scientific reports

OPEN

Arsenic causing gallbladder cancer disease in Bihar

Arun Kumar [1][✉], Mohammad Ali[1], Vivek Raj[1], Arti Kumari[2], Mahesh Rachamalla[3], Som Niyogi[3], Dhruv Kumar[4], Ashok Sharma[5], Amit Saxena[6], Ghanish Panjawani[1], Preeti Jain[1], Ajay Vidyarthi[1], Navin Kumar[1], Mukesh Kumar[1], Pintoo Kumar Niraj[1], Md. Samiur Rahman[1], Akhouri Bishwapriya[7], Ranjit Kumar[8], Maiko Sakamoto[9], Santosh Kumar[10], Manisha Singh[1] & Ashok Kumar Ghosh[1]

In recent times Gallbladder cancer (GBC) incidences increased many folds in India and are being reported from arsenic hotspots identified in Bihar. The study aims to establish association between arsenic exposure and gallbladder carcinogenesis. In the present study, n = 200 were control volunteers and n = 152 confirmed gallbladder cancer cases. The studied GBC patient's biological samples-gallbladder tissue, gallbladder stone, bile, blood and hair samples were collected for arsenic estimation. Moreover, n = 512 gallbladder cancer patients blood samples were also evaluated for the presence of arsenic to understand exposure level in the population. A significantly high arsenic concentration ($p < 0.05$) was detected in the blood samples with maximum concentration 389 µg/L in GBC cases in comparison to control. Similarly, in the gallbladder cancer patients, there was significantly high arsenic concentration observed in gallbladder tissue with highest concentration of 2166 µg/kg, in gallbladder stones 635 µg/kg, in bile samples 483 µg/L and in hair samples 6980 µg/kg respectively. Moreover, the n = 512 gallbladder cancer patient's blood samples study revealed very significant arsenic concentration in the population of Bihar with maximum arsenic concentration as 746 µg/L. The raised arsenic concentration in the gallbladder cancer patients' biological samples— gallbladder tissue, gallbladder stone, bile, blood, and hair samples was significantly very high in the arsenic exposed area. The study denotes that the gallbladder disease burden is very high in the arsenic exposed area of Bihar. The findings do provide a strong link between arsenic contamination and increased gallbladder carcinogenesis.

Abbreviations

GBC	Gall bladder cancer
ICMR	Indian Council of Medical Research
MMA	Monomethylarsonous acid
DMA	Dimethylarsinic acid
TMA	Trimethylarsine
ROS	Reactive oxygen species
USEPA	U.S Environmental Protection Agency
WHO	World Health Organisation
GF-AAS	Graphite Furnace Atomic Absorption Spectrophotometer
PCA	Principal component analysis

Gallbladder cancer (GBC) is one of the rarest biliary tract malignancies with high mortality rate with relatively less survivability of 5 years[1,2]. Globally in year 2020, out of total cancer cases of 19,292,789, new GBC cases reported in the year was 115,949 while number of deceased cases were 84,695[3]. India accounts for 10% of global GBC cases, about one million new cancer cases every year with mortality rate as high as 33% every year[4].

[1]Mahavir Cancer Sansthan and Research Centre, Patna, Bihar 801505, India. [2]Department of Biotechnology, Patna Women's College, Patna, Bihar, India. [3]University of Saskatchewan, Saskatoon, SK S7N 5E2, Canada. [4]School of Health Sciences & Technology, UPES University, Dehradun, Uttrakhand, India. [5]All India Institute of Medical Sciences, New Delhi, India. [6]Centre for Development of Advanced Computing (C-DAC), Pune, Maharashtra, India. [7]Geological Survey of India, Patna, Bihar, India. [8]Central University of Himachal Pradesh, Kangra, India. [9]University of Tokyo, Kashiwa, Japan. [10]Delft University of Technology, Delft, The Netherlands. [✉]email: drarunk31@gmail.com

According to Indian Council of Medical Research (ICMR), it is estimated that by the year 2025, there will be fivefold increase in cancer incidences in India, in which there will be 2.8-fold increase due to tobacco while 2.2-fold increase due to ageing. Such a drastic increase in cancer cases is more than the double which is predicted for the developed countries like USA[5–7].

GBC has a very unusual geographical distribution with high incidences seen in pockets as high incidences are seen in Chile, India, Japan, Poland, Israel, Bolivia, Thailand, South Korea etc.[1,8–10]. In India, the highest distribution of GBC has been detected in the states like Uttar Pradesh, Bihar, Orissa, West Bengal, Assam and Delhi (National Capital Territory)[11–14]. The cancer is two folds more common in women in comparison to men[15]. Population based data reveals that the incidences of this disease is very high in the northern cities of India (22 per 100,000) and low (0–0.7 per 100,000 women) in southern India[16–19]. The present institute caters the greatest number of cancer patients from the state of Bihar for the diagnosis and the treatment. Among the total cancer patients, 8.3% and 16.9% are GBC cases in male and female, respectively[5]. GBC patients exhibit various clinical symptoms, including benign biliary disease on one hand and incurable malignant on other hand[20–22]. The etiology of gallbladder cancer apart from the few reasons has been a mystery, however the interplay of gallstones, genetic susceptibility, changes in the lifestyle factors and infections lead to progression of the cancer disease[23,24].

Environmental pollutants in the recent times have become serious health hazards to the population. Due to ever increasing anthropogenic activities the level of the pollution effect has increased manyfold leading to eruption of the diseases in the population[25]. Human exposure to arsenic a non-essential metalloid, has increased in recent times due to geogenic changes and anthropogenic activities, which has caused severe health hazards in populations around the world[26–28]. The inorganic arsenic when enters the human body is metabolized to organic arsenic (Monomethylarsonous acid-MMA(III), dimethylarsinic acid-DMA (V) and trimethylarsine TMA) which is still a toxic carcinogen. These organic compounds bind with DNA molecules which can influence transcription and translation leading to cancer. However, it is eliminated by the body as a metabolic waste via kidney. It also gets deposited in the keratin containing tissues like hair, nail and skin[29] Arsenic is known to be a carcinogen (as carcinogen category-I) in all the forms—arsenic trioxide, arsenic pentoxide, arsenous acids, arsenic acids and their salts (arsenites and arsenates)[30–34]. Humans are exposed to this inorganic arsenic mainly by two ways—consumption of arsenic contaminated water and food[35–38].

Arsenic is known to cause cancer of skin, bladder, lungs, kidney, liver and prostate[39–41]. Arsenic causes cellular toxicity by inducing oxidative stress (ROS) which in turn leads to genotoxicity[42]. The continuous exposure of arsenic creates a very high oxidative stress that eventually leads to the appearance of symptoms like arsenicosis, skin lesions, black foot disease, vascular disease, hormonal imbalance and cancer[43–45]. Due to its physical characters (no odour, colour or flavour), arsenic exposure is often unnoticed until people develop some arsenicosis symptoms due to drinking of arsenic contaminated water. The World Health Organisation (WHO) and the U.S Environmental Protection Agency (USEPA) have recommended a threshold of 10 µg/L for inorganic arsenic concentration in drinking water. Unfortunately, millions of people are exposed to much higher toxic levels of arsenic and some populations are unaware of its ill effects, which in long term is causing the development of malignancies. Various molecular pathways, which has been deciphered in recent times, shows the progression of the disease by arsenic biotransformation process, that alters the methylation pattern, which is thought to play a key role in its carcinogenicity[46–51].

In the Gangetic plains of Bihar, a high percentage of population suffers from arsenic poisoning due to arsenic contamination of groundwater. The long duration exposure to this arsenic has magnified the toxicity by many folds causing deadly disease like cancer in human beings[52]. Recently, the gallbladder cancer disease burden in these areas have increased many times and the situation is catastrophic[53]. Hence, there is an urgent need to investigate the relationship between arsenic exposure and GBC in the local populace. There has not been any significant study which confirms the linkage between arsenic contamination in groundwater and increased incidence of gallbladder cancer till date. The present study throws new light on the association between gallbladder carcinogenesis and arsenic poisoning in the Gangetic plains of Bihar through the novel pathway.

Materials and methods

Ethical approval.
Ethical approval was obtained from the Institutional Ethics Committee of the Indian Council of Medical Research (ICMR)-Rajendra Memorial Research Institute and Medical Sciences, Patna, Bihar, India with IEC Ref No. 07/RMRI/EC/2019, dated 20/06/2019. All the patients were informed about the research study and written informed consent were obtained prior to their inclusion in the study.

Study group.
For the control study, volunteers (non-diseased subjects) were selected, while the gallbladder cancer patients who were undergoing gallbladder surgeries at the parent cancer institute with the assistance of the institute's Surgical Oncology Department were selected as diseased subjects. For the present study, volunteers were divided into confirmed Gallbladder Cancer cases (n = 152) and control cases (n = 200) after obtaining their consents. Moreover, n = 512 gallbladder cancer patients (of institute itself) blood samples were evaluated for arsenic contamination to understand the exposure level in the population. The demographic and clinical profiles of volunteers were obtained from the central records of the institute for the baseline information. The sample size numbers have been calculated through the formula and as per the previous studies carried out by our team[54].

Collection of samples.
After the surgery, all the biological samples—blood, gallbladder tissue, gallbladder stones, gallbladder bile and hair samples of the volunteers of gallbladder benign as well as the GBC cases were collected. The samples were subsequently stored in − 80 °C deep freezer prior to the analysis.

Analytical quality control.   For the analytical quality control, the blanks, procedural blanks, replicates, standards, correlation coefficient and detection limits were maintained for arsenic analysis in the atomic absorption spectrophotometer. The known standard concentration of arsenic was prepared from the standard arsenic stock solution (1000 µg/L) procured from PerkinElmer (CAS# As 7440-38-2; Lot# 20-85 ASX1; PE# N9300102), Singapore. The arsenic detection limit for blood was 0.05 µg/L in the employed method. The cancer patient's tissues, gallstones and bile samples were analysed for the first time in this part of the region, the protocol of National Institute for Occupational Safety and Health (NIOSH, 1994) was utilised[55].

Analysis of the samples through Atomic Absorption Spectrophotometer.

1. **Analysis of Gallbladder bile and blood samples:** After the collection, all the tissue samples were digested using concentrated HNO_3 (Merck analytical grade 69%) on hot plate under fume-hood and arsenic was estimated as per the protocol of National Institute for Occupational Safety and Health (NIOSH, 1994) by Graphite Furnace Atomic Absorption Spectrophotometer (GF-AAS) (Pinnacle 900T, Perkin Elmer, Singapore) at the studied institute. The blood and gallbladder bile samples were taken as 0.5 ml aliquots in 30 ml conical flask and 5 ml HNO_3 was added and left for overnight reaction. The samples were then digested on hotplate at 90–120 °C until the sample reached to 3 ml. Then 5 ml mixture of HNO_3:$HClO_4$ (6:1) ($HClO_4$— Merck analytical grade 70%) was added to the solution and again re-digested on hotplate until the volume reached to about 2 ml. Final volume was adjusted to 10 ml with 1% HNO_3. The solutions were filtered through Whatman filter paper no.41 and analysed through GF-AAS.
2. **Analysis of Gallbladder tissue and Gallstone samples:** The gallbladder tissue and gallbladder stone samples were taken as 0.5 g mass in the conical flask and 5 ml of Conc. HNO_3 was then added to the flask and left for overnight reaction. The samples were digested next day on hotplate at 60 °C for 2 h. After that, the samples were cooled and then added with 2 ml of $HClO_4$ and then re-digested at 90–120 °C for 5–10 min until the white fume of $HClO_4$ is emitted. The solution was cooled and diluted with 15 ml with deionized water and filtered with Whatman filter paper no.41 and analysed by GF-AAS.
3. **Analysis of hair samples:** For hair samples analysis, 0.120 g of hair samples were weighed and washed in 15 ml of 0.1% SDS solution (Sigma-Aldrich) and sonicated for 10 min. The solution was decanted and washed with ultrapure distilled water. Then the hair samples, were placed in 15 ml of high-grade acetone solution and sonicated for 10 min. The solution was decanted and rinsed with ultrapure distilled water three times. The hair samples were dried at 40 °C for 5 min in the oven and then 10 ml of HNO_3 was added and covered with watch glass and digested on hotplate 90–120 °C. When the solution was reduced to 3 ml, 1 ml of H_2O_2 was added and it was re-digested at the previous temperature until the bubbles came out and the solution was reduced to 2.5 ml. Then the heating was stopped, and the solution was rinsed with 1% HNO_3 and adjusted to a final volume of 10 ml with ultrapure distilled water. The sample was then filtered through Whatman filter paper no.41 and analysed through GF-AAS.

Molecular Docking analysis of Arsenic compounds with haemoglobin, cysteine and taurine.

1. **Retrieval and preparation of target protein and arsenic compounds**: The 3D-Crystal structure of human haemoglobin (PDB ID: 4HHB) was downloaded from Protein Data Bank (https://www.rcsb.org/structure/4HHB) in pdb format[56]. Human haemoglobin consists of 4 subunits (2 alpha subunit, chain A and chain C; 2 beta subunits, chain B and chain D). The structure was prepared for docking in Chimera X 1.1[57]. Haemoglobin alpha subunit chain A and beta subunit chain B was used for docking in separate experiments. Arsenous acid (Trihydroxy-arsenite (III), As(OH)$_3$, PubChem ID 545) 2-D structure was converted to 3D coordinates by MolView web based open-source application and was downloaded as mol file. The ligand molecule was also prepared in Chimera X 1.1 and saved as pdb. The 3D structure of Cysteine (PubChem CID: 5862) and taurine (PubChem CID: 1123) was downloaded from PubChem database as sdf file. These two molecules were also prepared and saved as pdb for further study.
2. **Docking analysis:** Docking was performed by HDOCK webserver[58]. The HDOCK server is to predict the binding complexes between two molecules using a hybrid docking strategy. Firstly, docking was performed between chain A of human haemoglobin and arsenous acid. Haemoglobin was uploaded as receptor and arsenous acid served as ligand molecule. In second experiment, human haemoglobin chain B served as receptor and docking was performed with arsenous acid ligand molecule. Similarly, cysteine as receptor was docked with ligand arsenous acid as a separate experiment and taurine as receptor was docked with arsenous acid at HDOCK webserver as another experiment to understand the interacting residues between two molecules. The results were visualised in Chimera X 1.1 and Pymol Viewer[59].

    GIS Mapping: The data of arsenic concentration in blood samples of the GBC patients (n = 512) were taken as input in ArcGIS Version 10.5.1 for spatial analysis and visualization. The location data of collected samples of the GBC patients were plotted and extracted as a GIS map layer for visualization. The study was conducted to understand the spatial pattern of the GBC patients geographically in the arsenic exposed area, and the resultant distribution map result is shown in Fig. 2. The arsenic concentration measured in the five biological samples (tissue, stone, bile, blood and hair) of the GBC patients from the arsenic exposed area were also analyzed in the ArcGIS environment. The output of five biological samples data, location of the GBC patients and the exposure rate pattern is shown in Fig. 6. The software used in the map layer generation was ArcGIS Version 10.5.1.

Statistical analysis.   Data were analysed with statistical software (GraphPad Prism 5) and values were expressed as Mean ± SEM. Differences between the groups were statistically analysed by one-way analysis of variance (ANOVA) followed by the Dunnett's test for comparison among study groups. The scatter plot graphs were designed using another statistical software SPSS-16.0 utilising the linear regression model as earlier used[60].

Principal component analysis.   Principal Component Analysis (PCA) was applied to construct a holistic evaluation index from the 152 patients' data. The used variables were arsenic concentrations in blood samples of gallbladder benign and gallbladder blood (BLOOD), gallbladder tissue (TISSUE), gallbladder stones (STONE), gallbladder bile (BILE), and hair (HAIR) samples. All the variables were standardized. The R function prcomp was used for computation[61].

Ethics approval and consent to participate.   The present research work was carried in accordance with the Declaration of Helsinki and the ethical approval were obtained from the Institutional Ethics Committee of the Indian Council of Medical Research (ICMR)-Rajendra Memorial Research Institute and Medical Sciences, Patna, Bihar, India with IEC Ref No. 07/RMRI/EC/2019, dated 20/06/2019. All the patients were informed about the research study and written informed consent were obtained prior to their inclusion in the study.

Results

In the present study, all the collected samples were analysed –

1. **Gender wise patients in control versus gallbladder cancer patients:** Out of total n = 200 control samples, there were n=76 male and n=124 female patients. Among the confirmed n=152 GBCs, there were n=37 male and n=115 female patients.
2. **Age group of patients in control and gallbladder cancer patients:** Out of total n = 200 control cases, the maximum age observed was 58 years, while the minimum was 21 years. Out of n=152 confirmed GBCs, the maximum age was 90 years, while the minimum was 25 years.
3. **Stagewise distribution of gallbladder cancer patients:** Out of n = 152 confirmed GBCs, there were only n = 3 in stage-I disease (1.9%), whereas n = 11 patients were in stage-II disease (7.2%), n = 15 were in stage-III disease (9.8%), and n = 123 patients in stage-IV (80.9%) disease.
4. **Arsenic concentration in tissue samples of gallbladder cancer patients:** Out of total n = 152 confirmed GBCs cases, the maximum arsenic concentration was 2166 µg/kg, while the minimum was 0.67 µg/kg (Table 1).
5. **Arsenic concentration in gallstone samples of gallbladder cancer patients:** Out of total n = 152 confirmed GBC patient's the maximum arsenic concentration was 634.9 µg/kg, while the minimum was 0.03 µg/kg (Table 1).
6. **Arsenic concentration in bile samples of gallbladder cancer patients:** Out of n = 152 confirmed GBCs, the maximum arsenic concentration was 483.2 µg/L, while the minimum was 0.17 µg/L (Table 1).
7. **Arsenic concentration in blood samples in control versus gallbladder cancer patients:** Out of the total n = 200 control volunteers, the maximum arsenic concentration observed was 11.7 µg/L, while minimum was zero µg/L. Out of n=152 confirmed GBCs cases, the maximum arsenic concentration was 389.3 µg/L, while the minimum was 0.2 µg/L (Table 1).
8. **Arsenic concentration in hair samples of Control versus gallbladder cancer patients:** Out of the total n = 200 control volunteers, the maximum arsenic concentration observed was 14.4 µg/kg, while the minimum was zero µg/kg. Out of n=152 confirmed GBC cases, the maximum arsenic concentration was 6980.4 µg/kg, while the minimum was 0.2 µg/kg (Table 1).

Biological samples studied	Arsenic concentration (mean ± S.E)	Maximum arsenic concentration	Minimum arsenic concentration	Lower 95% CI of mean	P value
GBC Patient's Tissue Arsenic Concentration (µg/Kg) (n = 152)	**340.6 ± 31.98**	2166	0.6700	277.4	< 0.0001
GBC Patient's Gallstone Arsenic Concentration (µg/Kg) (n = 152)	**78.41 ± 8.311**	634.9	0.0300	61.99	< 0.0001
GBC Patient's Bile Arsenic Concentration (µg/L) (n = 152)	**59.10 ± 6.554**	483.2	0.0230	46.15	< 0.0001
GBC Patient's Blood Arsenic Concentration (µg/L) (n = 152)	**52.28 ± 5.504**	389.3	0.2000	41.41	< 0.0001
GBC Patient's Hair Arsenic Concentration (µg/Kg) (n = 152)	**649.7 ± 93.88**	6980.1	0.2000	464.2	< 0.001
Control Subject's Blood Arsenic concentration (µg/L) (n = 200)	**0.8133 ± 0.1670**	11.70	0	0.4840	< 0.001
Control Subject's Hair Arsenic Concentration (µg/Kg) (n = 200)	**1.541 ± 0.2351**	14.40	0	1.077	< 0.0001

Table 1. Arsenic concentration in biological samples of gallbladder cancer patients Data are represented as Mean ± Standard Error using (ANOVA-Dunnett's Test, *P* < 0.05). Significant values are in bold.
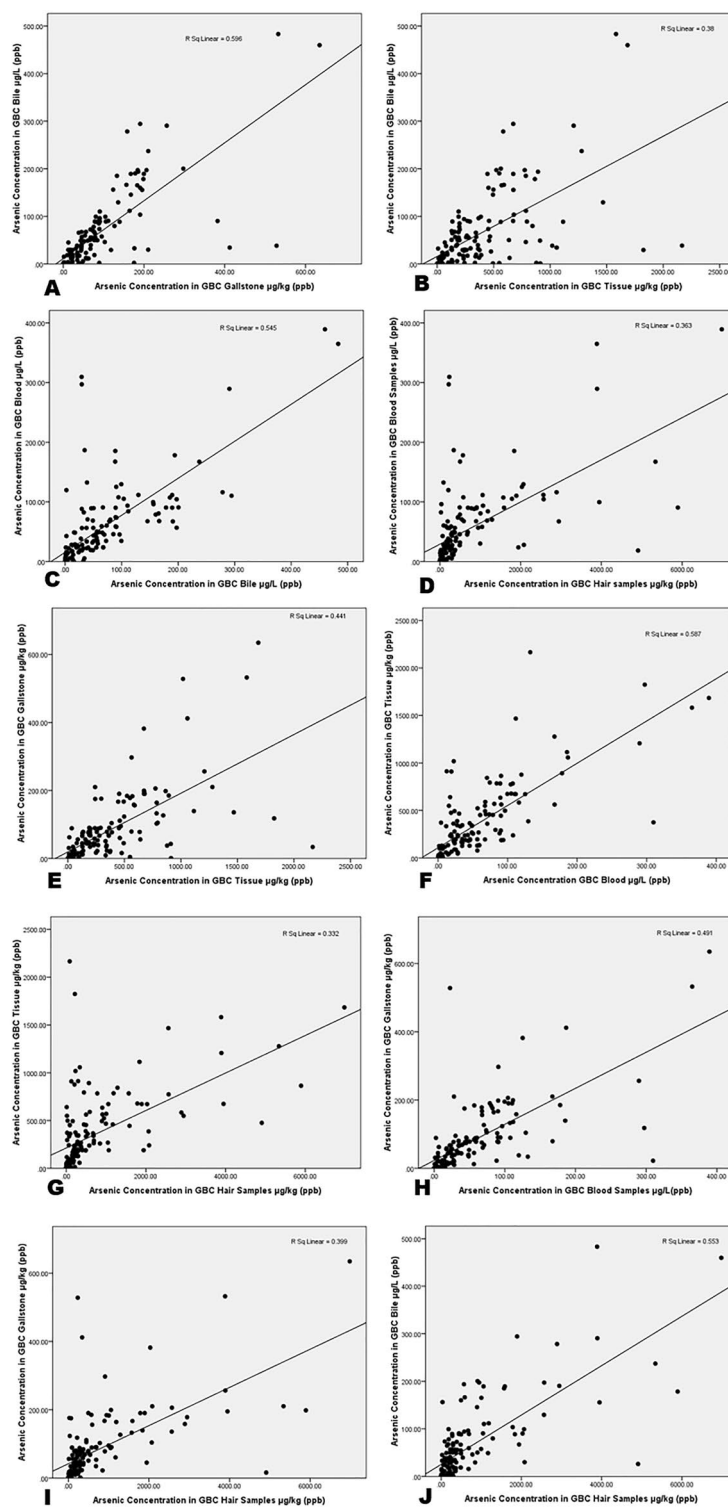
Figure 1. The scatterplot graphs of gallbladder cancer patients (**A**) GBC bile and gallstone (r = 0.596 & $p < 0.05$) (**B**) GBC bile and tissue (r = 0.38 & $p < 0.05$) (**C**) GBC bile and blood (r = 0.545 & $p < 0.05$) (**D**) GBC blood and hair (r = 0.363 & $p < 0.05$) (**E**) GBC gallstone and tissue (r = 0.441 & $p < 0.05$) (**F**) GBC blood and tissue (r = 0.587 & $p < 0.05$) (**G**) GBC hair and tissue (r = 0.332 & $p < 0.05$) (**H**) GBC gallstone and blood (r = 0.491 & $p < 0.05$) (**I**) GBC gallstone and hair (r = 0.399 & $p < 0.05$) (**J**) GBC bile and hair (r = 0.553 & $p < 0.05$).
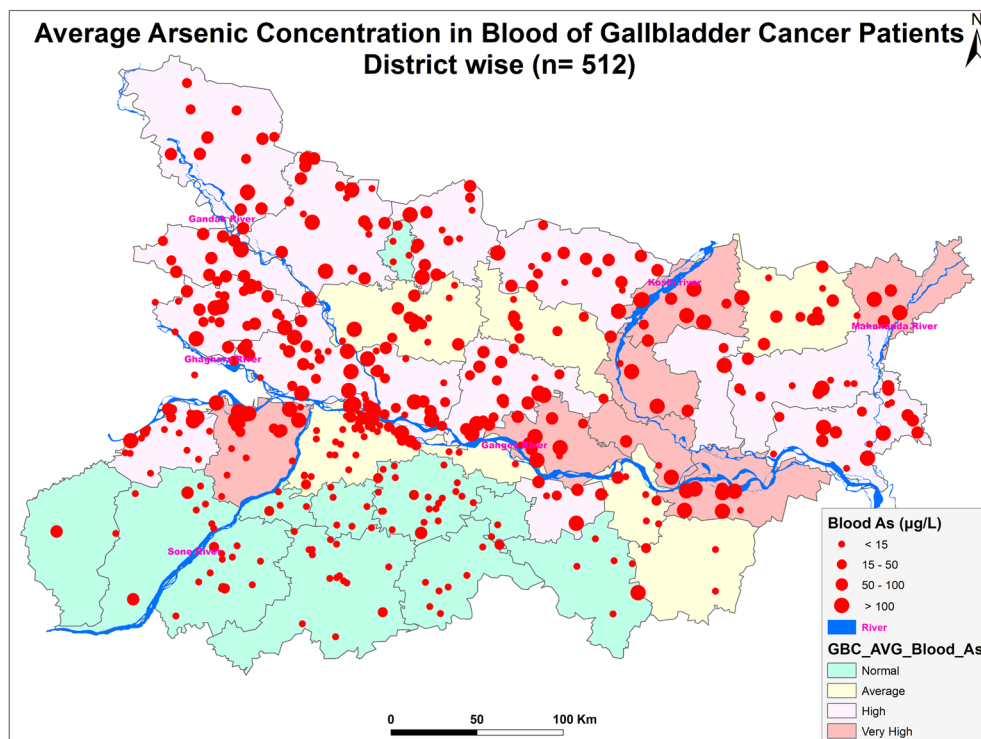
5

Figure 2. Figure showing distribution of GBC cancer patients (n = 512) (Arc-Gis-10.5.1).

9. **Correlation Coefficient between gallbladder cancer patient's tissue and bile:** There was a significant positive correlation showing an increasing trend between the arsenic concentrations in between the gallbladder tissue, gallstone, bile, blood and hair samples of confirmed GBC patients (Fig. 1).

10. **Arsenic Concentration in blood samples of n = 512 gallbladder cancer patients:** The blood arsenic concentration in the n = 512 gallbladder cancer patients showed significant raised levels in n = 274 (53.6%) patients having concentration more than 15 µg/L, while n = 238 (46.4%) had concentration less than 15 µg/L. The maximum arsenic concentration observed was 746.43 µg/L. In the study, n = 338 were female GBC patients while n = 174 were male GBC patients (Fig. 2).

11. **Principal Component Analysis (PCA):** (Table 2) shows the loadings and variances for each principal component.

   Only the first principal component (PC) had the eigenvalue larger than 1 and accounts for 76.4 % of the total variances. Thus, we picked only the PC 1 to construct a holistic evaluation index of the arsenic accumulation in biological samples. Equation (1) shows the constructed index, which is a formula to calculate the PC1 scores for each patient where the coefficients are based on the estimated eigenvector information. Interestingly, the coefficients (and loadings shown in Table 1) are more or less equivalent among the five variables. This means that each standardized variable similarly contributes to maximize the holistic index variance and suggests the heterogeneity in the arsenic accumulation process. (Fig. 3a) shows the PC1 scores ordered from the highest to the lowest with the cancer stage information (stage 1–4 represent the cancer stage, whereas stage 0 means non-cancer onset). (Fig. 3b-f) show the arsenic concentration for the biological samples ordered from the highest to the lowest for each sample.

$$\text{Holistic evaluation index} = 0.436 \times \text{TISSUE} + 0.457 \times \text{BLOOD} + 0.425 \times \text{HAIR}$$
$$+ 0.453 \times \text{STONE} + 0.464 \times \text{BILE} \tag{1}$$

   Figure 3b-f shows that the gallbladder cancer patients tended to have higher arsenic accumulations for each biological sample than non-cancer patients. Figure 3a made the accumulation characteristics more visible and demonstrates that a significant number of cancer patients are found on the right edge of the bar graph. Their overall arsenic accumulation in their bodies is low, so perhaps they have not been intensively exposed to arsenic in their daily lives. Therefore, in the case of these patients, the development of gallbladder cancer is less likely due to the arsenic intake but their inherent vulnerability to the cancer prevalence or other environmental factors.

12. **Proposed Novel Pathway of Gallbladder carcinogenesis:** The present study proposes following pathway for the gallbladder carcinogenesis caused due to arsenic poisoning. Since, humans in the Gangetic plains of Bihar are long term exposed to arsenic poisoning, hence their susceptibility to cause disease is very significant. The mode of exposure causing disease burden has been proposed below (Fig. 4).

	PC1	PC2	PC3	PC4	PC5
TISSUE	0.853	0.399	− 0.230	0.199	− 0.145
BLOOD	0.892	0.268	− 0.025	− 0.314	0.182
HAIR	0.830	− 0.443	− 0.315	0.075	0.105
STONE	0.885	− 0.013	0.376	0.244	0.124
BILE	0.907	− 0.221	0.163	− 0.185	− 0.259
Eigen value	**3.82**	**0.476**	**0.320**	**0.238**	**0.148**
Proportion of variance (%)	**76.4**	**9.53**	**6.41**	**4.76**	**2.96**

Table 2.. PCA Loadings applied to arsenic concentrations in biological samples. Significant values are in bold.

12. Molecular Docking to Establish the Novel Pathway: Association between arsenic and its conjugation with sulfhydryl group compounds causing carcinogenesis in Gallbladder Cancer

Protein ligand interaction between human hemoglobin chain A and chain B were studied by molecular docking where arsenous acid docked with human hemoglobin molecules and H-bonded with Cys104 and His 103 residues in chain A and His146 and Cys 93 residues in chain B Fig. 5A,B. The molecular interaction analyzed by ligplot has been presented in Fig. 5C for chain A and Fig. 5D for chain B which also shows the importance of SER 35 in the interaction. Interestingly, sulfhydryl group of taurine also interact with arsenic compound (Fig. 5D,E).

Geological aspect. In 18 out of 38 districts in Bihar, have been affected by arsenic contamination in groundwater. Major source of arsenic for human consumption is through arsenic contaminated water[62]. The predominance of gall bladder cases is clearly brought out in the arsenic affected districts of Bihar (Fig. 6). The source of arsenic is primarily geogenic as it is acquiring endemic proportions and affecting a large geographical area. The rivers originating from extra peninsular region contain sediments with high arsenic content. The rivers carry these sediments which under favourable conditions releases arsenic in groundwater. The residency time, litho framework of the recharge vis a vis discharge area, age of the water in the affected aquifer and climatic conditions also have a role to play in enhanced contamination. The secondary enrichment in favourable conditions of oxic and anoxic waters supplemented by the geomorphological components also play a decisive role in increased contamination of aquifers at various depth levels. Arsenic contaminated aquifers are mainly present in newer alluvium (Holocene age) which is also supported by a model which defines the primary and secondary source.

Discussion

Arsenic poisoning through drinking water and irrigation water in the recent times has caused serious health hazards in the population worldwide. Prolonged exposure to arsenic leads to many health outcomes. The diseases caused by this long-term arsenic exposure include skin manifestations, hypertension, ischemia, general body weakness, vascular disorders, auto-immune disorders (diabetes, arthritis, leukoderma), severe arteriosclerosis, neuropathies and many types of cancer[51,61,63]. In most of the cases if these symptoms are ignored, arsenic poisoning may lead to many types of cancer. One of highly correlated health outcome of arsenic poisoning is gallbladder cancer. In recent studies gallbladder cancer has been diagnosed in many patients coming from the arsenic hotspots of Gangetic plains of Bihar[64]. However, the etiology of aforementioned diseases leading to carcinogenesis has not been studied extensively. The exposure rate is so much high that one or more members in each exposed households exhibit arsenicosis symptoms[65–67].

Arsenic has been classified as a class I human carcinogen by International Agency of Research on Cancer (IARC), which denotes that arsenic possesses the property of not only changing the configuration of the cellular activity but also modifying the gene functions at the genome level in human beings. In the recent studies, skin and several types of internal cancers such as bladder, kidney, liver, prostate and lung have been associated with arsenic ingestion[51,68–71]. Arsenic usually affects the cellular activities mainly by two pathways –arsenic induced oxidative stress and also through epigenetic changes. The former is mediated by the biotransformation of arsenic leading to the cellular damage through the production of reactive oxygen species (ROS). The ROS typically consists of superoxide anions, hydrogen peroxides, and hydroxyl radicals, which can damage the DNA causing chromosomal aberrations. Enhanced arsenic induced production of ROS by the cells has been linked to the carcinogenicity of arsenic[30,72–77]. The other suggested mechanism of arsenic carcinogenicity is the epigenetic changes which occur due to the changes in DNA methylation, histone modification and microRNAs. The hypermethylation of aberrant DNA promoter is strongly linked with the silencing of the tumour suppressor genes such as gene *p53*. The gene *p53* is known to be the guardian of the genome, and silencing of this important gene can lead to the transformation of a normal cell into a cancer cell[78–81]. The modification of histone proteins by histone acylation is also known to be the cause of arsenic carcinogenicity. Arsenic metabolites have been shown to modulate the normal histone patterns in the cells[82,83]. The microRNAs are small noncoding RNAs that inhibit the expression of the genes, and they can also induce carcinogenicity by deregulating the normal functions of the cell. The chronic exposure of As[III] has been suggested to induce malignant transformation[84–87]. Gallbladder cancer in recent time has increased many folds in the Gangetic belt of India. The specific reasons for this increase have not been established yet, but few studies indicated the toxic metals or metalloids and pesticides as the causative agents[88–94].
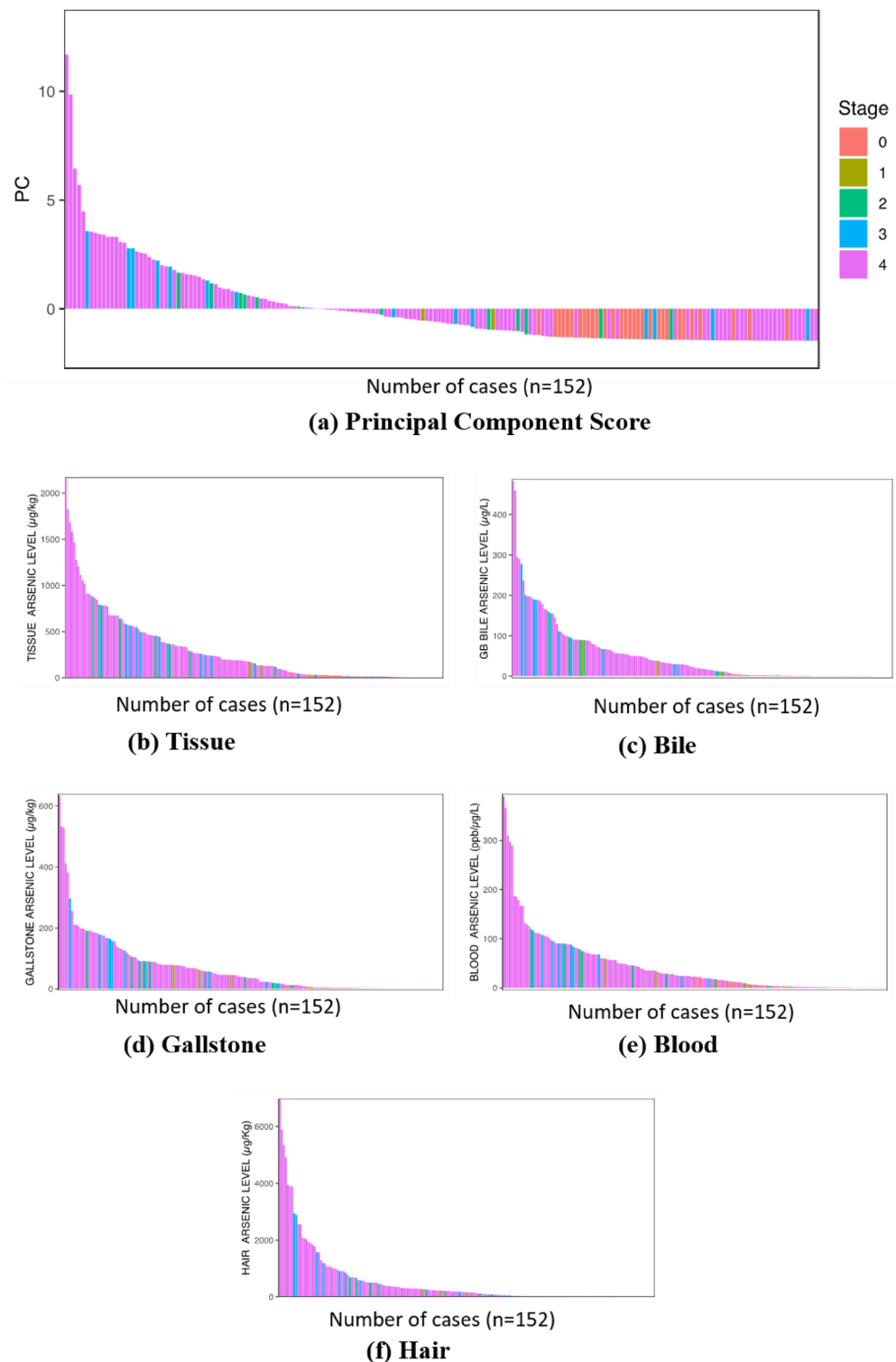
Figure 3. The bar graph on (**a**) PC scores and (**b**–**c**) arsenic concentration for each biological sample with the cancer stage information.

In the present study, consistently very high arsenic concentrations were recorded in the blood, gallbladder tissue, gallstones, gallbladder bile and the hair samples of gallbladder cancer patients. The comparison of the control (n = 200) and confirmed gallbladder cancer patients (n = 152) showed novel and significant findings. The study also confirms that the studied control subjects had non-significant arsenic concentration in their blood and hair samples. A significantly high arsenic concentration ($p < 0.05$) was detected in the blood samples with maximum concentration 389 µg/L in GBC cases in comparison to control. Similarly, in the gallbladder cancer
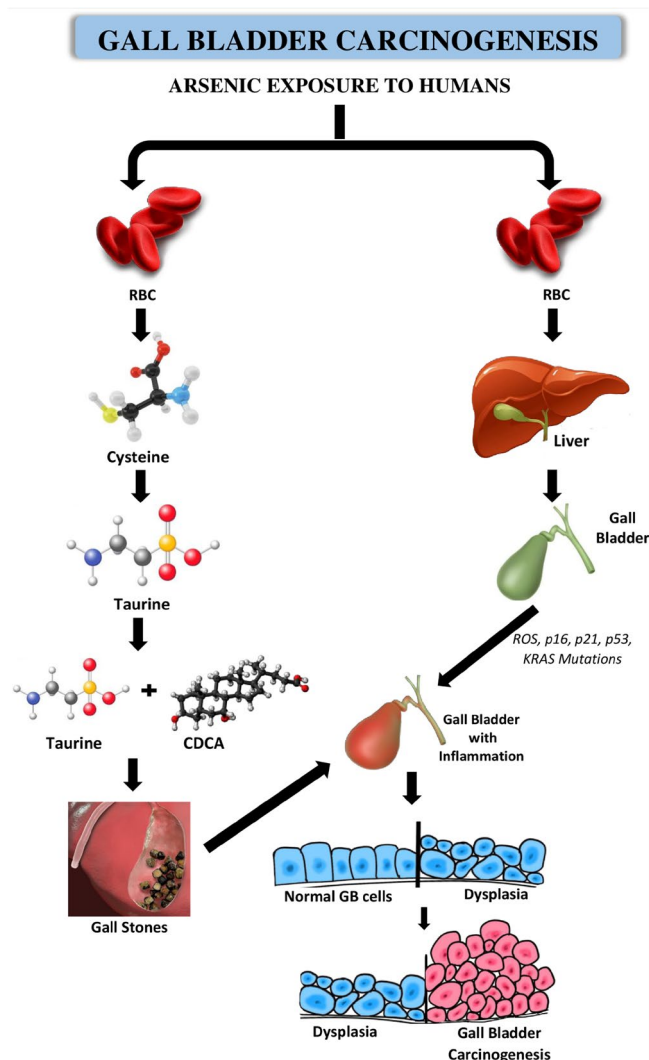
Figure 4. Schematic diagram showing the etiology of Gallbladder carcinogenesis.

patients, there was significantly high arsenic concentration observed in gallbladder tissue with highest concentration of 2166 µg/kg, in gallbladder stones 635 µg/kg, in bile samples 483 µg/L and in hair samples 6980 µg/kg respectively. There has been no benchmark range setup for the arsenic contamination in gall bladder tissue, bile, stones and blood, but for hair samples the normal levels of arsenic contamination in the unexposed human populaces ranges between 20 and 200 µg/kg. There was a significant positive correlation as increasing trend between the arsenic concentrations in between the gallbladder tissue, gallstone, bile, blood and hair samples of confirmed GBC patients were observed. Moreover, the n = 512 gallbladder cancer patient's blood samples study revealed very significant arsenic concentration in the population of Bihar with maximum arsenic concentration as 746 µg/L. Out of n = 512, n = 350 blood samples had relatively had arsenic concentration more than 2 µg/L. This means that 68% of the studied population had significant arsenic concentration in their blood. Most of the patients were form the Gangetic plain region in comparison to the non-Gangetic plain area. This positive trend indicates a strong association between arsenic contamination and the observed carcinogenesis.

No signalling pathways have yet found linking arsenic accumulation in the gallbladder tissue or gallstones or gallbladder bile to carcinogenicity. Nonetheless, the present study provides strong circumstantial evidence associating arsenic accumulation with the progression of the gallbladder carcinogenesis. Arsenic in the Gangetic plains of Bihar is also found in the form of arsenopyrite (FeAsS)[95]. It has very high affinity to bind with these two elements, hence it easily binds with endogenous sulfhydryl compounds including metallothioneins. Arsenic reaches the human blood after the gastrointestinal absorption and may cause gallbladder carcinogenesis in two ways. First, it is hypothesized that, arsenic binds with RBC since haemoglobin in RBC contains the iron, and then after the completion of RBC lifespan (120 days) reaches liver via spleen[96–98]. In liver, arsenic binds with cysteine residues which compounds with sulfhydryl group and accumulates further[99–103]. Subsequently, arsenic binds with the next sulfhydryl group containing compound taurine[104–111]. Taurine is the key factor which controls the formation of the gallstones, however arsenic likely supresses its function by lowering its levels. This initiates the formation of gallstones with sludge in the gallbladder due to the precipitation of the cholesterol[112,113]. The
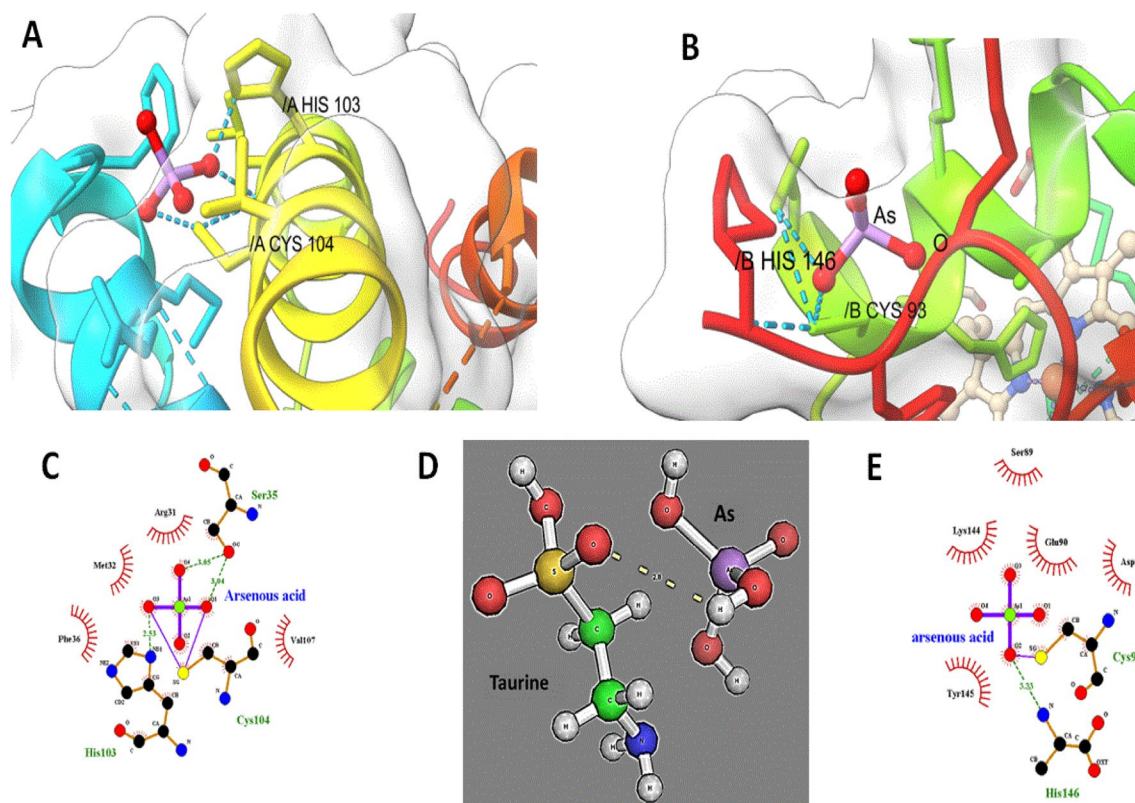
Figure 5. (**A**) Docking structure of Human Haemoglobin (PDB ID: 4HHB, bound with As III (Arsenous acid, PubChem ID: 545) at Cys 104 residue in chain A. (**B**) Docking structure of Human Haemoglobin (PDB ID: 4HHB, bound with As III (Arsenous acid, PubChem ID: 545) at Cys 93 residue in chain B (**C**) Hydrophilic and Hydrophobic interaction of the As III (arsenous acid) bound with chain A of human haemoglobin. (**D**) Interaction of taurine with Arsenic (**E**) Hydrophilic and Hydrophobic interaction of the As III (arsenous acid) bound with chain B of human haemoglobin.

taurine has the conjugation with the chenodeoxycholic acid (CDCA), that is why high arsenic contamination was observed in the bile of the GBC patients. Gallstones in due course of time causes inflammation in the inner wall of the gallbladder tissue, which after mutations transforms into a cancerous tissue[113]. Secondly, it is also hypothesized that arsenic in the gallbladder elicits rigorous production of ROS, leading to frequent mutations in the genes p53, p21, p16 and KRAS. Such mutations in turn induces inflammation in the epithelial cells of gallbladder causing conversion of normal cells to dysplasia, which in the long run leads to gallbladder carcinoma[112,114–131].

Molecular Docking studies of arsenic compounds with haemoglobin alpha chain and beta chain have revealed the efficient binding of arsenic compounds with cysteine residues. Arsenous acid was able to bind with Cys 104 residue of haemoglobin alpha chain globin chain A and also binds with the Cys 93 residue of haemoglobin beta chain globin chain B. Other important residues which can be involved in arsenic toxicity to haemoglobin are His, Ser which are efficiently involved in the interaction with arsenic. The toxicity of trivalent arsenicals likely to occur through the interaction of trivalent arsenic species with sulfhydryl groups in proteins. Trivalent arsenicals have high affinity for sulfhydryl groups and can bind to reduced cysteines in peptides and proteins. Several studies have reported the binding of trivalent arsenicals to the cysteine residues of proteins such as iAsIII binds at three cysteine residues, Cys32, Cys34 and Cys37 of ArsR repressor protein of *E. coli*. Cys156 and Cys206 of Arsenic methyltransferase (AS3MT) of human are the active site for arsenic binding whereas Cys72, Cys174 and Cys224 of AS3MT of thermophilic eukaryotic red algae *Cyanidioschyzon merolae* are believed to be involved in arsenic binding. Hence, the present study also strongly indicates strong association between the chronic arsenic contamination in the Gangetic plains of the Bihar (India) and gallbladder carcinogenesis.

A previous significant study also reported association between the arsenic exposure and gallbladder cancer incidences in the populations of India and Japan. However, the study was conducted using a very small number of gallbladder cancer patients (8 Indian and 5 Japanese GBC cases), and therefore a strong correlation could not be made. Moreover, they also suggested association of gallbladder carcinogenesis with the other heavy metals like chromium, cadmium, lead, zinc and mercury[132].

Similarly, in Chile (South America), the annual cancer incidences are approximately, 35,000 new cases per year and the cancer burden in the population in recent times has increased many folds similar to India[133–136]. The gallbladder cancer incidences in Chile have a higher ratio in comparison to the average data on a global scale[137,138]. Unfortunately, Chile also reports high arsenic concentration in the groundwater with concentrations as high as about 5,000 µg/L[139]. Not much is known about the association between arsenic exposure and increased
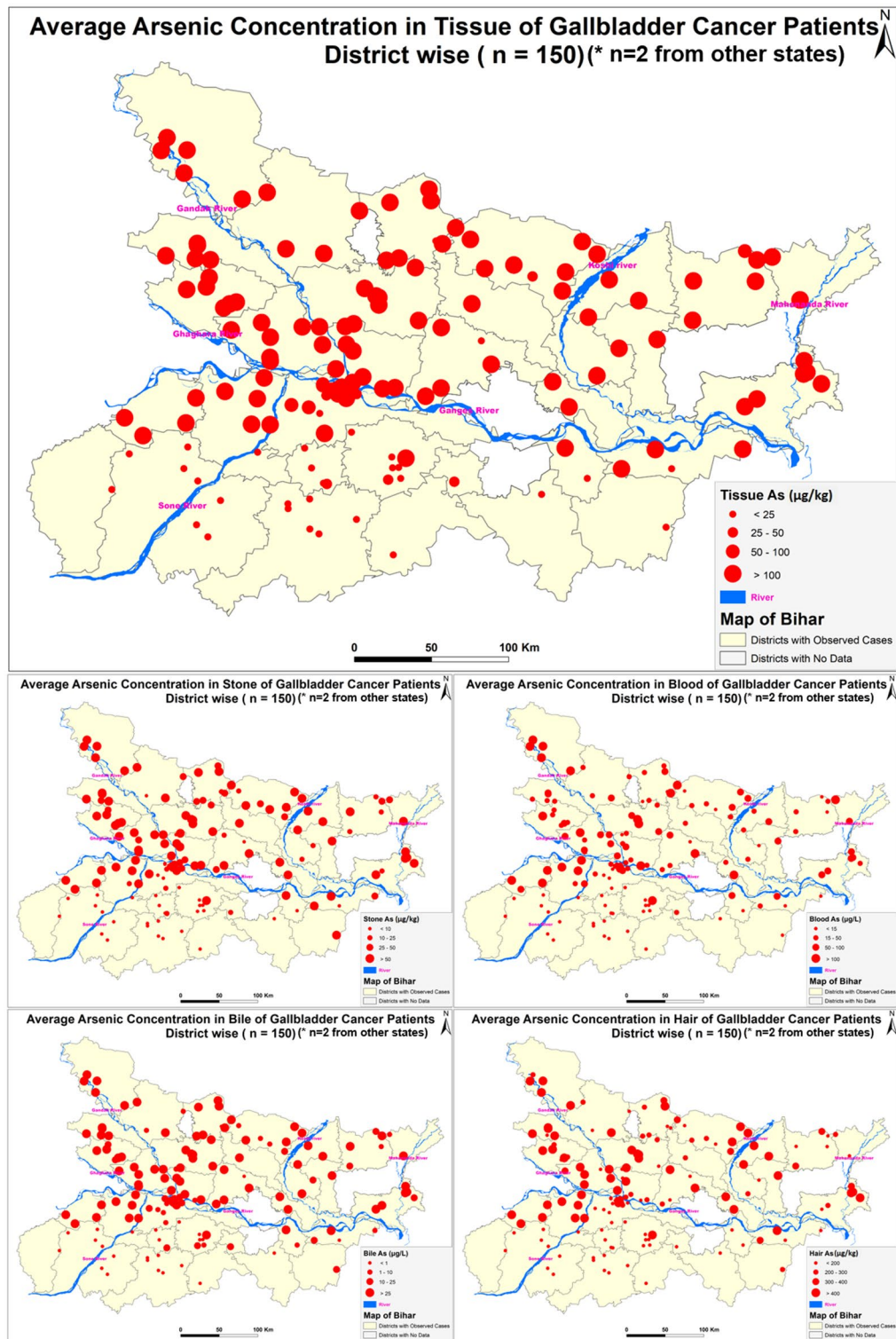
Figure 6. Map showing studied Gallbladder cancer patients with arsenic concentration in different biological samples with geomapping (Arc-Gis-10.5.1).

gallbladder cancer cases in Chile[140,141]. Recent studies, reported that if the exposed subjects are consuming arsenic contaminated water for long duration may cause risk of gallbladder cancer[142–144].

The PCA results demonstrated that the internal arsenic concentrations could not characterize all the severe gallbladder patients, but all the highly arsenic-loaded subjects were the severe gallbladder patients. This suggests the association between arsenic intake and gallbladder cancer. The high internal arsenic concentration may cause accelerating the progress of gallbladder cancer. Another suggestion was that a single biological medium

might not be sufficient to understand the real affection of the internal arsenic concentrations: e.g., some patients had high arsenic concentrations in gallbladder tissues, but hair samples showed low concentrations and vice versa. The way how arsenic is metabolized and accumulated in bodies may differ depending on the individuals. Moreover, the arsenic binding to cysteine of RBC, taurine in liver is well proved through our molecular docking which denotes the pathway of GBC disease carcinogenesis. This novel pathway of disease carcinogenesis strongly validates the hypothesis.

Conclusions

From the entire study, it can be concluded from the present study that gallbladder cancer is prevalent in the Gangetic plains of Bihar, where the arsenic contamination in the exposed population is very high increasing the disease burden of the state. The significantly high arsenic concentration observed in the gallbladder tissue, gallbladder stone, bile, blood and hair samples in GBC patients strongly indicates the linkage between high chronic exposure to arsenic and gallbladder carcinogenesis. The GBC disease burden is significantly increasing manyfolds in the arsenic exposed population. However, the other confounding factors can also add the disease burden of GBC manifolds. The novel pathway of GBC carcinogenesis validates that arsenic is one of the important toxicants which is responsible for causing the disease in this particular area. Hence, there is an urgent need to control the disease burden by developing policies and guidelines that will alleviate arsenic exposure to the impacted populations.

Data availability

The datasets used for the bioinformatic analysis is publicly available from RCSB Protein Data Bank (https://www.rcsb.org/). The Human Haemoglobin Molecule 3D -X-ray Crystallographic structure at PDB (PDB ID:4HHB) is freely accessible at https://www.rcsb.org/structure/4HHB (https://doi.org/10.2210/pdb4HHB/pdb). The software used in the map layer generation was ArcGIS Version 10.5.1. All the other data materials have been shared in the Zip form in the journal which on request can be obtained from the journal.

References

1. Randi, G., Franceschi, S. & La Vecchia, C. Gallbladder cancer worldwide: Geographical distribution and risk factors. *Int. J. Cancer* **118**, 1591–1602 (2006).
2. Are, C. *et al.* Global epidemiological trends and variations in the burden of gallbladder cancer. *J. Surg. Oncol.* **115**, 580–590 (2017).
3. GLOBOCAN 2020. https://gco.iarc.fr/today/data/factsheets/populations/900-world-fact-sheets.pdf. Accessed 11 Dec 2020.
4. Dutta, U., Bush, N., Kalsi, D., Popli, P. & Kapoor, V. K. Epidemiology of gallbladder cancer in India. *Chin. Clin. Oncol.* **8**, 33 (2019).
5. ICMR (Indian Council of Medical Research) 2020. https://ncdirindia.org/All_Reports/Report_2020/default.aspx. Accessed 11 Apr 2021.
6. Sharma, D. C. Cancer data in India show new patterns. *Lancet Oncol* **17**, e272 (2016).
7. American Cancer Society. Cancer Facts (ACS) and Fig.s 2016. Atlanta, GA: American Cancer Society; 2016.
8. Sharma, A., Sharma, K. L., Gupta, A., Yadav, A. & Kumar, A. Gallbladder cancer epidemiology, pathogenesis and molecular genetics: Recent update. *World J. Gastroenterol.* **23**, 3978 (2017).
9. Hundal, R. & Shaffer, E. A. Gallbladder cancer: Epidemiology and outcome. *Clin. Epidemiol.* **6**, 99 (2014).
10. Misra, S., Chaturvedi, A., Misra, N. C. & Sharma, I. D. Carcinoma of the gallbladder. *Lancet Oncol.* **4**, 167–176 (2003).
11. Kapoor, V. K. Advanced gallbladder cancer: Indian "middle path". *J. Hepatobiliary Pancreat. Surg.* **14**, 366–373 (2007).
12. Murthy, N. S. *et al.* Risk of cancer development in India. *Asian Pac. J. Cancer Prev.* **12**, 387–391 (2011).
13. Krishnatreya, M., Saikia, A., Kataki, A. C., Sharma, J. D. & Barnah, M. Variations in the spatial distribution of gall bladder cancer: A call for collaborative action. *Ann. Med. Health Sci. Res.* **4**, 329–331 (2014).
14. Bhattacharjee, P. K. & Nanda, D. Prospective observational study on cholelithiasis in patients with carcinoma gall bladder in a tertiary referral hospital of Eastern India. *J. Cancer Res. Ther.* **15**, 153 (2019).
15. Dhir, V. & Mohandas, K. M. Epidemiology of digestive tract cancers in India IV. Gall bladder and pancreas. *Indian J. Gastroenterol.* **18**, 24–28 (1999).
16. Hariharan, D., Saied, A. & Kocher, H. M. Analysis of mortality rates for gallbladder cancer across the world. *HPB* **10**, 327–331 (2008).
17. Barreto, S. G., Haga, H. & Shukla, P. J. Hormones and gallbladder cancer in women. *Indian J. Gastroenterol.* **28**, 126–130 (2009).
18. Jain, K., Sreenivas, V., Velpandian, T., Kapil, U. & Garg, P. K. Risk factors for gallbladder cancer: A case–control study. *Int. J. Cancer* **132**, 1660–1666 (2013).
19. Pilgrim, C. H., Groeschl, R. T., Christians, K. K. & Gamblin, T. C. Modern perspectives on factors predisposing to the development of gallbladder cancer. *HPB* **15**, 839–844 (2013).
20. Diehl, A. K. Epidemiology of gallbladder cancer: A synthesis of recent data. *J. Natl. Cancer Inst.* **65**, 1209–1214 (1980).
21. Pandey, M. Risk factors for gallbladder cancer: A reappraisal. *Eur. J. Cancer Prev.* **12**, 15–24 (2003).
22. Shaffer, E. A. Gallbladder cancer: The basics. *Gastroenterol. Hepatol. (N. Y.)* **4**, 737 (2008).
23. Guadagnolo, B. A., Petereit, D. G. & Coleman, C. N. Cancer care access and outcomes for American Indian populations in the United States: Challenges and models for progress. *Semin. Radiat. Oncol.* **27**, 143–149 (2017).
24. Nemunaitis, J. M. *et al.* Gallbladder cancer: Review of a rare orphan gastrointestinal cancer with a focus on populations of New Mexico. *BMC Cancer* **18**, 1–14 (2018).
25. Boelee, E., Geerling, G., van der Zaan, B., Blauw, A. & Vethaak, A. D. Water and health: From environmental pressures to integrated responses. *Acta Trop.* **193**, 217–226 (2019).
26. Shankar, S. & Shanker, U. Arsenic contamination of groundwater: A review of sources, prevalence, health risks, and strategies for mitigation. *Sci. World J.* **2014**, 304524 (2014).
27. Bhattacharya, P. *et al.* Temporal and seasonal variability of arsenic in drinking water wells in Matlab, southeastern Bangladesh: A preliminary evaluation on the basis of a 4 year study. *J. Environ. Sci. Health A Tox. Hazard. Subst. Environ. Eng.* **46**, 1177–1184 (2011).

28. Bundschuh, J., Armienta, B., Bhattacharya, P. & Matschullat, M. *Natural Arsenic in Groundwaters of Latin America. Arsenic in the Environment* (CRC Press 2009).
29. IPCS (International Programme on Chemical Safety). *Environmental Health Criteria 224 Arsenic and arsenic compounds* (World Health Organization, Geneva, 2001).
30. Straif, K. *et al.* A review of human carcinogens—part C: Metals, arsenic, dusts, and fibres. *Lancet Oncol.* **10**, 453–454 (2009).
31. Valko, M., Rhodes, C. J. B., Moncol, J., Izakovic, M. M. & Mazur, M. Free radicals, metals and antioxidants in oxidative stress-induced cancer. *Chem. Biol. Interact.* **160**, 1–40 (2006).
32. IARC (International Agency for Research on Cancer) Working Group on the Evaluation of Carcinogenic Risks to Humans. Some drinking-water disinfectants and contaminants, including arsenic. *IARC Monogr. Eval. Carcinog Risks Hum.* **84**, 1–477 (2004).
33. Karagas, M. R., Stukel, T. A. & Tosteson, T. D. Assessment of cancer risk and environmental levels of arsenic in New Hampshire. *Int. J. Hyg. Environ. Health* **205**, 85–94 (2002).
34. USEPA National Primary Drinking Water Regulations. Arsenic and clarifications to compliance and new source contaminants monitoring. *Fed. Regist.* **40**(9), 141–142 (2001).
35. Upadhyay, M. K., Shukla, A., Yadav, P. & Srivastava, S. A review of arsenic in crops, vegetables, animals and food products. *Food Chem* **276**, 608–618 (2019).
36. Sharma, A. K., Tjell, J. C., Sloth, J. J. & Holm, P. E. Review of arsenic contamination, exposure through water and food and low-cost mitigation options for rural areas. *Appl. Geochem.* **41**, 11–33 (2014).
37. Santra, S. C. *et al.* Arsenic in food chain and community health risk: A study in Gangetic West Bengal. *Procedia Environ. Sci.* **18**, 2–13 (2013).
38. Tuli, R., Chakrabarty, D., Trivedi, P. K. & Tripathi, R. D. Recent advances in arsenic accumulation and metabolism in rice. *Mol. Breed.* **26**, 307–323 (2010).
39. Rossman, T. G. Mechanism of arsenic carcinogenesis: An integrated approach. *Mutat. Res.* **533**, 37–65 (2003).
40. Yoshida, T., Yamauchi, H. & Sun, G. F. Chronic health effects in people exposed to arsenic via the drinking water: Dose–response relationships in review. *Toxicol. Appl. Pharmacol.* **198**, 243–252 (2004).
41. Tapio, S. & Grosche, B. Arsenic in the aetiology of cancer. *Mutat. Res. Rev. Mutat. Res.* **612**, 215–246 (2006).
42. Zhou, Q. & Xi, S. A review on arsenic carcinogenesis: Epidemiology, metabolism, genotoxicity and epigenetic changes. *Regul. Toxicol. Pharmacol. RTP* **99**, 78–88 (2018).
43. Gentry, P. R. *et al.* Analysis of genomic dose–response information on arsenic to inform key events in a mode of action for carcinogenicity. *Environ. Mol. Mutagen* **51**, 1–14 (2010).
44. Sharma, A. K., Tjell, J. C. & Mosbæk, H. Health effects from arsenic in groundwater of the Bengal delta: Effects of iron and water storage practices. *Environ. Geosci.* **13**, 17–29 (2006).
45. Smith, A. H. *et al.* Lung, bladder, and kidney cancer mortality 40 years after arsenic exposure reduction. *J. Natl. Cancer Inst.* **110**, 241–249 (2018).
46. Shi, H., Shi, X. & Liu, K. J. Oxidative mechanism of arsenic toxicity and carcinogenesis. *Mol. Cell. Biochem.* **255**, 67–78 (2004).
47. Wang, T. S. *et al.* Arsenite induces oxidative DNA adducts and DNA-protein cross-links in mammalian cells. *Free Radic. Biol. Med.* **31**, 321–330 (2001).
48. Suzuki, K. T., Kurasaki, K. & Suzuki, N. Selenocysteine β-lyase and methylselenol demethylase in the metabolism of Se-meth-ylated selenocompounds into selenide. *Biochim. Biophys. Acta Gen. Subj.* **1770**, 1053–1061 (2007).
49. Zhang, T. C., Schmitt, M. T. & Mumford, J. L. Effects of arsenic on telomerase and telomeres in relation to cell proliferation and apoptosis in human keratinocytes and leukemia cells in vitro. *Carcinogenesis* **24**, 1811–1817 (2003).
50. Benbrahim-Tallaa, L. *et al.* Molecular events associated with arsenic-induced malignant transformation of human prostatic epithelial cells: Aberrant genomic DNA methylation and K-ras oncogene activation. *Toxicol. Appl. Pharmacol.* **206**, 288–298 (2005).
51. Martinez, V. D., Vucic, E. A., Becker-Santos, D. D., Gil, L. & Lam, W. L. Arsenic exposure and the induction of human cancers. *J. Toxicol.* **2011,** 431287 (2011).
52. Kumar, A. *et al.* Arsenic exposure in Indo Gangetic plains of Bihar causing increased cancer risk. *Sci. Rep.* **11**, 1–16 (2021).
53. Madhawi, R. *et al.* Geographical pattern of carcinoma gallbladder in Bihar and its association with river Ganges and arsenic levels: Retrospective individual consecutive patient data from Regional Cancer Centre. *South Asian J. Can.* **7**, 167–170 (2018).
54. Daniel, W. W. *Biostatistics: A Foundation for Analysis in the Health Sciences* 7th edn. (Wiley, New York, 1999).
55. NIOSH (National Institute for Occupational Safety and Health). Elements in blood or tissue. Method: 8005. In *NIOSH Manual of Analytical Methods,* 4th edn 94–113 (Department of Health and Human Service, National Institute of Occupational Safety and Health, Cincinnati, DHHS publication, 1994)..
56. Berman, H. M. *et al.* The protein data bank. *Nucl. Acids Res.* **28**, 235–242 (2000).
57. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Prot. Sci.* **30**, 70–82 (2021).
58. Yan, Y., Tao, H., He, J. & Huang, S. Y. The HDOCK server for integrated protein-protein docking. *Nat. Protoc.* **15**, 1829–1852 (2020).
59. DeLano, W. L. Pymol: An open-source molecular graphics tool. *CCP4 Newslett. Prot. Crystallogr.* **40**, 82–92 (2002).
60. Kumar, A. *et al.* Assessment of arsenic exposure in the population of Sabalpur village of Saran District of Bihar with mitigation approach. *Environ. Sci. Pollut. Res. Int.* **28**, 43923–43934 (2021).
61. R Core Team R. *A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2022). https://www.R-project.org/.
62. Kumar, A. & Ghosh, A. K. Assessment of arsenic contamination in groundwater and affected population of Bihar. In *Arsenic Toxicity: Challenges and Solutions* (ed. Kumar, N.) (Springer, Singapore, 2021).
63. Engel, R. R., Hopenhayn-Rich, C., Receveur, O. & Smith, A. H. Vascular effects of chronic arsenic exposure: A review. *Epidemiol. Rev.* **16**, 184–209 (1994).
64. Kumar, A. *et al.* Assessment of arsenic exposure and its mitigation intervention in severely exposed population of Buxar district of Bihar, India. *Toxicol. Environ. Health Sci.* **13**, 287–297 (2021).
65. Chen, B. *et al.* Therapeutic and analytical applications of arsenic binding to proteins. *Metallomics* **7**, 39–55 (2015).
66. Kadono, T. *et al.* Skin manifestations of arsenicosis in two villages in Bangladesh. *Int. J. Dermatol.* **41**, 841–846 (2002).
67. Tseng, W. P. Effects and dose–response relationships of skin cancer and blackfoot disease with arsenic. *Environ. Health Perspect.* **19**, 109–119 (1977).
68. Marshall, G. *et al.* Fifty-year study of lung and bladder cancer mortality in Chile related to arsenic in drinking water. *J. Natl. Cancer Inst.* **99**, 920–928 (2007).
69. Chen, C. J. & Wang, C. J. Ecological correlation between arsenic level in well water and age-adjusted mortality from malignant neoplasms. *Cancer Res.* **50**, 5470–5474 (1990).
70. Bates, M. N. *et al.* Case-control study of bladder cancer and exposure to arsenic in Argentina. *Am. J. Epidemiol.* **159**, 381–389 (2004).
71. Ferreccio, C. *et al.* Lung cancer and arsenic concentrations in drinking water in Chile. *Epidemiology.* **11**(6), 673–679 (2000).
72. Li, L. & Chen, F. Oxidative stress, epigenetics, and cancer stem cells in arsenic carcinogenesis and prevention. *Curr. Pharmacol. Rep.* **2**, 57–63 (2016).

73. Lantz, R. C. & Hays, A. M. Role of oxidative stress in arsenic-induced toxicity. *Drug Metab. Rev.* **38**, 791–804 (2006).
74. Kitchin, K. T. & Ahmad, S. Oxidative stress as a possible mode of action for arsenic carcinogenesis. *Toxicol. Lett.* **137**, 3–13 (2003).
75. Vahter, M. Mechanisms of arsenic biotransformation. *Toxicology* **181**, 211–217 (2002).
76. Kessel, M., Liu, S. X., Xu, A., Santella, R. & Heil, T. K. Arsenic induces oxidative DNA damage in mammalian cells. In *Oxygen/Nitrogen Radicals: Cell Injury and Disease* 301–308 (Springer, Boston, MA, 2002).
77. Kitchin, K. T. Recent advances in arsenic carcinogenesis: Modes of action, animal model systems, and methylated arsenic metabolites. *Toxicol. Appl. Pharmacol.* **172**, 249–261 (2001).
78. Mass, M. J. & Wang, L. Arsenic alters cytosine methylation patterns of the promoter of the tumor suppressor gene p53 in human lung cells: A model for a mechanism of carcinogenesis. *Mutat. Res. Rev. Mutat. Res.* **386**, 263–277 (1997).
79. Cui, X., Wakai, T., Shirai, Y., Hatakeyama, K. & Hirano, S. Chronic oral exposure to inorganic arsenate interferes with methylation status of p16INK4a and RASSF1A and induces lung cancer in A/J mice. *Toxicol. Sci.* **91**, 372–381 (2006).
80. Marsit, C. J., Karagas, M. R., Schned, A. & Kelsey, K. T. Carcinogen exposure and epigenetic silencing in bladder cancer. *Ann. N. Y. Acad. Sci.* **1076**, 810–821 (2006).
81. Chanda, S. *et al.* DNA hypermethylation of promoter of gene p53 and p16 in arsenic-exposed people with and without malignancy. *Toxicol. Sci.* **89**, 431–437 (2006).
82. Zhou, X., Sun, H., Ellen, T. P., Chen, H. & Costa, M. Arsenite alters global histone H3 methylation. *Carcinogenesis* **29**, 1831–1836 (2008).
83. Jensen, T. J., Novak, P., Eblin, K. E., Gandolfi, A. J. & Futscher, B. W. Epigenetic remodeling during arsenical-induced malignant transformation. *Carcinogenesis* **29**, 1500–1508 (2008).
84. Filipowicz, W., Bhattacharyya, S. N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight?. *Nat. Rev. Genet.* **9**, 102–114 (2008).
85. Calin, G. A. & Croce, C. M. MicroRNA signatures in human cancers. *Nat. Rev. Cancer* **6**, 857–866 (2006).
86. Marsit, C. J., Eddy, K. & Kelsey, K. T. MicroRNA responses to cellular stress. *Cancer Res.* **66**, 10843–10848 (2006).
87. Wang, Z. *et al.* Reversal and prevention of arsenic-induced human bronchial epithelial cell malignant transformation by microRNA-200b. *Toxicol. Sci.* **121**, 110–122 (2011).
88. Jagannath, P., Dhir, V. & Mohandas, K. M. Geographic patterns in incidence of Gall Bladder cancer in India and the possible etiopathological factors. *HPB* **2**, 168–169 (2000).
89. Beyersmann, D. & Hartwig, A. Carcinogenic metal compounds: Recent insight into molecular and cellular mechanisms. *Arch. Toxicol.* **82**, 493–512 (2008).
90. Shukla, V. K., Prakash, A., Tripathi, B. D., Reddy, D. C. S. & Singh, S. Biliary heavy metal concentrations in carcinoma of the gall bladder: Case-control study. *BMJ* **317**, 1288–1289 (1998).
91. Singh, S. *et al.* Skeletal metastasis in gall bladder cancer. *HPB (Oxford).* **9**, 71–72 (2007).
92. Singh, V. K., Singh, K. P. & Mohan, D. Status of heavy metals in water and bed sediments of river Gomti—a tributary of the Ganga river, India. *Environ. Monit. Assess* **105**, 43–67 (2005).
93. Unisa, S. *et al.* Population-based study to estimate prevalence and determine risk factors of gallbladder diseases in the rural Gangetic basin of North India. *HPB* **13**, 117–125 (2011).
94. Chakraborti, D. *et al.* Fate of over 480 million inhabitants living in arsenic and fluoride endemic Indian districts: Magnitude, health, socio-economic effects and mitigation approaches. *J. Trace Elem. Med. Biol.* **38**, 33–45 (2016).
95. Haldar, S. K. *Introduction to Mineralogy and Petrology. Minerals and Rocks* 1–51 (2020).
96. Rael, L. T., Ayala-Fierro, F., Bar-Or, R., Carter, D. E. & Barber, D. S. Interaction of arsine with hemoglobin in arsine-induced hemolysis. *Toxicol. Sci.* **90**, 142–148 (2006).
97. Yamauchi, T. *et al.* Possible production of arsenic hemoglobin adducts via exposure to arsine. *J. Occup. Health.* **57**(2), 161–168 (2015).
98. Lu, M. *et al.* Evidence of hemoglobin binding to arsenic as a basis for the accumulation of arsenic in rat blood. *Chem. Res. Toxicol.* **17**, 1733–1742 (2004).
99. Shen, S., Li, X. F., Cullen, W. R., Weinfeld, M. & Le, X. C. Arsenic binding to proteins. *Chem. Rev.* **113**, 7769–7792 (2013).
100. Li, J., Packianathan, C., Rossman, T. G. & Rosen, B. P. Nonsynonymous polymorphisms in the human AS3MT arsenic methylation gene: Implications for arsenic toxicity. *Chem. Res. Toxicol.* **30**, 1481–1491 (2017).
101. Zhou, X. *et al.* Arsenite interacts selectively with zinc finger proteins containing C3H1 or C4 motifs. *J. Biol. Chem.* **286**, 22855–22863 (2011).
102. Zhou, X. *et al.* S-nitrosation on zinc finger motif of PARP-1 as a mechanism of DNA repair inhibition by arsenite. *Oncotarget* **7**, 80482 (2016).
103. Zhao, X., Li, Q., Meng, Q., Yue, C. & Xu, C. Identification and expression of cysteine sulfinate decarboxylase, possible regulation of taurine biosynthesis in Crassostrea gigas in response to low salinity. *Sci. Rep.* **7**, 1–10 (2017).
104. Strom, B. L. *et al.* Biochemical epidemiology of gallbladder cancer. *Hepatology* **23**, 1402–1411 (1996).
105. Thomas, C. E. *et al.* Association between pre-diagnostic serum bile acids and hepatocellular carcinoma: The Singapore Chinese Health Study. *Cancers* **13**, 2648 (2021).
106. Zhao, Z. *et al.* Downregulation of the long non-coding RNA taurine-upregulated gene 1 inhibits glioma cell proliferation and invasion and promotes apoptosis. *Oncol. Lett.* **15**, 4026–4032 (2018).
107. Ueki, I. & Stipanuk, M. H. Enzymes of the taurine biosynthetic pathway are expressed in rat mammary gland. *J. Nutr.* **137**, 1887–1894 (2007).
108. Wang, Q. Q. *et al.* Methylated arsenic metabolites bind to PML protein but do not induce cellular differentiation and PML-RARα protein degradation. *Oncotarget* **6**, 25646 (2015).
109. Miyazaki, T. *et al.* Impaired bile acid metabolism with defectives of mitochondrial-tRNA taurine modification and bile acid taurine conjugation in the taurine depleted cats. *Sci. Rep.* **10**, 1–15 (2020).
110. Kondo, Y., Shinjo, K. & Katsushima, K. Long non-coding RNA s as an epigenetic regulator in human cancers. *Cancer Sci.* **108**, 1927–1933 (2017).
111. Das, J., Ghosh, J., Manna, P. & Sil, P. C. Protective role of taurine against arsenic-induced mitochondria-dependent hepatic apoptosis via the inhibition of PKCδ-JNK pathway. *PLoS ONE* **5**, e12602 (2010).
112. Ganesan, N., Bambino, K., Boffetta, P. & Labgaa, I. Exploring the potential carcinogenic role of arsenic in gallbladder cancer. *Eur. J. Cancer Prev.* **29**, 100–109 (2020).
113. Schmidt, M. A., Marcano-Bonilla, L. & Roberts, L. R. Gallbladder cancer: Epidemiology and genetic risk associations. *Chin. Clin. Oncol.* **8**, 31 (2019).
114. Espinoza, J. A. *et al.* The inflammatory inception of gallbladder cancer. *Biochim. Biophys. Acta Rev. Cancer* **1865**, 245–254 (2016).
115. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–339 (2013).
116. Li, M. *et al.* Whole-exome and targeted gene sequencing of gallbladder carcinoma identifies recurrent mutations in the ErbB pathway. *Nat. Genet.* **46**, 872–876 (2014).
117. Nakamura, H. *et al.* Genomic spectra of biliary tract cancer. *Nat Genet* **47**, 1003–1010 (2015).
118. Moreno, M., Pimentel, F., Gazdar, A. F., Wistuba, I. I. & Miquel, J. F. TP53 abnormalities are frequent and early events in the sequential pathogenesis of gallbladder carcinoma. *Ann. Hepatol.* **4**, 192–199 (2005).

119. Jain, K. *et al.* Sequential occurrence of preneoplastic lesions and accumulation of loss of heterozygosity in patients with gallbladder stones suggest causal association with gallbladder cancer. *Ann. Surg.* **260**, 1073–1080 (2014).
120. Gouas, D., Shi, H. & Hainaut, P. The aflatoxin-induced TP53 mutation at codon 249 (R249S): Biomarker of exposure, early detection and target for therapy. *Cancer Lett.* **286**, 29–37 (2009).
121. Meissonnier, G. M. *et al.* Immunotoxicity of aflatoxin B1: Impairment of the cell-mediated response to vaccine antigen and modulation of cytokine expression. *Toxicol. Appl. Pharmacol.* **231**, 142–149 (2008).
122. Slack, J. M. W. Metaplasia and somatic cell reprogramming. *J. Pathol.* **217**, 161–168 (2009).
123. Espinoza, J. A., García, P., Bizama, C. & Roa, J. C. Molecular genetics of gallbladder cancer. In *eLS* 1–7 (John Wiley & Sons, Ltd, Chichester, 2015).
124. Colotta, F., Allavena, P., Sica, A., Garlanda, C. & Mantovani, A. Cancer-related inflammation, the seventh hallmark of cancer: Links to genetic instability. *Carcinogenesis* **30**, 1073–1081 (2009).
125. Bieging, K. T., Mello, S. S. & Attardi, L. D. Unravelling mechanisms of p53-mediated tumour suppression. *Nat. Rev. Cancer* **14**, 359–370 (2014).
126. Roa, I. *et al.* Incipient gallbladder carcinoma. Clinical and pathological study and prognosis in 196 cases. *Rev. Med. Chile* **129**, 1113–1120 (2001).
127. Roa, I., De Aretxabala, X., Araya, J. C. & Roa, J. Preneoplastic lesions in gallbladder cancer. *J. Surg. Oncol.* **93**, 615–623 (2006).
128. Mathur, S. K. *et al.* Correlation of gallstone characteristics with mucosal changes in gall bladder. *Trop. Gastroenterol.* **33**, 39–44 (2012).
129. Jones, S. *et al.* Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* **321**, 1801–1806 (2008).
130. Buckles, D. C., Lindor, K. D., LaRusso, N. F., Petrovic, L. M. & Gores, G. J. In primary sclerosing cholangitis, gallbladder polyps are frequently malignant. *Am. J. Gastroenterol.* **97**, 1138–1142 (2002).
131. Said, K., Glaumann, H., Björnstedt, M. & Bergquist, A. The value of thioredoxin family proteins and proliferation markers in dysplastic and malignant gallbladders in patients with primary sclerosing cholangitis. *Dig. Dis. Sci.* **57**, 1163–1170 (2012).
132. Chhabra, D. *et al.* Chronic heavy metal exposure and gallbladder cancer risk in India, a comparative study with Japan. *Asian Pac. J. Cancer Prev.* **13**, 187–190 (2012).
133. Jemal, A. *et al.* Global cancer statistics. *CA Cancer J. Clin.* **61**, 69–90 (2011).
134. Vallebuona C, Alfaro T. Vigilancia Epidemiológica del Cáncer, Taller de Epidemiología. *Ministerio de Salud Gobierno de Chile* 2011.
135. DEIS (Departamento de Estadísticas e Información en Salud, Ministerio de Salud de Chile) (2014). http://www.deis.cl/. Accessed 18 Dec 2014.
136. Kumar, A. *et al.* Assessment of disease burden in the arsenic exposed population of Chapar village of Samastipur district, Bihar, India, and related mitigation initiative. *Environ. Sci. Pollut. Res.* https://doi.org/10.1007/s11356-021-18207-6 (2022).
137. Heise, K., Bertran, E., Andia, M. E. & Ferreccio, C. Incidence and survival of stomach cancer in a high-risk population of Chile. *World J. Gastroenterol.* **15**, 1854 (2009).
138. Bertran, E., Heise, K., Andia, M. E. & Ferreccio, C. Gallbladder cancer: Incidence and survival in a high-risk area of Chile. *Int. J. Cancer* **127**, 2446–2454 (2010).
139. Bundschuh, J. *et al.* One century of arsenic exposure in Latin America: A review of history and occurrence from 14 countries. *Sci. Total Environ.* **429**, 2–35 (2012).
140. Jiménez Paneque, R. & Pavés Carvajal, J. R. Occupational hazards and diseases among workers in emergency services: A literature review with special emphasis on Chile. *Medwave.* **15**, e6239 (2015).
141. Fraser, B. Cancer cluster in Chile linked to arsenic contamination. *Lancet* **379**, 603 (2012).
142. Barahona Ponce, C. *et al.* Arsenic and gallbladder cancer risk: Mendelian randomization analysis of European prospective data. *Int. J. Cancer* **146**, 2648–2650 (2020).
143. Mhatre, S. *et al.* The role of gallstones in gallbladder cancer in India: A Mendelian randomization study. *Cancer Epidemiol. Biomark. Prev.* **30**, 396–403 (2021).
144. Kumar, A., Ravi, C., Dhingra, S., Krishna Murti, M. A. & Ghosh, A. K. Arsenic causing gallbladder cancer disease near the Himalayan bound rivers in Bihar: A case study of gallbladder cancer. *J. Cancer Sci. Clin. Ther.* **6**, 388–391 (2022).

Author contributions
A.K and A.K.G conceptualized the entire work. A.K. is the principal author and had the major contributions in writing the manuscript but support was also provided by A.K.G, A.B, D.K, P.J, G.P, V.R, Arti.K, M.R, S.N, A.S, Maiko.S, S.K and M.S. Literature search was done by V.R, M.R, M.S.D. Figures were developed by A.K, Arti.K, M.S.D, A.B, S.K and Maiko.S. The study design was carried out by A.K, A.K.G, and A.B. Human samples for the study were provided by Surgical oncologists-P.S, G.P, A.V, N.K. The samples were collected by P.K.N, M.K. The experimentation and data analysis were carried out by P.K.N, M.K, V.R, M.S.D, M.R and R.K. The ArcGIS maps were prepared by S.K and A.B. The statistics and data interpretation were done by A.K, M.A, R.K, D.K, N.K., A.B, Maiko.S, A.S, Amit.S, S.K and S.N. The final manuscript writing was done by A.K., A.K.G, D.K, A.B, Maiko.S, Arti.K, P.J, A.S, S.N and M.S. All authors read and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-30898-0.

Correspondence and requests for materials should be addressed to A.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.