SUPPLEMENTARY INFORMATION

Unveiling microbial risks in Chinese household dust: a comprehensive analysis from absolute abundance to virulence unit

Lu Long[1,2], Yan An[3], Li-Ting Zhu[1], Xue-Li Xu[1,2], Jing-Jing Lin[1], Wen-Juan Xu[1,2], Jing-Yu Chen[1], Fang-Yu Liu[1,2], Xu-Yi Liu[1], Qiansheng Huang[1]*

[1] Xiamen Key Laboratory of Indoor Air and Health, State Key Laboratory of Regional and Urban Ecology, State Key Laboratory of Advanced Environmental Technology, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, 361021, China.

[2] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, 100049, China.

[3] College of Life Sciences, Institute of Life Sciences and Green Development, Hebei University, Baoding, 071002, China.

*Corresponding author: Tel/fax: +86-592-6190542

E-mail address: qshuang@iue.ac.cn

The authors declare no actual or potential competing interests.

This file includes:

Number of pages 16

Supplementary text

Number of figures 11

Other Supplementary Materials for this manuscript include the following:

Number of tables 9 as separate Excel files

Supplementary text

Indoor microbiological DNA copy-based exposure

The average daily dose (ADD) of dust DNA copies through inhalation (ADD_{inh}) and dermis (ADD_{der}) were calculated, as follows,

$$ADD_{inh} = \frac{C \times IR \times EF \times ED}{BW \times AT} \quad (1)$$

$$ADD_{der} = \frac{C \times SA \times SAF \times ABS \times EF \times ED \times GI \times CF}{BW \times AT} \quad (2)$$

the variables represent the concentration of bacteria or fungi in household dust (C, DNA copies/g), the inhalation rate (IR, mg/day), the indoor exposure frequency (EF, day/year), exposure duration (ED, year), body weight (BW, kg), average lifetime (AT, day), human exposure skin surface area (SA, m^2), human skin adherence factor (SAF), the dermal absorption factor (ABS), gastrointestinal adsorption rate (GI, m^2/kg),and the conversion factor (CF). Referenced values of these parameters (Table S2) were from the Chinese exposure factors handbook(*1, 2*). Here, we adapted the classical ADD equation by replacing CFU concentrations with DNA copy numbers in dust, with the sole purpose of obtaining an order-of-magnitude index of potential exposure. We emphasize that this DNA-based ADD index is not a quantitative microbial risk assessment, because it does not account for viability, resuspension dynamics, or dose–response relationships. It should therefore be interpreted only as a comparative indicator rather than a measure of actual health risk.

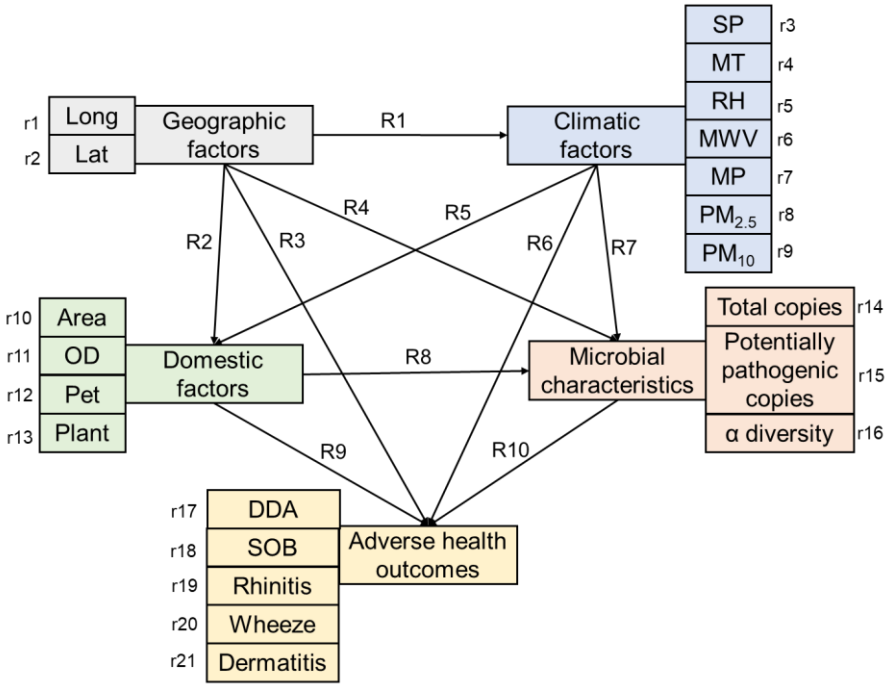

Figure S1. A conceptual model of Partial Least Squares Structural Equation Modeling (PLS-SEM). Variables included geographic factors (Latitude (Lat) and Longitude (Long)), climatic factors (Surface pressure (SP), Mean temperature (MT), Relative humidity (RH), Mean wind velocity (MWV), Mean precipitation (MP), $PM_{2.5}$ concentration ($PM_{2.5}$) and PM_{10} concentration (PM_{10})), domestic factors (Area, Occupancy density (OD), Pet and Pant), microbial characteristics (The absolute DNA copy number of the microbiota (Total copies), The absolute DNA copy number of the potential pathogens (Potentially pathogenic copies) and Microbial α diversity)), and adverse health outcomes (Doctor-diagnosed allergies (DDA), Shortness of breath (SOB), Rhinitis, Wheeze and Dermatitis).

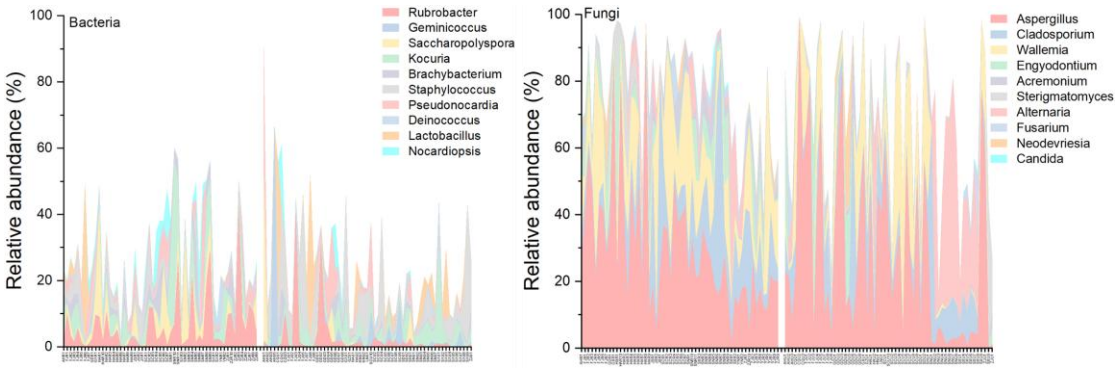

Figure S2. Composition and abundance of the top 10 genera of household dust bacteria (left) and fungi (right).

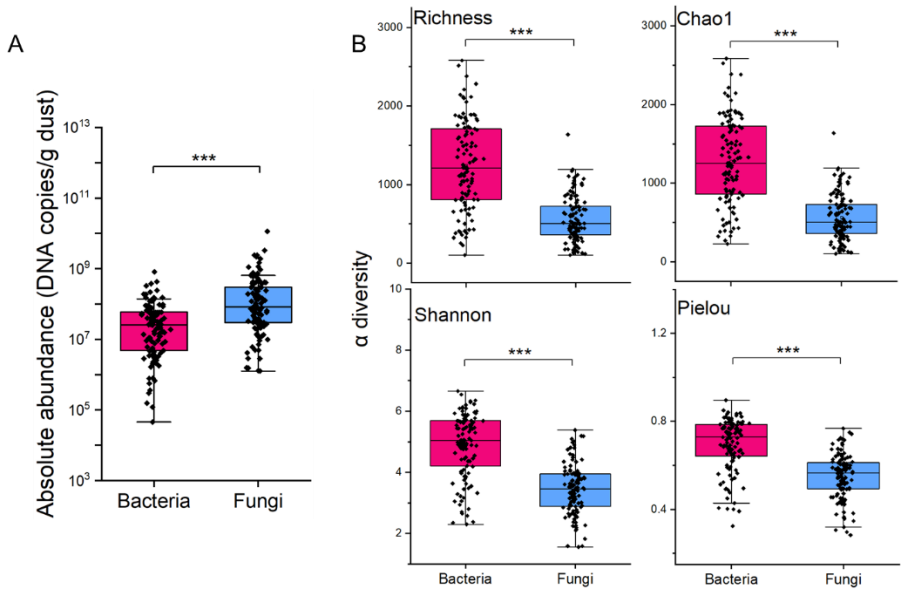

Figure S3. Differences in the diversity of bacteria and fungi in household dust. (A) Absolute abundance of the total household bacteria and fungi. (B) α diversity (Observed Richness, Chao1, Shannon and Pielou) of bacterial and fungal communities across dust samples. *** $P < 0.001$.

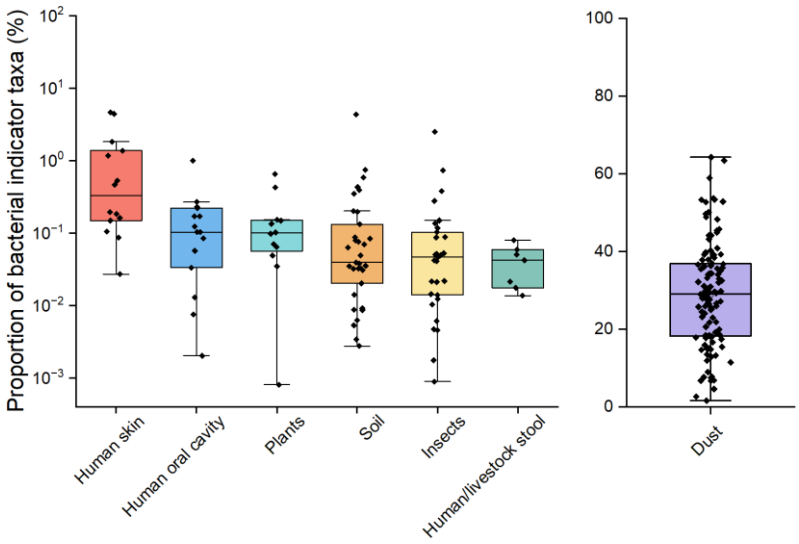

Figure S4. Average proportions of bacterial indicator taxa of different potential source environments in household dust (left). Proportions of the indicator communities from the six potential sources in household dust (right).

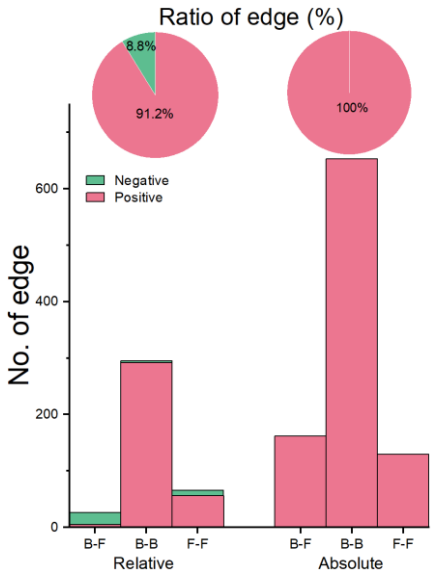

Figure S5. The number and ratio of negative and positive correlations between core microbes in relative and absolute abundance. B-F: Interaction between microbes (bacteria and fungi), B-B: Interaction between bacteria, F-F: Interaction between fungi.

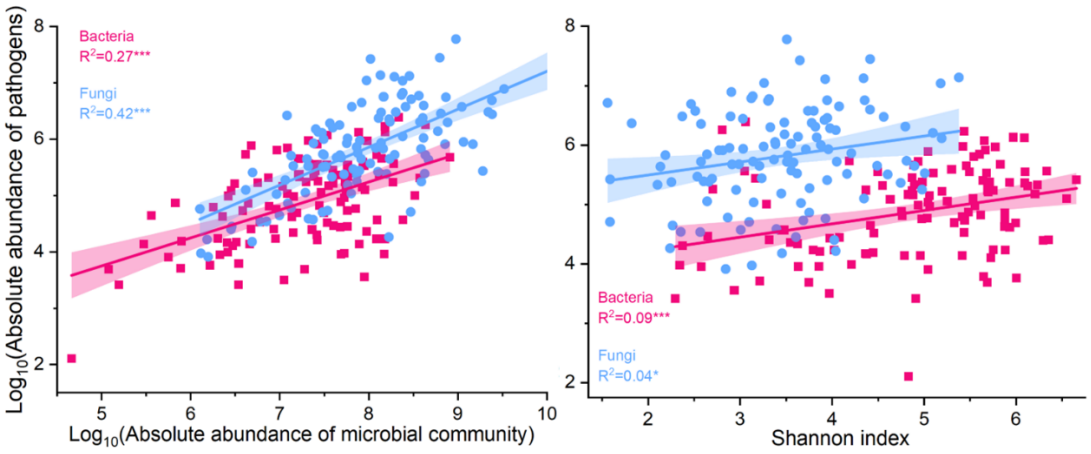

Figure S6. The absolute load of microorganisms and the Shannon index correlate with potentially pathogenic bacteria (red) and fungi (blue). * $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$.

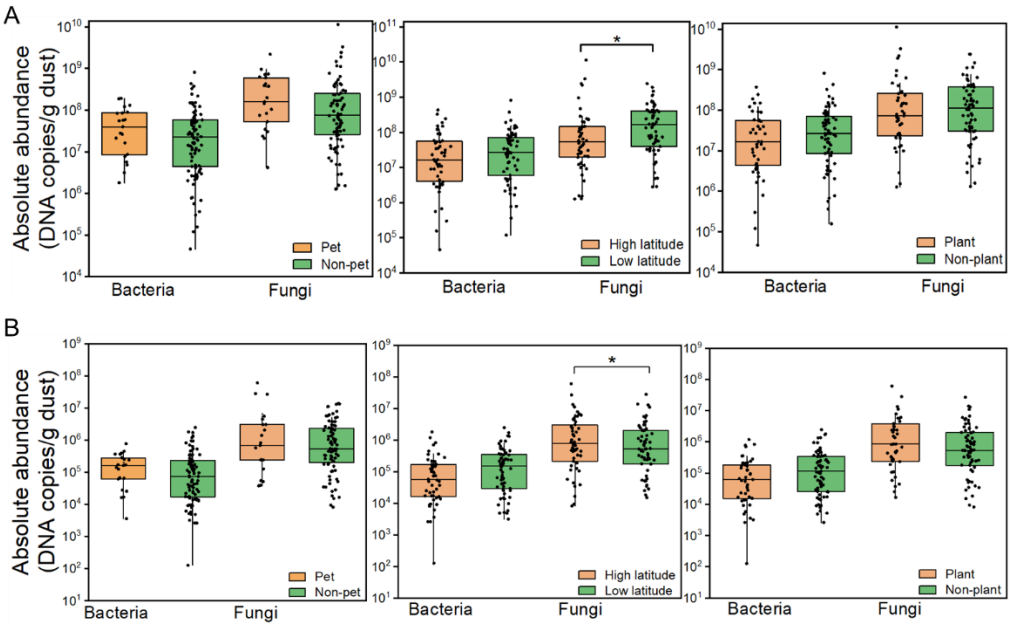

Figure S7. Compare the abundances of microbiota and potential pathogens in different groups. (A) A comparison of absolute abundance of total microbiota community in different groups. (B) A comparison of absolute abundance of potential pathogens in different groups. * $P < 0.05$.

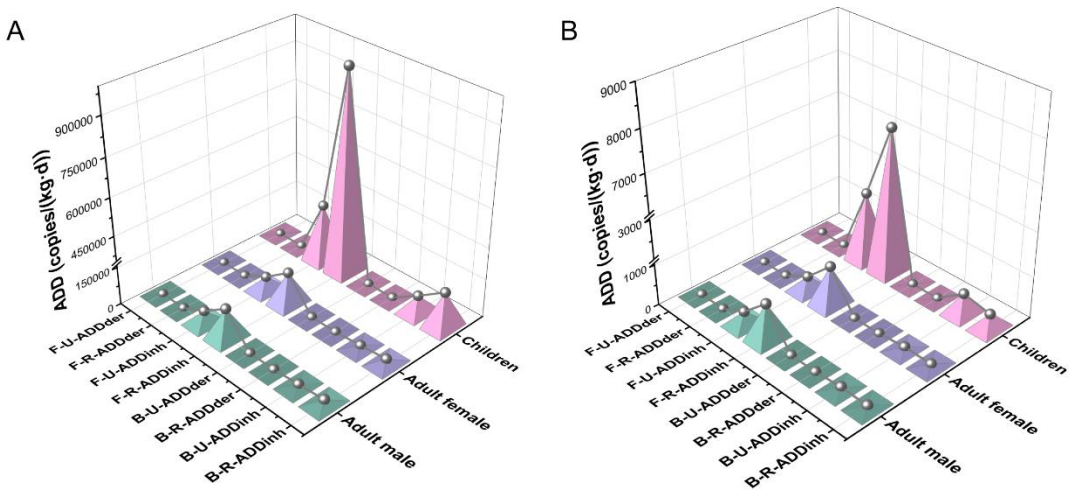

Figure S8. ADDs of total microbes (A) and potential pathogens (B) from household dust (B: bacteria; F: fungi; R: rural; U: urban).

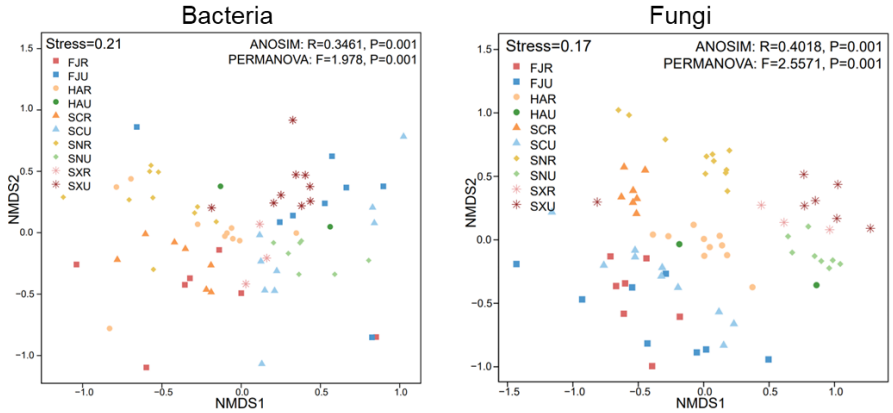

Figure S9. Non-metric multidimensional scaling (NMDS) of bacterial (left) and fungal (right) community compositions in household dust, as determined by DNA amplicon sequencing, was used to assess the source microbiomes of EVs in different areas. Dust samples were collected from rural and urban areas in five provinces including Fujian (FJ, n=15), Henan (HA, n=12), Sichuan (SC, n=16), Shaanxi (SN, n=18) and Shanxi (SX, n=12). Different colors and shapes indicate varied areas.

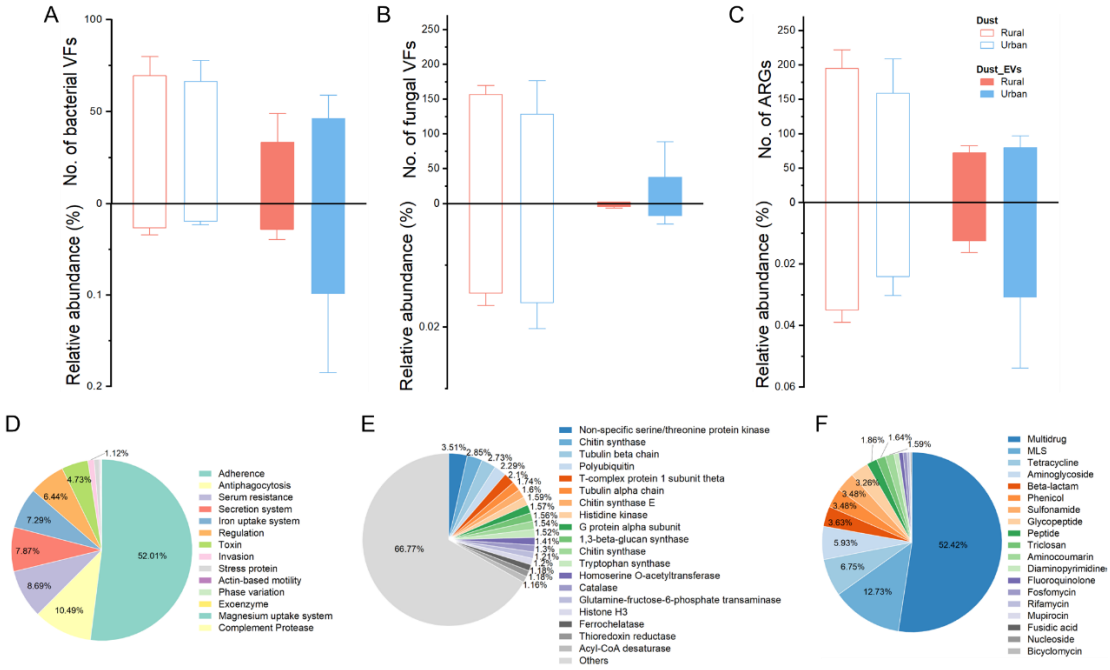

Figure S10. The VF profiles and ARG profiles of household dust and its EVs. The number and abundance of bacterial VFs (A), fungal VFs (B) and ARGs (C) in dust and EVs in rural and urban areas. (D) The distribution of the bacterial VF types in household environments. (E) The pie chart presents top 20 fungal VFs, and the VFs of which normalized abundances were below those of 20 VFs were classified as “Others”. (F) The distribution of ARG types in household environments.

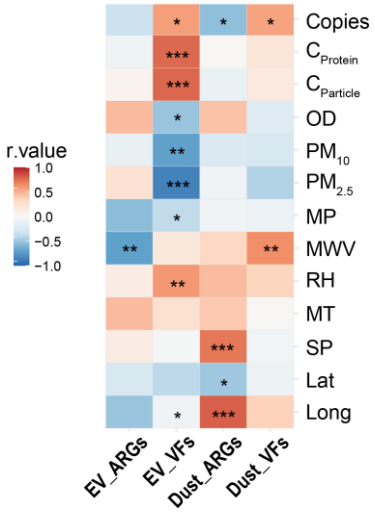

Figure S11. Heat map of Spearman's correlation between varies factors and dust- and EV-associated VFs. Variables included particle concentration ($C_{Particle}$) and protein concentration ($C_{Protein}$) of EVs, absolute abundance of dust microbiome (Copies), occupancy density (OD), $PM_{2.5}$ concentration ($PM_{2.5}$), PM_{10} concentration (PM_{10}), Mean precipitation (MP), Mean wind velocity (MWV), Relative humidity (RH), Mean temperature (MT), Surface pressure (SP), Latitude (Lat) and Longitude (Long). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Tables S1 to S9 (separate Excel file).

Table S1. Samples information of indoor household dust collected across China. The samples were divided into two groups based on latitude: the high-latitude group (latitude > 32°) and the low-latitude group (latitude < 32°). This distinction aligns with the Qinling Mountains-Huaihe River Line (approximately at latitude 32°), a significant geographical boundary in China, separating regions with differing landforms, climates, and soils.

Table S2. Parameters for risk assessment of children and adults.

Table S3. Summary of important pathogens-derived EVs proteomic datasets used in the study.

Table S4. Overview of Illumina sequencing results targeting bacterial and fungal communities in all dust samples. The samples marked by NA were removed from sequencing due to failed amplification.

Table S5. Proportions and number of bacterial indicator taxa of different potential source environments across dust samples.

Table S6. Prevalence of adverse health outcomes among occupants (N = 266) in different area households in China. *P* values were calculated by Chi-square test. Significant *P* values ($P < 0.05$) were formatted with bold font.

Table S7. Associations between the number of household microbial species and allergy confirmed by doctors (Benjamini-Hochberg FDR correction).

Table S8. Detailed information of species that were enriched in EVs. Species that were significantly enriched in EVs were selected as efficient producers (one-tailed Wilcoxon

rank-sum test with Benjamini-Hochberg FDR correction). The information includes relative abundance (mean and median) and *P* values for each species in both dust and EVs samples.

Table S9. The presence of EVs carrying VFs from 16 key pathogenic bacterial taxa in household environments.

Reference

1. D. o. E. A. (DEA), The Framework for the Management of Contaminated Land, South Africa. *Department of Environmental Affairs (DEA)* **http://sawic.environment.gov.za/documents/562.pdf**, (2010).
2. U. S. E. P. A. (USEPA), Exposure Factors Handbook: 2011 Edition. **United States Environmental Protection Agency (USEPA)**, (2011).