Computational Biology and Chemistry 60 (2016) 9–20

Contents lists available at ScienceDirect

Computational Biology and Chemistry

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

Research article

CDH1/E-cadherin and solid tumors. An updated gene-disease association analysis using bioinformatics tools

María Florencia Abascal[a,1,2], María José Besso[a,1,2], Marina Rosso[a,2], María Victoria Mencucci[a,2], Evangelina Aparicio[a,2], Gala Szapiro[a,2], Laura Inés Furlong[b], Mónica Hebe Vazquez-Levin[a,c,*]

[a] Laboratory of Cell–Cell Interaction in Cancer and Reproduction, Instituto de Biología & Medicina Experimental (IBYME), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Fundación IBYME (FIBYME), Vuelta de Obligado 2490, Zip Code C1428ADN, Buenos Aires, Argentina
[b] Research Programme on Biomedical Informatics (GRIB) (IMIM), DCEXS, Universitat Pompeu Fabra, C/Dr Aiguader 88, Zip Code 08003, Barcelona, Spain
[c] Laboratory of Cell-Cell Interaction in Cancer and Reproduction, Instituto de Biología y Medicina Experimental (IBYME), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Fundación IBYME (FIBYME), Vuelta de Obligado 2490, Zip Code C1428ADN, Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 24 May 2015
Received in revised form 17 October 2015
Accepted 19 October 2015
Available online 7 November 2015

Keywords:
CDH1
Epithelial Cadherin
Solid tumors
Bioinformatics
Somatic mutations

ABSTRACT

Cancer is a group of diseases that causes millions of deaths worldwide. Among cancers, Solid Tumors (ST) stand-out due to their high incidence and mortality rates. Disruption of cell–cell adhesion is highly relevant during tumor progression. Epithelial-cadherin (protein: E-cadherin, gene: *CDH1*) is a key molecule in cell–cell adhesion and an abnormal expression or/and function(s) contributes to tumor progression and is altered in ST. A systematic study was carried out to gather and summarize current knowledge on *CDH1*/E-cadherin and ST using bioinformatics resources. The DisGeNET database was exploited to survey *CDH1*-associated diseases. Reported mutations in specific ST were obtained by interrogating COSMIC and IntOGen tools. *CDH1* Single Nucleotide Polymorphisms (SNP) were retrieved from the dbSNP database.

DisGeNET analysis identified 609 genes annotated to ST, among which *CDH1* was listed. Using *CDH1* as query term, 26 disease concepts were found, 21 of which were neoplasms-related terms. Using DisGeNET ALL Databases, 172 disease concepts were identified. Of those, 80 ST disease-related terms were subjected to manual curation and 75/80 (93.75%) associations were validated. On selected ST, 489 *CDH1* somatic mutations were listed in COSMIC and IntOGen databases. Breast neoplasms had the highest *CDH1*-mutation rate. *CDH1* was positioned among the 20 genes with highest mutation frequency and was confirmed as driver gene in breast cancer. Over 14,000 SNP for *CDH1* were found in the dbSNP database.

This report used DisGeNET to gather/compile current knowledge on gene-disease association for *CDH1*/E-cadherin and ST; data curation expanded the number of terms that relate them. An updated list of *CDH1* somatic mutations was obtained with COSMIC and IntOGen databases and of SNP from dbSNP. This information can be used to further understand the role of *CDH1*/E-cadherin in health and disease.
© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Cancer is a group of diseases characterized by an uncontrolled cell proliferation resistance to cell death, induction of angiogenesis, activation of invasion and metastasis and growth suppressor evasiveness (Negrini et al., 2010; Hanahan and Weinberg 2011). Among different cancer types, Solid Tumors (ST) stand-out due to their high incidence and mortality rates. Over 90% of ST start in the epithelium, a tissue composed of cells interconnected by intercellular junctions, among them the adherent junctions (Cooper 2000). Disruption of cell–cell adhesion is a very relevant event during tumor progression and metastasis in ST. In the transition to malignancy, down-regulation of cell–cell adhesion

* Corresponding author at: Instituto de Biología y Medicina Experimental (IBYME), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Fundación IBYME (FIBYME), Vuelta de Obligado 2490 (C1428ADN), Buenos Aires Argentina. Fax: +5411 4786 2564.
E-mail addresses: abascalflorencia@gmail.com (M.F. Abascal), majobesso_88@hotmail.com (M.J. Besso), rossomarina84@gmail.com (M. Rosso), viki_mvm89@hotmail.com (M.V. Mencucci), aparicio.evangelina32@gmail.com (E. Aparicio), galaszapiro@gmail.com (G. Szapiro), lfurlong@imim.es (L.I. Furlong), mhvazl@gmail.com, mhvaz@levindigital.com.ar (M.H. Vazquez-Levin).
[1] These authors contributed equally to this work.
[2] Fax: +54 11 47862564.

http://dx.doi.org/10.1016/j.compbiolchem.2015.10.002
1476-9271/© 2015 Elsevier Ltd. All rights reserved.

molecules, cytoskeleton reorganization and signaling pathway activation that avoid adherent junction assembly accompanies an increase of proliferation and migration (Conacci-Sorrell et al., 2002; Le Bras et al., 2012; Vasioukhin 2012).

Adherent junctions participate in processes involved in keeping cellular organization. Among their functions are maintenance of cell polarity and tissue integrity, cytoskeletal dynamics and movement within epithelium proliferation, transcription, differentiation and survival (Ivanov and Naydenov, 2013; Perez-Moreno et al., 2003). They are composed of classical cadherins, a vast superfamily of membrane proteins that form mainly Ca^{2+}-dependent homophilic interactions to maintain cell–cell contact (Angst et al., 2001; Kemler 1992; Nishimura and Takeichi, 2009; Oda and Takeichi, 2011). Epithelial cadherin (E-cadherin) is the founder member of the cadherin superfamily (Takeichi 1977). It is a Ca^{2+}-dependent cell-adhesion molecule mainly expressed in epithelial cells, essential for development, cell differentiation and tissue homeostasis, as well as for maintenance of epithelial polarity and structural integrity (van Roy and Berx, 2008). E-cadherin localization is restricted to cell–cell contact sites, and part of the cell surface-located E-cadherin is subjected to endocytosis and recycling (Bryant and Stow, 2004; Mosesson et al., 2008). The human E-cadherin gene, named *CDH1,* is located on chromosome 16q22.1, spans a region of approximately 100 kb, and comprises 16 exons and 15 introns (Berx et al., 1995).

The E-cadherin mature protein is a 120 kDa glycoprotein organized in an extracellular domain (ectodomain) of five tandem cadherin motifs, a single transmembrane domain, and a highly conserved cytoplasmic domain. The E-cadherin extracellular domain mediates mainly homophilic cell–cell adhesions between adjacent cells (Nose et al., 1990; Ozawa et al., 1990). On the other hand, specific sequences of the E-cadherin intracellular domain participate in regulation of its adhesive activity (Ozawa and Kemler, 1998) and interact with several proteins, among them α-, β-, δ-, p120- and γ- (plakoglobin) catenins (official symbols CTNNA1, CTNNB1,CTNND2, CTNND1 and CTNNG) that form a complex link to the actin cytoskeleton, which regulates the strength of the cadherin-mediated cell adhesion, and are involved in signal transduction pathways (Nagafuchi et al., 1993; Hong et al., 2013).

The functional roles of E-cadherin anticipate that genetic and epigenetic alterations on the *CDH1* gene have great implications on tumor invasion and metastasis, with a loss or reduced expression of E-cadherin, resulting in a more invasive tumor (Gall and Frampton, 2013). E-cadherin has been defined as an invasion tumor suppressor since it has been frequently found down-regulated in epithelial tumors, a process that leads to cell motility and invasion. In fact, a reduced/lack of E-cadherin expression or/and loss of function contributes to cancer progression by increasing proliferation, invasion and metastasis (Berx and van Roy, 2009; Gheldof and Berx, 2013; Schneider and Kolligs, 2014; van Roy, 2014).

Disruption of E-cadherin expression and loss of its function(s) has been extensively documented in several ST. Examples are breast (Sinn et al., 2014), ovarian (Cowden Dahl et al., 2008), gastric (Schildberg et al., 2014), endometrial (Wójcik-Krowiranda et al., 2013), colorectal (Deng et al., 2014) and bladder (Bryan 2015) cancers. Several mechanisms of E-cadherin inactivation have been reported, among them are loss of heterozygosity at the 16q22.1 chromosome region (Chalmers et al., 2001), presence of inactivating mutations (Berx et al., 1998; Corso et al., 2014), CpG-island hypermethylation of the *CDH1* promoter (Caldeira et al., 2006; Gall and Frampton, 2013; Kanazawa et al., 2002), gene expression silencing by binding of specific transcription factors to sequences in the *CDH1* promoter (Zhang et al., 2014), and post-translational modifications (i.e. proteinase processing/phosphorylation/glycosylation) that negatively regulate E-cadherin functions (Rashid et al., 2001).

As a result of over 30 years of research since its identification, *CDH1*/E-cadherin has been the subject of numerous studies that led to a vast number of reports in scientific journals (over 21,000 publications using "E-cadherin" keyword, 11,000 publications using "E-cadherin AND cancer"; PubMed search on January 2015). This exceptional growth of information requires integrative approaches such as translational bioinformatics to transform the deluge of data into knowledge and, more importantly, to enable a deeper understanding of disease mechanisms and provide actionable information for the clinical practice (Altman, 2012; Sarkar et al., 2011). Publicly available comprehensive knowledge sources on disease genes are an important asset. The "big data" phenomenon in biomedical information is also observed in the scientific literature. Nowadays, the increasing size of literature repositories makes imperative the use of computational tools to identify relevant information and place it in the context of current biomedical knowledge. Several bioinformatics tools were developed to survey/gather information. Among them is DisGeNET, a knowledge platform on human diseases and their genes plugin for Cytoscape to query and analyze human gene-disease networks (Bauer-Mehren et al., 2010; Piñero et al., 2015). In some cases, data curation is done, and involves the identification, review and organization of the gathered information by a human expert to make it accessible to both other experts and computer systems, and it is particularly important to filter/prioritize information provided by automatic text-mining approaches (Howe et al., 2008). DisGeNET has been used for the analysis of mechanisms underlying adverse drug reactions (Bauer-Mehren et al., 2011; Grosdidier et al., 2014), the association between diabetes and Parkinson disease (Santiago and Potashkin, 2014), the prediction of disease associations for ncRNAs (Alaimo et al., 2014) and the analysis of disease-relevant nodes in metabolic pathways (Galhardo et al., 2013), among other studies.

Since tumor development has been related to the presence of gene mutations in numerous tissues, in particular in the case for the *CDH1* gene in which mutations have been reported in several publications mainly related to breast and gastric cancers (Berx et al., 1998; Corso et al., 2013, 2014; Valente et al., 2014). During the past decades, the number of reported mutations has largely increased, mainly from high-throughput approaches using next generation sequencing technologies (Pastrello et al., 2014). This information can be found in the scientific publications, and is being systematically compiled in specific databases that gather and organize the data. Among these resources are the COSMIC (Catalogue-Of-Somatic-Mutations-In-Cancer) (Forbes et al., 2010) and the IntOGen (Integrated-Onco-Genomics) (Perez-Llamas et al., 2011; Gonzalez-Perez et al., 2013) tools to search for gene mutations.

Based on the relevance of *CDH1* in human physiopathology, a systematic search was carried out to gather/summarize current knowledge on the *CDH1*/E-cadherin gene/protein and its role in human disease, in particular in cancer, using a selection of bioinformatic resources. The information contained in DisGeNET was exploited to gather diseases associated to *CDH1*, and this information was complemented with knowledge on mutations described in specific cancer samples by interrogating COSMIC and IntOGen and on SNP from dbSNP database.

2. Materials and methods

2.1. Bioinformatics tools

2.1.1. DisGeNET

Discovery platform integrating information on human diseases and their genes from expert–curated databases and the scientific literature discovered by text-mining approaches (Piñero et al.,

Download English Version:

https://daneshyari.com/en/article/14958

Download Persian Version:

https://daneshyari.com/article/14958

Daneshyari.com