

Identifying Cross-Species Responses to UV Exposure using Artificial Intelligence

Proposed by *UPRI : ERIT PSII - Plant Science, Interactions and innovation*

This project represents a unique opportunity to harness AI for comprehending cross-species responses to UV exposure. It promises valuable insights with important ecological and scientific implications.

- **Background:** UV is a high-energy that affects living organisms. It can cause DNA damage, oxidative stress, and cellular responses.
- **Objective:** Our study aims to identify common genes and pathways involved in UV responses across different species.
- **Significance:** Understanding these shared mechanisms can inform biology, agriculture, and medicine and reveal UV effects on different species.
- **Overview:** Discuss the impact of UV exposure on living organisms, including:
 - DNA damage and repair mechanisms.
 - Activation of stress response pathways.
 - Cellular and physiological changes.
 - Long-term effects on health and adaptation.
 - PCD and HR and other defensive mechanisms

Introduction

UV (ultraviolet radiation) is a broad spectrum of light encompassing UV-A (315-400 nm), UV-B (280-315 nm), and UV-C (100-280 nm) wavelengths. While UVA and UVB penetrate the Earth's atmosphere, UV-A radiation (95%) and, to a lesser extent, UV-B (5%) (Correa et al., 2023), UV-C is largely filtered out, reaching us only in minimal amounts. The ozone layer efficiently filters UV-C radiation, but UV-B penetrates deeper into the Earth's surface due to ozone depletion. UVR has been shown similar impacts on different species. These UV impacts are DNA damage and repair mechanisms, protein unintegration and destabilization, cellular and physiological functions and cause long-term effects on health and adaptation in all living organisms. While valuable for disinfection by its germicidal properties, it can also cause detrimental effects on living organisms, including plants, animals and humans. In human and animals, it causes health disorders, immunity disruption, cancer, skin lesions, optic tumors, caloric stress or even death (Londero et al., 2019; Neale et al., 2023). In plants similar hazards and damages are caused by UV exposure.

UV spectra has been shown to have benefit like germicidal properties and hazards like cancer. While valuable for disinfection, it can also cause detrimental effects on living organisms, including human, plants and animals. Although research has addressed UV impacts within individual species, understanding cross-species responses remains limited. This project proposes a novel approach utilizing artificial intelligence (AI) to identify genes and pathways involved in UV response that are conserved across humans, animals, and plants.

Objectives

- Develop a comprehensive database of existing literature on UV effects on different species.
- Train and optimize AI models to identify patterns and relationships within the data, including:
 - Similarities and differences in morphological, physiological, and molecular responses to UV across species.
 - Key biological pathways activated or suppressed by UV exposure.
 - Predictive models for estimating species sensitivity based on specific characteristics.
- Validate the identified patterns and models through experimentation on selected species.

Materials and Methods

Data Collection and Preparation

- **Compilation of human, animal, and plant data related to UV radiation:** Collect data of living organisms exposed to UV radiation.
- **Data extraction and classification:** classify the UV type and the genes involved in biological processes and pathways.

Bioinformatic Tools and AI Approaches

- **Transcriptome Analysis:**
 - **Differential Expression Analysis optional, or can be replaced by data extraction):** Identify genes differentially expressed after UV radiation. This can be done by using the GEO/SAR database of NCBI.
 - **Pathway Enrichment Analysis:** Determine enriched pathways associated with UV responses.
- **Machine Learning and AI:**
 - **Deep Learning Models:** Train neural networks to predict UV-C-responsive genes.
 - **Feature Selection Algorithms:** Identify critical features for classification.
 - **Clustering Algorithms:** Group genes based on expression patterns.

Statistical Methods

- **ANOVA:** Compare gene expression levels between control and UV-exposed samples.
- **Multiple Testing Correction:** Adjust p-values for multiple comparisons.
- **Correlation Analysis:** Explore relationships between gene expression and UV exposure duration.

Data Acquisition: Collect and curate text and data from published literature, databases, and public repositories related to UV effects on various species. Utilize Natural Language Processing (NLP) techniques to extract relevant information.

AI Model Development: Employ supervised and unsupervised learning approaches, such as Deep Learning with convolutional neural networks (CNNs) and recurrent neural networks (RNNs), to analyze the gathered data. Optimize the models for identifying patterns and relationships relevant to the project objectives.

Experimental Validation: Select representative species with diverse characteristics and perform controlled UV exposure experiments. Measure morphological, physiological,

and molecular responses. Compare these findings with AI predictions and refine the models accordingly.

Data Acquisition:

- Publicly available transcriptomic and genomic datasets will be collected from humans, animals (model organisms like rodents, zebrafish), and plants (various crop species) exposed to varying UV doses.
- Datasets should be carefully curated for quality control and standardization.

Preprocessing and Feature Engineering:

- Datasets will be preprocessed through normalization, filtering, and data cleaning techniques.
- Feature engineering might involve extracting relevant features like gene expression levels, sequence motifs, or protein-protein interaction data.

AI Approaches:

- **Network Analysis:** Constructing co-expression networks and applying network topology analysis (e.g., community detection) to identify interconnected genes potentially involved in common pathways.
- **Machine Learning:** Training supervised learning models (e.g., random forest, support vector machines) to predict UV response based on gene expression data, allowing for identification of key genes and pathways.
- **Deep Learning:** Implementing deep learning algorithms like convolutional neural networks (CNNs) to analyze sequence data (DNA/RNA) and homology and identify sequence motifs indicative of UV response across species.

Statistical Methods:

- Enrichment analysis will be used to identify significantly enriched pathways (e.g., KEGG, GO) associated with UV response across species.
- Comparative genomics techniques can be employed to identify conserved genes and regulatory elements across species.

Validation:

- Comparative analysis of protein structures and interaction networks across species for further insights into conserved mechanisms.

Expected Results

- **Common Genes and Pathways:**

- **Gene Families:** Identify conserved gene families (e.g., DNA repair, stress response) across species.
- **Pathways:** Highlight shared pathways (e.g., MAPK, p53, Nrf2) across species.
- **Functional Annotations:**
 - **Perception and Transduction:** Genes involved in UV sensing and signal transduction.
 - **Functional Responses:** Genes related to cell cycle regulation, apoptosis, and repair.

Conclusion

- **Biological Implications:**
 - **Conservation:** Understanding shared UV responses aids conservation efforts.
 - **Crop Resilience:** Apply findings to enhance UV tolerance in crops.
- **Medical Applications:**
 - **Skin Cancer:** Insights into human UV responses may inform skin cancer prevention.
 - **Drug Discovery:** Target UV-responsive pathways for therapeutic interventions.

Impact and Significance:

The project's findings will contribute to understanding the ecological and evolutionary implications of UV exposure. It will:

- Inform environmental risk assessments and conservation strategies for species impacted by UV, both natural and artificial.
- Facilitate the development of UV-based technologies with minimized ecological impact.
- Advance our understanding of the general principles of stress response and adaptation across diverse organisms.