolysis, DNA synthesis, and cholesterol transporter** genes in comparison with the reoxygenation population.

• The reoxygenation population shows over-expressions in cell- extracellular matrix and cell-cell adhesion, immune response, nuclear receptors, and growth factor genes.
Characteristics of Subpopulations

• The hypoxia population shows over-expressed superoxide reductases versus peroxide reductases in reoxygenation population, revealing that the two populations have different sources of ROS
  • superoxide reductases: such as XDH, SOD2 and IDO1
  • peroxide reductases: such as GPX2, GPX3, and GPX5

• Over-expressed mesenchymal marker genes (CDH2, FN1, VIM and VTN) and matrix metalloproteinase genes were observed in the reoxygenation population, suggesting the invasiveness nature of the population
Characteristics of Subpopulations

• Up-regulated apoptosis genes are observed in the hypoxia population

• Over-expression of the ATM/ATR DNA repair pathway are observed in reoxygenation populations
Characteristics of Subpopulations

Increased reoxygenation population with reoxygenation from newly constructed blood vessels

- ↑ Cell proliferation
- ↑ Glycolysis and DNA synthesis
- ↑ Glucose, amino acids, and cholesterol uptake
- ↑ Superoxide reductases
- ↑ Apoptosis
- ↑ DNA damage check
- ↑ Epithelial marker genes
- ↑ Stem cell marker PROM1

Increased hypoxia population with more severe hypoxia

- ↑ Extracellular matrix proteins
- ↑ Cell-cell adhesion
- ↑ Cell-cell signaling
- ↑ Immune responses
- ↑ Nuclear receptors and growth factors
- ↑ Peroxide reductases
- ↑ ATM/ATR DNA repair pathway
- ↑ Mesenchymal marker genes
- ↑ Matrix metalloproteinases
- ↑ Stem cell marker ALDH1A1
Speculations

- The two subpopulations seem to have conjugative relations in support of tissue growth.

- The reoxygenation population builds extracellular matrix, hence laying a foundation for further tissue growth; also cell adhesion to keep the tissue together; it also provides growth factors.

- The hypoxia population grows and also dies rapidly; it also heavily uses glycolysis, hence creating an acidic environment to encroach into the neighboring region; and also provides protection against apoptosis and T-cells.
Population Dynamics

• The fractions of the two subpopulations clearly change with the availability of oxygen throughout the repeated hypoxia and reoxygenation cycles.

• This strongly suggests that the two populations have different growth and death rates under specific oxygenic condition, hence changing relative population sizes.

• Such changes in population sizes can be accomplished through cell-cell competition or other means.
Population Dynamics

- The overall mutation rate goes up when a tissue becomes more hypoxic and it goes down when a tissue becomes more oxygenic.
Overall, the fraction of hypoxia population goes up as a cancer advances, although the actual fraction varies across different cancer types.
Some Thoughts

• The reoxygenation population has strong cell-ECM and cell-cell adhesion, hence termed **sticky cells** while the hypoxia population has substantially lower expression levels in such adhesions, hence termed **loose cells**

• We speculate that the sticky cells, along with the ECM built by them, hold the tumor cells together

• The relative sizes between the two populations may play a role in determination of whether a cancer forms a single tumor, or multiple tumors or a diffused form
Some Thoughts

• Specifically, the larger a hypoxia population, the more diffused a cancer biomass may become.

• Our data suggests that as a cancer advances, it tends to become increasingly more diffused.
Metastatic Cancer

- Metastatic cancers are responsible for 90% of cancer-related death.
- Currently metastatic cancer is considered as a “terminal” illness.
- While it represents a highly important health issue, virtually nothing is known about the unique biology of metastatic cancer.
- Some of the very basic questions remain unanswered, including:
  - What drives primary cancers to metastasize?
  - What drives metastatic cancer grow so much more aggressively than its primary cancer counterpart?
Key Steps in Cancer Metastasis

• Local invasion: moving across basement membrane to enter stromal compartment (ECM)

• Interactions with stromal cells (fibroblasts and pericytes):
  • **Fibroblasts** to synthesize ECM proteins, glycoproteins and glycosaminoglycans, and function in wound healing
  • **Pericytes** are to regulate capillary blood flow and clearance of cellular debris, and they serve as a gatekeeper in preventing cancer cells from spreading
  • Cancer tissues indeed tend to have lower level of pericytes
Key Steps in Cancer Metastasis

- **Intravasation**: penetrating through blood vessels to get into circulation

- **In circulation**: cancer cells typically stay in circulation for a few hours

- **Extravasation**: lodging on to the inner wall of blood vessel and then penetrate the wall to get inside the stromal compartment of the new location

- **Adaptation** to the new microenvironment: arriving cancer cells may stay in dormancy for months or even years before they are reactivated
Key Steps in Cancer Metastasis
Molecular Level Activities

• To break away, cancer cells need to cut their ties with the local connection, including cell-cell adhesion and cell-extracellular matrix (ECM) binding

• Cancer cells tend to reduce the expression levels of genes involved in cell-cell adhesion and cell-ECM connection (via synthesis of hyaluronic acids and wrapping around)
Molecular Level Activities

• In addition, cancer cells may form and release complex structures called invadopodium, which delivers MMPs to cut through ECMs to enable cancer cells to enter the stromal compartments.

• It is essential for cancer cells to enter stromal compartment since that is where the blood vessels, the channel for escaping, are.

• In addition, it is essential for the cancer cells to interact with the local stromal cells for them to get inside the blood vessels.

• E.g., pericytes have to be recruited to relax their safeguard to cancer cells to pass them to get inside blood vessels.
Molecular Level Activities

• The current understanding is that tumor associated macrophages play key roles in the intravasation process of cancer cells to break the endothelial cells to get inside blood vessels.

• It is estimated that for each gram of tumor cells ($10^8$-$10^9$ cells) accessible to blood circulation, about one million ($10^6$) tumor cells can actually get into circulation, and 1 out of 10,000 cancer cells in circulation can ultimately settle in a distant location, overall 1 out of $10^6$ – $10^7$ cells.

• It is noteworthy that anchorage-free cells cannot survive since a programmed death process called anokis will be activated.

• One speculation is that these cells may be wrapped around by hyaluronic acids and glued together with platelets in circulation.
Possible Projects

• What are the key processes involved in the intravasation?

• What roles are played by hyaluronic acids in intravasation?

• Do metastasizing cancer tissues have specially high ROS levels?
Molecular Level Activities

• Extravation (getting out of circulation) is not a trivial process as compared to intravasation through often a leaky blood vessels at the tumor site, the extravasation sites are normal and healthy blood vessels.

• One speculation is that cancer cells may have activated the same process that leukocytes use when they extravasation process into the inflammatory tissues.

• The most challenging step for cancer metastasis is to survive the new and very hostile environment!
Molecular Level Activities

• It has been proposed that cancer cells activate a process called epithelial-to-mesenchymal (EMT) to enable them to travel

• ... and once the cells have arrived the new location, they activate another process called mesenchymal-to-epithelial transition (MET) for them to get settled in the new location
Possible Projects

• Is EMT indeed involved in cancer cell circulation?

• What are processes are active in circulating tumor cells and why?
Molecular Level Activities

• The new environment is profoundly different from the original
  • Well-organized tissue structures
  • Oxygen-rich versus hypoxic condition in their old environment
  • Substantially less acidic
  • Healthy and more aggressive immune cells
  • Possibly lower level of Fenton reactions
Molecular Level Activities

• The current theory is that the arriving cancer cells will remain dormant for weeks, months or even years before they become aggressive again.

• During this dormancy period, cancer cells are speculated to gradually change the local environment through releasing their material.

• One proposal is that they release something called exosomes, which consist of cytoplasmic RNA and functional proteins, to make the local environment inflammatory, causing tissue damages.
Possible Projects

• Through comparative analyses of transcriptomic data of primary cancer tissues with metastasis versus without, one can possibly learn substantial amount of information regarding
  • drivers of metastasis
  • detailed mechanisms of metastasis

for different cancer types

• Through analyses of transcriptomic data of circulating tumor cells, one can possibly answer various questions regarding how circulating cells protect themselves among other questions
Issues to Think About

• The previous studies have been mostly focused on
  • observing unique molecular and cellular activities by various involved cell types throughout the metastasis process and
  • figuring out how each activity may have been executed at the molecular level
• But very little is known about
  • the driving forces of any of these activities; and
  • the survival programs encoded somewhere, genome or epigenome, since all cancer metastases seem to go through similar escaping, survival and re-activation processes
Two Major Theories about Metastasis

• One theory posits that cancer starts to metastasize once it reaches certain size

• A competing theory suggests that cancer may start spreading as it advances to form many micro-metastases

• Speculation: it is probably feasible to test using available omic data of DTC, which of the two hypotheses may be closer to the reality
What Drive Cancer to Metastasize

• One hypothesis is that cancer metastasizes to escape from the too high ROS level

• This is consistent with our “theory” that
  • ROS continues to increase as a cancer advances
  • During progression, cancer cells reduce ROS-induced pressure through division to reduce the ROS level
  • Once cell-division could not keep up with the increase of ROS, i.e., when ROS goes over some threshold, it will be trigger a number of mechanisms to set the cells free so they can move to new locations
Possible Questions to Address

• Knowing that cancer cells typically stay in circulation for only a few hours (∼6 hours), one can possibly study the key functional characteristics of such cells, namely CTC

  • What are the main differences between the basic metabolisms of such cells and the cancer tissue cells they left behind?

• Similar questions can be asked regarding cancer cells that just arrived in the new locations, called disseminated tumor cells (DTC), as well as cancer cells that have been residing in the new locations for a while
Possible Questions to Address

• One particularly important issue to study is whether metastatic cells indeed go through a dormancy period – need to find relevant data.

• If such data are available, one interesting question to study could be: what do the dormant cancer cells do?

• A related, but potentially easier to address question could be: how cancer cells alter the local environment in their new locations?

• Hypothesis: our genomes may have encoded programs to enable traveling cells to adjust to a new environment. This may provide some useful guidance for searching for such programs.
A Longer List of Questions

• The available omic data of cancer tissue at different development stages may allow us to address many fundamentally important questions regarding cancer metastasis

• Do cancer cells travel alone or with some tumor-associated stromal or immune cells?

• Do easy-to-metastasize cancers have common characteristics which are different from hard-to-metastasize cancers?

• Do cancers metastasizing to the same organs tend to have common properties? What are they?
A Longer List of Questions

• When cancer cells circulate in blood flow, what programs do they use, to prevent them to be killed by anokis program?

• What programs encoded in our genomes do they use to help them to settle down in their new locations?
A Longer List of Questions

• Are there some thresholds of certain environmental factors, beyond which cancer cells start to leave their base? What are these factors and what the potentially relevant mechanisms encoded in our genome that may facilitate this?
  • A combination of population sizes of sticky versus loose cells and ROS levels?

• The metastatic process seems to be a rather challenging process and the relevant cells must have learned to activate these programs very rapidly. Do these cells have learned something during the “long” period of primary cancer development? What are they?
The Deadliest Stage of Cancer

• For reasons that are not understood, once a cancer is metastasized to a new location, it becomes substantially more aggressive and more difficult to control.

• Drugs designed for primary cancers generally lose their effectiveness on metastatic cancers, and currently no drugs specifically designed for metastatic cancers.

• Question: why cancer becomes so deadly once it is metastasized to new locations.
How to Address this Issue?

• We have collected all the gene-expression data of metastatic cancer tissue samples with matching primary cancer tissues from public database on the Internet

• Resulting in: 16 sets of genome-scale transcriptomic data, covering 11 types of primary-to-metastatic cancers and 858 tissue sample
  • prostate-to-bone metastases;
  • breast-to-brain metastases;
  • breast-to-liver, colon-to-liver, pancreas-to-liver and prostate-to-liver metastases;
  • bone-to-lung, breast-to-lung, colon-to-lung, kidney-to-lung and pancreas-to-lung metastases
Questions to Ask

• **Question #1**: are there genes (pathways) consistently up-regulated across all the metastatic cancer samples in comparison with the matching primary cancer tissues

• The answer is yes, there are a few dozen such genes?

• **Question #2**: what are the functions of these genes?

• Answer: A substantial fraction of these genes are involved cholesterol influx and metabolism
Cholesterol?
More Questions to Ask

• Question #3: what do cholesterols do in metastatic cancers?

• Question #4: why do metastatic cancers need cholesterol?

• Both questions can be answered through mining omic data!
- Cholesterol can be oxidized to oxysterols, bile acids, and further metabolized to steroids.
Data Mining: Step-by-Step

• Cells obtain cholesterol through either biosynthesis (HMG-CoA) or uptake from circulation of cholesterol-carrying lipoprotein particles: high, low and very low-density lipoproteins (HDL, LDL and VLDL) and chylomicrons.

• The average percentages of cholesterol (or cholesterol ester) in these particles are: 5% in chylomicron, 25% in VLDL, 47% in HDL and 61% in LDL.

• By checking these genes’ expression levels, one can infer if cholesterol uptake or synthesis is up-regulated.
• 14 out of 16 datasets show up-regulation in SRB1, LDLR, or LRP5, one show up-regulation in cholesterol biosynthesis and the other shows over-expression of CD36, indicating that all metastatic cancers have increased cholesterol influx

• Interestingly, all 16 datasets show up-regulation in cholesterol efflux or esterification genes

• Question 5: do cholesterols do anything when going through the cells?
A number of CYP genes are up-regulated across all the metastatic cancers, which can have cholesterols oxidized to oxysterols and bile acids

- CYP3A and CYP27 can oxidize cholesterols to oxysterols
- CYP11, CYP17A1, CYP19A1 and CYP21A2 can oxidize them to steroids

A few additional enzymes such as HSD3B1, 2, HSD17B1,3,7, SRD5A1-2 are up-regulated in 14 out of the 16 datasets, which can convert oxidized cholesterols to steroidogenic products
Data Mining: Step-by-Step

• It is known that oxysterols and steroids can bind and activate various nuclear receptors and even growth factor receptors directly.

• **Nuclear receptors** are generally responsible for controlling development, homeostasis and metabolism.
• 15 datasets each have at least one up-regulated nuclear receptors such as LXR, FXR, ESR1-2, AR, PGR, NR5A1, ESRRA,B,G and PPARA,D,G

• Activated nuclear receptors can directly or indirectly activate various growth factor receptors such as EGFR, HER2 and ERBB3

• All 16 datasets each have at least one growth factor receptor up-regulated
The following seems to serve “growth engine” for all metastatic cancers with the detailed oxidized cholesterol metabolites, nuclear receptors, growth factor receptors varying across different metastatic cancers.
• All 16 datasets show up-regulated DNA synthesis genes and cell cycle genes
• Various experiments have been carried out to validate the key steps of this oxidized cholesterol-based driver model for accelerated cell proliferation

The abundances of 27-OHC in normal and primary colon cancer tissues; normal and primary gastric cancer tissues; and normal, primary liver cancer and liver metastases tissues
Why Cholesterol?

• Among the many functions of cholesterol, a key function is that they are a key ingredient of cell membrane.

• It is known that cholesterol concentration affects O2 permeability: the higher the cholesterol concentration, the lower the permeability.
Cholesterol as a Barrier of O2

• Published studies have shown red blood cells tend to have lower cholesterol concentration in more hypoxic environment

• All these lead to our main hypothesis: cancer cells need to increase their membrane cholesterol when they move from hypoxic environment to O2-rich environment
O2-Rich versus Cholesterol

• Metastatic cancers tend to have higher O2 level and oxidative stress compared to their primary counterparts, revealed by expression levels of HIF, SOD and other oxidative stress response genes

• Multiple evidences show that membranes of metastatic cancer cells tend to be damaged by increased oxidative stress
  • genes in response to membrane damages are up-regulated
  • the catabolism of the oxidized products of phospholipids, a key component of cell membrane, is up-regulated
  • indication of lipid peroxidation
Lipid Peroxidation

LIPID PEROXIDATION

Initiation:

Unsaturated lipid + OH \rightarrow Lipid radical

Propagation:

Lipid peroxide + \cdotOOH \rightarrow Lipid peroxyl radical

Inside
O₂-Rich versus Cholesterol

• These lead to the (continuous) loss of membrane cholesterol, hence creating a need for (continuous) influx of cholesterol

• Published studies have established that (1) SREBP is the main regulator for cholesterol influx; (2) O₂ can regulate SREBP, and PDZK1 is another regulator

• **Strong positive correlations are observed in all datasets between the responses to membrane damages and SREBP or PDZK1**
Driver Model for Accelerated Proliferation
• **Question #6**: is it possible that increased cholesterol influx is not the driver of increased cell proliferation, instead it is the result of cell proliferation since cell proliferation needs more cholesterol.

• The answer is **No** since...
Any Supporting Evidence?

- Several recent cancer-epidemiology studies have found that taking cholesterol-lowering drugs, such as Statins, can prolong the survival time of cancer patients, hence providing indirect but strong evidence that cholesterol has an important role in cancer-related death in general.
Implications

• If the model is correct, cholesterol, in any form (HDL, LDL or VLDL) is BAD for cancer patients.

• Some “incurable” primary cancers may be the result of utilization of cholesterol in their cells as we have observed although the reason for using cholesterol may be different from here.

• While both primary and metastatic cancers proliferate, the dominating reasons for their cell proliferation are fundamentally different from each other.
Summary of the Course

• Cancer tissue omic data provided unprecedented opportunities to study cancer in its actual environment.

• To study the biology of cancer development, it is vitally important to understand: cancer represents a very rapid evolution process;

• Hence it is critically important to know: what pressures the evolving cells try to adapt to; or what pressures drive the evolution of the disease;

• According to Darwin’s Theory of Evolution and the recent Red Queen model: it is the pressures that give rise to the observed mutations, some of which are selected to assist the cells to adapt to the pressures.
• Mining of cancer tissue omic data lead to our discovery that Fenton reaction may play a key role in driving the disease evolution while keeping the pH stability may represent a defining pressure of cancer.

• A key and constant pressure for the affected cells is that the cells must continue to divide in order to survive; and divide at a rate comparable with the OH- production rate by cytosolic Fenton reactions.

• Various mutations are selected to assist the cells to keep up with the required division rate and to avoid being killed.
Summary of the Course

- Epigenomic level activities might be continuously called upon to help the cells overcome very stressful conditions.

- Metastatic cancers might represent a disease with fundamentally different drivers from those of primary cancers; and hence they should be treated with different treatment paradigms.

- Cancer tissue omic data probably have ALL the information one needs to fully understand cancer biology and find more effective ways to treat cancer.
• Happy mining!

• Hope that you find the course useful!

• Don’t forget about the Acknowledgement