

90125_Auto_Edited.docx

Name of Journal: *World Journal of Clinical Cases*

Manuscript NO: 90125

Manuscript Type: EDITORIAL

Metabologenomics and network pharmacology to understand the molecular mechanism of cancer research

Tutar Y. Metabologenomics

Yusuf Tutar

Abstract

In this editorial I comment on the article “Network pharmacological and molecular docking study of the effect of Liu-Wei-Bu-Qi capsule on lung cancer” published in the recent issue of the *World Journal of Clinical Cases* 2023 November 6; 11 (31): 7593-7609. Almost all living forms are able to manufacture particular chemicals-metabolites that enable them to differentiate themselves from one another and to overcome the unique obstacles they encounter in their natural habitats. Numerous methods for chemical warfare, communication, nutrition acquisition, and stress prevention are made possible by these specialized metabolites. Metabolomics is a popular technique for collecting direct measurements of metabolic activity from many biological systems. However, confusing metabolite identification is a typical issue, and biochemical interpretation is frequently constrained by imprecise and erroneous genome-based estimates of enzyme activity. Metabolite annotation and gene integration uses a biochemical reaction network to obtain a metabolite-gene association so called metabologenomics. This network uses an approach that emphasizes metabolite-gene consensus *via* biochemical processes. Combining metabolomics and genomics data is beneficial. Furthermore, computer networking proposes that using metabolomics data may improve annotations in sequenced species and provide testable hypotheses for specific biochemical processes.

Key Words: Network pharmacology; Metabologenomics; Genome; Pathways; Cancer

Tutar Y. Metabologenomics and network pharmacology to understand the molecular mechanism of cancer research. *World J Clin Cases* 2023; In press

Core Tip: Regulation of biochemical pathways is similar to Le Chatelier chemical dynamic equilibrium principle. It states that the equilibrium is disrupted by changing the conditions, and the position of dynamic equilibrium shifts to counteract the change to reestablish an equilibrium. Metabolites disrupt pathways either themselves or by enhancing gene expression levels. The equilibrium or “homeostasis” always reestablishes itself to survive the organism. The relationship between genes and metabolites may be discovered by understanding the link through computational networking.

INTRODUCTION

Molecules and macromolecules coordinate to regulate metabolism. At cellular level the coordination seems predictable as the hormones and organs do not bring complex interaction. Recent advances in omics technologies have been useful to understand the molecular mechanism of biological processes through distinct algorithms^[1].

There are several parameters to be considered when predicting coordination among metabolites, macromolecules, genes and hormones. However, wet lab experiments may not fully compensate the *in silico* experiments since extraction of metabolites with different solutions provides distinct metabolite contents, and organism may transform the molecule to a totally different structure, *i.e.* hydroxylation but the overall information provides general concept of the molecular action. Therefore, databases and software attract scientists to find a relationship between molecules and macromolecules.

Sites for OMICs databases

Remember that the databases contain more than 220000 metabolites (The Human metabolome database, KEGG, PubChem, MetaCyc, UniProt, and GenBank). Further human genome contains more than 24000 genes (NCBI, EMBL) thus combining metabolite data, genomic data, molecular biology/biochemistry models (molecular docking, simulations, metabolite binding to protein/DNA) and clinical information provides descent information^[2]. Linking genomic content to chemical content provides new insights about the pathways involved and crosslink between these pathways. Computational methods provide link to understand correlation of expressed genes and metabolites. Understanding the combinatorial complexity of the biosynthesis of gene clusters and metabolites requires neat strategies to tweak them. *In silico* methods devise strategies to elucidate the essential factors and deciphering metabolite to gene or gene to metabolite approaches and both may provide insights to analyze critical factors on the pathways. It should be noted that none of the methods provide precise results due to limitations of detection methods as well as correlated straight forward mathematical algorithm between metabolites and genes.

Linking genes and metabolites: Metabologenic integration

Bioactivity guided fractionation of chemical extracts enabled scientist to elucidate functional metabolites and their corresponding structures. However, lag in acquiring all associated biochemical factors and methodological inadequacies render bioinformatics solutions. Therefore, data mining provides identification of pathways for both known and unknown metabolites^[3].

Acquiring metabolic data and assembly technologies of genome sequencing provide the basis of metagenome^[4]. Furthermore, recent advances in untargeted metabolomics and analysis of metabolic extracts enabled to determine identity and corresponding gene function. Universally conserved gene and metabolite function helped our understanding of continuity of life and coordinated metabolism through these molecules. One key reason for the coordination of genes and metabolites is genome sequencing and revolutionary assembling/reconstruction of genome from a variety of organisms. These

large-scale mining of metabolomes and genomic data not only provided data from distinct organisms but also integrated the metabolites and genes^[4] Computational networking or metabogenomics approaches analyze different organisms and determine common interactions and functions.

Genome mining and metabolome mining may provide details of transcriptional regulation mechanism, although abiotic and biotic environment can be different. However, annotation of metabolites is not an easy task even with tandem mass spectrometers and/or Nuclear Magnetic Resonance spectroscopies (NMRs). Public databases and analysis tools may increase annotation of metabolites, but spectral matching and molecular networking cannot provide structural information. Utilizing *in silico* annotation techniques, fragmentation patterns of a vast array of structures can be computationally learned from known metabolites. These spectra produced *in silico* are then compared to experimental tandem mass spectroscopy data. While this has been successful for certain minor metabolites, it is still too difficult to obtain high-quality annotations by spectral matching for larger natural products^[5].

Separation of metabolites by using polar/apolar solvents is another problem as polar solvent extracts polar metabolites or apolar solvents solve apolar metabolites. Also, ionization of the compounds is a must to be detected in mass spectrometry. Thus, some metabolites cannot be detected readily. Although NMR determines these compounds, the technique is not sensitive but the technique is powerful in determining structures of the metabolites.

Despite advances in technology, current analytical platforms cannot measure the whole metabolome entirely. Achieving sufficient metabolome coverage is still a significant problem as mix approaches required. The critical step in metabolomics research is to quench biochemical pathways along with metabolite isolation at a predetermined time to reflect proper metabolite proportion and concentration to mimic original living cell condition. Furthermore, sample collection and extraction must maintain the original metabolite concentration, maximize instrument repeatability, and reduce the matrix influence in the assay.

Metabologenomic integration links datasets and enables structural and functional annotation. Linking the datasets in practice is not an easy task as one needs to integrate multi omics data. Sample handling can be affected by target metabolites and is the source of the most fluctuation. Thus, standardization of these techniques is required for accurate analysis. Mass spectrometry and nuclear magnetic resonance spectroscopy are high-resolution spectroscopic techniques with distinct and interrelated applications in metabolite determination. Each technique may depend on distinct ion sources and prior separation. Standardization is important; however, this is not a demanding but not a feasible process as structures of several metabolites are not commercially available. Furthermore, despite tremendous efforts, the metabolic reprogramming of cancer cells is still not fully understood due to its mechanistic complexity and limitations of current techniques.

Metabologenomics and network pharmacology at cancer research

Reports of the metabologenomics and network pharmacology in literature indicate how the approach could be leveraged in the future to improve cancer treatment.

In China, the Xihuang pill is a popular adjuvant medication for cancer patients. One hundred and ninety overlapping candidate targets were used to build the network and enrichment analysis. ¹ The target *NR3C2* and its compound naringenin were chosen for additional validation by text mining and molecular docking. A high expression of *NR3C2* was found to support a greater survival probability regarding overall survival, as per the TCGA database. The main active ingredient of Xihuang pill, naringenin, has the ability to reduce the stemness of triple negative breast cancer cells *via* controlling the *NR3C2* gene. This study examined the impact of Xihuang pill on treating advanced triple negative breast cancer cells^[6].

Modified nucleosides and ribosylated metabolites in urine samples were the subject of an intriguing metabolomics investigation on breast cancer conducted by Hennes *et al*^[6] For this network analysis, a collection of 35 trimmed metabolites was divided into 44 pairwise combinations of metabolite characteristics. Among the metabolite pairs, S-

adenosylhomocysteine was the most frequently occurring molecule, indicating its significance for RNA methylation in the pathophysiology of cancer^[7].

Untargeted lipidomics was used in a more recent investigation to determine the characteristics that are most crucial for the early identification of lung cancer. Using mass spectrometry, lung plasma lipidomic profiling was performed on 311 individuals. A targeted assay based on mass spectrometry and liquid chromatography was developed using nine lipids. The scientists confirmed the efficacy of these nine lipids in identifying early-stage cancer in many independent cohorts, including a prospective clinical cohort and a hospital-based lung cancer screening cohort^[8].

Possible metabolic indicators for oral cavity squamous cell carcinoma were discovered by Hsu *et al*^[8] group. The group employed a method to discriminate cancer tissues from nearby non-cancerous tissues precisely by receiver operating characteristic analysis through putrescine, glycyl-leucine, and phenylalanine marker panel^[9].

Future developments

Machine learning techniques are effective solutions to examine metabolomics unstructured data relationships. Metabolomics also complements genomics and transcriptomics. These vast volumes of data form complex network for analysis. Establishing training sets for attribute selection and cross validation provides learning algorithms to get a validated model.

Choosing proper machine learning algorithm is important to determine metabolite-gen link from omics data set. However, machine learning methods have not been developed for reliable outcomes yet. Enrichment studies and pathway analysis define new roles to the identified metabolites. New and larger database, increased computation power, and contemporary machine learning techniques may provide precise diagnosis and treatment^[10].

CONCLUSION

The genome and metabolites of biological organisms are not fully characterized with current technologies. However, increasing high-throughput metabolomics and genomics data provide promising generation of paired data sets to understand the molecular mechanism of biochemical processes as well as determining targets for pharmaceutical drug design. Contemporary network infrastructures to integrate omics analysis can provide molecular mechanism of biochemical pathways. Furthermore, clinical data may be integrated to gene expression – metabolite expression by system genetics approach. Calculating pair-wise correlations and weighted correlation network analysis provide the basis of this integration^[11-13]. The occurrence of strong correlations between classified metabolites and co-expression transcripts implies either various roles of metabolites or linkages between metabolic pathways and the immune system.

1 %

SIMILARITY INDEX

PRIMARY SOURCES

1

Yu-Zhu Zhang, Jia-Yao Yang, Rui-Xian Wu, Chen Fang et al. "Network Pharmacology-Based Identification of Key Mechanisms of Xihuang Pill in the Treatment of Triple-Negative Breast Cancer Stem Cells", Frontiers in Pharmacology, 2021

Crossref

10 words — 1 %

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	< 1%
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 10 WORDS