Biochemical and Biophysical Research Communications 465 (2015) 777–783

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Shared and unique mutational gene co-occurrences in cancers

Junqi Liu [a, b], Di Zhao [c], Ruitai Fan [a, *]

[a] Department of Radiation Oncology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China
[b] Department of Radiation Oncology, Universitätsmedizin Mannheim, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany
[c] The Institute of Experimental and Clinical Pharmacology and Toxicology, Universitätsmedizin Mannheim, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

ARTICLE INFO

Article history:
Received 12 August 2015
Accepted 20 August 2015
Available online 24 August 2015

Keywords:
Cancer treatment
Gene mutation
Network
Mutation frequency

ABSTRACT

Cancers are often associated with mutations in multiple genes; thus, studying the distributions of genes that harbor cancer-promoting mutations in cancer samples and their co-occurrences could provide insights into cancer diagnostics and treatment. Using data from the Catalogue of Somatic Mutations in Cancer (COSMIC), we found that mutated genes in cancer samples followed a power-law distribution. For instance, a few genes were mutated in a large number of samples (designated as high-frequent genes), while a large number of genes were only mutated in a few samples. This power-law distribution can be found in samples of all cancer types as well as individual cancers. In samples where two or more mutated genes are found, the high-frequent genes, i.e., those that were frequently mutated, often did not co-occur with other genes, while the other genes often tended to co-occur. Co-occurrences of mutated genes were often unique to a certain cancer; however, some co-occurrences were shared by multiple cancer types. Our results revealed distinct patterns of high-frequent genes and those that were less-frequently mutated in the cancer samples in co-occurring and anti-co-occurring networks. Our results indicated that distinct treatment strategies should be adopted for cancer patients with known high-frequent gene mutations and those without. The latter might be better treated with a combination of drugs targeting multiple genes. Our results also suggested that possible cross-cancer treatments, i.e., the use of the same drug combinations, may treat cancers of different histological origins.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Cancer is one of the leading killers of humans and is characterized by uncontrolled growth and spread of malignant cells; worldwide, there are approximately 14.1 million new cases annually. In recent years, neoadjuvant therapies, in which 'drugs' are used to act on target genes, proteins or cells, have shown some promising signs in cancer therapy. The 'drugs' may be genes, proteins or cells [1,2]. All cancers begin with genetic mutations in the genome; cells accumulate mutations and eventually proliferate uncontrollably. Thus, identifying and characterizing mutations in cancer samples is one of the first critical steps of cancer diagnosis and treatment. Until recently, most efforts toward the characterization of the mutational patterns have focused on the nucleotide level, i.e., mutational asymmetries in the four nucleotides, *ATGC*. For example, early studies revealed mutational asymmetries of A → G vs. T → C in some cancers [3] or pre-selected gene sets [4]; these findings were latter expanded to various mutational patterns and signatures [5,6] and extended to other cancer types [7–12]. Although identifying reoccurring mutations across cancer samples is one of the prime tasks of cancer research, the distributions of mutated genes in various cancers are relatively less characterized. Their co-occurrence and anti-co-occurrence patterns are even less well understood. These issues are equally important because they would help us prioritize drug targets and device drug combinations in cases where two or more genes are targeted. Most cancers, if not all, can be caused by mutations in multiple genes. To date, only a few studies have focused on disentangling gene co-occurrence and anti-co-occurrence networks. For example, Cui assembled such a network (designated as the CCA network, which stands for *C*ancer gene network with *C*o-occurring and *A*nti-co-occurring mutations) using large-scale and whole-genome sequencing of samples from multiple cancer types and argued that such a network consisted of two modules, each characterized by the enrichment and depletion

Abbreviations: CCA network, cancer gene network with co-occurring and anti-co-occurring mutations; COSMIC, the Catalogue of Somatic Mutations in Cancer.

* Corresponding author.
E-mail addresses: Junqi.Liu@medma.uni-heidelberg.de (J. Liu), di.zhao@medma.uni-heidelberg.de (D. Zhao), fanruitai@126.com (R. Fan).

http://dx.doi.org/10.1016/j.bbrc.2015.08.086
0006-291X/© 2015 Elsevier Inc. All rights reserved.

of genes of certain functions; genes in each module are tightly connected, but interactions between modules are rare [13]. However, it is still unclear if the same pattern applies to samples of individual cancer types and whether the conclusions still hold in expended datasets. In this study, we tried to address these issues by taking advantage of the comprehensive collection of cancer samples available at COSMIC, the Catalogue of Somatic Mutations in Cancer. Our results revealed a power-law distribution of mutated genes in cancer samples, i.e., the majority of the collected cancer samples contained a mutation in at least one of the five most frequently mutated genes, such as *JAK2*, *BRAF*, *KRAS*, *TP53* and *EGFR*. Similar distributions could be found in individual cancer types, although the frequencies of the high-frequent genes varied. High-frequent genes (those that were frequently mutated in samples of all cancers or an individual cancer type) often anti-co-occurred with other genes, while co-occurrences were more likely to be found among less frequently mutated genes. Co-occurrence and anti-co-occurrence were often unique to specific cancer types, but some could be found in multiple cancers.

2. Results and discussion

2.1. A power-law distribution of mutated genes in cancer samples

In COSMIC (ver. 70), 6295 genes were found to contain cancer-promoting mutations ('mutations' hereafter) in 171,724 cancer samples (see Methods). On average, one gene was found to be mutated in ~40 samples, and one sample contained ~1.5 mutated genes (Tables S1 and 2). However, the distribution of disease-causing genes in the cancer samples was extremely heterogeneous, with a few that were mutated in a large number of samples, while most were only mutated in a few samples, indicating a possible power-law relationship between the two factors. One of the characteristics of a power-law relationship is that it can be converted to a linear relationship by logarithmic transformation of the data. This is indeed the case: as shown in Fig. 1A, after logarithmic transformation, a linear relationship can be observed between the number of samples in which a gene is mutated and the number of mutated genes found in a sample; in addition, we found that and this power-law distribution also applied to samples of individual cancer types (Fig. 1).

However, the top-ten most frequently mutated genes, such as *JAK2*, *BRAF*, *KRAS*, *TP53* and *EGFR* and *etc.*, seemed to be outliners of the power-law distribution (Fig. 1A, colored dots). These genes were mutated in a large number of samples, and in total, most samples collected in COSMIC contained at least one of these mutated genes (Table S3), suggesting they could be ideal targets for cancer diagnostics and/or treatment.

Interestingly, we found that patients whose samples contained more cancer-promoting-mutation containing genes (designated as cancer-promoting genes) did not have significantly shorter survival times, at least not in the cancer types we investigated (Fig. 2 from Ref. [14] and TCGA, The Cancer Genome Atlas, see Methods). This finding indicates that although cancers tended to accumulate mutations over time, the effects of cancer-promoting genes on patient survival might not simply be additive.

DNA integrity and cell-cycle checkpoint and DNA repair related genes are known to allow cells to deal with DNA damage and control cell reproduction [15,16]. We thus check in our data if related genes were prone to be modified, especially in samples where multiple mutated genes could be found and driver-mutations were difficult to define. Grouping samples that contained more than five mutated genes according to their primary histology sites, we found that up to 12% of samples in the resulting groups contain mutated cell-cycle and DNA repair related genes (Tables S4 and S5). However, these percentages did not differ significantly from random expectation ($p > 0.05$ in all tissues, Fisher's Exact Test; Table S4; see also Methods). These results are consistent with our observation that none of the top ten most-frequently mutated genes are cell cycle and DNA repair related (Fig 1A and Table S5). Our results suggested that the DNA integrity and cell-cycle checkpoints were likely to be bypassed indirectly in cancer samples.

2.2. Co-occurrence and anti-co-occurrence network of mutated genes in cancers

A recent study by Cui suggested that co-occurrences of mutated genes in cancers were more abundant than anti-co-occurrences [13]; although this finding generally held true in our study, the importance of anti-co-occurrences was underestimated. We found that the high-frequent genes, i.e., those that were mostly frequently mutated in cancers, were often involved in the CCA network and tended to anti-co-occur with other genes. Significant co-occurrences were indeed the most abundant, but they often occurred in genes that were mutated in a small number of cancer samples. For example, we found 76 pairs of significant co-occurrences and anti-co-occurrences in cancer samples of the urinary tract. Of those, 79% were co-occurrences (Fig. 3; green lines), and only 21% were anti-co-occurrences (Fig. 3; red lines). In addition, all of the top five most frequently mutated genes (hereafter referred to as 'high-frequent genes'; colored nodes) were involved in the CCA network, and the anti-co-occurrences among them and between them and other genes (grey nodes) were significantly enriched ($p = 0.04$, Odds Ratio (OR) = 3.19; Fisher's Exact Test (FET)). Similar results could be found in samples that originated from other tissues such as hematopoietic and lymphoid tissue ($p = 0.04$, OR = 3.19; FET), as shown in Fig. 4 as well as the central nervous system (Fig. S1) and breast (Fig. S2). Overall, among the 19 tissues (primary histology sites) of cancer samples that were collected in COSMIC, none showed enrichment of co-occurrences in the top five high-frequent genes. As shown in Figs. 3 and 4, the genes that are less frequently mutated in cancer samples (grey nodes) tend to be co-mutated (connected with green edges).

2.3. Shared and unique co-occurrence and anti-co-occurrence in cancer samples

Most of the 4023 CCA pairs (Table S6) were either unique to a certain tissue or shared by a few tissues, but their distribution could be attributed to randomness. In total, only 46 CCA pairs were shared by at least two tissues, and the sharing was statistically significant. For example, *PTCH1* and *ERBB2* were mutated in seven tissue samples, and the two genes were significantly co-mutated in five of them, including urinary tract tissue (Fig. 3) and hematopoietic and lymphoid tissue (Fig. 4) ($p = 0.04$, OR = 10.5; FET). Both genes are known to be involved in tumorigenesis. For example, *PTCH1* is a tumor suppressor; mutations in *PTCH1* could cause Gorilin syndrome [17]. *ERBB2* is a proto-oncogene and a known marker of breast cancer [18]. However, the co-occurrence of mutations in these two genes has not been reported. Genes in shared co-occurring pairs often did not have significant co-expression patterns, nor did they share more similar Gene Ontology terms than those in unique pairs (see Methods). These findings indicate that the co-occurrence was likely determined by the genes' functions, not their broad functional categories.

In this study, we revealed strikingly distinct patterns of high-frequent and other genes in the CCA network. Similar to previous results [13], we found that co-occurrences of genes with cancer-promoting mutations are much more prevalent than anti-co-

Download English Version:

https://daneshyari.com/en/article/10750139

Download Persian Version:

https://daneshyari.com/article/10750139

Daneshyari.com