Applied & Translational Genomics 2 (2013) 48–54

Contents lists available at ScienceDirect

Applied & Translational Genomics

journal homepage: www.elsevier.com/locate/atg

Review

Phylooncogenomics: Examining the cancer genome in the context of vertebrate evolution☆

GuangJun Zhang [a,b,*], Tracy H. Vemulapalli [a], Jer-Yen Yang [b,c]

[a] Department of Comparative Pathobiology, Purdue University, College of Veterinary Medicine, United States
[b] Purdue University Center for Cancer Research, United States
[c] Department of Basic Medical Sciences, Purdue University, College of Veterinary Medicine, United States

ARTICLE INFO

Article history:
Received 12 May 2013
Received in revised form 20 June 2013
Accepted 22 June 2013

Keywords:
Phylooncogenomics
Comparative oncogenomics
Evolution
Vertebrate
Copy number alterations
Cancer genome

ABSTRACT

Currently, human cancer genomics is making great progress, and many mutations of new cancer driver genes have been detected at an unprecedented rate in a variety of human cancers. Many details of the genetic alterations in cancer cell genomes have been revealed by the massively parallel sequencing. Long-lasting aneuploidy caused large-scale somatic copy number alterations remains a difficulty as there are too many genes located on such big chromosomal fragments, and this cannot simply be solved by increasing sequencing depth and tumor sample numbers. Comparative oncogenomics may provide us with a solution to this problem. Here, we review some of the common animal cancer models and propose to analyze cancer cell genomics in vertebrate phylogenetic backgrounds. Thus phylooncogenomics may provide us with a unique perspective on he nature of cancer biology unattainable by single species studies.

Contents

1. Introduction

Cancer is essentially an aging-related disease, with most of adult cancers found in the latter half of the lifespan. With the increase of the expected average lifespan of human beings, cancer will continue to be one of the major health threats in the future due to its worldwide prevalence and the lack of effective treatments (Yancik, 2005). In the past several decades, substantial research endeavors have been made, because the effectiveness of controlling cancer depends on our knowledge of the nature of the disease. For example, upon learning that cancer cells usually divide more rapidly compared to normal cells, chemotherapy targeting cellular proliferation was created and it remains the most common treatment regime of cancer afterwards (Varmus, 2006). With advances in molecular and cellular mechanisms of tumorigenesis, targeting therapies using less toxic agents such as hormones, antibodies, and enzyme inhibitors were invented. The best example is imatinib, a kinase inhibitor which specifically inhibits the chimera protein, ABL–BCR, in some forms of chronic myeloid leukemia (O'Brien et al., 2003). More importantly, this specific targeting approach provides great hope in conquering this notorious disease.

One of the prerequisites of cancer targeting therapy is knowing the specific alterations that are only or mainly present in cancer cells, thus we can specifically target cancer cells in effective ways. Towards this rationale, many levels of alterations have been explored, such as histology, biochemistry, metabolism, and genetics (Pierce et al.,

* Corresponding author at: VPTH Building 103A, 725 Harrison Street, West Lafayette, IN 47907-2027, United States. Tel.: +1 765 496 1523; fax: +1 765 494 9830.
E-mail address: gjzhang@purdue.edu (G. Zhang).

2212-0661/$ – see front matter
http://dx.doi.org/10.1016/j.atg.2013.06.002

1978). With the advances of nucleic acid sequencing and microarrays, genetic changes in the cancer genome have been discovered at an unprecedented rate. These changes include: single nucleotide point mutations; frame reading shifting mutations caused by both insertions and deletions; chromosomal rearrangements including translocations, inversions, and copy number changes of protein coding and regulatory regions and may occur simultaneously in a single tumor (Stratton et al., 2009). Additionally, the mutation repertoires can be different among cancer cells in a single tumor; this is even evident by the different chromosome numbers. It has been estimated that a tumor can generate billions of mutations (Klein, 2006). Clearly, not all genetic alterations equally contribute to cancer progression. Mutations that are positively selected for and are advantageous in growth, tissue invasion and metastasis are defined as "drivers." Mutations that are byproducts of genomic instability, are not selected for, and do not confer cancer development are named "passengers." Very recently, a "mut-driver" definition has been proposed to precisely describe the genes whose mutations could cause cancers (Vogelstein et al., 2013).

Identifying cancer drivers is one of the central goals of current cancer research (Stratton et al., 2009), as not only are they essential for understanding the molecular mechanisms of cancer biology, but they may also serve as potential therapy targets and markers for diagnosis and prognosis. The current list of known "cancer genes" is about 488, according to the Cancer Gene Consensus (Futreal et al., 2004; Santarius et al., 2010). This list appears far from complete, as many new cancer driver genes are constantly being discovered with the completion of more and more cancer genomes. The International Cancer Genome Consortium (ICGC) network of cancer genome projects was initiated to target the mutational repertoire in 50 of the most common human cancer types (Hudson et al., 2010). This was only possible because of the completion of the Human Genome Project and advances in massive parallel sequencing technologies. To date, many important discoveries have been made or confirmed using these high-throughput technologies. However, even with these new technologies, identifying cancer driver genes still remains challenging due to the heterogeneities and mutational hierarchies.

2. Heterogeneity in cancer genome and somatic evolution

Tumor development is thought to occur as a somatic evolutionary process in which mutations are accumulated in a sequential manner (Merlo et al., 2006; Nowell, 1976). Like evolution on the whole organism level, the mutation process in cancer is stochastic. Very recently, evidence of Darwinian evolution has been confirmed in human pancreatic cancer and leukemia using sequencing and microarray (Campbell et al., 2008, 2010; Notta et al., 2011; Sisman and Geyikoglu, 2008; Yachida et al., 2010). Selections from the micro- and macro-environments of the cells determine which mutation(s) are retained and which ones are eliminated (Gillies et al., 2012; Merlo et al., 2006). The mutations giving cell growth advantages over surrounding cells are generally selected, and thus they are likely to be cancer drivers that give rise to the tumor cells' hallmarks (Hanahan and Weinberg, 2011; Stratton, 2011; Stratton et al., 2009). Conversely, many other bystander gene mutations cannot be eliminated rapidly enough and thus they stay in the cancer genome as passengers. Therefore, multiple levels of genetic heterogeneity (intra- and inter-tumoral, inter- and intra-metastatic, and inter-patient heterogeneities) were frequently revealed by traditional cytogenetics and recent genomic sequencing analysis (Almendro et al., 2013; Heppner, 1984; Marusyk et al., 2012). For example, it has long been recognized from earlier cytogenetic studies that there is almost no consistent karyotype in different cancer cells within the same solid tumor (Wolman, 1986). Similarly, the genes and genomes have recently been noticed in a similar situation. If tumorigenesis really is an evolutionary process, the evolutionary biological approaches, such as phylogenetic analysis, should be able to be applied to trace the natural history of cancer cells. Indeed, recently genomic sequencing and copy number analysis methods successfully tracked the cancer cell development process and the relationships between the original tumor and subsequent metastatic tumors (Campbell et al., 2008; Gerlinger et al., 2012; Navin et al., 2010; Tsao et al., 2000). This kind of information on cancer cell nature history not only is very important for us to understand the dynamic processes of tumor formation, but also might serve as guidance for therapeutic strategies.

3. Animal models of human cancers

Animal models play a very important role in our understanding of cancer biology, such as in the identification of novel cancer drivers, validating potential oncogenes and tumor suppressor genes, investigation of molecular mechanisms, and testing new cancer therapy strategies. Currently, there are several popular cancer models in the cancer research community.

The mouse model has a long history in cancer research as the most extensively used model system due to its mature genetic manipulations, relative short breeding time and the availability of inbred strains. For example, inducible and tissue-specific gene manipulation can be achieved using mouse embryonic stem cell and advanced cre-recombinase mediated knockout and knockin technologies (Cheon and Orsulic, 2011; Frese and Tuveson, 2007). In addition, N-ethyl-N-Nitrosourea (ENU) chemical, murine retrovirus-mediated (murine leukemia virus and mouse mammary tumor virus), and transposon-based mutagenesis (Sleeping Beauty and PiggyBac) have also been utilized in the mouse to identify novel cancer driver genes (Ding et al., 2005; Dupuy et al., 2005; Hrabe de Angelis et al., 2000; Kool and Berns, 2009; Nolan et al., 2000). Moreover, chromosomal engineering strategies have been successfully applied to the mouse model in order to mimic bigger regions of chromosomal deletion and duplication that frequently occur in human cancer genomes (Yu and Bradley, 2001; Yu et al., 2006). Since the mouse is a different species, it is not surprising that many differences between human and mouse tumor biology have been reported in the literature. For example, the common laboratory mouse (*Mus musculus*) possesses more active telomerase, and thus the tumors in genetically engineered mice generally possess fewer genetic alterations, including aneuploidy, when compared to corresponding tumors in humans. In order to make the mouse tumor cells aneuploid, like those of human tumors, more genes have to be manipulated in multiple pathways (Maser et al., 2007; Moens, 2008). Differences in human and mouse tumor biology are also evident with regard to the tumor type spectrums within the same orthologous cancer gene mutations. For example, mouse mutations to p53 result in multiple sarcomas and lymphomas, while human p53 mutations result predominantly in carcinomas and some sarcomas (Jacks et al., 1994). In regard to the number of essential genetic alterations to convert normal fibroblasts to tumorigenic cells, a minimum of six alterations is needed for human cells; only two are sufficient for mouse cell transformation (Rangarajan et al., 2004). Interestingly, this fits the Peto's paradox hypothesis, which suggests that there are stronger tumor repressor mechanisms in larger longer-lived animals than in smaller sized animals with shorter lifespans due to natural selection (Peto et al., 1975). Recent comparative genomics revealed that the mouse and human share about 15,213 genes. The mouse has 2785 unique genes that do not have homologous genes in human; conversely, human has 3189 genes that the mouse does not possess (Howe et al., in press). Though greater knowledge has been achieved using the mouse model, clearly human cancers cannot be completely recapitulated using this model. Thus, caution should be made when general conclusions are extrapolated from single species data.

The zebrafish is rapidly becoming a popular model organism for studying cancer and a number of tumor models have been made by the transgenic expression of oncogenes or via the mutation of tumor suppressor genes (Liu and Leach, 2011; Mione and Trede, 2010). The evidence that fish can mimic human cancer comes from multiple sources. First, human oncogenes and tumor suppressor genes can induce tumors