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

in zebrafish. Human oncogene active mutants (e.g. *c-myc*, *b-raf*, *k-ras*) have been successfully used for creating zebrafish cancer models (Dovey et al., 2009; Langenau et al., 2003; Patton et al., 2005). Likewise, loss of function in the tumor suppressor gene, *tp53*, resulted in the formation of cancer in the zebrafish although it gave rise to a different tumor spectrum (Berghmans et al., 2005; Parant et al., 2010). This phenomenon that hereditary mutations in the same orthologous cancer gene may give rise to different tumors spectrums in different species is supported in the use of the mouse model. For example, certain *MET* gene mutations lead to hereditary renal carcinoma in human, while in the mouse they give rise to various sarcomas depending on the types of mutations (Graveel et al., 2005). This may be caused by the varied tissue specificity of the cancer genes in different animal models. Second, zebrafish tumors have been shown to have similar gene expression signatures to human tumors. For example, the chemically induced fish liver tumor, and conditional *k-ras* induced melanomas and rhabdomyosarcoma have been reported to share similar transcriptomes with human corresponding tumors, respectively (Dovey et al., 2009; Lam et al., 2006; Langenau et al., 2007). Third, the fish tumors are cytogenetically aneuploidy like humans. Very recently, after strictly examining zebrafish malignant peripheral nerve sheath tumors (MPNST) that were induced by *ribosomal protein* (*rp*) and *tp53* mutations, we revealed that these tumors are highly aneuploid, and that the aneuploid chromosomes show preferential presentations within a tumor and among fish tumors. For example, within a single fish tumor there is often preferential overrepresentation of chromosome 25 with simultaneous underrepresentation of chromosome 15 (Zhang et al., 2010). This finding is very similar with the aneuploidy chromosomes in human solid tumors (Wolman, 1986). Moreover, there are also recurrent focal copy number alterations (CNA) as well as human solid tumors. We succeeded in finding cancer driver genes within recurrent focal CNAs using array-comparative genomic hybridization (CGH) and Illumina deep sequencing technology (Zhang et al., 2010). Similar to the mouse cancer model, genetic manipulation such as forward genetic screening with ENU, retroviruses, and targeted mutagenesis have also been successful in zebrafish (Amsterdam et al., 1999; Mullins et al., 1994). It is not surprising that many developmental processes and disease mechanisms were found to be similar to humans. Zebrafish share 12,897 strict homologous genes with humans. This gene number is just slightly less compared with the orthologous genes between mouse and human (Howe et al., in press). Thus, it is reasonable to expect that certain properties of human cancer biology can be recapitulated using zebrafish cancer models.

While the mouse and zebrafish compose the two major cancer model organisms, there exist other vertebrate cancer models. Among them, the dog is the most prominent one. Compared to tumors of mouse and zebrafish, most dog tumors are spontaneous, like the ones in the human population (Rowell et al., 2011). Similar to humans, the cancer rate in dogs has increased in recent years due, in part, to longer life expectancy resulting from better nutrition and veterinary care. Many human cancer gene mutations, like *TP53* and *KIT*, have subsequently been identified in canine cancers (Paoloni and Khanna, 2008). Genetic variations within each dog breed have been greatly reduced compared to humans and this makes them an ideal model system for genetic association studies of disease mapping and pharmacogenomics. One disadvantage of the canine models is the few genetic manipulation approaches available to test causal effects of candidate genes. Human, mouse, and dog genomes share about 13,816 genes with a 1:1:1 orthology (Lindblad-Toh et al., 2005). Thus, similar to the mouse and zebrafish, human cancer biology cannot be completely recapitulated in the dog.

Again, there is no single species disease model that completely mimics human disease pathogenesis, as every species has a unique ecological and evolutionary niche. Employing models from multiple species and proper comparative analysis are essential for understanding the mechanisms of human diseases including cancer.

4. Comparative oncogenomics

Currently, the method predominately used for finding cancer drivers is limited to finding statistical significance among certain number of tumors, usually within a single species. The approach is greatly affected by the very small signal-to-noise ratio as well as the choice of sampling sites within a given tumor. As such, this method requires relatively large sample sizes (e.g. 250 samples per tumor in human cancer genome projects) (Hudson et al., 2010) and it is very inefficient when dealing with large-scale genetic alterations. This is especially true for large and whole chromosomal copy number changes, which are mainly caused by aneuploid chromosomes in cancer cells. Aneuploidy is one of the most common genetic changes in human tumors, especially solid tumors (Mitelman et al., 2013). The discovery of the chromosome changes can be traced back to a century ago. Theodor Boveri and David von Hanseman proposed the first hypothetical cancer origin based on chromosome number changes (Hardy and Zacharias, 2005). However, the biological meaning of this aneuploidy remains unclear. Whether aneuploidy is causative or simply a byproduct of genomic instability has been actively debated as the etiology of human cancers (Duesberg, 2005). Accumulating evidence shows that these chromosomal abnormalities are indeed not random and suggests that they play certain roles during tumorigenesis. In terms of cytogenetics, aneuploidy is synonymous with whole chromosome or arm-level copy number changes that are used for modern genomic level studies, usually with array comparative genomic hybridization (aCGH). Recent large-scale investigations showed that there were indeed common chromosome changes in human solid tumors and some alteration patterns are even shared by many solid tumors (Baudis, 2007; Beroukhim et al., 2010). In most of the published CNA analysis using a single animal model species, only the small-sized focal CNAs (or minimum common regions) are informative, as they contain manageable numbers of genes for candidate gene functional validations. In contrast, identification of cancer drivers on the large-sized CNAs, usually chromosome or chromosomal arm lengths, remains a challenge. These regions often contain hundreds to thousands of genes, each requiring functional studies — well beyond the ability of most average-size laboratories. It has been proposed that high-throughput screening using RNAi and ORF libraries could be combined for this purpose in order to narrow down the scope (Boehm et al., 2007; Hahn et al., 2009). Others have shown that integrating the gene expression data with CNAs could lead to this same purpose (Akavia et al., 2010). Here, we only focus on an alternative approach, comparative genomics and genetics, for identification and biological validation of cancer driving genetic events, using copy number alterations as an example.

Animal models, especially the mouse cancer model, have been extensively used for many years to study many aspects of human cancers, from the discovery of tumor initiation genes' to testing cancer therapeutics (Frese and Tuveson, 2007). Recently, it has been shown that some mouse tumors also contain many CNAs, including both those at the whole chromosome level and more focal changes (Kim et al., 2006; Zender et al., 2006). Like humans, recurrent focal changes in the mouse genome are useful in identifying cancer drivers. More importantly, the mouse models have helped pinpoint cancer driver genes by comparing the gene contents of recurrent CNAs in both human and mouse tumors. Likewise, it has been shown effective with the focal CNAs in mouse liver tumors, melanoma, and osteosarcoma (Kim et al., 2006; Maser et al., 2007; Molyneux et al., 2010; Zender et al., 2006). However, this approach reaches potential limitations when applied to larger regions, especially at whole chromosome levels. This is because the gene locations in humans and mice are so highly conserved that the majority of the gene linkages are nearly identical. Thus it is still not efficient in eliminating the passenger genes. Other currently available mammal cancer models, such as the dog, are similar to the mouse for comparative purposes due to their close evolutionary distance.

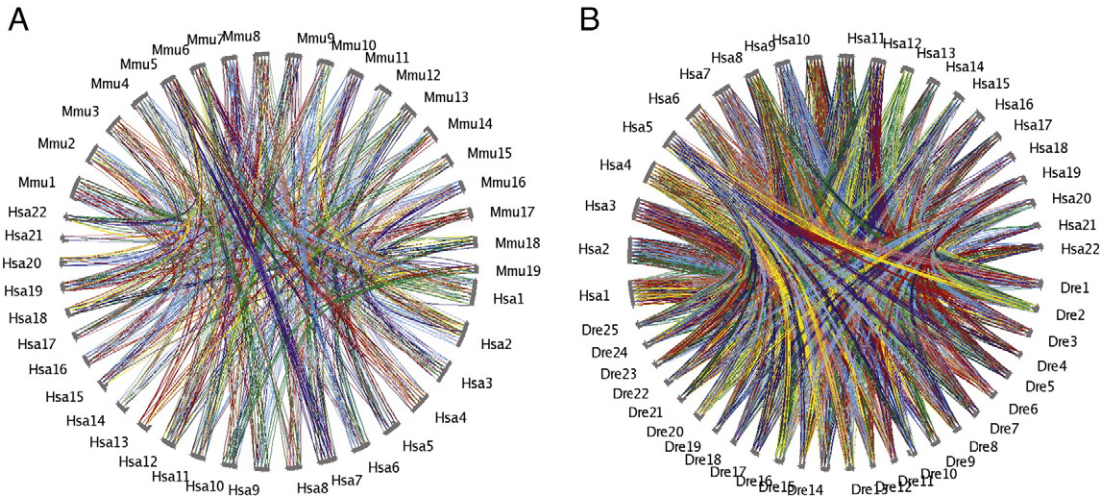

Fig. 1. Gene positions were reshuffled in human, zebrafish and mouse genomes. (A) Circos scatter plot for the chromosomal locations of orthologous genes between mouse (Mmu) and human (Hsa). (B) The syntenic relationship between zebrafish (Dre) and human (Hsa). The plot was generated using the published synteny database (Catchen et al., 2009). The orthologous gene counts in the shared chromosome fragments are connected with lines.

Due to these common features among fish, human and mouse tumors and the conservation of genes' functions, it is reasonable to assume that most driver genes within the preferential whole-chromosome and focal-CNAs are also shared among vertebrates. As these and other data show that, some preferences are specific to one type of tumor while some preferences are shared by many different tumor types. Because the orthologous gene locations along the chromosomes were reshuffled during evolution in different vertebrates, phylogenetically close species share large blocks of conserved syntenies (Kasahara et al., 2007; Kirkness et al., 2003). For example, the mouse and human genomes share large regions of conserved chromosome fragments containing similar gene contents (Fig. 1A). This is also the rationale for using the mouse to model human aneuploid syndromes (e.g., Down syndrome) (Reeves et al., 1995). On the other hand, zebrafish and humans share conserved regions to a much lesser degree; instead, most of the conserved syntenies contain much smaller fragments (Fig. 1B). It is apparent from numerous phylogenetic studies that fish genes were widely dispersed among the 23 pairs of human chromosomes during evolution (Kasahara et al., 2007; Schartl et al., 2013; Woods et al., 2000). Conversely, genes that lie on a particular human chromosome are often found together on a single mouse chromosome. The small syntenies between the zebrafish and human genome make the zebrafish cancer models uniquely useful in identifying cancer drivers on large CNAs by comparative oncogenomics. This is because the overlapping of genes between fish and human preferential whole-chromosome and focal-CNAs is much less.

In human malignant peripheral nerve sheath tumors and many other human cancers, the most frequently over-represented chromosome is 7,

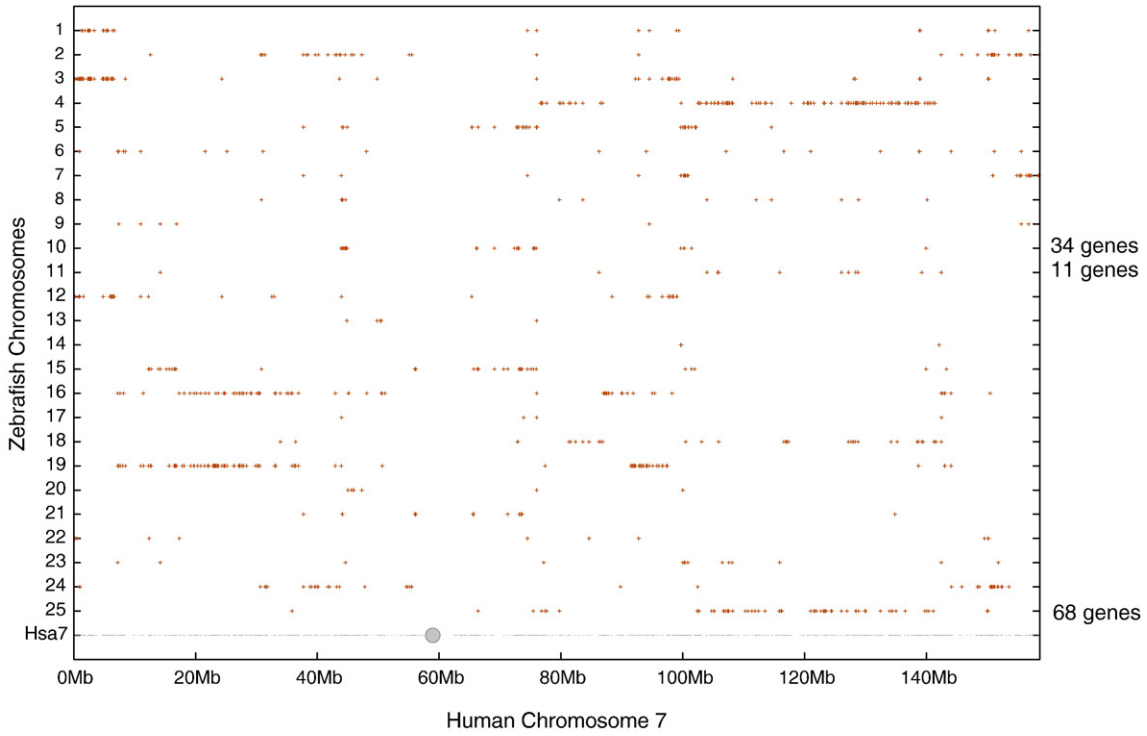

Fig. 2. Dot plot (generated from the synteny database, Catchen et al., 2009) shows the gene position relationships between human chromosome 7 and all zebrafish chromosomes.

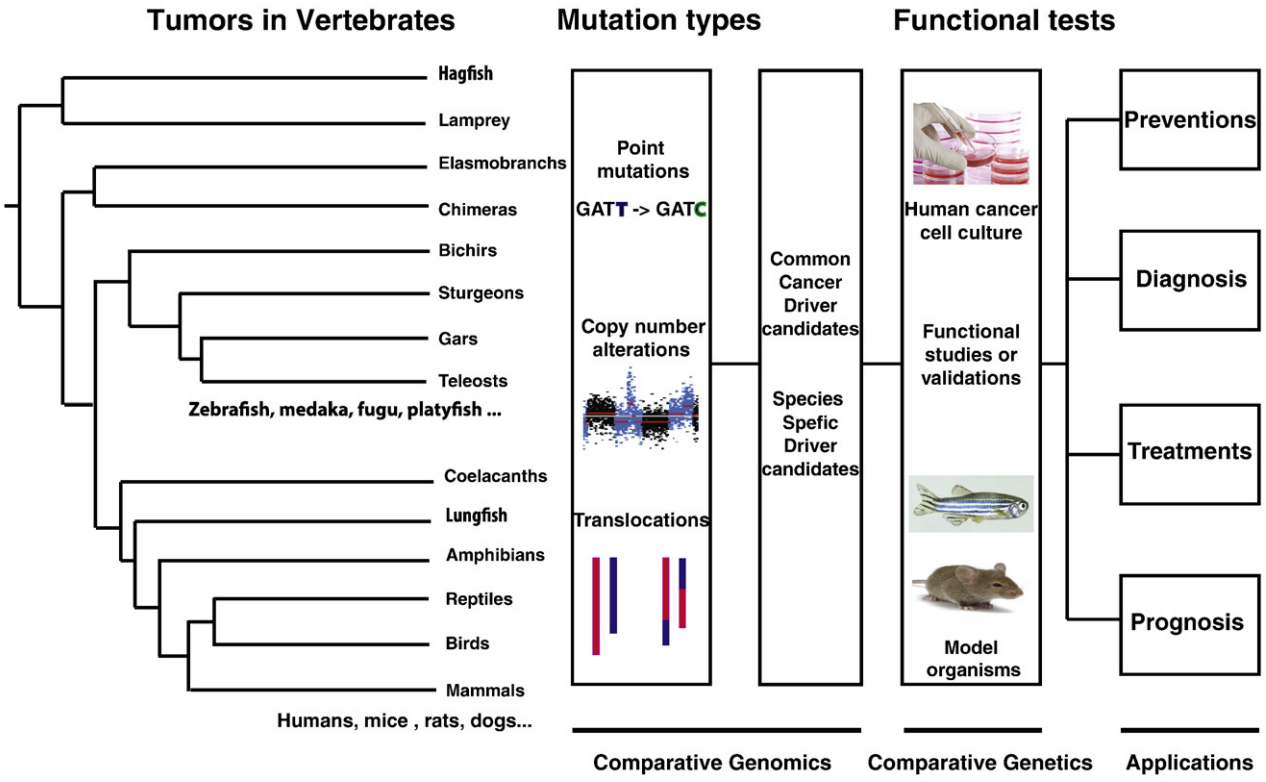

Fig. 3. Phylooncogenomic approach for cancer driver identification, and understanding the nature of cancer biology.

followed by 8q, and 17q (Beert et al., 2011; Brekke et al., 2010). In zebrafish MPNSTs, the most frequently over-represented chromosome is 25, followed by 11, 10, 23 and 22 (Zhang et al., 2010). Human chromosome 7 alone possesses ~1750 genes. In contrast, there are only 68 on fish chromosome 25, 18 on chromosome 11, and 34 on chromosome 10 (Fig. 2). As the presence of unbalanced chromosomes is the result of tumor evolution selection, these common genes could be candidates that underlie the growth advantage conferred by extra copies of human chromosome 7. The same logic can be applied to genes that lie on chromosomes that are under-represented in both fish and human tumors and generating a list of genes whose under-representation favors tumor growth. In humans, the *MET* oncogene lies on chromosome 7. Amplification of chromosome 7 carrying a mutated form of *MET* underlies an inherited cancer in humans (Lee et al., 2000; Schmidt et al., 1997). It is thought that the presence of wild-type *MET* may explain, at least in part, the frequent over-representation of chromosome 7 in many human cancers (Grabellus et al., 2010; Jenkins et al., 1998; Zhuang et al., 1998). Strikingly, the zebrafish *met* gene lies on fish chromosome 25. One of other genes also presented in the overrepresented zebrafish chromosome 25 is *fgf6a*. We have shown that *fgf* over-expression accelerates the onset of MPNSTs in combination with a p53 mutation, demonstrating that *fgf* is probably a driver in fish MPNSTs (Zhang et al., 2010). This result indicates that there are at least two drivers on chromosomes that are over-represented in fish tumors and it is possible that there are more. This discovery is consistent with the report of multiple tumor suppressors in chromosomal region 8p22 (Solimini et al., 2012; Xue et al., 2012). Combined with genetic functional validations, the zebrafish model provides us a unique model to catalogue human cancer drivers on the recalcitrant large scale CNAs.

5. Phylooncogenomics, beyond the CNAs

Although mouse–human and zebrafish–human cancer genomic comparisons are helpful to pinpoint the cancer driver genes on large CNAs in corresponding human tumors, this comparative approach should not be limited to two species only. Comparison of cancer genomics from multiple vertebrate species will be more powerful to narrow down the scope of cancer driver genes on the large scales CNAs. The early vertebrates, chondrichthyes, may be particularly interesting, as their genomes were not duplicated as teleosts (Venkatesh et al., 2007). The increased phylogenetic distance of these cartilaginous fishes to human could be useful to breakdown the chromosomally linked drivers and passengers. However, at the current time, the zebrafish is particularly useful as it is a model organism whose genomic and other biological infrastructures are much more mature compared to other non-model organisms. With advances in current sequencing technologies, we expect that other fish tumor models will play their roles in the future. For example, platyfish and lamprey genomes have recently been published (Schartl et al., 2013; Smith et al., 2013). Ultimately, multiple species tumor comparison could further reduce the passenger genes in a given tumor, thus we coined the word "phylooncogenomics" to accommodate this approach (Fig. 3). Common evolutionarily conserved cancer drivers and species-specific drivers can be identified and functionally tested in this way. For example, cancer genomes of MPNSTs from human (Beert et al., 2011; Brekke et al., 2010), mouse (Cichowski et al., 1999; Vogel et al., 1999), Tasmanian devil (Murchison et al., 2010, 2012) and zebrafish (Zhang et al., 2010) DNA could be compared in the future to identify common shared cancer drivers and biological tumorigenic mechanisms. Another potential candidate for tumor phylooncogenomic studies could be melanomas from human (Gast et al., 2010; Stark and Hayward, 2007), mouse (O'Hagan et al., 2003), zebrafish (Patton et al., 2005), medaka (Schartl et al., 2010) and platyfish (*Xiphophorus*) (Meierjohann and Schartl, 2006). The platyfish itself is a long used melanoma model as certain hybrid species naturally develop a variety of degrees of melanoma (Meierjohann and Schartl, 2006).

Though we use CNAs as the basis for phylooncogenomic analysis, this approach should not be only limited to CNAs. Other genetic abnormalities, including point/frameshift mutations, indels, chromosome translocations, even gene expression and methylation and other epigenetic patterns that present in the tumors could also be investigated in a

similar way. With phylooncogenomics, cancer genomes may be examined and compared in vertebrate phylogenetic background for common and distinct features. This is particularly important, as it will not only deepen our understanding of cancer biology, but will also give us hints to the discovery of novel cancer drugs. We have seen, on the cellular level, that a phylogenetic approach has been applied to intra-tumor heterogeneity and mutational hierarchies.

Acknowledgment

This work is supported by a grant from John T. and Winifred M. Hayward Foundation to GuangJun Zhang. Jer-Yen Yang is supported by American Cancer Society Institutional Research Grant, #58-006-53.

References

Akavia, U.D., Litvin, O., Kim, J., Sanchez-Garcia, F., Kotliar, D., Causton, H.C., Pochanard, P., Mozes, E., Garraway, L.A., Pe'er, D., 2010. An integrated approach to uncover drivers of cancer. Cell 143, 1005–1017.

Almendro, V., Marusyk, A., Polyak, K., 2013. Cellular heterogeneity and molecular evolution in cancer. Annual Review of Pathology 8, 277–302.

Amsterdam, A., Burgess, S., Golling, G., Chen, W., Sun, Z., Townsend, K., Farrington, S., Haldi, M., Hopkins, N., 1999. A large-scale insertional mutagenesis screen in zebrafish. Genes & Development 13, 2713–2724.

Baudis, M., 2007. Genomic imbalances in 5918 malignant epithelial tumors: an explorative meta-analysis of chromosomal CGH data. BMC Cancer 7, 226.

Beert, E., Brems, H., Daniels, B., De Wever, I., Van Calenbergh, F., Schoenaers, J., Debiec-Rychter, M., Gevaert, O., De Raedt, T., Van Den Bruel, A., et al., 2011. Atypical neurofibromas in neurofibromatosis type 1 are premalignant tumors. Genes, Chromosomes & Cancer 50, 1021–1032.

Berghmans, S., Murphey, R.D., Wienholds, E., Neuberg, D., Kutok, J.L., Fletcher, C.D., Morris, J.P., Liu, T.X., Schulte-Merker, S., Kanki, J.P., et al., 2005. tp53 mutant zebrafish develop malignant peripheral nerve sheath tumors. Proceedings of the National Academy of Sciences of the United States of America 102, 407–412.

Beroukhim, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., et al., 2010. The landscape of somatic copy-number alteration across human cancers. Nature 463, 899–905.

Boehm, J.S., Zhao, J.J., Yao, J., Kim, S.Y., Firestein, R., Dunn, I.F., Sjostrom, S.K., Garraway, L.A., Weremowicz, S., Richardson, A.L., et al., 2007. Integrative genomic approaches identify IKBKE as a breast cancer oncogene. Cell 129, 1065–1079.

Brekke, H.R., Ribeiro, F.R., Kolberg, M., Agesen, T.H., Lind, G.E., Eknaes, M., Hall, K.S., Bjerkehagen, B., van den Berg, E., Teixeira, M.R., et al., 2010. Genomic changes in chromosomes 10, 16, and X in malignant peripheral nerve sheath tumors identify a high-risk patient group. Journal of Clinical Oncology 28, 1573–1582.

Campbell, P.J., Pleasance, E.D., Stephens, P.J., Dicks, E., Rance, R., Goodhead, I., Follows, G.A., Green, A.R., Futreal, P.A., Stratton, M.R., 2008. Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. Proceedings of the National Academy of Sciences of the United States of America 105, 13081–13086.

Campbell, P.J., Yachida, S., Mudie, L.J., Stephens, P.J., Pleasance, E.D., Stebbings, L.A., Morsberger, L.A., Latimer, C., McLaren, S., Lin, M.L., et al., 2010. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature 467, 1109–1113.

Catchen, J.M., Conery, J.S., Postlethwait, J.H., 2009. Automated identification of conserved synteny after whole-genome duplication. Genome Research 19, 1497–1505.

Cheon, D.J., Orsulic, S., 2011. Mouse models of cancer. Annual Review of Pathology 6, 95–119.

Cichowski, K., Shih, T.S., Schmitt, E., Santiago, S., Reilly, K., McLaughlin, M.E., Bronson, R.T., Jacks, T., 1999. Mouse models of tumor development in neurofibromatosis type 1. Science 286, 2172–2176.

Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y., Xu, T., 2005. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. Cell 122, 473–483.

Dovey, M., White, R.M., Zon, L.I., 2009. Oncogenic NRAS cooperates with p53 loss to generate melanoma in zebrafish. Zebrafish 6, 397–404.

Duesberg, P., 2005. Does aneuploidy or mutation start cancer? Science 307, 41.

Dupuy, A.J., Akagi, K., Largaespada, D.A., Copeland, N.G., Jenkins, N.A., 2005. Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. Nature 436, 221–226.

Frese, K.K., Tuveson, D.A., 2007. Maximizing mouse cancer models. Nature Reviews. Cancer 7, 645–658.

Futreal, P.A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Rahman, N., Stratton, M.R., 2004. A census of human cancer genes. Nature Reviews. Cancer 4, 177–183.

Gast, A., Scherer, D., Chen, B., Bloethner, S., Melchert, S., Sucker, A., Hemminki, K., Schadendorf, D., Kumar, R., 2010. Somatic alterations in the melanoma genome: a high-resolution array-based comparative genomic hybridization study. Genes, Chromosomes & Cancer 49, 733–745.

Gerlinger, M., Rowan, A.J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., et al., 2012. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. The New England Journal of Medicine 366, 883–892.

Gillies, R.J., Verduzco, D., Gatenby, R.A., 2012. Evolutionary dynamics of carcinogenesis and why targeted therapy does not work. Nature Reviews. Cancer 12, 487–493.

Grabellus, F., Konik, M.J., Worm, K., Sheu, S.Y., van de Nes, J.A., Bauer, S., Paulus, W., Egensperger, R., Schmid, K.W., 2010. MET overexpressing chordomas frequently exhibit polysomy of chromosome 7 but no MET activation through sarcoma-specific gene fusions. Tumour Biology 31, 157–163.

Graveel, C.R., London, C.A., Vande Woude, G.F., 2005. A mouse model of activating Met mutations. Cell Cycle 4, 518–520.

Hahn, W.C., Dunn, I.F., Kim, S.Y., Schinzel, A.C., Firestein, R., Guney, I., Boehm, J.S., 2009. Integrative genomic approaches to understanding cancer. Biochimica et Biophysica Acta 1790, 478–484.

Hanahan, D., Weinberg, R.A., 2011. Hallmarks of cancer: the next generation. Cell 144, 646–674.

Hardy, P.A., Zacharias, H., 2005. Reappraisal of the Hansemann–Boveri hypothesis on the origin of tumors. Cell Biology International 29, 983–992.

Heppner, G.H., 1984. Tumor heterogeneity. Cancer Research 44, 2259–2265.

Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., et al., 2013. The zebrafish reference genome sequence and its relationship to the human genome. Nature. http://dx.doi.org/10.1038/nature12111 (in press).

Hrabe de Angelis, M.H., Flaswinkel, H., Fuchs, H., Rathkolb, B., Soewarto, D., Marschall, S., Heffner, S., Pargent, W., Wuensch, K., Jung, M., et al., 2000. Genome-wide, large-scale production of mutant mice by ENU mutagenesis. Nature Genetics 25, 444–447.

Hudson, T.J., Anderson, W., Artez, A., Barker, A.D., Bell, C., Bernabe, R.R., Bhan, M.K., Calvo, F., Eerola, I., Gerhard, D.S., et al., 2010. International network of cancer genome projects. Nature 464, 993–998.

Jacks, T., Remington, L., Williams, B.O., Schmitt, E.M., Halachmi, S., Bronson, R.T., Weinberg, R.A., 1994. Tumor spectrum analysis in p53-mutant mice. Current Biology 4, 1–7.

Jenkins, R.B., Qian, J., Lee, H.K., Huang, H., Hirasawa, K., Bostwick, D.G., Proffitt, J., Wilber, K., Lieber, M.M., Liu, W., et al., 1998. A molecular cytogenetic analysis of 7q31 in prostate cancer. Cancer Research 58, 759–766.

Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., Ahsan, B., Yamada, T., Nagayasu, Y., Doi, K., Kasai, Y., et al., 2007. The medaka draft genome and insights into vertebrate genome evolution. Nature 447, 714–719.

Kim, M., Gans, J.D., Nogueira, C., Wang, A., Paik, J.H., Feng, B., Brennan, C., Hahn, W.C., Cordon-Cardo, C., Wagner, S.N., et al., 2006. Comparative oncogenomics identifies NEDD9 as a melanoma metastasis gene. Cell 125, 1269–1281.

Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K., Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M., et al., 2003. The dog genome: survey sequencing and comparative analysis. Science 301, 1898–1903.

Klein, C.A., 2006. Random mutations, selected mutations: a PIN opens the door to new genetic landscapes. Proceedings of the National Academy of Sciences of the United States of America 103, 18033–18034.

Kool, J., Berns, A., 2009. High-throughput insertional mutagenesis screens in mice to identify oncogenic networks. Nature Reviews. Cancer 9, 389–399.

Lam, S.H., Wu, Y.L., Vega, V.B., Miller, L.D., Spitsbergen, J., Tong, Y., Zhan, H., Govindarajan, K.R., Lee, S., Mathavan, S., et al., 2006. Conservation of gene expression signatures between zebrafish and human liver tumors and tumor progression. Nature Biotechnology 24, 73–75.

Langenau, D.M., Traver, D., Ferrando, A.A., Kutok, J.L., Aster, J.C., Kanki, J.P., Lin, S., Prochownik, E., Trede, N.S., Zon, L.I., et al., 2003. Myc-induced T cell leukemia in transgenic zebrafish. Science 299, 887–890.

Langenau, D.M., Keefe, M.D., Storer, N.Y., Guyon, J.R., Kutok, J.L., Le, X., Goessling, W., Neuberg, D.S., Kunkel, L.M., Zon, L.I., 2007. Effects of RAS on the genesis of embryonal rhabdomyosarcoma. Genes & Development 21, 1382–1395.

Lee, J.H., Han, S.U., Cho, H., Jennings, B., Gerrard, B., Dean, M., Schmidt, L., Zbar, B., Vande Woude, G.F., 2000. A novel germ line juxtamembrane Met mutation in human gastric cancer. Oncogene 19, 4947–4953.

Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas III, E.J., Zody, M.C., et al., 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438, 803–819.

Liu, S., Leach, S.D., 2011. Zebrafish models for cancer. Annual Review of Pathology 6, 71–93.

Marusyk, A., Almendro, V., Polyak, K., 2012. Intra-tumour heterogeneity: a looking glass for cancer? Nature Reviews. Cancer 12, 323–334.

Maser, R.S., Choudhury, B., Campbell, P.J., Feng, B., Wong, K.-K., Protopopov, A., O'Neil, J., Gutierrez, A., Ivanova, E., Perna, I., et al., 2007. Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers. Nature 447, 966–971.

Meierjohann, S., Schartl, M., 2006. From Mendelian to molecular genetics: the *Xiphophorus* melanoma model. Trends in Genetics 22, 654–661.

Merlo, L.M., Pepper, J.W., Reid, B.J., Maley, C.C., 2006. Cancer as an evolutionary and ecological process. Nature Reviews. Cancer 6, 924–935.

Mione, M.C., Trede, N.S., 2010. The zebrafish as a model for cancer. Disease Models & Mechanisms 3, 517–523.

Mitelman, F., Johansson, B., Mertens, F.E., 2013. In: Mitelman, F., Johansson, B., Mertens, F. (Eds.), Mitelman Database of Chromosome Aberrations in Cancer (2013) (http://cgap.nci.nih.gov/Chromosomes/Mitelman).

Moens, C., 2008. Whole mount RNA in situ hybridization on zebrafish embryos: mounting. CSH Protoc 2008, pdb prot5038.

Molyneux, S.D., Di Grappa, M.A., Beristain, A.G., McKee, T.D., Wai, D.H., Paderova, J., Kashyap, M., Hu, P., Maiuri, T., Narala, S.R., et al., 2010. Prkar1a is an osteosarcoma tumor suppressor that defines a molecular subclass in mice. The Journal of Clinical Investigation 120, 3310–3325.

Mullins, M.C., Hammerschmidt, M., Haffter, P., Nusslein-Volhard, C., 1994. Large-scale mutagenesis in the zebrafish: in search of genes controlling development in a vertebrate. Current Biology 4, 189–202.

Murchison, E.P., Tovar, C., Hsu, A., Bender, H.S., Kheradpour, P., Rebbeck, C.A., Obendorf, D., Conlan, C., Bahlo, M., Blizzard, C.A., et al., 2010. The Tasmanian devil transcriptome reveals Schwann cell origins of a clonally transmissible cancer. Science 327, 84–87.

Murchison, E.P., Schulz-Trieglaff, O.B., Ning, Z., Alexandrov, L.B., Bauer, M.J., Fu, B., Hims, M., Ding, Z., Ivakhno, S., Stewart, C., et al., 2012. Genome sequencing and analysis of the Tasmanian devil and its transmissible cancer. Cell 148, 780–791.

Navin, N., Krasnitz, A., Rodgers, L., Cook, K., Meth, J., Kendall, J., Riggs, M., Eberling, Y., Troge, J., Grubor, V., et al., 2010. Inferring tumor progression from genomic heterogeneity. Genome Research 20, 68–80.

Nolan, P.M., Peters, J., Strivens, M., Rogers, D., Hagan, J., Spurr, N., Gray, I.C., Vizor, L., Brooker, D., Whitehill, E., et al., 2000. A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. Nature Genetics 25, 440–443.

Notta, F., Mullighan, C.G., Wang, J.C., Poeppl, A., Doulatov, S., Phillips, L.A., Ma, J., Minden, M.D., Downing, J.R., Dick, J.E., 2011. Evolution of human BCR–ABL1 lymphoblastic leukaemia-initiating cells. Nature 469, 362–367.

Nowell, P.C., 1976. The clonal evolution of tumor cell populations. Science 194, 23–28.

O'Brien, S.G., Guilhot, F., Larson, R.A., Gathmann, I., Baccarani, M., Cervantes, F., Cornelissen, J.J., Fischer, T., Hochhaus, A., Hughes, T., et al., 2003. Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. The New England Journal of Medicine 348, 994–1004.

O'Hagan, R.C., Brennan, C.W., Strahs, A., Zhang, X., Kannan, K., Donovan, M., Cauwels, C., Sharpless, N.E., Wong, W.H., Chin, L., 2003. Array comparative genome hybridization for tumor classification and gene discovery in mouse models of malignant melanoma. Cancer Research 63, 5352–5356.

Paoloni, M., Khanna, C., 2008. Translation of new cancer treatments from pet dogs to humans. Nature Reviews. Cancer 8, 147–156.

Parant, J.M., George, S.A., Holden, J.A., Yost, H.J., 2010. Genetic modeling of Li–Fraumeni syndrome in zebrafish. Disease Models & Mechanisms 3, 45–56.

Patton, E.E., Widlund, H.R., Kutok, J.L., Kopani, K.R., Amatruda, J.F., Murphey, R.D., Berghmans, S., Mayhall, E.A., Traver, D., Fletcher, C.D.M., 2005. BRAF mutations are sufficient to promote nevi formation and cooperate with p53 in the genesis of melanoma. Current Biology 15, 249–254.

Peto, R., Roe, F.J., Lee, P.N., Levy, L., Clack, J., 1975. Cancer and ageing in mice and men. British Journal of Cancer 32, 411–426.

Pierce, G.B., Shikes, R., Fink, L.M., 1978. Cancer, a Problem of Developmental Biology. Prentice-Hall Inc., Eaglewood Cliffs, New Jersey.

Rangarajan, A., Hong, S.J., Gifford, A., Weinberg, R.A., 2004. Species- and cell type-specific requirements for cellular transformation. Cancer Cell 6, 171–183.

Reeves, R.H., Irving, N.G., Moran, T.H., Wohn, A., Kitt, C., Sisodia, S.S., Schmidt, C., Bronson, R.T., Davisson, M.T., 1995. A mouse model for Down syndrome exhibits learning and behaviour deficits. Nature Genetics 11, 177–184.

Rowell, J.L., McCarthy, D.O., Alvarez, C.E., 2011. Dog models of naturally occurring cancer. Trends in Molecular Medicine 17, 380–388.

Santarius, T., Shipley, J., Brewer, D., Stratton, M.R., Cooper, C.S., 2010. A census of amplified and overexpressed human cancer genes. Nature Reviews. Cancer 10, 59–64.

Schartl, M., Wilde, B., Laisney, J.A., Taniguchi, Y., Takeda, S., Meierjohann, S., 2010. A mutated EGFR is sufficient to induce malignant melanoma with genetic background-dependent histopathologies. The Journal of Investigative Dermatology 130, 249–258.

Schartl, M., Walter, R.B., Shen, Y., Garcia, T., Catchen, J., Amores, A., Braasch, I., Chalopin, D., Volff, J.N., Lesch, K.P., et al., 2013. The genome of the platyfish, *Xiphophorus maculatus*, provides insights into evolutionary adaptation and several complex traits. Nature Genetics 45, 567–572.

Schmidt, L., Duh, F.M., Chen, F., Kishida, T., Glenn, G., Choyke, P., Scherer, S.W., Zhuang, Z., Lubensky, I., Dean, M., et al., 1997. Germline and somatic mutations in the tyrosine kinase domain of the MET proto-oncogene in papillary renal carcinomas. Nature Genetics 16, 68–73.

Sisman, T., Geyikoglu, F., 2008. The teratogenic effects of polychlorinated naphthalenes (PCNs) on early development of the zebrafish (*Danio rerio*). Environmental Toxicology and Pharmacology 25, 83–88.

Smith, J.J., Kuraku, S., Holt, C., Sauka-Spengler, T., Jiang, N., Campbell, M.S., Yandell, M.D., Manousaki, T., Meyer, A., Bloom, O.E., et al., 2013. Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. Nature Genetics 45, 415–421.

Solimini, N.L., Xu, Q., Mermel, C.H., Liang, A.C., Schlabach, M.R., Luo, J., Burrows, A.E., Anselmo, A.N., Bredemeyer, A.L., Li, M.Z., et al., 2012. Recurrent hemizygous deletions in cancers may optimize proliferative potential. Science 337, 104–109.

Stark, M., Hayward, N., 2007. Genome-wide loss of heterozygosity and copy number analysis in melanoma using high-density single-nucleotide polymorphism arrays. Cancer Research 67, 2632–2642.

Stratton, M.R., 2011. Exploring the genomes of cancer cells: progress and promise. Science 331, 1553–1558.

Stratton, M.R., Campbell, P.J., Futreal, P.A., 2009. The cancer genome. Nature 458, 719–724.

Tsao, J.L., Yatabe, Y., Salovaara, R., Jarvinen, H.J., Mecklin, J.P., Aaltonen, L.A., Tavare, S., Shibata, D., 2000. Genetic reconstruction of individual colorectal tumor histories. Proceedings of the National Academy of Sciences of the United States of America 97, 1236–1241.

Varmus, H., 2006. The new era in cancer research. Science 312, 1162–1165.

Venkatesh, B., Kirkness, E.F., Loh, Y.H., Halpern, A.L., Lee, A.P., Johnson, J., Dandona, N., Viswanathan, L.D., Tay, A., Venter, J.C., et al., 2007. Survey sequencing and comparative analysis of the elephant shark (*Callorhinchus milii*) genome. PLoS Biology 5, e101.

Vogel, K.S., Klesse, L.J., Velasco-Miguel, S., Meyers, K., Rushing, E.J., Parada, L.F., 1999. Mouse tumor model for neurofibromatosis type 1. Science 286, 2176–2179.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz Jr., L.A., Kinzler, K.W., 2013. Cancer genome landscapes. Science 339, 1546–1558.

Wolman, S.R., 1986. Cytogenetic heterogeneity: its role in tumor evolution. Cancer Genetics and Cytogenetics 19, 129–140.

Woods, I.G., Kelly, P.D., Chu, F., Ngo-Hazelett, P., Yan, Y.L., Huang, H., Postlethwait, J.H., Talbot, W.S., 2000. A comparative map of the zebrafish genome. Genome Research 10, 1903–1914.

Xue, W., Kitzing, T., Roessler, S., Zuber, J., Krasnitz, A., Schultz, N., Revill, K., Weissmueller, S., Rappaport, A.R., Simon, J., et al., 2012. A cluster of cooperating tumor-suppressor gene candidates in chromosomal deletions. Proceedings of the National Academy of Sciences of the United States of America 109, 8212–8217.

Yachida, S., Jones, S., Bozic, I., Antal, T., Leary, R., Fu, B., Kamiyama, M., Hruban, R.H., Eshleman, J.R., Nowak, M.A., et al., 2010. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature 467, 1114–1117.

Yancik, R., 2005. Population aging and cancer: a cross-national concern. Cancer Journal 11, 437–441.

Yu, Y., Bradley, A., 2001. Engineering chromosomal rearrangements in mice. Nature Reviews Genetics 2, 780–790.

Yu, Y.E., Morishima, M., Pao, A., Wang, D.Y., Wen, X.Y., Baldini, A., Bradley, A., 2006. A deficiency in the region homologous to human 17q21.33–q23.2 causes heart defects in mice. Genetics 173, 297–307.

Zender, L., Spector, M.S., Xue, W., Flemming, P., Cordon-Cardo, C., Silke, J., Fan, S.T., Luk, J.M., Wigler, M., Hannon, G.J., et al., 2006. Identification and validation of oncogenes in liver cancer using an integrative oncogenomic approach. Cell 125, 1253–1267.

Zhang, G., Hoersch, S., Amsterdam, A., Whittaker, C.A., Lees, J.A., Hopkins, N., 2010. Highly aneuploid zebrafish malignant peripheral nerve sheath tumors have genetic alterations similar to human cancers. Proceedings of the National Academy of Sciences of the United States of America 107, 16940–16945.

Zhuang, Z., Park, W.S., Pack, S., Schmidt, L., Vortmeyer, A.O., Pak, E., Pham, T., Weil, R.J., Candidus, S., Lubensky, I.A., et al., 1998. Trisomy 7-harbouring non-random duplication of the mutant MET allele in hereditary papillary renal carcinomas. Nature Genetics 20, 66–69.