Completion Report on
Marine Ecology Enhancement Fund (MEEF) 2024-25

PART I: Project information

(i) Project Title: Developing a non-invasive method to estimate the age of Chinese White Dolphins and Finless Porpoises using the DNA methylation rate

Project Reference: MEEF2024010

Project Leader: Prof. Masayuki USHIO

Name of the Applicant Organization: The Hong Kong University of Science and Technology

Project Start Date: 01/07/2024

Project End Date: 30/06/ 2025

Project Duration (in months): 12

Reporting Period: 1 July 2024 – 30 June 2025

Brief description of the Project:

This project aimed at developing a non-invasive method to estimate the age of Chinese White Dolphins (CWDs) and Finless Porpoises (FPs) in Hong Kong waters. Age is critical information for the study and conservation of cetaceans. One conventional method of estimating the age of toothed whales is by counting dental growth layer groups, which is both highly invasive and time- and labor-consuming. This project aims to develop a non-invasive method to estimate age based on the DNA methylation rate. The project will be divided into three phases. We will (1) **develop an experimental protocol to measure the DNA methylation rate and estimate the age of CWDs using tissue-derived DNA**, (2) **apply the age estimation method to FPs using tissue-derived DNA**, and (3) **test the possibility of estimating the age of CWDs and FPs based on DNA methylation using environmental samples such as environmental DNA from the water and fecal DNA**. In the first and second phases, we will establish a protocol using the "methylation-sensitive high-resolution melting" method to achieve cost-effective quantification of the DNA methylation rate. The relationship between the DNA methylation rate and the age of CWDs will be analyzed using a machine learning approach, enabling us to estimate the age from the methylation rate. In the third phase, we will determine the technical limitations of the methylation-based age estimation method for environmental samples and apply it. This project will contribute to the development of cost-effective, non-invasive, and efficient methods to gain essential information to enable CWD and FP conservation in Hong Kong waters.

PART II: Executive Summary

In the first phase of the project, our objectives were twofold: (1) to develop a cost-effective experimental protocol for analyzing DNA methylation rates and (2) to examine the relationship between DNA methylation rates and actual ages.

For the first objective, we fine-tuned Methylation-Sensitive High-Resolution Melting (MS-HRM) analysis to improve its sensitivity. Initially, this process was not included in the schedule, but the quality of DNA derived from stranded individuals was low, making effective amplification difficult, especially after bisulfite conversion (this chemical reaction can further degrade DNA quality). To address this issue, we included a pre-amplification step for bisulfite-converted DNA using a new PCR master mix. This improved MS-HRM sensitivity, allowing us to successfully develop standard curves for our target genetic regions (GRIA2, CDKN2A, and TET2) and measure DNA methylation rates for most samples.

For the second objective, we analyzed the relationship between DNA methylation rates and actual ages, determined through dental growth layer group (GLG) analysis. We employed support vector regression (SVR) to ensure robust age predictions and found a significant positive correlation between actual ages and predicted ages based on DNA methylation rates.

In summary, we successfully established an improved version of MS-HRM analysis and measured DNA methylation rates in stranded Chinese white dolphin tissue samples. Additionally, we developed an age prediction model using these rates. Based on these achievements, we believe the objectives of the first phase were successfully met. Nonetheless, as the sample size is still small, we aim to increase the number of samples in the age prediction model to enhance its reliability for Chinese white dolphins. Simultaneously, we will establish an age prediction model for the finless porpoise in the second phase, as scheduled in the original plan.

PART III: **Completed activities against the proposed Work Schedule**

Proposed period	Proposed tasks/activities	Achievements	Status
Jul 2024	Detailed scheduling	A meeting was held on 12 July to discuss the detailed scheduling for DNA methylation measurements and tooth sample analysis	Completed
Jul – Aug 2024	Prepare standard DNA from CWD tissue-derived DNA and DNA extraction from additional tissue samples	DNA extraction from CWD tissues was completed in Oct 2024. The Standard DNAs were prepared in March 2025.	Completed
Sep – Oct 2024	Determine accurate ages for stranded CWD individuals. Prepare the standard curve for the methylation analysis	Two RPG students were recruited for the project. One student is primarily working on the collection of Finless porpoise (FP) tissues and DNA extraction (FP analysis was initially scheduled for the second year, but due to obtaining permission, we began some preliminary studies on the FP component of the project), while the other student is focusing on the CWD aspect of the project. Accurate ages were determined by our collaborators using tooth samples from stranded CWD individuals in April 2025. Note that some stranded individuals have only tooth samples. We are collecting muscle samples to determine the methylation rate and use them to improve the age prediction model. Standard curves for the three genes were prepared in April and May 2025.	Completed
Sep 2024		Collection of Finless Porpoise tissue samples from Ocean Park	Completed
Oct 2024		DNA extraction from the Finless Porpoise tissue samples	Completed
Nov – Dec 2024	Measure the DNA methylation rate for CWD	Completed in April – May 2025	Completed

Jan – Apr 2025	Analyze the relationship between the DNA methylation rate and the actual age	The relationship between DNA methylation rates and actual ages was analyzed using Support Vector Regression in May–June 2025.	Completed
May – Jun 2025	Wrap-up 2024/2025 and submit a completion report for the first phase	Completed in May–June 2025	Completed

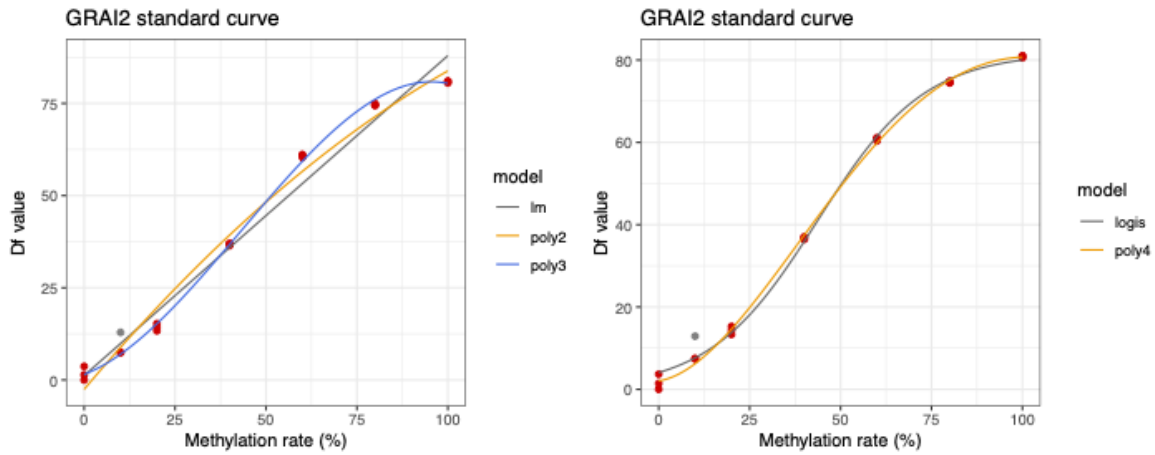
PART IV: **Results/descriptions on the completed activities with appropriate analysis**

Key outcomes:

- We modified the original protocol to enhance the analysis sensitivity. Low-concentration samples were pre-amplified using inhibitor-tolerant DNA polymerase, increasing the overall sensitivity of the MS-HRM analysis.
- Using the modified protocol, we successfully generated standard curves of DNA methylation rates for three genetic regions (**Figure 1**).
- The DNA methylation rates were determined for three genetic regions, GRIA2, TET2, and CDKN2A (**Figure 2**).
- The actual ages of some of the stranded CWD individuals were determined using dental growth layer group (GLG) analysis.
- A machine learning method, Support Vector Regression (SVR), was employed to analyze the relationship between the DNA methylation rates and actual ages (**Figure 3**).

<u>Standard curves for MS-HRM (Methylation-Sensitive High Resolution Melting) analysis</u>

The following figures are standard curves for the three genetic regions (GRIA2, CDKN2A, and TET2). For all the standards, we tested different regression methods, namely, linear regression, 2nd order polynomial regression, 3rd order polynomial regression (left figure), 4th order polynomial regression, and logistic regression (right figure) (**Figure 1**). We chose the best regression method to estimate the methylation rate of our samples.



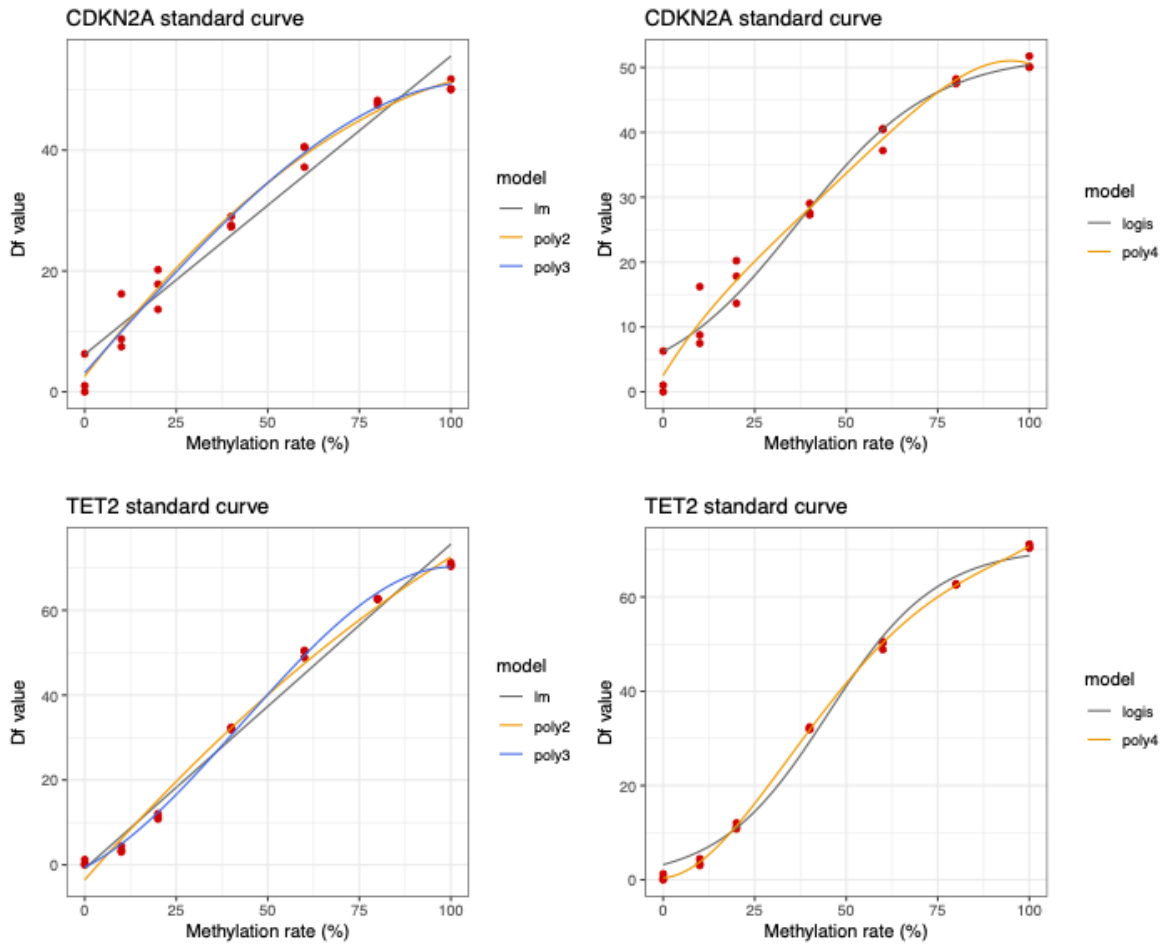


Figure 1. Standard curves for GRIA2, CDKN2A, and TET2. Points represent the Df values (the difference in normalized fluorescence from 0% methylated DNA) for each standard DNA. Lines in different colors indicate the various regression methods.

In general, the 4th-order polynomial regression best explains the standard curve, so we used it to calculate the methylation rate for CWD samples.

The DNA methylation rates for GRIA2, CDKN2A, and TET2

We analyzed 23 tissue samples from stranded CWD individuals. Preliminary analysis revealed that some tissue types (e.g., blubber and skin) exhibited variations in DNA methylation rates. Therefore, we removed the blubber and skin samples and constructed the age prediction model using the remaining samples. **Figure 2** shows the relationship between DNA methylation rates and actual ages determined by dental GLG analysis. The sample size remains small, as dental GLG analysis has not been completed for some stranded CWD individuals.

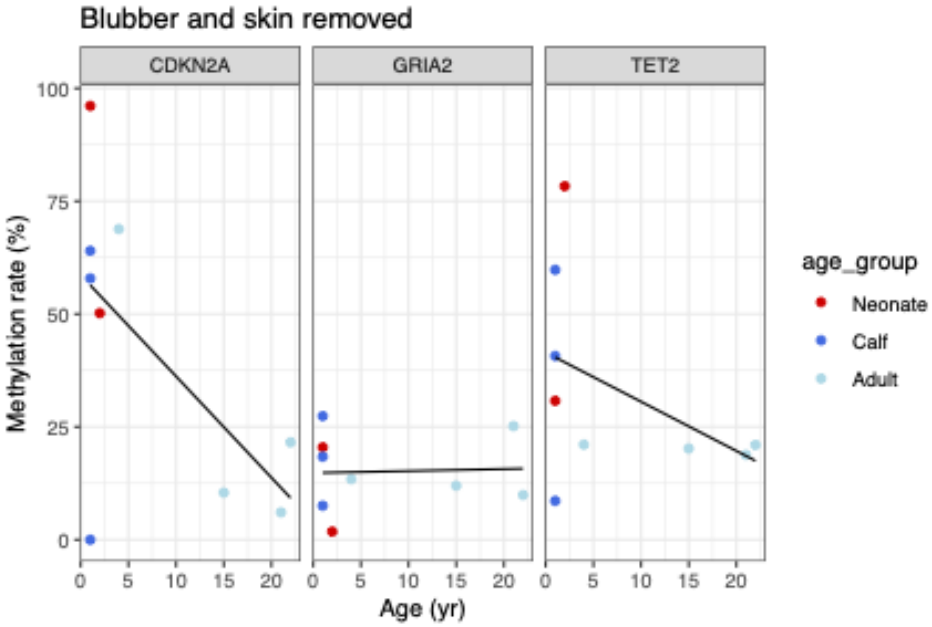


Figure 2. The relationship between the DNA methylation rates and ages. Different colors represent different age groups (Neonate, Calf, and Adult). Solid lines indicate linear regression.

The age prediction model using GRIA2, CDKN2A, and TET2 and Support Vector Regression

Although the sample size remains small, we successfully constructed the age prediction model using the three genetic regions and Support Vector Regression (SVR). As shown in **Figure 3**, we found a clear positive correlation between chronological age (actual age, *x*-axis) and predicted age based on DNA methylation rate (*y*-axis). The correlation coefficient between predicted and actual ages was 0.957.

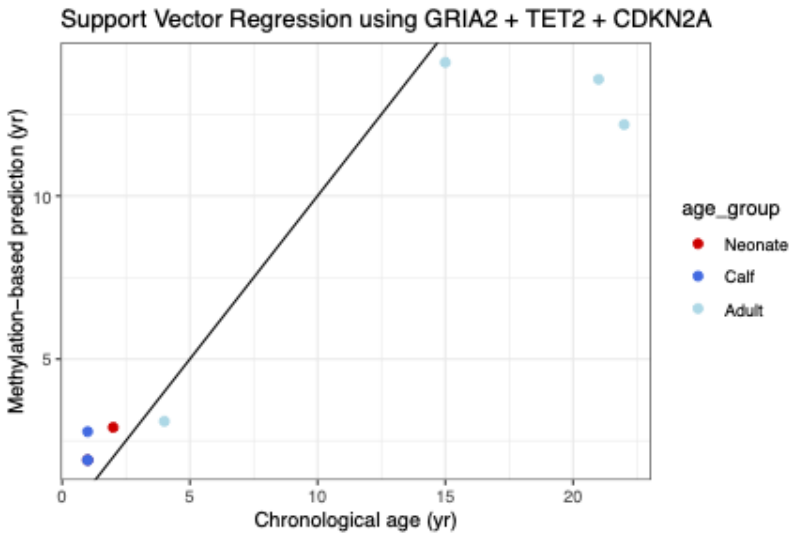


Figure 3. The relationship between the actual ages (*x*-axis) and that predicted based on the DNA methylation rate (*y*-axis). Different colors represent different age groups (Neonate, Calf, and Adult). Solid lines indicate the 1:1 line (i.e., perfect prediction).

Promoting activities including scientific publications, conferences, and press release

We introduced our DNA methylation project in an external seminar held on 23 June 2025 at the University of Malaysia, Sabah (UMS), Malaysia. The presentation title is "*Detecting vertebrates in aquatic and terrestrial ecosystems using environmental DNA*," and there were approximately 20 audience there. Due to a delay in establishing the protocol until March/April 2025, we could not publish a scientific paper or press release in the first phase of the project. However, we wrote a review paper summarizing the current status and application of the DNA methylation technique from May to June 2025. The paper has been posted as a preprint (Liu C, Xia F, Hirayama IT, Yagi G, Ushio M (2025) "Epi-eDNA: From methylation signal detection to functional ecological monitoring" https://doi.org/10.32942/X2R340) and is currently under in-depth review in the prestigious journal *Communications Biology*.

PART V: Evaluation of the project effectiveness in achieving the proposed objectives as well as the impact (benefits) of the Project

In the first phase of the project, we achieved the two major objectives of the project: we successfully established an experimental protocol to measure DNA methylation rates in three genetic regions of CWD tissue samples (Achievement 1). Based on the actual ages of stranded CWD individuals, we constructed an age prediction model based on the DNA methylation rates of the three genetic regions, with a correlation coefficient exceeding 0.9 (Achievement 2). Based on these accomplishments, we believe the objectives of the first phase were successfully met.

The cost and handling time of the age estimation methods

In this project, we established a method to estimate the age of CWD based on DNA methylation patterns using the MS-HRM method. Once the estimation model is established, the cost for age estimation is approximately HK$200 per sample, including DNA extraction, bisulfite conversion, and MS-HRM analysis (note that this is a rough estimate and may depend on how we control the detailed conditions of the experiment). An experienced researcher can analyze 20–30 samples within 2–3 days, but further modifications to the protocol will improve the number of samples that can be analyzed in the same time frame. In contrast, the dental GLG analysis typically costs approximately HK$50–100 per sample, which includes manual cleaning of surface tissue, epoxy resin embedding, sectioning with a slowly rotating diamond saw, hematoxylin staining, and microscopic examination. An experienced researcher can prepare and analyze about 20 samples over 4–5 days. Importantly, it takes a long time to train an experienced researcher for the dental GLG. At present, MS-HRM is more costly but more time-efficient than dental GLG analysis.

PART VI: Summary and Way Forward

In the first phase of the project, we successfully established an experimental protocol to measure the DNA methylation rate of CWD tissue samples. We also established an age prediction model which showed a significant positive correlation between actual and predicted ages. One limitation is the small number of samples. In the second year, we repeated the same analysis for finless porpoise samples. At the same time, we continued to collect stranded CWD samples and conduct dental GLG analysis to improve the reliability of the prediction model.

PART VII: Financial statement of the project (enclose as an appendix to the completion report) in the suggested format as set out in Appendix 2

Financial statement is not disclosed due to confidentiality reasons.

PART VIII: Staff attendance record in accordance with the attendance monitoring plan (enclosed as an appendix) (see Section 5.17)

Staff attendance record is not disclosed due to confidentiality reasons.

Declaration

I hereby irrevocably declare to the MEEF Management Committee and the Steering Committee of the relevant Funds including the Top-up Fund, that all the dataset and information included in the completion report has been properly referenced, and necessary authorization has been obtained in respect of information owned by third parties.

Any opinions, findings, conclusions or recommendations expressed in this report do not necessarily reflect the views of the Marine Ecology Enhancement Fund or the Trustee.

Signature of Project Leader:

Masayuki USHIO

Department of Ocean Science, HKUST