Insilico Investigation of CYP1A1 Mediated Pathway Contribution in Liver Hepatocellular Carcinoma: A Drug Discovery Approach

Zafar Abbas Shah1 and Asima Tayyeb2*
1Department of Bioinformatics, Hazara University Mansehra, Pakistan
2Department of Biological Sciences, University of the Punjab, Pakistan

Abstract

Liver hepatocellular carcinoma has diverse nature on molecular grounds due to its multi-dimensional physiology. Cytochrome P450 family has universal driver of various metabolic and regulatory functions in liver. Aim of the current study is to access the role of extra-hepatic CYP1A1 protein mediated pathway in liver malignancy development. For this purpose, novel insilico 3-multi-layered protocol has been designed for extensive identification of drug targets of pathway responsible for tumorigenesis. STRING database (www.string-db.org) has been used for retrieval of CYP1A1 mediated pathway of high confidence score and construct functionally associated clusters. WebGestalt toolkit (www.webgestalt.com) for gene enrichment and characterization analysis of CYP1A1 network was carried out. Finally, using cBioPortal (www.cbioportal.org), a cancer genome protocol for analysis of CYP1A1 pathway contribution in liver carcinoma was also performed. Results identified more than 45 proteins (AHR, EPHX1, GSTP1, HPGDS, UGT1A6, GST1, HELZ2, NCOA2, CREBBP, TBL1XR1, RXRA, PPARA, TP53, RB1, BCL9, CCNE2, HSF1, SNRPE, ARID1A, PYGO2 and Histone family) that have unique interactions which play comprehensive role in progression of carcinoma. Identified proteins were proposed as a drug co-targets for advanced personalized therapy. Moreover, this study provides a repository of drug markers in context of CYP1A1 mediated signaling in liver carcinoma. It also opens window for further research to characterize the binding sites, interaction sites, signaling flow quantification and drug designing to break these interactions for better survival.

Keywords: CYP1A1; Hepatocellular carcinoma; Bioinformatics

Introduction

The Hepatocellular Carcinoma (HCC) has the status of starting neoplastic condition of hepatocytes and largely affected the already cirrhotic and liver disease patients [1]. Globally more than 500,000 cases are recorded by HCC and third most cancer causing death agent. In Asia and Africa, the ratio of HCC cases are extremely elevated due to Hepatitis B and Hepatitis C [2]. HBV is the main causing agent of HCC due to DNA genome behavior that produces HBV X proteins which are vital promoter of HCC [3,4]. Earlier studies confirm the molecular mutational role in HCC including mostly are ARIDA, PIKCA, P53 and β-catenin genes [5,6]. The in-depth knowledge for identification of drug co-targets in complex signaling pathways contribution in HCC initiation we perform multi-layered Cytochrome P450 group analysis.

In 1961 the name Cytochrome P450 (CYP450) were assigned due to its pigment nature of 450 nm spectral peak absorption upon reduction with CO. They are haemoproteins superfamily located on endoplasmic reticulum bilayer lipid membrane [7]. Due to advances in sequences alignment techniques CYP450 showed close homology among bacteria and human which is the vital sign of its 3 billion year old ancestral gene origination [8]. The CYP450 classification system has logical foundations of amino acid homology for construction of families and subfamilies that has 40% and 55% similarity in amino acid sequence respectively [9,10]. The CYP450 has above 270 gene families including 18 families in mammals. The A. thaliana has 1% CYP450 genes of total genome in which 249 actives genes and 24 pseudogenes. The rice plant has 324 active genes of this superfamily [11,12]. The CYP450 has 57 functional and 58 non-functional pseudogenes mostly located on autosomal chromosomes in humans [13]. They are involved in various processes of metabolism,
In this study we design novel 3-multi-layered approach to identify inter/intra cluster connections which serves as a drug target sites in therapeutics (Figure 1.2).

In this work our basic intention to comprehensively understand the mechanistic participation of CYP1A subfamily member CYP1A1 mediated pathway in liver hepatocellular carcinoma promotion. We design novel insilico 3-multi-layered approach to identify proliferative and anti-apoptotic activators in context of CYP1A1 mediated pathway interactors. The CYP1A1 mainly occurred in extrahepatic areas so this study is unique in nature regarding its global explorative search view for advance therapy.

Methodology

In this study we design novel 3-multi-layered approach to identify wide range of CYP1A1 associated co-targets for drug designing. In 1st multi-layered approach we gain CYP1A1 pathway from STRING database (www.string-db.org) which is the huge systematic collection of protein interaction networks with confidence scoring mechanism for reliability and authenticity of interactions. STRING database has user friendly interface with several chosen parameters for desired network retrieval. The database also allowed performing clustering in network based on functional association via K-MEANS algorithm. We perform gene ontology and enrichment analysis of CYP1A1 pathway from WebGestalt toolkit (www.webgestalt.com) that gives extensive services to scientific societies for better understanding of genes/proteins involvement in physiology, patho-physiology, pathways and lot of other vital information’s freely. Its interface demanded slight skilled abilities for ideal and correct operations. We analyzed the whole CYP1A1 pathway participation in liver carcinoma through cBioPortal a cancer genome portal (www.cbioportal.org) that contain vast amount of carcinoma datasets with analysis tools. It has amazing graphical representation of results with distinguished coloration for alterations. In 2nd multi-layered approach we used cBioPortal network tab that retrieve CYP1A1 expression associated protein interactors. Finally in 3rd multi-layered we reused the cBioPortal network option tab that construct state change interaction network with key association of CYP1A1 regulatory protein interactors.

Results

CYP1A1 mediated pathway retrieval from STRING database: As a 1st multi-layered approach.

CYP1A1 pathway was retrieved from STRING database that contains 15 nodes (proteins), 39 edges (interactions), 5.2 average node degree (ratio of interactions), 0.952 average local clustering coefficient (ability of functional association among interactors) and 0.900 confidence score (accuracy of interactions). The red color node indicates query protein with other interactors of various colors regarding their source (Figure 1.1).

CYP1A1 mediated pathway clusters from STRING database

STRING database allowed to construct clusters in pathway/network based on functional links that play key role in understanding of their co-expression. We draw 2 clusters via K-MEAN algorithm for identification of inter/intra cluster connections which serves as a drug target sites in therapeutics (Figure 1.2).

CYP1A1 mediated pathway gene enrichment analysis by WebGestalt toolkit

The CYP1A1 mediated pathway involved in metabolism of degradation and redox reactions of retinoids, biogenic amines, steroids, bile acids, prostaglandins, drugs, anesthetic agents, alcohols, dyes and environmental pollutants [14-19]. The superfamily mainly consists on enzymes which has biochemically one protein one haem group concept for diversified reactionary behavior [20]. The CYP450 enzymes are dispersed in liver, brain, kidney, gonads, skin, pancreas, digestive system and have key associations with menstrual cycle, pregnancy and stress conditions [21-23]. Several studies reported the participation of CYP450 enzymes in toxicity effects, birth defects and carcinogenesis [24,25]. In cancer many metabolism specific enzymes undergo down regulation due to TNF-α, IL-6 and other interleukins in liver [26]. The CYP1A subfamily has mainly CYP1A1 and CYP1A2 which are situated on chromosome 15q24.1 consisting of 7 exon and 6 intron [27]. There are strong evidences of the participation of CYP1 family in promotion and progression of multiple cancer studies [28-32].

Table 1.1: CYP1A1 pathway biological processes characterization analysis.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Biological Processes</th>
<th>No. of Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cellular response to xenobiotic stimulus</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Response to xenobiotic stimulus</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Xenobiotic metabolic process</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>Cellular hormone metabolic process</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>Flavonoid metabolic process</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>Phenylpropanoid metabolic process</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>Retinoic acid metabolic process</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Flavonoid glucuronidation</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>Xenobiotic glucuronidation</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>Hormone metabolic process</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 1.2: CYP1A1 pathway molecular function enrichment analysis.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Molecular Functions</th>
<th>No. of Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Retinoic acid binding</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Glutathione transferase activity</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Glucuronosyl transferase activity</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>Retinoid binding</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>Isoprenoid binding</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>Transferase activity, transferring alkyl or aryl (other than methyl) groups</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>Monocarboxylic acid binding</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>Transferase activity</td>
<td>11</td>
</tr>
<tr>
<td>9</td>
<td>UDP-glycosyl transferase activity</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>Protein homodimerization activity</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 1.3: CYP1A1 pathway cell localization exploration analysis.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Cellular Localization</th>
<th>No. of Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Endoplasmic reticulum membrane</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>Nuclear outer membrane-endoplasmic reticulum membrane network</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>Endoplasmic reticulum</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>Endoplasmic reticulum part</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>Cytoplasm</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>Organelle membrane</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>Cytoplasm part</td>
<td>13</td>
</tr>
<tr>
<td>8</td>
<td>Endomembrane system</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>Intracellular part</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>Intracellular</td>
<td>15</td>
</tr>
</tbody>
</table>
endogenous and exogenous chemicals, detoxification, distribution, elimination and proper response in stress conditions (Table 1.1).

The CYP1A1 pathway involved in several molecular bindings with GSTs, UGTs, Histones and AHR- mediated receptors for transcription, biosynthesis of vital compounds, viral response, glycosylation, post-translational modifications (Table 1.2).

The CYP1A1 pathway primarily located in endoplasmic reticulum and involved in communication of metabolites in intra-cellular microenvironment. The pathway components are dispersed in overall cell periphery to notice the slightest effect of stimulus and properly respond them (Table 1.3).

The CYP1A1 pathway components has diverse links with other fundamental pathways of metabolism, cell survival, growth, division, differentiation, migration, metastasis and wide range of anti-apoptotic activities regulating pathways (Table 1.4).

The CYP1A1 pathway in deregulated form heavily caused drug toxification, metabolic disorders and wide range of carcinomas that badly affect the immune microenvironment which is the initiation of metastasis (Table 1.5).

**CYP1A1 mediated pathway differential expression analysis from cBioPortal a cancer genomic data portal**

CYP1A1 pathway participation in liver carcinoma was analyzed by cBioPortal. Liver Hepatocellular Carcinoma (TCGA, Provisional) dataset was selected. It contained 366 samples from which 19% of dataset (68 samples) showed alteration. The dataset has molecular mutations, alterations, amplifications and deletions with comprehensive color presentation of relevant modification (Figure 1.3).

**CYP1A1 mediated pathway expression controller interactors by cBioPortal: 2nd multi-layered approach**

CYP1A1 expression regulatory protein interactions having extensive impact on linked proteins expression were attained. The red dark color showed the elevated modification in maximum cases (Figure 1.4).
CYP1A1 expression regulatory protein interactions are analyzed by cBioPortal a cancer genomic data portal

Expression regulatory proteins participation in liver carcinoma was also analyzed. Most of the proteins showed amplification that is up-regulated at cellular level. They are observed in 110 altered cases which is 30% of dataset (Figure 1.5).

CYP1A1 mediated expression controlling interactors induced state change interaction network by cBioPortal: 3rd multi-layered approach

A unique state change protein interaction network through association with CYP1A1 expression regulatory proteins was retrieved. It depicted the effect of one protein modification on the regulatory nature of next protein (Figure 1.6).

CYP1A1 mediated state change interaction network differential expression analysis by cBioPortal a cancer genomic data portal

We also analyzed the state change interaction network participation in liver cancer that is observed in 279 altered cases which is 76% of dataset. Most of the proteins which has cell proliferatory nature showed overexpression while apoptotic proteins showed loss of function (Figure 1.7).

Discussion

The CYP1A1 is the initiator of several carcinomas by transformation of highly reactive products that trigger alteration in cellular signaling networks which leads to carcinogenesis [37,38]. The CYP1A1 gene regulation is indirectly induced by pollutants via AHR receptors [39]. The harmful impact of CYP1A subfamily is the conversion of metabolites that promote tumorigenesis but their parent compounds lack the carcinogenic background [40,41]. The CYP1A1 has key role in metabolism and activation of highly oncogenic agent Benzo[a]Pyrene B[a]P which is via epoxide and diol epoxide metabolism [42]. The heterocyclic amines like PhIP that occurred in fried meat undergo hydroxylation at N2 position by CYP1A1 which is initiation of carcinogenesis [43]. In this study we design novel insilico 3-multi-layered protocol to identify large number of CYP1A1 associations as a co-target for drug designing. In 1st multi-layered approach CYP1A1 showed overexpression that is the evidence of tumorigenesis. Our CYP1A1 differential expression result confirmed the earlier studies regarding CYP1A1 in lung adenocarcinomas and bronchioalveolar carcinomas by the correlation with smoking [44]. The CYP1A1 overexpression is observed in breast, urinary bladder, endothelial cells under stress and esophageal malignances [45-
52]. In our study CYP1A1 has interactions with AHR that showed overexpression in certain cases. Several genes are regulated by AHR which has ligand triggered activation. It is involved in all CYP1 members overexpression in various studies. The benzoflavones, tryptophan, indoles, dioxins and PAHs are the mostly accounted ligand for AHR transcription. It is involved in several cell survival regulatory processes. In various researches initiation of tumorigenesis prescribed through CYP1A1 gene regulation by AHR pathway. The elevated oxidative stress is the outcome AHR signaling that promote carcinogenesis [53-59]. The CYP1A1 has diverse interactions with GST family members that are involved in endo/exogenous detoxification, leukotrienes, steroids, prostaglandins biosynthesis, and lipid oxidation that regulate several cellular regulatory processes [60-71]. The GSTs and specifically GSTP1 showed overexpression in this study that authenticate its relevant studies overexpression in ovarian, lung, kidney, breast and colon carcinomas [72,73]. The CYP1A1 has various interactions with UGT family members that are involved in anti-malignant drug resistant pathways [74,78]. Several members of UGT family showed overexpression in this study. The CYP1A1 has link of HPGDS which showed overexpression that is involved in tissue development, inflammation and malignances [79]. Diverse studies confirm the up-regulation of COX-2 and mPGES-1 that are main regulators of prostaglandins [80-85].

The most significant protein interaction of 1st multi-layered approach is EPHX1 that showed highest overexpression. The EPHX1 perform various functions including biotransformation, detoxification of deleterious compounds and also activate PAHs like procarcinogens [86-88]. The EPHX1 has differential expression in breast, liver, lungs, colorectal and ovarian malignancies in previous studies [89-92]. This is the ideal therapeutic marker in this 1st approach.

In 2nd multi-layered approach we extract CYP1A1 expression level regulators in context of HCC. The expression of such regulators has wide impact on whole microenvironment of CYP1A1 mediated cascade. In this study TGS1/PIMT showed up-regulation that are involved in protection of cell from apoptosis due to H2O2 stress and prevent the development of ROS [93]. They have well known participation in atherosclerosis and diabetes by repairing of proteins [94-96]. It is also observed in several studies that it reduced the regulatory activity of p53 [97]. The CREBBP/CRB is ubiquitously expressed transmembrane protein showed amplification in few cases in this study which are situated in lipid rafts. Its overexpression is also reported in B-cell lymphomas and follicle derived lymphoma [98]. The CYP1A1 expression regulator PPARA also has amplification in few cases and it has extensive role in progression of malignancies [99]. The TBL1XR1 showed overexpression that performs key contribution in transcription of several genes [100,101]. Its overexpression reported in HCC, breast, cervical, lungs, esophageal, nasopharyngeal, gastric and lymphatic malignancies [102-107]. Interestingly our study showed amplification of NCOA2 that verify the unique study such behavior in aggressive prostate cancer [108]. The HELZ2 showed overexpression that is involved in DNA repairing, RNA translation and gene regulatory mechanisms [109,110]. Several studies reported its participation in type 3 familial lipodystrophy due to alteration in PPARC [111]. The RXRA showed overexpression in this work that is involved in metabolism, differentiation and embryogenesis [112,113]. Various researches confirm its up-regulation in thyroid, prostate, breast, skin and cholangiocarcinoma [114-118]. In 2nd advancement we identify various novel CYP1A1 regulatory interactions in liver carcinoma for advanced personalized therapy.

In this study we apply 3rd multi-layered approach for retrieval of multiple drug co-targets that are induced by CYP1A1 regulatory interactors specifically CREBBP, TGS1 and TBL1XR1. These co-targets perform state change interactions in which one protein expression has wider impact on state of next linked protein regulation. The HIST2H2AC showed overexpression that is involved in nucleosomal stability and chromatin remodeling. Its amplification observed in breast cancer [119]. The PYGO2 showed overexpression that is involved in multitdrug resistance by the induction of Wnt/β-catenin pathway in breast cancer [120]. The HIST2H3A has overexpression that induced abnormal gene expression and reduced the chromatin condensation that transforms the normal into malignancies [121]. The HIST2H4B showed overexpression which is interesting and further research appealing finding. Various studies indicate its similar expression in both normal and malignant cells [122]. The HIST2H3C showed amplification that is involved in the regulation of mitosis by phosphorylation at several residues [123]. The HSFI1 showed overexpression that triggered rapid cell division, invasion, metastasis, migration and anti-apoptotic activities. The HSFI1 stimulate MAPK, NF-KB, PI3K/ AKT and PKC signaling cascades that participate heavily in carcinogenesis [124]. The SNRPE showed overexpression interestingly. Its amplification reported in prostate cancer [125]. The CCNE2 showed overexpression that is involved in regulation of cell cycle by the activation of S phase [126]. The BCL9 showed overexpression which is well known oncoprotein that speedup β-catenin pathway components transcription performance which enhanced cell proliferation, differentiation and metastasis [127]. The ARID1A showed loss of function due to mutation that is involved in tumor suppressing activities and has strong association with p53/PI3K/AKT pathways [128]. The CTNNB1 showed loss of function that confirmed its mutation in hepatocellular malignancies in previous studies [129]. The RB1 showed loss of function due to mutation that is key tumor repressor in several studies. Finally the great apoptotic activator and security officer of the cellular microenvironment p53 showed high ratio loss of function which is the major sign of tumorigenesis in various researches.

In this approach we acquire a comprehensive set of protein interactors that are key regulatory members of cell survival, growth, differentiation, migration and stress response. There are mostly anti-apoptotic proteins undergo functionally inactivation and proliferating agents showed up-regulation which is the logical justification of research work.

Conclusion

Cancer is complex heterogeneous disease that has complicated inter-link molecular arrangements which showed combined phenotypic effect in response of oxidative stress, viral attack and harmful chemical stimuli. In bioinformatics signaling pathway analysis gain high esteem attention due to easy access with low cost understandable approach to uncover the secrets of cancer molecular biology. In this study, we apply novel insilico design that explores large number of proteins with association of CYP1A1 mediated pathway. We found more than 45 proteins (AHR, EPHX1, GSTP1, HPGDS, UGT1A6, GST1, HELZ2, NCOA2, CREBBP, TBL1XR1, RXRA, PPARA, TP53, RB1, BCL9, CCNE2, HSFI1, SNRPE, ARID1A, PYGO2 and Histone family) that have unique interactions which play comprehensive role in progression of carcinoma. In future this study is proven as the repository of drug markers in context of CYP1A1 mediated signaling in liver carcinoma. We identify these proteins
as a drug co-targets for advanced personalized therapy. Our study open the window for further research to identify the binding sites, interaction sites, signaling flow quantification and drug designing to break these interactions for better survival. The experimental validation will make our study more visionary and rational.

References

1. Alison MR. Liver stem cells. Stem cell reviews. 2005;1(3):253-60.


Asima Tayyeb, et al., Annals of Stem Cells and Regenerative Medicine


