Biochemical and Biophysical Research Communications 445 (2014) 734–738

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Charting the molecular links between driver and susceptibility genes in colorectal cancer

Rodrigo Arroyo [a], Miquel Duran-Frigola [a], Clara Berenguer [a], Montserrat Soler-López [a], Patrick Aloy [a,b,*]

[a] Joint IRB-BSC Program in Computational Biology, Institute for Research in Biomedicine (IRB Barcelona), c/Baldiri Reixac 10-12, Barcelona 08028, Spain
[b] Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, Barcelona 08010, Spain

ARTICLE INFO

Article history:
Available online 9 January 2014

Keywords:
Colorectal cancer
Network biology
Protein interaction networks

ABSTRACT

Despite significant advances in the identification of specific genes and pathways important in the onset and progression of colorectal cancer (CRC), mechanistic insight into the relationship between driver and susceptibility genes is needed. In this paper, we systematically explore physical interactions between causative and putative CRC susceptibility genes to reveal the molecular mechanisms involved in tumor biology. In total, we identify 622 high-confidence protein–protein interactions between 42 CRC causative and 65 candidate susceptibility genes. Among the latter, 28 are located in the CRCS9 loci, related to the etiology of CRC, and 17 are co-expressed with well-established CRC drivers, which makes them excellent candidates for further functional studies. Moreover, we find a high degree of functional coherence between connected driver and susceptibility genes, which indicates that our network-based strategy is useful to gain insight into the underlying mechanisms of those proteins with unknown roles in CRC.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Colorectal cancer (CRC) is a major cause of cancer mortality, with over 600,000 recorded deaths in 2008 [1]. It is the third most common cancer worldwide and the fourth in number of caused deaths, after lung, stomach, and liver cancers [2]. The etiological factors and pathogenetic mechanisms underlying CRC development are complex and heterogeneous, and have been traditionally classified into three types: sporadic (representing 70%), familial (more than 25%) and inherited (less than 10%) [3].

There has been a significant advance in identifying the specific driver genes and pathways important in the initiation and progression of CRC, as well as the constellation of somatic alterations that are present in sporadic CRCs, which include the WNT, RAS-MAPK, PI3K, TGF-β, P53 and DNA mismatch-repair pathways [4]. Furthermore, large-scale sequencing and expression analyses have identified numerous genes recurrently mutated or whose expression is dysregulated in colorectal tumors [5,6]. However, it is not clear what is the relationship between these altered genes and CRC drivers, and further mechanistic insights into these relationships may enable deeper understanding of the pathophysiology of CRC and help identifying novel therapeutic targets.

Recent studies have shown that disease-causing genes for many disease phenotypes often work together within the same biological module [7], be it a protein complex, a pathway or a protein interaction sub-network, highlighting a strong link between protein connectivity and disease [8,9]. Indeed, protein connectivity has been used as a powerful property for the identification of novel relevant genes in complex diseases such as breast cancer [10], Huntington's [11], schizophrenia [12], cerebral ataxias [13] or Alzheimer's [14].

Accordingly, in this paper, we set up an interaction discovery strategy to unveil novel interactions between CRC causative (driver) and putative susceptibility genes, with the aim to provide mechanistic details and understand their relatedness.

2. Materials and methods

2.1. Subcloning of human cDNAs into Y2H plasmids

55 Clones from the human ORFeome v1.1 [15], 27 from Life Technologies Ultimate™ ORF Clones [16] and 5 from the Human ORFeome Collaboration Clone were obtained as a Gateway® cloning adapted plasmids and sequence verified. 3 cDNA clones from human verified full length cDNA repository (ImaGenes) were amplified by PCR with primers containing additional nucleotides (CACC). These cDNAs were cloned into a pENTR™/D-TOPO® vector (pENTR Directional TOPO cloning kit; Life Technologies) and sequenced.

All 90 ORFs were individually transferred into yeast two-hybrid (Y2H) destination vectors by Gateway© recombinational cloning (ProQuest System, Life Technologies). Seed genes were cloned into

* Corresponding author at: Joint IRB-BSC Program in Computational Biology, Institute for Research in Biomedicine (IRB Barcelona), c/Baldiri Reixac 10-12, Barcelona 08028, Spain.
E-mail address: patrick.aloy@irbbarcelona.org (P. Aloy).

0006-291X/$ - see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bbrc.2013.12.012

pDEST32 to generate bait plasmids and candidate genes were cloned into pDEST22 to obtain prey plasmids.

2.2. Y2H co-transformation screens

Bait and prey plasmids were pair-wise co-transformed into MaV203 yeast strain in 96-well arrays, plated onto selective SD2 (lacking Leu and Trp amino acids) agar media and incubated for 48 h at 30 °C to detect colony growth. Co-transformant arrays were then replica plated onto different selective media for interaction screening. To assay the activation of the HIS3 reporter gene, SD3 (lacking Leu, Trp and His) agar plates were supplemented with 12–100 mM of 3-aminotriazole (3AT, Sigma–Aldrich), being 50 mM 3AT the optimal concentration for positive HIS3 activation colonies. Similarly, we assayed the activation of the URA3 reporter gene by plating onto SD3 (lacking Leu, Trp and Ura). Double reporter HIS3/URA3 activation was evaluated by SD4 (lacking Leu, Trp, His and Ura) agar plates supplemented with 20 mM of 3AT and LacZ reporter gene was tested by the β-galactosidase assay on a nylon membrane placed onto a SD2 agar plate. All configurations were tested in duplicate.

2.3. Y2H library screens

Y2H library screens were performed using an adult human brain cDNA prey library (ProQuest, Life Technologies). The overlapping transcriptome ratio between brain and colon/rectum tissues is 0.93, according to the TissueInfo database [17]. Yeast cells expressing individual baits were transformed with the cDNA library using the LiAc/SS carrier DNA/PEG method [18] and were screened onto selective agar media for reporter gene activation. We checked HIS3 and URA3 reporter gene activation by 7 day incubation of transformed cells at 30 °C in selective agar media (HIS- and/or URA-). We picked up positive growing colonies and cultured them in prey selective liquid medium (lacking Trp). In each screen, approximately 5×10^5 auxotrophic transformants were tested on selective plates, obtaining 20–40 positive colonies on average. The prey plasmid DNA was extracted from the cultures and the bacterial transformation of each plasmid was carried out in order to enable DNA sequencing and subsequent gene identification by BLAST search. The preys identified by the library were tested with their respective baits for activation of reporter gene expression in co-transformation assays, in a similar procedure as explained above.

2.4. Correlation in gene expression profiles

We used microarray data from [19], a compendium of gene expression profiles from 73 normal tissue and cell types. As input, we used this compendium of gene expression profiles, a list of known disease-related genes (OMIM, COSMIC and literature based seed genes) and a list of candidate disease genes. We then applied a mixture model in order to obtain correlation coefficients that were robust to the presence of noise and we fit the model using the Expectation-Maximization (EM) algorithm [20]. The procedure computes gene expression correlation coefficient among disease genes and also between known and candidate disease genes. We defined that known disease gene and a candidate disease gene as co-expressed if their EM correlation coefficient was greater than 2 and the probability of noise less than 2.

2.5. Functional coherence of the HC interactions

Functional similarity between pairs of proteins was computed using best-match average of G-SESAME semantic similarities among annotated GO biological processes [21] [UniProt annotation file (April 30, 2013), including electronically inferred terms; GO ontology file (May 6, 2013), with "cross-products" and "has part" relationships removed]. For each seed and all of its newly identified interactors, we defined functional coherence as the average of pair-wise functional similarities. The significance of this measure was assessed after assigning for 10,000 times an equal number of random proteins from the human binary interactome [22]. Analogously, we calculated a global functional coherence degree by simply averaging the functional similarities of our experimental interactome. To supply compelling evidence that functional coherence around seeds can indeed provide mechanistic insights, we conducted the same procedure with seed direct interactors that are annotated in Reactome pathways (May 6, 2013) [23].

3. Results and discussion

It has been described that, in complex diseases, causative and susceptibility genes tend to be highly interconnected [7]. Based on this observation, we sought to exploit this high interconnectivity to reveal novel direct relationships between CRC causative genes, namely drivers, and a set of putative CRC susceptibility genes. We propose that, by discovering interactions between drivers and susceptibility genes, we can help unraveling the role of the latter in the CRC phenotype and provide mechanistic details for such relationships.

To this end, we compiled a comprehensive list of 45 driver genes. Eleven of these drivers were manually selected for being well-established and characterized genes in CRC development [24] and/or being key genes in the main signaling pathways involved in colorectal cancer syndromes [25]. In addition, we included 24 genes whose defects can strongly contribute to an inherited predisposition to CRC, according to OMIM database [26]. In particular, mutations in any of those 24 genes result in a marked predisposition to colorectal cancer in any of the 2 distinct syndromes: familial adenomatous polyposis (FAP) and hereditary nonpolyposis colorectal cancer (HNPCC). Finally, by examining COSMIC database [27,28], we selected another 10 genes, whose somatic alterations are known to be crucial in sporadic CRCs. A complete list of the selected driver genes is given in Table S1A.

There is a large number of genes potentially involved in CRC, which have been associated with CRC by several assays such as expression arrays or gene linkage analyses and genome-wide association studies, but whose role in the disease is barely understood. Therefore, in order to reduce the number of putative genes and maximize the biological relevance of the discovered interactions, we selected 45 susceptibility genes after applying the following criteria. First, we prioritized genes with coordinated expression changes with CRC drivers across a compendium of normal tissues and cell types. We estimated co-expression in terms of correlation coefficients computed using an Expectation-Maximization EM algorithm [20]. Besides, in all the cases, we forced a co-expression in the colon and rectum related tissues to obtain the most relevant correlation for CRC (see Section 2). This procedure identified 17 candidate genes that did significantly co-express with the known CRC-causative genes. In addition, we considered genes within chromosomal loci identified by gene linkage analyses and genome-wide association studies as related to CRC. In particular, we included the region 9q22.32-31.1, which contains a susceptibility locus (CRCS9) involved in the development of known hereditary colorectal cancer syndromes, familial adenomatous polyposis (FAP) and hereditary non-polyposis colorectal cancer (HNPCC) [29–31]. It is known that these loci are linked to CRC, but the specific genes responsible for this linkage remain to be discovered. Accordingly, we selected 28 additional genes from this locus

Download English Version:

https://daneshyari.com/en/article/10756901

Download Persian Version:

https://daneshyari.com/article/10756901

Daneshyari.com