Computational Toxicology 5 (2018) 16–24

Contents lists available at ScienceDirect

Computational Toxicology

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

BCScreen: A gene panel to test for breast carcinogenesis in chemical safety screening

Rachel G. Grashow[a], Vanessa Y. De La Rosa[a,b], Sean M. Watford[c], Janet M. Ackerman[a], Ruthann A. Rudel[a,*]

[a] Silent Spring Institute, 320 Nevada Street, Newton, MA 02460, United States
[b] Social Science Environmental Health Research Institute, Northeastern University, Boston, MA, United States
[c] Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, UNC-Chapel Hill, Chapel Hill, NC, United States

ARTICLE INFO

Keywords:
Breast cancer
Chemical carcinogenesis
Xenoestrogen
Endocrine disrupting chemical
Normalized pointwise mutual information

ABSTRACT

Targeted gene lists have been used in clinical settings to specify breast tumor type, and to predict breast cancer prognosis and response to treatment. Separately, panels have been curated to predict systemic toxicity and xenoestrogen activity as a part of chemical screening strategies. However, currently available panels do not specifically target biological processes relevant to breast development and carcinogenesis. We have developed a gene panel called the Breast Carcinogen Screen (BCScreen) as a tool to identify potential breast carcinogens and characterize mechanisms of toxicity. First, we used four seminal reviews to identify 14 key characteristics of breast carcinogenesis, such as apoptosis, immunomodulation, and genotoxicity. Then, using a hybrid data and knowledge-driven framework, we systematically combined information from whole transcriptome data from genomic databases, biomedical literature, the CTD chemical-gene interaction database, and primary literature review to generate a panel of 500 genes relevant to breast carcinogenesis. We used normalized pointwise mutual information (NPMI) to rank genes that frequently co-occurred with key characteristics in biomedical literature. We found that many genes identified for BCScreen were not included in prognostic breast cancer or systemic toxicity panels. For example, more than half of BCScreen genes were not included in the Tox21 S1500 + general toxicity gene list. Of the 230 that did overlap between the two panels, representation varied across characteristics of carcinogenesis ranging from 21% for genes associated with epigenetics to 82% for genes associated with xenobiotic metabolism. Enrichment analysis of BCScreen identified pathways and processes including response to steroid hormones, cancer, cell cycle, apoptosis, DNA damage and breast cancer. The biologically-based systematic approach to gene prioritization demonstrated here provides a flexible framework for creating disease-focused gene panels to support discovery related to etiology. With validation, BCScreen may also be useful for toxicological screening relevant to breast carcinogenesis.

Introduction

Breast cancer represents a significant public health concern, with approximately 250,000 new diagnoses in US women each year [2]. While heritable genetic mutations like *BRCA1* and *BRCA2* have been shown to greatly increase risk in a subset of the population [28,27,26], known non-heritable risk factors for breast cancer include exposure to pharmaceutical hormones, medical radiation, age of first birth and other aspects of reproductive history, post-menopausal body mass index (BMI), reduced physical activity, alcohol consumption, and smoking [34,1,46,25,23,38]. To date, most breast cancer research has focused on treatment, heritable gene mutations, and the behavioral factors mentioned above. However, hormone-relevant risk factors and other medical and scientific evidence suggest additional influences on non-heritable breast cancer risk, including environmental chemical exposure [14,18,38]. At least three environmental chemical classes that are likely to increase breast cancer risk have been identified: 1) chemicals that cause mammary gland tumors via DNA damage pathways, 2) endocrine disrupting chemicals (EDCs) that alter mammary tumor growth, 3) toxicants that alter susceptibility by disrupting mammary gland development [7,41,39,38]. Investigating environmental chemical risk and associated mechanisms of carcinogenesis could inform prevention efforts, shape public health policy, and also illuminate avenues for new treatments.

Currently, the *Mammary Carcinogens Review Database* names over 200 chemicals considered to be mammary carcinogens (MCs) based on

---

* Corresponding author.
E-mail address: Rudel@silentspring.org (R.A. Rudel).

https://doi.org/10.1016/j.comtox.2017.11.003
Received 26 July 2017; Received in revised form 16 November 2017; Accepted 18 November 2017
Available online 21 November 2017
2468-1113/ © 2017 Elsevier B.V. All rights reserved.

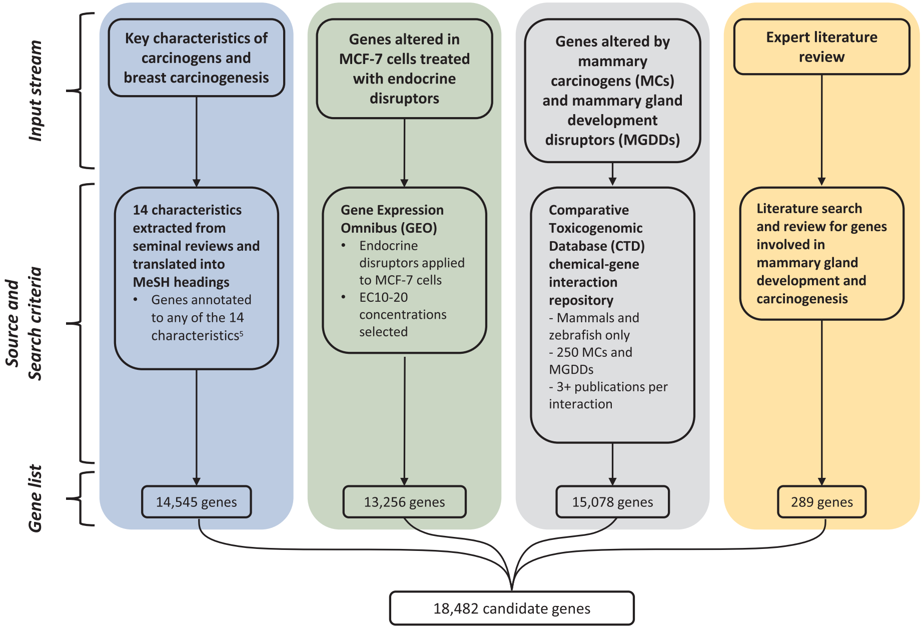

Fig. 1. Conceptual data inputs and selection criteria for candidate gene list. Legend: MeSH: Medical Subject Headings; NPMI: normalized pointwise mutual information; MCF-7: Michigan Cancer Foundation-7 mammary tissue cell line; EC: effective concentration; MC: mammary carcinogens; MGDD: mammary gland developmental disruptors.

increased mammary gland tumors in animal studies [40,39]. However, many commercial and industrial chemicals have not been evaluated for breast cancer risk at any level. With thousands of untested chemicals in use and many more being introduced to the market each year, current toxicological approaches are not sufficient to identify chemicals that may increase breast cancer risk.

A paradigm proposed by the Interagency Breast Cancer and Environmental Research Coordinating Committee (IBCERCC) and others suggests working “ … backward from a disease to identify the changes caused by chemicals that could serve as early indicators of toxicity” [9,24,43]. Such changes may serve as a link between cellular events relevant to breast cancer that are also responsive to environmental chemical exposures. For example, *in vivo* and *in vitro* studies have shown that environmental exposure to carcinogens and endocrine disruptors may exert influence via changes in gene expression [20,53,49,47]. These and other gene expression changes could be used to build a centralized list of environmentally susceptible genes that are also important in breast cancer. Such a gene list would serve as a critical tool in the evaluation of chemicals for carcinogenicity, and advance our mechanistic understanding of mammary carcinogens and mammary gland developmental disruptors.

Gene prioritization is broadly defined as the process by which the most promising genes or proteins are selected or targeted from a larger pool using systematic methods [32]. Some involve the use of a “seed” or training genes that are already associated with the endpoint or process of interest [52], while others are developed through the manual review of public biomedical and scientific databases. Targeted gene panels have previously been used in a variety of prediction contexts: to identify breast tumor subtype [10], therapeutic response [11], and likelihood of tumor recurrence [37]. Separately, a number of toxicology initiatives have sought to create sentinel or representative gene sets that can serve as markers or predictors of systemic toxicity including Tox21′s S1500+ [31], and the LINCS L1000 list [36,16]. Finally, gene panels have also been created to predict estrogen activity as a tool to identify xenoestrogens [42]. However, none of the currently available gene panels specifically target diverse biological processes relevant to breast development and carcinogenesis. This gap may be due to the difficulty in integrating multiple sources of gene data, as well as the heterogeneity of data quality and database curation.

To this end, we have developed a framework to prioritize and curate a panel of 500 genes to serve as a biomarker of mammary toxicity and breast carcinogenesis. Known as the Breast Carcinogen Screen (BCScreen), this approach represents a departure from conventional breast cancer gene platforms which focus on biomarkers for diagnosis and response to therapeutics. Instead, BCScreen is intended as a tool to identify potential breast carcinogens and the key molecular initiating events and pathways that may increase breast cancer risk in the context of chemical exposure. It can be applied in experimental studies in animals or *in vitro* systems, including high-throughput chemical screening. BCScreen synthesizes information from multiple sources including seminal papers on carcinogenesis, primary biomedical literature, whole transcriptome data from the publicly available GEO database, chemical-gene interactions from the Comparative Toxicogenomics Database (CTD) and expert literature review (ELR). In addition, this study introduces a novel application of normalized pointwise mutual information (NPMI), a data-mining technique that assigns a co-occurrence score between genes in PubMed and 14 key characteristics of breast carcinogenesis [55]. This framework combines data and knowledge driven approaches in that it relies on expert judgment for input selection and weighting, and subsequently applies a systematic approach to select genes based on those criteria and weights. This hybrid methodology allows us to integrate heterogeneous data, while maintaining flexibility to accommodate alternate input streams, model systems and weighting schema relevant to breast carcinogenesis or other disease etiologies.

Methods

In order to select genes for BCScreen, we first integrated four informative data streams including: 1) genes annotated in the scientific literature to biological processes established as key characteristics of carcinogenesis using normalized pointwise mutual information (NPMI); 2) gene expression data from endocrine disruptor exposures *in vitro*; 3) genes associated with rodent mammary carcinogens or chemicals that alter rodent mammary gland development; and 4) genes identified as involved in breast carcinogenesis through traditional expert literature review (ELR). Genes from these data sources were united into a candidate list of 18,482 genes, intended to capture all genes that may be relevant for carcinogenesis and mammary cancers. Fig. 1 shows the sources and selection criteria for candidate genes. Criteria and methods for ranking genes and identifying 500 high priority genes are described below.

Download English Version:

https://daneshyari.com/en/article/8376779

Download Persian Version:

https://daneshyari.com/article/8376779

Daneshyari.com