

## **SUPPLEMENTARY INFORMATION**

### **Methodology**

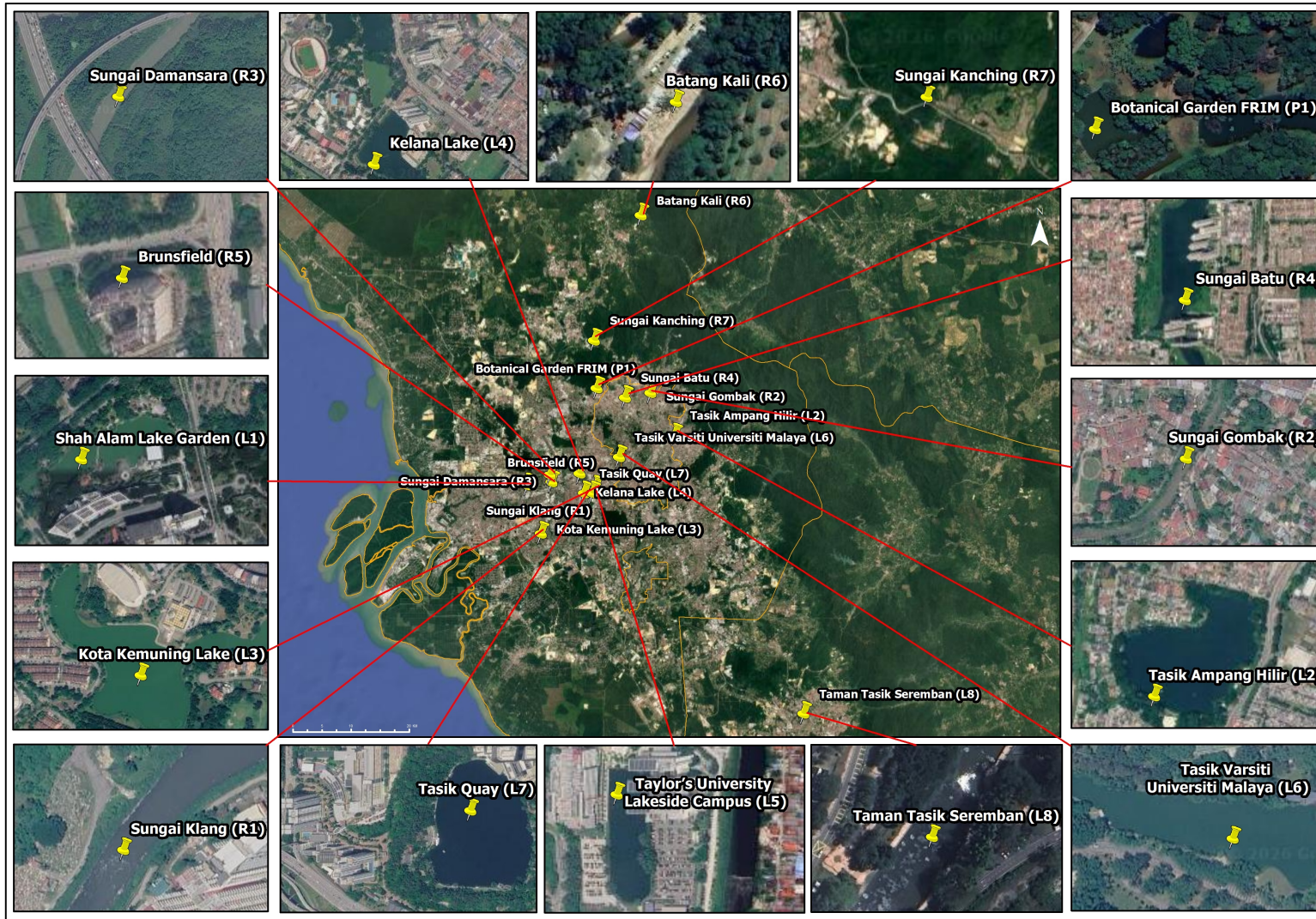
#### **Study Design and Objectives**

This study aimed to characterize diatom assemblages from freshwater bodies in Selangor, Malaysia, through systematic field sampling and laboratory extraction of frustules. The methodology encompassed site-specific collection of water samples, acid digestion of particulate matter, and microscopic identification of diatom frustules. All steps were conducted under strict contamination control and chain-of-custody protocols to ensure forensic reliability.

#### **Field Sampling**

Representative freshwater samples were collected independently from diverse aquatic habitats (ponds, lakes, rivers, creeks, and streams) within Selangor, excluding the vicinity of MSU. Each sampling site was geo-referenced using latitude and longitude coordinates obtained from Google Maps, and photographic documentation of the surrounding environment was performed to provide ecological context. For each site, triplicate water samples (approx. 500 mL each) were collected just below the water surface using sterile mineral water bottles or sealed zip-lock bags. Samples were transported to the laboratory within 5–7 days prior to analysis and stored at ambient temperature. Chain of custody was meticulously documented to preserve evidentiary integrity.

# Study Area (Mapping)



The study was conducted in selected freshwater bodies in the Klang Valley, Selangor, and Negeri Sembilan. The geographical coordinates of Selangor are 3°20'N 101°30'E, and Negeri Sembilan are 2°43'33" N, 101°56'33" E. Both states are located on the west coast of Peninsular Malaysia. Diatom research focuses on microscopic, eukaryotic plants that live in diverse aquatic environments such as rivers, lakes, oceans, and ponds. However, in this study, we focus on three sample points: ponds, lakes, and rivers. Figure 1 shows that a total of nine lakes were included: Shah Alam Lake Garden, Tasik Ampang Hilir, Kota Kemuning Lake, Kelana Lake, Taylor's University Lakeside Campus, Tasek Varsiti Universiti Malaya, Tasik Quay, and Taman Tasik Seremban. Meanwhile, River samples were collected from Sungai Klang, Sungai Gombak, Sungai Damansara, Sungai Batu, and a river near Brunsfield. A pond sample was collected from the Botanical Garden at the Forest Research Institute of Malaysia (FRIM).

The study area encompasses both **lentic (standing water)** and **lotic (flowing water)** systems. A pond sample was collected from the Botanical Garden at the Forest Research Institute of Malaysia (FRIM), an area surrounded by secondary lowland dipterocarp forest that serves as a managed research and conservation site.

A total of nine lakes were included: Shah Alam Lake Garden, Tasik Ampang Hilir, Kota Kemuning Lake, Kelana Lake, Taylor's University Lakeside Campus, Tasek Varsiti Universiti Malaya, Tasik Quay, and Taman Tasik Seremban. These lakes are primarily urban or peri-urban and serve recreational, aesthetic, and stormwater management functions. They are subject to varying degrees of human activity such as jogging, fishing, boating, and community events, making them representative of anthropogenically influenced freshwater ecosystems.

River samples were collected from Sungai Klang, Sungai Gombak, Sungai Damansara, Sungai Batu, and a river near Brunsfield, all of which traverse highly urbanized catchments. These rivers are known to receive multiple inputs from domestic, commercial, and light industrial activities, leading to elevated levels of pollutants and organic loading. To provide contrast, a river sample was also taken from Batang Kali, a glamping and recreational site located in Hulu Selangor. Unlike the heavily urbanized rivers of the Klang Valley, Batang Kali lies within a semi-rural, forest-fringe area, and therefore experiences relatively lower anthropogenic pressure.

Collectively, these sites capture a **gradient of environmental conditions**, from minimally disturbed forested and semi-rural rivers to highly urbanized lakes and rivers embedded within metropolitan areas. This range allows for meaningful comparisons of water quality, ecological status, and potential microbial diversity across different freshwater ecosystems under varying degrees of human influence in Peninsular Malaysia.

### **Acid Digestion and Extraction of Diatom Frustules**

All laboratory work was performed using sterile equipment disinfected with 70% ethanol. Water samples were homogenized in 500-mL beakers using sterile glass rods. Aliquots of 10 mL were transferred to 50 mL centrifuge tubes and centrifuged at  $14,000 \times g$  for 3 minutes. The supernatant was discarded, and the sediment was resuspended in 150  $\mu\text{L}$  of distilled water by vortexing. The suspension was transferred into a clean 25 mL beaker, and the centrifuge tube was rinsed with an additional 150  $\mu\text{L}$  of distilled water to recover residual sediment. Concentrated nitric acid (15 mL) was added to the pooled suspension. The mixture was heated to boiling using a Bunsen burner while held in a wooden clamp, until the solution became clear, indicating digestion of organic material. The digest was cooled, transferred to a centrifuge tube, diluted to the original sample volume with distilled water, and centrifuged again. This washing step was repeated twice to remove residual acid. The final sediment pellet was mounted onto glass slides, covered with coverslips, and examined under a light microscope at magnifications of  $100\times$  and  $400\times$ . Diatom frustules were identified morphologically, and representative sketches and photomicrographs were prepared.

### **Data Recording and Analysis**

Site metadata (habitat type, coordinates, environmental photographs) and laboratory observations were documented systematically. The presence, diversity, and morphological features of diatoms were recorded at each site, providing a basis for comparative ecological and forensic interpretation.