Comparative investigation to analyse the critical role of NFE2L2 gene in heavy metal induced toxicity through in silico approaches

Prekshi Garg¹, Neha Srivastava¹, Prachi Srivastava*¹

¹Amity Institute of Biotechnology, Amity University, Lucknow, Uttar Pradesh, India
²Pharma Analytics Department, Excelra knowledge Solution Pvt Ltd, Hyderabad, Telangana, India

Abstract

Background: Accumulation of heavy metals like copper, nickel, arsenic, cobalt, and cadmium are increasing day by day. Accumulation of these heavy metals poses serious threats to human health. Elemental toxicity percentage has been found to be increased day-by-day and creating different minor to major problems from tissue to the gene level. Different important gene expressions are altered due to the contamination of such heavy metals. The present study aimed to identify a common target gene that is involved in the toxicity of major heavy metals and to study the major impact of the gene on the concerned biological system.

Methods: In the progression of the work, major genes involved in copper, nickel, arsenic, cobalt, and cadmium toxicity were listed through intense data curation, and a pathway showing the correlation and physical interaction of all the genes that were constructed using in-silico tools STRING and Gene Mania database. Further, functional and expression analysis of the discovered gene was done using in-silico tools like genome-wide association study (GWAS), genotype-tissue expression (GTEx), and RegulomeDB.

Results: According to the network analysis, NFE2L2 was recognized as a common target involved in the above-mentioned heavy metals toxicity. Expression analysis revealed that the highest expression of NFE2L2 was observed in tissues of oesophagus, ovary, bladder, vagina, thyroid, and skin. Detailed investigation at the pathway level revealed that the involvement was importantly observed in immunodeficiency and developmental delay.

Conclusion: The study opened a wide vision that the major target of such toxicants is various pathways of neurobiological distress and biological processes, and hence, it can be considered as a susceptible target for heavy metals-induced toxicity.

Keywords: Heavy metals, Brain, Gene expression, Environmental pollutants

Introduction

Metals are substances that have properties like high electrical conductivity, luster, and malleability. The metals that have a specific density above 5 g/cm³ are regarded as heavy metals (1). Heavy metal contamination is widespread around the world due to various anthropogenic, technogenic, and geogenic activities (2). Heavy metals are evident environmental pollutants that pose increasing toxicity due to evolutionary, environmental nutritional, and ecological reasons (3, 4). Toxicity induced by heavy metals is dependent upon the route of administration, absorbed dose, and the duration of exposure to the pollutant. These days, heavy metal toxicity is a topic of major concern due to its fatal effects on the human body (Figure 1) (5). Such toxicities can lead to severe disorders that cause increased cellular damage resulting largely due to oxidative stress. Heavy metal-induced toxicity has the potential to lower the energy levels that subsequently damage the function of the lungs, brain, liver, and kidney. They may also alter blood composition. Long-term exposure to heavy metals can lead to gradual degeneration of muscular, neurological, and physical processes that mimic diseases like multiple sclerosis (MS), muscular dystrophy, Alzheimer's, and Parkinson's diseases. In some cases, continuous exposure to heavy metals can also lead to cancer. Although heavy metals have been reported to impose several serious health effects when exposed for a long time, the usage of these metals is still prevalent.
in various parts of the world. Sources of heavy metals in the environment include soil erosion, mining, urban runoff, industrial effluents, sewage discharge, and many more (6). The most commonly found heavy metals in the environment are arsenic, cadmium, cobalt, copper, and nickel (7). The present study aimed to identify a common susceptible target that is being expressed in all five heavy metal-induced toxicities. This will, therefore, help in bridging the gap between various heavy metal toxicities by linking them to a common gene target. Table 1 presents major toxic heavy metals, their source, and side effects generated by their accumulation (8).

Materials and Methods

Intensive data mining and use of various in silico tools were done to form a proper workflow (Figure 2) for the computational identification of a biomarker for heavy metal-induced toxicity.

Literature study

An extensive study of various research papers was carried out to identify the genes whose expression is altered due to the accumulation of heavy metals like copper, nickel, arsenic, cobalt, and cadmium. The main focus was to study such genes that can be used as a susceptible target for various body tissues. Thus, curative data mining was required to create a list of genes involved in heavy metal toxicity.

Pathway construction

Pathways of all the genes involved in heavy metal-induced toxicity were studied using an in-silico tool STRING (31). STRING is a database that shows protein-protein interactions of both known and predicted protein molecules. The database shows both direct and indirect interactions but functional interactions of the protein. The information about the interactions of the protein is obtained from various sources like experimental, co-expression, co-occurrence, databases, text mining, neighborhood, and gene fusion. In the present study, the STRING database was used to construct a pathway involving all 102 genes at 0.700 high confidence interaction score.

Functional analysis using genome-wide association study (GWAS)

Functional analysis of the NFE2L2 gene, the identified common target of heavy metal-induced toxicity, was done using the PhenoScanner package. The package includes publicly available large-scale GWAS data, associations, unique single nucleotide polymorphisms (SNPs), and comprehensive phenotypes data (32). In the present research, the GWAS of diseases and traits and gene expression analysis at p-value 1E-5 were performed.

Gene expression analysis in genotype-tissue expression (GTEx)

Gene expression datasets of the NIH GTEx project were used to determine the relationship between the NFE2L2 gene and gene expression in multiple human tissues (33). The GTEx project includes median gene expression levels in 2 cell lines and 51 tissues. The project is based on the data collected from 11,688 tissue samples obtained from 714 adult post-mortem individuals (34).

Results

Literature study

A literature survey was conducted in order to identify major toxic heavy metals that are adversely affecting humans. Heavy metals like copper, nickel, arsenic, cobalt, and cadmium were found to be highly accumulated in humans posing certain fatal effects (Table 1). Furthermore, an intense literature study was performed to identify the

Figure 1. Effects of heavy metal-induced toxicity.

Figure 2. A brief workflow of the study.
genes whose expressions get altered due to this toxicity of these heavy metals. A list of 102 genes was prepared that involved 20 genes for copper-induced toxicity, 24 for nickel-induced toxicity, 20 for arsenic-induced toxicity, 18 for cobalt-induced toxicity, and 20 for cadmium-induced toxicity. Table 2 presents a list of key genes involved in the toxicity induced by five different heavy metals (35-37).

Pathway analysis
The expression of genes was analyzed based on text mining, databases, neighborhood, gene fusion, co-occurrence, experimental, and co-expression data sources using the STRING database at 0.700 high confidence interaction score. The pathway obtained contained 68 nodes and 79 edges with a P value < 1.0e-16. The average node degree was obtained to be 11.7. The interaction of all the genes with each other was analyzed (Figure 3), and it was observed that the genes play an important role in biological processes like positive regulation of neuron maturation, negative regulation of synaptic transmission, cardiac neural crest cell development, and many more.

Gene expression analysis of NFE2L2 in GTEx
Gene expression of NFE2L2 was studied in GTEx. The gene was analyzed to be expressed in various parts of the body like the brain, adipose tissue, heart, kidney, liver, skin, thyroid, and more. The maximum expression of the gene was observed in esophagus tissue having a mean TPM of 174.3. The minimum expression of the gene was seen in the pancreas with a mean TPM of 14.12 (Figure 4).

Phenotypic expression analysis of NFE2L2 gene in GWAS
The NFE2L2 gene was studied and analyzed for its phenotypic expression in the GWAS central database and it was observed that the gene shows its phenotypic expression in various neuro-related disorders like stroke, MS, amyotrophic lateral sclerosis, Kuru disease, Alzheimer’s disease, Parkinson’s disease, etc. This depicts a strong correlation between NFE2L2 and esophagus, ovary, bladder, vagina, thyroid, skin and, brain tissues and its link in cancer, metabolic diseases, liver, kidney, and neurodegenerative diseases, therefore, it can be used as a susceptible target for various degenerative diseases (Figure 5).

Discussion
Nuclear factor, erythroid-2-like 2 known as NFE2L2/ Nrf2 is a transcription factor that is mostly present in the cytoplasm of the cell. The transcription factor is redox-sensitive that regulates various antioxidant enzymes (38). The gene is known to play a vital role in providing resistance to oxidative stress (39). Under oxidizing conditions, the gene controls the upregulation of several antioxidant proteins. The factor also acts as a primary system that...
Table 2. Key genes involved in heavy metal-induced toxicity

<table>
<thead>
<tr>
<th>Entry</th>
<th>Genes for Copper Toxicity</th>
<th>Genes for Nickel Toxicity</th>
<th>Genes for Arsenic Toxicity</th>
<th>Genes for Cobalt Toxicity</th>
<th>Genes for Cadmium Toxicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ATP7B</td>
<td>EGFR</td>
<td>AS3MT</td>
<td>TP53</td>
<td>EGFR</td>
</tr>
<tr>
<td>2</td>
<td>SLC31A1</td>
<td>MTHFR</td>
<td>TP53</td>
<td>TNF</td>
<td>TP53</td>
</tr>
<tr>
<td>3</td>
<td>APP</td>
<td>TGFB1</td>
<td>ABCB1</td>
<td>NFE2L2</td>
<td>AKT1</td>
</tr>
<tr>
<td>4</td>
<td>TP53</td>
<td>ABCB1</td>
<td>GST1M1</td>
<td>PTGS2</td>
<td>MAPK1</td>
</tr>
<tr>
<td>5</td>
<td>SNCA</td>
<td>AKT1</td>
<td>ABCB1</td>
<td>IL1B</td>
<td>SOD1</td>
</tr>
<tr>
<td>6</td>
<td>SOD1</td>
<td>NFE2L2</td>
<td>NFE2L2</td>
<td>CYP3A4</td>
<td>IL6</td>
</tr>
<tr>
<td>7</td>
<td>PRNP</td>
<td>NFKB1</td>
<td>GSTP1</td>
<td>CXCL8</td>
<td>PON1</td>
</tr>
<tr>
<td>8</td>
<td>BIRC5</td>
<td>BCL2</td>
<td>AKT1</td>
<td>CASP3</td>
<td>MMP9</td>
</tr>
<tr>
<td>9</td>
<td>ABCB1</td>
<td>PTGS2</td>
<td>VEGFA</td>
<td>HMOX1</td>
<td>NFE2L2</td>
</tr>
<tr>
<td>10</td>
<td>NFE2L2</td>
<td>CDKN2A</td>
<td>HMOX1</td>
<td>CASP8</td>
<td>NOS3</td>
</tr>
<tr>
<td>11</td>
<td>APOE</td>
<td>CXCL8</td>
<td>PON1</td>
<td>ALB</td>
<td>NFkB1</td>
</tr>
<tr>
<td>12</td>
<td>MAPT</td>
<td>IFNG</td>
<td>MMP9</td>
<td>RH0A</td>
<td>HMOX1</td>
</tr>
<tr>
<td>13</td>
<td>VEGFA</td>
<td>GSTT1</td>
<td>BRCA1</td>
<td>MCL1</td>
<td>ESR1</td>
</tr>
<tr>
<td>14</td>
<td>PON1</td>
<td>CCND1</td>
<td>BCL2</td>
<td>POU5F1</td>
<td>PRNP</td>
</tr>
<tr>
<td>15</td>
<td>GSTP1</td>
<td>MAPK14</td>
<td>ESR1</td>
<td>GDF15</td>
<td>CXCL8</td>
</tr>
<tr>
<td>16</td>
<td>CYP1A1</td>
<td>MGMT</td>
<td>PTEN</td>
<td>EPO</td>
<td>TERT</td>
</tr>
<tr>
<td>17</td>
<td>CRP</td>
<td>APEX1</td>
<td>CDKN1A</td>
<td>ALAD</td>
<td>PTGS2</td>
</tr>
<tr>
<td>18</td>
<td>PARK7</td>
<td>POU5F1</td>
<td>MAPK1</td>
<td>SLC40A1</td>
<td>MAPK3</td>
</tr>
<tr>
<td>19</td>
<td>CXCL8</td>
<td>HAMP</td>
<td>CYP1A1</td>
<td>-</td>
<td>UGT1A1</td>
</tr>
<tr>
<td>20</td>
<td>HTT</td>
<td>SQSTM1</td>
<td>MTOR</td>
<td>-</td>
<td>XRPC1</td>
</tr>
<tr>
<td>21</td>
<td>-</td>
<td>CAT</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>-</td>
<td>CLDN1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>23</td>
<td>-</td>
<td>MT2A</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>-</td>
<td>ATP13A2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
counteracts reactive oxygen species (ROS) derived from mitochondria. Firstly, the NEF2l2 gene maintains the mitochondrial glutathione (GSH) pool by elevating the inducible expression of enzymes responsible for GSH biosynthesis along with GSH reductase (regeneration enzyme) (40). In the mitochondria, detoxification of superoxide-derived hydrogen peroxide to water is done using GSH peroxidases (GPx1 and GPx4) and NADPH (41). Secondly, Nrf2 upregulates several enzymes involved in the pentose phosphate pathway that increases the level of an essential reducing molecule, i.e., NADPH in the cell. NADPH is used for the production of GSH reduced by the activity of GSH reductase and for removal of hydrogen peroxide caused by GPx (42). Also, Nrf2 is directly involved in the expression of antioxidant enzymes present in mitochondrial like peroxiredoxin 3 (Prdx3 and Prdx5), mitochondrial superoxide dismutase 2 (Sod2), thioredoxin reductase 2 (TrxR2), and GPx1 (43-46). The combined effect of all these roles of NFE2L2 makes the cell resistant to oxidative stress (47).

At greater concentrations and persistent exposure, ROS can damage DNA, proteins, lipids, and cellular macromolecules. This may lead to cell death through apoptotic or necrotic pathways. Any alterations in the biochemical attributes of these macromolecular components can lead to several different pathological conditions and human diseases, especially cancer, metabolic disorders, and neurodegenerative diseases (48).

The results of various wet-lab studies and in silico validation are consistent with the results of the present study. In one of the studies, knockout of the Nrf2 gene in mice, increased the susceptibility of mice to chemical toxicity and disease conditions associated with oxidative pathology (49,50). In another genomic study, several ARE-containing genes were identified as Nrf2 target genes that are involved in the control of oxidant homeostasis (51). In the other study, mouse embryo fibroblast of NFE2L2 knockout mice was seen to exhibit reduced expression of autophagy genes (52). This further validates the role of the NFE2L2 gene in the process of autophagy.

Therefore, it can be concluded that the NFE2L2 gene can be used as a susceptible common target for heavy metal-induced toxicity after detailed network analysis. Further, expression analysis reveals the expression of the gene in most of the oesophagus, ovary, bladder, vagina, thyroid, skin, and brain tissues. Thus, the study opened a wide vision that the major target of such toxicants is showing important involvement in cancer, metabolic diseases, neurobiological distress, and hence, it can be considered as a reference target for such issues.

The accumulation of heavy metals also leads to cancer. The accumulation of heavy metals generates ROS as mentioned above. ROS plays a dual role in carcinogenesis by acting as an oncogenic in the early stages of metal carcinogenesis and as an antioncogenic in the late stages of metal carcinogenesis. NFE2L2 plays a key role in differentiating the two stages of metal carcinogenesis. During the early stages, NFE2L2 acts as an antioxidant...
reducing the elevated levels of ROS whereas, during the late stage, NFE2L2 acts as an oncogene by manipulating the reduced levels of ROS in order to obtain apoptosis resistance (53).

Therefore, NFE2L2 is an important gene that affects several biological processes due to the accumulation of heavy metals leading to several fatal conditions like cancer and neurodegeneration.

**Conclusion**

The NFE2L2 gene was recognized as a common gene in the toxicity induced by different heavy metals like copper, nickel, arsenic, cobalt, and cadmium using *in silico* meta-analysis based on the data retrieved from the intense literature study. The high expression of the gene in the oesophagus, ovary, bladder, vagina, thyroid, skin, and brain establishes an expectable link between the NFE2L2 gene and cancer, metabolic disorders, neuro-related disorders, ROS, and disruption of cell cycle and cell signaling as well. Also, various wet-lab studies and the role of NFE2L2 in the generation of ROS on the accumulation of heavy metal further make it an important target for study. Therefore, these findings will help in understanding the metal-induced toxicity in humans and can act as a susceptible target for heavy metal-induced toxicity.

**Acknowledgments**

The authors would like to acknowledge Amity Institute of Biotechnology, Amity University, Lucknow Campus for supporting the study. They would also acknowledge all individuals who supported the work directly or indirectly. Special thanks to all Bioinformatics tools, software, and databases for providing support in the work.

**Ethics Issues**

Not applicable.

**Competing Interests**

The authors declare that there is no conflict of interest.

**Authors’ Contributions**

PG conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper. NS conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft. PS conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

**References**


