The Somatic Nature of Cancer Allows It to Affect Highly Constrained Genes

Sheli L. Ostrow and Ruth Hershberg*

Rachel & Menachem Mendelovitch Evolutionary Processes of Mutation & Natural Selection Research Laboratory, Department of Genetics and Developmental Biology, The Ruth and Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel

*Correspondence author: E-mail: ruthersh@tx.technion.ac.il.

Accepted: May 1, 2016

Abstract

Cancer is special among genetic disorders in two major ways: first, cancer is a disease of the most basic of cellular functions, such as cell proliferation, differentiation, and the maintenance of genomic integrity. Second, in contrast to most genetic disorders that are mediated by germline (hereditary) mutations, cancer is largely a somatic disease. Here we show that these two traits are not detached and that it is the somatic nature of cancer that allows it to affect the most basic of cellular functions. We begin by demonstrating that cancer genes are both more functionally central (as measured by their patterns of expression and protein interaction) and more evolutionarily constrained than non-cancer genetic disease genes. We then compare genes that are only modified somatically in cancer (hereinafter referred to as "somatic cancer genes") to those that can also be modified in a hereditary manner, contributing to cancer development (hereinafter referred to as "hereditary cancer genes"). We show that both somatic and hereditary cancer genes are much more functionally central than genes contributing to non-cancer genetic disorders. At the same time, hereditary cancer genes are only as constrained as non-cancer hereditary disease genes, while somatic cancer genes tend to be much more constrained in evolution. Thus, it appears that it is the somatic nature of cancer that allows it to modify the most constrained genes and, therefore, affect the most basic of cellular functions.

Key words: cancer, somatic evolution, constraint, disease genes.

Introduction

Hereditary disease mutations modify genes in a manner that is harmful to the organism. The magnitudes of the effects of disease-causing germline mutations need to be sufficiently high to cause a phenotype (disease). Fitting with this, it has been demonstrated that disease genes tend to be more "important" than genes that have not been shown to be involved in genetic disorders (Cai et al. 2009; Cai et al. 2010). Gene importance can be considered from two angles:

1. Centrality, which can be defined as the extent to which the function of the gene is central to the organism, and which can be measured by such parameters as how many protein–protein interactions (PPIs) a gene has and/or in how many tissues it is expressed. It is reasonable to predict that mutations to more central genes will be more likely to lead to stronger phenotypic effects.
2. Constraint, which can be defined as how much a gene is limited by natural selection in its evolution. More "important" genes from the perspective of constraint will tend to be more conserved and have lower levels of functional variation, as measured by such estimates as the ratio of rates of non-synonymous and synonymous substitutions (dN/dS) (Nei and Gojobori 1986; Fay and Wu 2003). It is likely that genes in which mutations cause stronger phenotypic effects will be more constrained in evolution, as such strong phenotypic effects will most often be deleterious.

The levels of constraint and centrality of genes are presumably often correlated, as more functionally central genes are likely to also often be more constrained. However, this correlation is probably not perfect. Some genes that are more central will tend to be less constrained than others.

Disease genes were shown to have a higher number of PPIs than non-disease genes (Cai et al. 2010), implying higher functional centrality, and were also shown to be more conserved (Cai et al. 2009), implying higher constraint. At the same time, since mutations leading to hereditary diseases cannot harm the individual carrying them to the point that it

will not survive to birth, there may be a limit on the genes in which such mutations may occur. In other words, one may hypothesize that hereditary disease genes, while more central and constrained than the average gene, will not tend to be the most central and constrained genes.

Cancer is often a somatic disease, meaning many of the genes modified to enable the initiation and progression of cancer are not modified in a hereditary manner that affects the entire organism. Rather, they are modified somatically in a manner that affects only the cells in which the modification occurs and their decedents. Somatic alterations within tumors were shown to be subject to less stringent constraint, compared with germline alterations (McFarland et al. 2013, 2014; Ostrow et al. 2014). This implies that cancer may be allowed to somatically modify genes that tend to be more constrained in evolution, compared with other genetic hereditary disorders that are determined by germline mutations.

Additionally, cancer is a disease of extremely basic functions, such as cellular proliferation, differentiation, and maintenance of chromosomal integrity (reviewed in Hanahan and Weinberg 2000, 2011). The genes that are modified via driver mutations within cancer cells are often highly central genes which are involved in housekeeping functions such as replication timing, maintenance of DNA integrity, and apoptosis. It is, therefore, plausible that genes involved in cancer would tend to be central. Indeed, it has been shown that cancer genes tend to have more PPIs than genes that have not been associated with cancer (Rolland et al. 2014) and that they tend to more often be globally expressed across all human tissues (Ostrow et al. 2014). However, no comparison was made to date between cancer genes and genes involved in other diseases to examine whether cancer genes tend to be more central than genes contributing to non-cancer hereditary diseases.

Here, we show that cancer genes tend to be both more constrained and more conserved, not only when compared with non-disease genes but also when compared with genes that are involved in non-cancer hereditary diseases. We further provide evidence that the reason cancer can affect genes that are more constrained in evolution is the largely somatic nature of the cancer disease. Our results suggest that it is the somatic nature of cancer that allows cancer to be a disease of the most basic of cellular functions

Materials and Methods

A list of the genes currently known to be associated with cancer was downloaded from the Catalogue of Somatic Mutations in Cancer (COSMIC) on June, 2015 (Forbes et al. 2015). Cancer genes were classified as hereditary if COSMIC contained information of germline cancer mutations occurring within them (irrespective of whether they were also known to contain somatic cancer mutations). Cancer genes were classified as somatic only if no known cancer germline mutations were indicated for them.

The COSMIC database cancer gene census is broadly considered the gold standard of cancer gene datasets. It is a manually curated dataset that collects genes that were shown to be somatically mutated within tumors, more often than expected by chance (Futreal et al. 2004; Forbes et al. 2006). While the manual curation of this dataset provides clear advantages, it can also contribute to an ascertainment bias for our study. It is possible that the curators will be more likely to include genes that are more functionally central in the dataset (although it is important to note that according to the COSMIC annotation no curation was intentionally performed based on any aspect of functional importance, Futreal et al. 2004; Forbes et al. 2006). Such a curation bias would be impossible for us to control for. We, therefore, decided to use, in addition to the COSMIC dataset, three additional cancer gene datasets extracted from Kandoth et al. (2013), Lawrence et al. (2014), and Vogelstein et al. (2013). These datasets again collect genes that were identified solely based on their patterns of somatic mutation within tumors, and were not manually curated according to any parameter of functional centrality or germline constraint.

Data on disease-causing mutations were downloaded from the Online Mendelian Inheritance in Man (OMIM) database (Amberger et al. 2015) website on December, 2014. We downloaded the OMIM Morbid Map dataset, which has a list of diseases followed by single gene associations from published studies. We considered only genes annotated as "disease" but not as "susceptibility" or as "nondisease" in the OMIM Morbid Map. We removed from consideration all cancer genes annotated within OMIM (that were not annotated as cancer genes within COSMIC and that were, therefore, not included in the cancer gene groups), as well as all genes involved in non-cancer somatic diseases, by filtering out genes annotated as "cancer" or as "somatic".

The resulting data contain 487 known cancer somatic genes, 79 known cancer hereditary genes, 2,750 known non-cancer disease hereditary genes and 18,620 other genes (47 genes were removed from consideration entirely as they were classified as cancer genes according to OMIM but not according to COSMIC, or because they were involved in somatic diseases other than cancer).

Conservation scores for each gene were generated by the program phyloP (phylogenetic *P* values) (Pollard et al. 2010). The phyloP score is based on a 46-way alignment of placental genomes. We considered the average of phyloP position scores for each gene.

Data of d*N*/d*S* human–mouse per-gene ortholog values were downloaded from the Ensembl biomart (Cunningham et al. 2015).

In order to parse the different datasets, gene name conversion tables were extracted from the Ensembl biomart (Cunningham et al. 2015) and ANNOVAR (Wang et al. 2010). Gene expression data were extracted as described in TissueNet (Barshir et al. 2013). Data on protein–protein

interaction were downloaded from the MyProteinNet site (Basha et al. 2015). Data were combined into a database written in an adaptation of Simple Query Language (SQL) using the sqlite3 Python module.

Statistical tests were performed using in-house scripts and the statistical functions (stats) module of the python-based open-source software—SciPy (Jones et al. 2001).

Results

Cancer Genes Are More Constrained than Hereditary Disease Genes

We classified all human genes into those in which somatic mutations were previously associated with cancer (referred to as "cancer genes"), those that were shown to be involved in hereditary non-cancer diseases, and all other genes (Materials and Methods section). Classifications of cancer genes were carried out either based on the COSMIC database (Forbes et al. 2015) or based on three additional cancer gene datasets (Kandoth et al. 2013; Vogelstein et al. 2013; Lawrence et al. 2014). For each gene, we calculated two measures of constraint. The measures of constraint that we quantified were degree of conservation in 46 placental vertebrates, and the ratio of the rates of non-synonymous and synonymous substitutions (d*N*/d*S*) between human and mouse (Materials and Methods section).

As expected from previous results (Cai et al. 2009), hereditary non-cancer disease genes were found to score as more

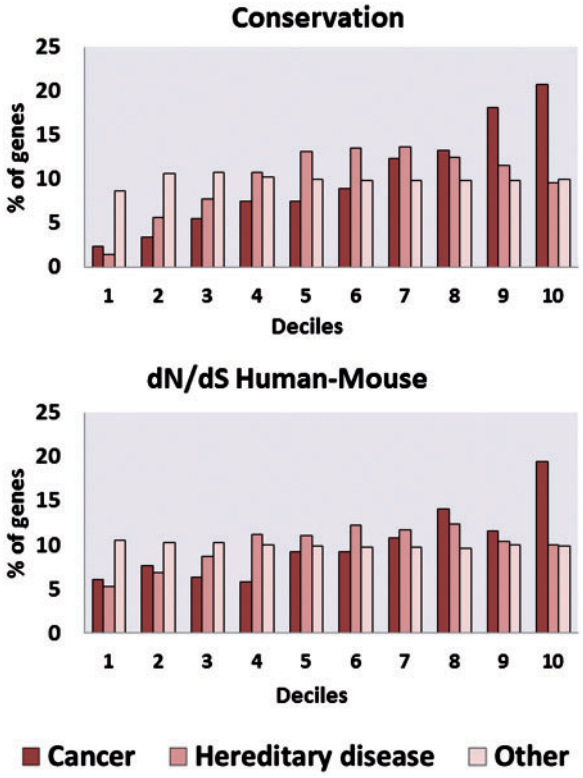

Fig. 1.—Cancer genes tend to be more constrained than non-cancer genetic disease genes. Non-cancer disease genes are in turn more constrained than genes that are not known to be involved in any diseases. For each of the two measures of constraint (conservation and human–mouse dN/dS), genes were binned into 10 equally populated bins (deciles) using all analyzed genes. This means that when one considers all analyzed genes (irrespective of their grouping) the first decile contains the 10% of genes that score as least constrained according to that measure, while the tenth decile contains the most constrained 10% of genes. We depict for each gene group (cancer vs. hereditary disease vs. other) the distribution of numbers of genes falling within each decile. Annotation of cancer genes was taken from the COSMIC database (Forbes et al. 2015).

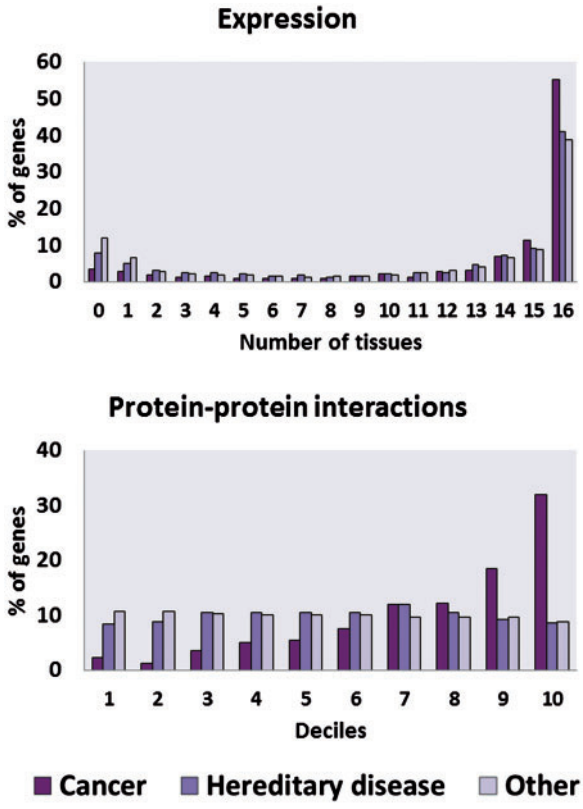

Fig. 2.—Cancer genes tend to be more functionally central than non-cancer genetic disease genes. Non-cancer disease genes are in turn more functionally central than genes that are not known to be involved in any diseases. For the tissue-expression centrality measure, genes were binned according to the number of tissues in which they were found to be expressed (out of 16 examined tissues). For the protein–protein interaction (PPI) measure, genes were binned into 10 equally populated bins (deciles) using all analyzed genes. This means that when one considers all analyzed genes (irrespective of their grouping) the first decile contains the 10% of genes that have the lowest number of PPIs, while the tenth decile contains the 10% of genes with the highest number of PPIs. We depict for each gene group (cancer vs. hereditary disease vs. other) the distribution of numbers of genes falling within each bin. Annotation of cancer genes was taken from the COSMIC database (Forbes et al. 2015).

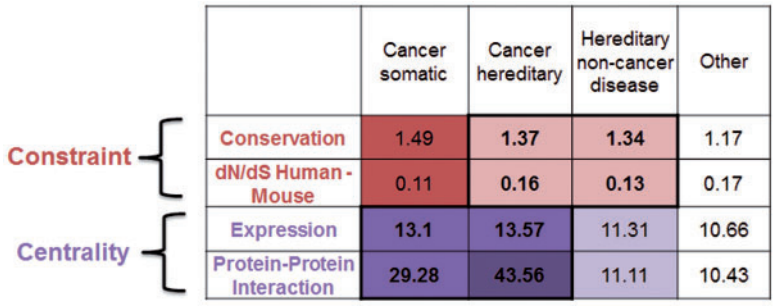

Fig. 3.—Somatic cancer genes are more constrained than both cancer and non-cancer hereditary disease genes. Hereditary cancer genes are as functionally central as somatic cancer genes but are only as constrained as other hereditary disease genes. Average values of each centrality/constraint measure are given for each gene group (somatic cancer genes vs. hereditary cancer genes vs. hereditary disease genes vs. all remaining genes). Different shades indicate that differences between values are significant ($P \ll 0.05$ according to a two-tailed Mann–Whiney test). Darker shading indicates higher centrality/constraint.

constrained, using both measures, when compared with genes that were not previously associated with diseases or cancer ($P \leq 0.0001$ for both comparisons using a two-tailed Mann–Whitney test, fig. 1). At the same time, cancer genes scored as significantly more constrained than hereditary non-cancer disease genes, according to both measures ($P \ll 0.0001$ for all comparisons). Similar results were obtained when using the COSMIC dataset to classify cancer genes ($P \ll 0.0001$, fig. 1), or when using the additional three datasets ($P \ll 0.0001$, supplementary fig. S1, Supplementary Material online).

Cancer Genes Are More Functionally Central than non-Cancer Hereditary Disease Genes

One possible explanation for the increased constraint on cancer genes relative non-cancer disease genes may be that cancer is a disease of more central functions. To test this hypothesis, we examined whether cancer genes tend to display higher functional centrality than non-cancer hereditary disease genes. Two measures of functional centrality were considered: the number of protein–protein interactions (PPIs) a gene has and the number of tissues a gene is expressed in (out of 16 examined tissues). Data were extracted from Barshir et al. (2013). We found that cancer genes tend to be expressed in a higher number of tissues and have more PPIs than non-cancer hereditary disease genes ($P \ll 0.0001$, for all comparisons, fig. 2, when using COSMIC; supplementary fig. S2, Supplementary Material online, when using the additional cancer gene datasets).

Hereditary Cancer Genes Are as, or Even More, Central than Somatic Cancer Genes, Yet They Are Significantly Less Constrained in Evolution

Our results so far demonstrate that cancer genes tend to be both more functionally central and more constrained than non-cancer hereditary disease genes. It is possible that the increased functional centrality of cancer genes is what drives their increased constraint, and that the somatic nature of cancer does not play a role. In order to examine whether the somatic nature of cancer also plays a role in explaining the increased constraint observed on cancer genes, we compared levels of functional centrality and constraint between somatic and hereditary cancer genes. Somatic cancer genes were defined as genes in which only somatic mutations were shown to contribute to cancer. In contrast, hereditary cancer genes are those genes that were also shown to sometimes harbor hereditary mutations contributing to cancer.

If the somatic nature of cancer contributes to the increased constraint observed on cancer genes, we would expect higher constraint on somatic compared with hereditary cancer genes. At the same time, we could think of no reason to expect that somatic cancer genes will be more functionally central than hereditary cancer genes, as both are likely involved in the same type of central functions. Fitting with these expectations, we found that levels of constraint were significantly lower in the hereditary cancer genes ($P \leq 0.0012$ for both measures according to a two-tailed Mann–Whitney test, fig. 3). At the same time, levels of constraint were not significantly different between the hereditary cancer genes and the non-cancer hereditary disease gene ($P \gg 0.05$ for both comparisons, fig. 3). In contrast, also fitting our expectations, functional centrality was not found to be lower in the hereditary, compared to the somatic cancer genes. Specifically, hereditary and somatic cancer genes do not significantly differ with regard to the number of tissues in which they are expressed ($P = 0.31$, fig. 3). When it comes to PPIs, hereditary cancer genes tend to have an even higher number of interactions than somatic cancer genes ($P = 0.0008$, fig. 3).

We can, therefore, conclude that when it comes to constraint, hereditary cancer genes behave more like other, non-cancer hereditary disease genes and are less constrained in evolution than somatic cancer genes. However, when it comes to measures of functional centrality, hereditary and somatic cancer genes are both significantly more functionally

central than non-cancer hereditary disease genes ($P \ll 0.0001$, fig. 3).

Discussion

Our results show that both somatic and hereditary cancer genes tend to be more functionally central than other hereditary disease genes (i.e., they tend to have higher numbers of PPIs and be expressed in more tissues). The observed higher functional centrality of cancer genes likely stems from the special nature of cancer as a disease of the most central gene pathways. At the same time, only those cancer genes that can only be modified somatically are also more evolutionarily constrained, when compared with non-cancer hereditary disease genes (as reflected by their levels of conservation in placental mammals and their patterns of sequence divergence between human and mouse). In contrast, the hereditary cancer genes are only as evolutionarily constrained, as other hereditary (non-cancer) disease genes. This demonstrates that it is the largely somatic nature of the cancer disease (i.e., the fact that most genes involved in cancer are only modified somatically and not in a hereditary manner) that allows cancer to modify genes that are more constrained in evolution.

Why would genes that are highly constrained in organismal evolution be more amenable to somatic compared with hereditary (germline) modification? Genes that are highly constrained in organismal evolution are, per definition, not very amenable to germline modification, due to strong purifying selection applied on germline mutations occurring within such genes. However, it has previously been demonstrated that the effect of purifying selection on somatic mutations within tumors is relaxed relative to what is observed in organismal evolution (McFarland et al. 2013, 2014; Ostrow et al. 2014). Such relaxation in selection is thought to be due to a combination of factors, including prevalent hitchhiking, small effective population sizes of stem cell pools, and the fact that somatic mutations affect only a small subset of cells while germline substitutions affect the entire organism (McFarland et al. 2013; Ostrow et al. 2014). Our observation that somatic cancer genes are significantly more constrained in organismal evolution compared with hereditary disease and cancer genes is likely explained by the relaxed purifying selection on tumor somatic substitutions. Thus, it appears that the highly somatic nature of cancer allows it to modify genes that are constrained enough as to not be readily accessible to hereditary diseases. This in turn implies that it may be the somatic nature of cancer that allows it to affect the very basic functions that constitute the "hallmarks of cancer" (Hanahan and Weinberg 2000, 2011). After all, genes involved in such functions will often be highly constrained in evolution. In other words, our results imply that a disorder such as cancer that affects functions that are so very basic and important may only be allowed to occur somatically.

We show that hereditary cancer genes tend to be more functionally central, yet no more constrained than genes involved in non-cancer hereditary diseases. This demonstrates the existence of a group of genes that are, on one hand, extremely functionally central, and in which, mutations can alter the most basic cellular functions in a manner that leads to cancer, but that are nevertheless not all that constrained in organismal evolution. A good example of such a gene is *TP53*, a major regulator of cell division that has long been considered to be one of the most important cancer genes, involved in almost all types of tumors (Levine and Oren 2009). Both somatic and hereditary modifications to TP53 can contribute to cancer (Petitjean et al. 2007). Fitting its important role, the *TP53* gene scores among the highest when it comes to measures of functional centrality; it has hundreds of PPIs and is expressed across all examined human tissues. In contrast, it scores quite low when it comes to measures of constraint; it is not that well conserved among placental animals and seems to tolerate a relatively high proportion of functional variation in human–mouse comparisons.

The relative high proportion of functional variation of the extremely central *TP53* gene may be partially explained by the existence of intrinsically disordered domains within its protein structure (Wells et al. 2008). These domains tend to be more diverged among species than the more conserved "ordered" parts of *TP53* (Xue et al. 2013). The presence of such domains is thought to be important for allowing disordered proteins such as *TP53* to interact with large numbers of other proteins (Babu et al. 2011). Thus, it is possible that the function of *TP53* requires that it be relatively vulnerable to alterations. The existence of disordered regions may explain why the *TP53* gene would appear to be relatively unconstrained, despite its central role.

At the same time, it is also possible that some hereditary substitutions may be allowed to persist within *TP53* because they are actually under positive selection. For example, it has been suggested that a polymorphism found in codon 72 of *TP53* may be subject to balancing selection (Beckman et al. 1994), allowing it to persist within humans despite its potential role in cancer (Storey et al. 1998). While this suggestion was later put into question (Ojeda et al. 2003), it nevertheless raises an interesting possibility that some cancer-related human polymorphisms may be maintained within the human population by positive selection. A selection trade-off by which a germline substitution in a cancer gene is advantageous early on in embryonic development or prior to the age of cancer onset, yet increases the chance of acquiring cancer later on, could lead to certain hereditary cancer genes appearing to be less constrained in organismal evolution.

It is also important to note that our ability to identify germline mutations that contribute to hereditary diseases or hereditary cancer is affected by the frequency with which these mutations segregate within the human population. More

frequent disease alleles will be easier to identify than ones that appear at extremely low frequencies. It is, therefore, possible that some alleles that are known to be somatic drivers of cancer and that are not known to contribute to hereditary cancer can in fact be hereditary drivers of cancer as well. However, since these alleles are subject to relatively strong purifying selection when they occur as germline mutations, they may be extremely rare, reducing the likelihood that they would be discovered. These alleles may be identified as somatic cancer variants because, due to the relaxed purifying selection on somatic mutations, they appear at much higher frequencies when they occur somatically. This implies that it may be useful to screen data of hereditary polymorphism among cancer patients for rare instances in which a known somatic cancer mutation appears as a germline mutation. This may allow for the identification of unknown hereditary cancer mutations that can be added to genetic screens. Additionally, given that we show that cancer genes tend to be more functionally central and constrained, it may prove useful to take gene functional centrality and constraint into account when attempting to identify novel somatic cancer genes.

In summary, we show here that cancer genes tend to be both more functionally central and more evolutionarily constrained than genes contributing to non-cancer genetic disorders. We further show that the somatic nature of cancer allows it to modify genes that are more constrained in evolution. This in turn allows cancer to affect the most basic and constrained cellular functions that could not be easily modified in a hereditary manner.

Supplementary Material

Supplementary figures S1–S3 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

R.H. was supported by an ERC FP7 CIG grant (No. 321780), by a BSF grant (No. 2013463) by a Yigal Allon Fellowship awarded by the Israeli Council for Higher Education, by the Rappaport Family Institute for Research in the Medical Sciences, and by the Robert J. Shillman Career Advancement Chair. The described research was carried out in the Rachel & Menachem Mendelovitch Evolutionary Process of Mutation & Natural Selection Research Laboratory.

Literature Cited

Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. 2015. OMIM.org: Online Mendelian Inheritance in Man (OMIM(R)), an online catalog of human genes and genetic disorders. Nucl Acids Res. 43:D789–D798.

Babu MM, van der Lee R, de Groot NS, Gsponer J. 2011. Intrinsically disordered proteins: regulation and disease. Curr Opin Struct Biol. 21:432–440.

Barshir R, et al. 2013. The TissueNet database of human tissue protein–protein interactions. Nucl Acids Res. 41:D841–D844.

Basha O, et al. 2015. MyProteinNet: build up-to-date protein interaction networks for organisms, tissues and user-defined contexts. Nucl Acids Res. 43:W258–W263.

Beckman G, et al. 1994. Is p53 polymorphism maintained by natural selection? Human Heredity 44:266–270.

Cai JJ, Borenstein E, Chen R, Petrov DA. 2009. Similarly strong purifying selection acts on human disease genes of all evolutionary ages. Genome Biol Evol. 1:131–144.

Cai JJ, Borenstein E, Petrov DA. 2010. Broker genes in human disease. Genome Biol Evol. 2:815–825.

Cunningham F, Amode MR, Barrell D, et al. 2015. Ensembl 2015. Nucl Acids Res. 43:D662–D669.

Fay JC, Wu CI. 2003. Sequence divergence, functional constraint, and selection in protein evolution. Annu Rev Genomics Human Genet. 4:213–235.

Forbes S, Clements J, Dawson E, et al. 2006. Cosmic 2005. Br J Cancer 94:318–322.

Forbes SA, Beare D, Gunasekaran P, et al. 2015. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucl Acids Res. 43:D805–D811.

Futreal PA, et al. 2004. A census of human cancer genes. Nat Rev Cancer 4:177–183.

Hanahan D, Weinberg RA. 2000. The hallmarks of cancer. Cell 100:57–70.

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

Jones E, Oliphant T, Peterson P. 2001. SciPy: open source scientific tools for python. Online.

Kandoth C, et al. 2013. Mutational landscape and significance across 12 major cancer types. Nature 502:333–339.

Lawrence MS, et al. 2014. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 505:495–501.

Levine AJ, Oren M. 2009. The first 30 years of p53: growing ever more complex. Nat Rev Cancer 9:749–758.

McFarland CD, Korolev KS, Kryukov GV, Sunyaev SR, Mirny LA. 2013. Impact of deleterious passenger mutations on cancer progression. Proc Natl Acad Sci U S A. 110:2910–2915.

McFarland CD, Mirny LA, Korolev KS. 2014. Tug-of-war between driver and passenger mutations in cancer and other adaptive processes. Proc Natl Acad Sci U S A. 111:15138–15143.

Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol. 3:418–426.

Ojeda JM, et al. 2003. p53 codon 72 polymorphism and risk of cervical cancer. Biol Res. 36:279–283.

Ostrow SL, Barshir R, DeGregori J, Yeger-Lotem E, Hershberg R. 2014. Cancer evolution is associated with pervasive positive selection on globally expressed genes. PLoS Genetics 10:e1004239.

Petitjean A, Achatz MI, Borresen-Dale AL, Hainaut P, Olivier M. 2007. TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes. Oncogene 26:2157–2165.

Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. 2010. Detection of non-neutral substitution rates on mammalian phylogenies. Genome Res. 20:110–121.

Rolland T, Tasan M, Charloteaux B, et al. 2014. A proteome-scale map of the human interactome network. Cell 159:1212–1226.

Storey A, Thomas M, Kalita A, et al. 1998. Role of a p53 polymorphism in the development of human papillomavirus-associated cancer. Nature 393:229–234.

Vogelstein B, et al. 2013. Cancer genome landscapes. Science 339:1546–1558.

Wang K, Li M, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucl Acids Res. 38:e164.
Wells M, et al. 2008. Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. Proc Natl Acad Sci U S A. 105:5762–5767.
Xue B, Brown CJ, Dunker AK, Uversky VN. 2013. Intrinsically disordered regions of p53 family are highly diversified in evolution. Biochim Biophys Acta. 1834:725–738.

Associate Editor: Partha Majumder