Journal of Bioinformatics and Computational Biology
Vol. 13, No. 2 (2015) 1571002 (11 pages)

DOI: 10.1142/S021972001571002X

Imperial College Press
www.icpress.co.uk

Stresses drive a cancer's initiation, progression and metastasis: Critical comments on the book "Cancer Bioinformatics"

Chao Liu[*,†,§], Chi Zhang[†], Jing Su[‡], Dongsheng Zhang[*,§,¶,**] and Sha Cao[†,‖,**]

[*]*Department of Oral and Maxillofacial Surgery*
School of Stomatology, Shandong University, Jinan, China

[†]*Computational Systems Biology Lab*
Department of Biochemistry and Molecular
Biology and Institute of Bioinformatics
University of Georgia, Athens, Georgia, USA

[‡]*School of Basic Medicine*
Jilin University, Changchun, Jilin, China

[§]*Department of Oral and Maxillofacial Surgery*
Shandong Provincial Hospital
Affiliated to Shandong University, Jinan, China
[¶]*ds63zhang@sdu.edu.cn*
[‖]*shacao@uga.edu*

Received 20 January 2015
Accepted 21 January 2015
Published 27 February 2015

"Cancer Bioinformatics" is a new book published in 2014 by Springer. This 14-chapter book offers a quite unique and potentially controversial view about what drives a cancer to initiate, progress, metastasize and develop in an accelerated manner in metastatic sites. The book treats cancer as an evolutionary process of a diseased tissue (rather than cells) in an increasingly more challenging microenvironment; and discusses the various stresses encountered by a neoplastic tissue and their roles in (driving) cancer initiation, progression, metastasis and post-metastatic development. Most of the discussions are made based on discoveries through mining cancer tissue omic data. In contrast to the on-going theories that cancers are the result of genomic mutations, the book clearly downplays the roles of genomic mutations, particularly oncogenic mutations, in cancer formation and progression. Throughout the book, the authors made special efforts in conveying their overarching view that cancer is a pathway to cell survival under certain stresses, and cell proliferation is either the result or a side-effect of survival processes when evolving to overcome the stresses. While the book is presented in an informatics style, it is actually a book of cancer biology focused on how information can be derived from cancer omic data to address a variety of basic cancer biology questions. Compared to other cancer biology books, this book is clearly less detail-oriented but more holistic and spans the entire range of cancer evolution.

Keywords: Cancer drivers; cancer evolution; cancer bioinformatics; cancer biology.

[**]Corresponding authors.

1. A Book about Fundamental Cancer Biology, Presented from an Informatics Perspective

The book[1] starts with an introductory chapter to cancer biology, followed with a chapter on cancer omic data and information potentially derivable from such data. These two chapters (Chapters 1 and 2) provide good introductory material for a reader to follow the discussions in the following chapters. Chapter 3 is another introductory chapter on the basic informatics approaches for analysis of cancer omic data. The authors used cancer classification in terms of both cancer staging and grading as examples to introduce a few data mining techniques and the general idea of information discovery using such techniques.

Chapter 4 is on computational analyses of cancer genomes, focused on genomic mutations. A rather intriguing reanalysis of genomic mutations in cancers published by other groups was conducted but focused on mutations called *nondriver mutations* by the authors of the original studies. A few fascinating points were made based on the analysis results, including that: (i) precancerous tissues tend to harbor mutations of genes related to extracellular matrices; (ii) cancer tissues at different developmental stages tend to be enriched with mutations in genes with specific functions; and (iii) cancer-associated mutation-harboring genes tend to cover a substantially wider range of functions than oncogenic and tumor suppressor genes only, which is typically done in published cancer genome studies, hence suggesting that the current analysis strategy may be too limiting for gaining a full understanding of what factors contribute to cancer development. The authors continue and further suggest that mutations associated with nonproliferation related genes may have equal or even more important roles in cancer development compared to oncogenic and tumor-suppressor mutations. Then they further extended their arguments to suggest that cell-based cancer studies may have ignored the most important factors associated with cancer evolution, i.e. the tissue-level constraints that cancer evolution must follow, especially in their early stage, a view that may not necessarily fare well with cancer biologists who have based their research solely on cell lines.

Chapters 5 and 6 are two closely related chapters, focused on drivers of cancer initiation specifically for cancers exhibiting the Warburg effect. These two chapters start to set the book apart from other cancer biology books currently available by examining the fundamental issue of why human beings, along with mice and rats, can develop cancer. The starting question raised by the authors is: What happens to energy demand and supply under persistent hypoxia? The authors presented fascinating data to show that there is an intrinsic gap between ATP demand and supply in human (and mouse and rat) cells under persistent hypoxia, which is determined by evolution according to the authors. This gap leads to the altered glucose metabolism, referred to as *energy metabolism reprogramming* in the literature, as well as the widely observed accumulation of glucose metabolites in a large class of cancers (all those with the Warburg effect), the basis for PET-CT-based cancer diagnosis.[2]

According to the authors, another consequence of human evolution, the limited capacity of the glycolytic fermentation pathway, also contributes to the formation of this gap, which has never been used (and trained) to supply all the ATPs needed by human cells for extended periods of time; instead it has been mainly used as a temporary supplement to other major ATP producers such as oxidative phosphorylation. Their limited capacity could not keep up with the increased glucose influx induced to meet the ATP demand under persistent hypoxia, resulting in the accumulation of glucose metabolites and congestions along the fermentation pathway. Interestingly, a few other vertebrates such as frogs and turtles do not have this gap when the living condition changes from normoxia to hypoxia, also a consequence of their evolution as they have been trained to live in both hypoxic and normoxic conditions by switching on and off certain ATP-consuming cellular processes when the condition changes. The authors went further to suggest that this intrinsic gap between energy demand and supply may be one fundamental reason for humans to develop cancer while frogs and turtles do not.

The discussion becomes even more intriguing when the authors proposed that the production and fragmentation of hyaluronic acids in a persistent manner, which is a key component of the extracellular matrix and has been found to play roles in cancer metastasis,[3] represents a fundamental facilitator for cancer development. The authors surveyed a large literature body about what is known regarding the functionalities of (fragmented) hyaluronic acids. In summary, when a tissue is injured, hyaluronic acids, long chains consisting of hundreds of thousands of di-sugars as the basic units, in the extracellular matrix will be fragmented and released into the extracellular space. Scientists who study tissue injuries have long known that fragmented hyaluronic acids, ranging from 2 to 1,000 di-sugars, serve as tissue injury and repair signals of different sorts, including: (i) cell survival, (ii) cell-cycle control, (iii) cell growth, (iv) anti-apoptosis, (v) recruitment of immune cells, (vi) angiogenesis, (vii) facilitation of anchorage-free proliferation, and (viii) inactivation of contact-inhibition signals among others, basically covering all the signals needed for a cancer to start and develop. These signals typically appear only when a tissue is injured — this is the way it is supposed to be, but under persistent hypoxia along with plentiful G6P (the first metabolite of the glycolysis pathway) and inflammatory signals, the biosynthesis pathway of hyaluronic acids can be activated to generate and release the di-sugar chains into the extracellular matrix, where they are degraded by a specific set of enzymes known as hyaluronidases. The authors suggest that persistent hypoxia, along with inflammatory signals, leads to persistent production of hyaluronic acid fragments, which result in a continuous wound healing process that lasts as long as the inflammatory and hypoxia conditions continue. The authors then pointed out that hyaluronic acids can regulate and/or mediate a number of developmental programs, including those regulated by *MYC* and *HSF*1 (a heat shock protein). These two drive the early cell proliferation, mainly to overcome the stress incurred by the continuous accumulation and congestion of glucose metabolites, which will be diverted towards DNA synthesis and cell division.

The authors conclude that cell proliferation, at this point still being controlled, is a way to overcome a specific stress. They went on to speculate that oncogenic mutations may be selected as more efficient replacements for on-going functionalities that contribute to cell proliferation, rather than starting new ones, a theme that the authors maintain on other cancer-contributing mutations throughout the entire evolution of a cancer.

After these discussions on sporadic cancers, the authors went on to state that hereditary cancers may take a similar route to initiate. Specifically they observed from literature searches that virtually all the cancer-inducing mutated genes have one commonality: their loss-of-function mutations all lead to the accumulation of reactive oxygen species (ROS), including *BRCA* for breast cancer, *APC* for colon cancer, and *RB*1 for retinoblastoma. From here, the authors suggest that the accumulation of ROS over time will lead to cellular repression of the mitochondrial activities, including the electron transfer chain for ATP production, which will lead to the up-regulation of the glycolytic fermentation pathway to supplement the reduced ATP production in mitochondria, hence leading to the same energy metabolism reprogramming as in the case of inflammation/hypoxia-induced reprogramming.

Overall these two chapters are densely written, with multiple thought-provoking discussions consisting of a few rather unexpected and potentially controversial as well as transforming points of view.

Chapter 7 discusses how cancer cells evade apoptosis. Compared to the previous two chapters, this chapter is substantially more traditional, essentially presenting the existing knowledge with an emphasis on how omic data can be used to address various issues related to cancer's evasion of apoptosis, given towards the end of the chapter.

Chapter 8 presents a set of discussions regarding the evolving microenvironments and their relationships with cancer evolution, focused on three types of stresses associated with a cancer environment, namely cell–cell competition among dividing cells, cancer cells coping with lactic acidosis and interactive relationships between cancer and immune cells. The authors proposed a mechanism for cell–cell competition within a cancer tissue, based on published studies of cell–cell competition in Drosophila and cancer omic data. This model can be potentially used to address a number of issues associated with population dynamics within a cancer tissue to cope with the changing microenvironments.[4] On the topic of how cancer cells cope with lactic acidosis (created by the cells), the authors presented an interesting model regarding how cancer cells evade acidosis-induced apoptosis, to explain why lactic acidosis can trigger apoptosis in normal cells but not in cancer cells. Basically, the model suggests that a number of mechanisms for pumping out and/or neutralizing protons become available only to dividing cells, possibly in conjunction with hypoxic conditions.

Only limited material was presented on cancer immunity, which is somewhat disappointing, knowing that immunity plays fundamental roles in every stage of a cancer development. It would be most beneficial to the readers of the book if a much

more expanded version on this topic can be included in a future edition, particularly on the complex relationships between different types of immune cells and cancer cells, as well as possible relationships between early inflammation types and future evolutionary trajectories of a cancer evolution, which, we suspect, could be revealed through mining omic data of cancer tissues containing immune cells.

Chapter 9 focuses on interactions between the evolving cancer cells and also evolving microenvironment(s) and how such interplay may ultimately drive the levels of stress so high that the diseased cells start to rely on epigenomic level responses to cope with the stresses. An intriguing model is proposed by the authors regarding key roles played by cancer epigenomes, based on a recent study on Drosophila. The authors proposed that the epigenome may encode a last-resort survival program under extreme stresses, which could not be overcome by any program encoded in the Drosophila genome or triggered by the current (stress) conditions. The proposal suggests that an epigenomic level program will be triggered once the stress level goes beyond certain thresholds, which systematically go through all response programs and directly activate them but by-passing their usual triggering conditions; any responses that lead to the survival of the underlying cells in conjunction with the current stress will be recorded in the epigenome of the cancer cell. This information can be passed on to the next generation; hence, when cells of the new generation face the same stress, the recorded response program will be directly initiated, without the need for systematic searches. The authors speculate that this proposal can be computationally checked for its validity with the proper omic data.

Chapters 10 and 11 are two closely related chapters focused on cancer metastasis and post-metastatic development. We consider them as another high point of the book. Like the previous chapters, the authors continue their thought-provoking and potentially controversial discussions. After a brief introduction to the current knowledge of cancer metastasis, the authors focused their presentation on a fundamentally novel model to explain why metastatic cancers behave so differently from their primary cancer counterparts. Interestingly the authors attribute the fundamental change, referred to as the second transformation by them, to the difference in the oxygen level between the primary cancer and the metastatic sites, with the former being hypoxic and the latter being O_2 rich. This change in O_2 level, along with the fact that primary cancer cells have adapted to the hypoxic environment, which makes them less equipped to cope with normoxic conditions, leads to at least two major changes in the migrating cells: (a) peroxidation of the cholesterols in the cell membrane, leading to their damage on a continuous basis, and (b) up-regulation of all anti-oxidant enzymes, as revealed by the omic data of metastatic cancer tissues. Unfortunately this is a very bad combination since the cells respond to the membrane damage by a substantial increase in cholesterol uptake from circulation (via cholesterol-carrying lipoproteins) and/or *de novo* synthesis; and some of the new cholesterol is apparently oxidized by one class of anti-oxidant enzymes of the cytochrome P450 family (*CYP* enzymes), giving rise to a variety of oxidized cholesterol derivatives, namely oxysterols and bile acids. Even worse, some of these oxidized

cholesterols are further metabolized to various steroidal hormones, including estrogens and androgens. All these can bind with and activate a variety of nuclear receptors, which in turn, activate multiple types of growth-factor receptors, hence accelerating cell proliferation in metastatic cancers in general. This is clearly a fascinating, potentially transforming proposal. Based on their model, the authors suggest that terminating cholesterol influx is a key to stopping the explosive growth of a metastatic cancer. Going even further, they suggest that the drugs currently being used to inhibit metastatic cancer are not particularly relevant to the main driver of the disease, which is fundamentally different from what drives a primary cancer. From here, the authors made a bold statement that calling metastatic cancers as a terminal illness based on previous treatment results is premature.

Chapter 12, not as heavy as some of the earlier chapters, presents possible ways to computationally predict biomarkers in human body fluids for cancer detection. Reading of this chapter should be considerably easier compared to the previous seven chapters. The chapter consists of presentations of three computational techniques for predicting proteins that can travel from the diseased cells to blood circulation, urine and even saliva, all based on computational classification techniques trained on proteins that have been detected in each of these three human body fluids *versus* those the authors consider as unlikely to get into these body fluids. In contrast to the previous chapters that are mostly focused on fundamental cancer biology, this chapter deals with how such knowledge can be used to effectively search for biomarkers for cancer diagnosis.

Chapter 13 showcased a few examples of how to use the relevant public databases to conduct basic cancer research through data analysis and computational prediction. Both the research questions posed by the authors and the list of public databases can serve as a useful starting point for someone to enter this type of research through mining cancer omic data.

Chapter 14, the last chapter of the book, is fun to read although some of the views presented may potentially be controversial. Overall it presents the authors' views about omic data-based cancer research ahead.

By going through this information dense and intellectually unique book, we feel that this book could benefit from including some additional material as a comprehensive cancer biology book. Our wish list includes: (1) more discussion on specific cancer types; (2) comparative analyses across different classes of cancers including blood cancers and pediatric cancers; (3) discussions on cancers without an obvious Warburg effect, such as prostate cancer; (4) cancer immunity as mentioned earlier; and (5) some discussions on biology related to cancer recurrence, which is completely lacking now.

2. What a Reader Can Possibly Learn from this Book

This is a fascinating book, but we must say that it is not the easiest book to read because of the wide spectrum of topics covered, the very unique views about a variety

of cancer biology questions, and the unique presentation style to discuss cancer biology issues from a very informatics perspective. However, we also must say that after reading the book in its entirety, we feel that our overall understanding of fundamental cancer biology is elevated to a new level as a result of thinking about the long list of thought-provoking discussions throughout this book. We believe that a reader can benefit from reading the book in the following areas, regardless of whether his or her background is in bioinformatics or cancer biology.

Understanding cancer as an evolving system at a tissue level: Throughout the book, the authors tried repeatedly to convey their message that cancer is an evolutionary process driven, like any evolving system, by its changing environment, which may have cast stress on the underlying tissue (not just cells), leading to altered metabolism by the affected cells as their adaptation to the various stresses. One point that one needs to remember is that increased production of ROS is often part of the response when an organism is under attack,[5] which leads to increased mutation rates. Neoplastic cells generally tend to be under increasing stress and clearly have increased ROS, which leads to increased mutations and hence opportunities for survival. An intriguing point made by the authors in multiple sections is that cell proliferation is part of the process to reduce stress. To help to convince the reader that it is stressors that drive cancer evolution, the authors cited the study by evolutionary biologist Leigh van Valen on the Red Queen Hypothesis[6] multiple times, as well as a more recent study on drivers of antagonistic co-evolution between a bacterial population and bacterial phages.[7] The main conclusion drawn from this assessment is that antagonistic coevolution is a cause of rapid and divergent evolution and is likely to be a major driver of evolutionary changes within species.[7]

A practical benefit to the reader of studying cancer using an evolutionary framework is that it helps to untangle the exceedingly complex relationships among many relevant factors contributing to and resulting from the disease. Knowing that cancer tends to involve multiple vicious cycles, it becomes very challenging to sort out which are the causes *versus* the results among the relevant factors. The stress-based model proposed by the authors focuses on stresses and cellular responses, offering a framework, particularly when applied in conjunction with evolutionary stage information, for untangling the complex relationships and thus enabling data mining-based validation of proposed hypotheses regarding causal relationships. While the current framework may need further development, it provides a foundation for omics-based cancer researchers to study cancer as an evolving system.

Another key point that has been repeatedly emphasized by the authors is that cancer is a tissue level rather than cell level problem alone. The difference is that if cancer is indeed a survival process under certain stresses, the survival pathways available to cells in tissues may be quite different from those of cells in culture plates as cells in a tissue environment need to follow certain constraints forced on them by the tissue, such as cell proliferation in a tissue requires change in their extracellular environment, or cell growth can only take place when anchored on a flat surface.

Such a consideration clearly helps a reader to appreciate better the need for cancer studies in more realistic environments than cell culture.

Holistic versus reductionist approaches to cancer study: The authors attributed the slow progress in the current ability to cure cancer in the past few decades to the highly reductionist approaches used in cancer research. This has led to the reality that virtually all cancer drugs are targeted on proliferation and/or survival pathways rather than the (stress-associated) conditions that drive or inhibit such pathways, hence the fundamental problem of drug resistance by cancer cells. This is clearly an interesting perspective. Reductionist approaches have proved to be highly successful when studying human anatomy, which laid the foundation of modern surgical science, one of the main procedures for treating many human illnesses. However, unlike human anatomical relations, cancer seems to be an issue not amenable to divide and conquer strategies. Modern system theories have discovered that many complex systems, particularly living systems with components that have substantial interplay, can exhibit properties that intrinsically cannot be explained in terms of the properties of the components; this is referred to *emerging properties*.[8] These properties are predominantly the results of specific interactions among the components of the system, rather than properties of the individual components. While studies on individual pathways and their activation conditions are clearly needed for a better understanding of cancer, it is probably the interplay among the pathways that largely determine the observed properties of the cancer. This clearly suggests that we will need more advanced techniques as well as more sophisticated ways to design experiments for information collection rather than simply treating cancer behaviors as linear combinations of its relevant pathways. The reality that cancers tend to quickly become drug resistant clearly suggests that cancers are much more sophisticated than the way we have been modeling them using overly simplistic, generally linear, experimental models. Hence, more realistic tissue-based models may become essential for cancer studies.

As the authors point out, cancer tissues and the public availability of their omic data offer an opportunity to study cancer in a more realistic manner, but this requires cancer biologists to become quantitatively educated and/or bioinformaticians to be biologically informed. A new generation of cancer researchers is clearly needed who are capable of studying cancer through analyses of cancer omic data and computational modeling of cancer evolution.

Hypothesis-driven omic data mining: A number of large databases for cancer tissue omic data have been developed and made publicly available, such as TCGA (The Cancer Genome Atlas).[9] These data contain a substantial amount of information, possibly all the information needed to understand the driving forces and the associated mechanisms of a cancer. However, extraction of such information from the BIG omic data clearly represents a nontrivial matter. The authors proposed a scheme for information discovery through hypothesis-driven data analysis. Specifically, the authors suggest that one asks questions, related to solving a specific cancer related problem, which can possibly be addressed through analyzing the available omic data.

A large number of examples are offered by the authors regarding the types of questions that can be answered through omic data analyses, which clearly gives useful guidance to a new cancer omic data analyst studying cancer biology questions. It is noteworthy that to be an effective omic data miner for cancer studies, one will need to have a general understanding about cancer biology and an in-depth understanding about a specific research topic under consideration. In addition, familiarity with the basic analysis capabilities such as data clustering, pathway enrichment analysis and inference of regulatory relationships is clearly needed. Fortunately, the book covers all these topics.

One specific strategy that the authors used in examples throughout the book is to use marker genes for each biological process under consideration and from these to infer relationships among different biological processes through detection of co-expression relationships among the relevant marker genes. Then the detected associations are used to infer possible causal relationships based on known relationships in the literature or public databases, an approach that seems to be quite effective as shown in their model building throughout the book.

Basic cancer biology, research topics and history: Some portions of the book can be used as basic introductory material to omic-age cancer biology as the first three chapters along with the first few pages of each chapter of the book have substantial coverage of the basics of cancer biology and various current research topics in cancer, along with relevant historical events. Hence, the book could potentially be used as two-semester cancer systems biology course, with the first semester covering the

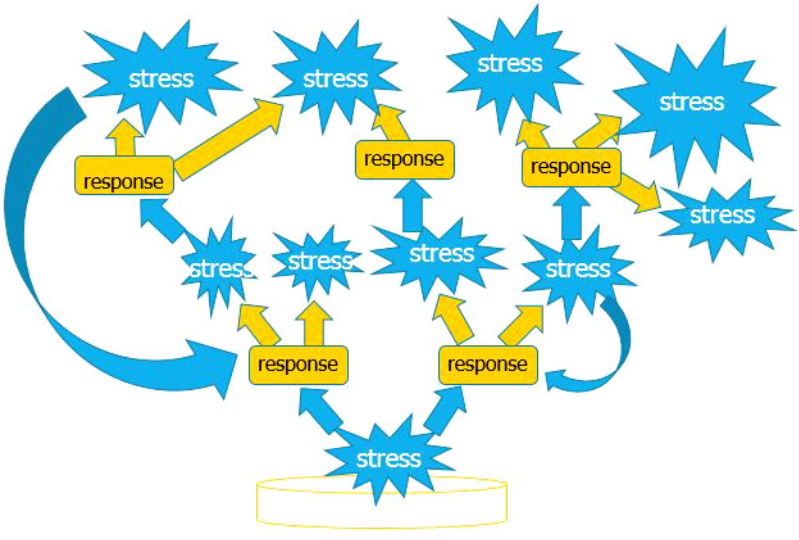

Fig. 1. A cartoon model of self-propelled vicious cycles, which lead to increasingly more challenging stresses to the underlying neoplastic cells and responses in an increasingly less normal ways, driving the evolving system deeper into an irreversible pathway to becoming a monster.

basic material and the second focused on the more research-oriented material of the book.

In summary, this is an information-packed cancer biology book, but one written in a context of how to tackle challenging fundamental biology questions of cancer through omic data mining and computational prediction and modeling. Since the book is more focused on the "big-picture" issues of cancer evolution but less detail-oriented, it is probably reasonable to read this book while having a more traditional cancer biology book handy to help to fill the gaps. Figure 1 summarizes our interpretation of a key message conveyed throughout this book: cancer is the result of cell survival from stresses, which are largely induced by the cells' adaptations to earlier stresses, via alterations in their metabolisms, forming a self-propelling vicious cycle.

Acknowledgment

Chao Liu, Chi Zhang and Jing Su made equal contributions to this work.

References

1. Xu Y, Cui J, Puett D, *Cancer Bioinformatics*, Springer, 2014.
2. Townsend DW (ed.), Combined positron emission tomography-computed tomography: The historical perspective, *Semin Ultrasound CT MRI* **29**(4):232–235, 2008.
3. Bourguignon LY, Zhu H, Shao L, Chen YW, CD44 interaction with tiam1 promotes Rac1 signaling and hyaluronic acid-mediated breast tumor cell migration, *J Biol Chem* **275**(3):1829–1838, 2000.
4. Zhang C, Cao S, Xu Y, Population dynamics inside cancer biomass driven by repeated hypoxia-reoxygenation cycles, *Quant Biol* 1–15, 2014, doi:10.1007/s40484-014-0032-8.
5. Finkel T, Signal transduction by reactive oxygen species, *J Cell Biol* **194**(1):7–15, 2011, doi:10.1083/jcb.201102095.
6. Valen LV, A new evolutionary law, *Evolut Theory* **1**:1–30, 1973.
7. Paterson S, Vogwill T, Buckling A *et al.*, Antagonistic coevolution accelerates molecular evolution, *Nature* **464**(7286):275–278, 2010, doi:10.1038/nature08798.
8. Hanahan D, Weinberg Robert A, Hallmarks of cancer: The next generation, *Cell* **144**(5):646–674, 2011, doi:10.1016/j.cell.2011.02.013.
9. Weinstein JN, Collisson EA, Mills GB *et al.*, The cancer genome atlas pan-cancer analysis project, *Nat Genet* **45**(10):1113–1120, 2013.

Chao Liu is a Ph.D. candidate at Shandong University (China). Her current research is about the relationship between inflammation and development of cancer using omic data mining and biological experiments.

Chi Zhang is a Ph.D. student in the Institute of Bioinformatics at the University of Georgia and she has broad research interests on cancer microenvironmental stresses and related cellular responses.

Jing Su, Ph.D., is an Associate Professor at Jilin University (China). Her current research is about the metastatic tumors using omic data mining and biological experiments.

Dongsheng Zhang, Ph.D./M.D., is a Professor at Department of Oral and Maxillofacial Surgery, Shandong Provincial Hospital, Affiliated to Shandong University (China). His current research is on the metastasis of adenoid cystic carcinoma.

Sha Cao is a Ph.D. student in the Department of Statistics at the University of Georgia and she has broad research interests on cancer metabolism, methylation and related cellular responses.