Development of epigenetic clocks in multiple felid species

—from small to big, domestic to wild

Huiyuan Qi

1 | Introduction

Knowledge of individual age can help both *in-situ* and *ex-situ* conservation programs to design more efficient and suitable management plans for targeted wildlife species. However, for most species, it is not easy to tell age only from observation. DNA methylation is one of the epigenetic aging markers that has emerged as a promising tool that can estimate age with high accuracy using only a tiny amount of biological material, which can be collected in a minimally invasive way. It is an epigenetic process in which 5-methylcytosine is formed via the transfer of a methyl group, usually onto the C5 position of cytosine in the cytosine-guanine dinucleotide (CpG) sites in mammals.

Here, I developed DNA methylation-based epigenetic clocks for multiple felid species, from small to big species and domestic to wild species. Felids have received much conservation attention, but a few studies have focused on improving their age assessment for conservation implications. I aimed to reach epigenetic clocks that yield acceptable age estimation accuracies using only a few gene regions and lower-cost methods to increase the ease of age estimation in conservation applications. With blood and fecal samples, and DNA methylation detection methods suitable for studying either large sample sizes or small sample sizes, the models in this thesis will be useful in investigating the age of rescued individuals (*ex-situ*, only a small sample size each time, both blood and feces could be used) and the age structure of wild populations (*in-situ*, a large sample size each time, feces).

2 | Materials and Methods

The study species were seven cat species—two small and five big cat species. Small cat species were domestic cats (*Felis catus*) and Tsushima leopard cats (*Prionailurus bengalensis euptilurus*). Big cat species included jaguars (*Panthera. onca*), Amur leopards (*P. pardus orientalis*), African lions (*P. leo leo*), snow leopards (*P. uncia*), and Amur tigers (*P. tigris altaica*). Individuals with diverse health conditions were included, since it is the case in the context of conservation.

Homogeneous regions of previously reported age-correlated gene regions in related species were searched against the reference genomes of each target species. Methylation-

sensitive high-resolution melting (MS-HRM) and next-generation targeted bisulfite sequencing are used to detect the methylation rates. MS-HRM is a real-time polymerase chain reaction (RT-PCR)-based method, which is easily conductible for most genetic labs and cost-effective for small sample sizes. On the other hand, targeted bisulfite sequencing is more accurate and able to obtain data from large sample sizes cost-effectively at once.

In this thesis, I first built epigenetic clocks for domestic cats (79 blood samples)—a vital companion animal, and also a model of feline animals with access to large numbers of samples, and also had a preliminary test of the model on a few samples of endangered snow leopards (11 blood samples), based on the methylation rates of two gene regions (*ELOVL2* and *RALYL*) measured with MS-HRM.

Secondly, to further improve model accuracy and explore the possibility of the application in other felid species within an acceptable cost, I shifted the method to targeted bisulfite sequencing, with increased target gene regions (*TCF21*, *DLX5*, *PRMT8*, *ELOVL2*, and *RALYL*) and samples from all seven felid species mentioned: domestic cats (139 blood samples), Tsushima leopard cats (84 blood samples), and five *Panthera* species (96 blood samples).

Finally, to further broaden the application in a conversation context, I built epigenetic clocks on non-invasive fecal samples (27 fecal samples) of Tsushima leopard cats, one endangered population of the Amur leopard cat (*Prionailurus bengalensis euptilurus*) that only lives on Tsushima Island in Japan. Since candidate markers have been found previously, I returned to MS-HRM. The methylation rate was investigated in *DLX5*, *ELOVL2*, and *SLC12A5*. The successful rate of methylation detection is lower in fecal samples compared to blood samples; therefore, I also investigated the reasons behind this and provided possible improvement.

3 | Results

3.1 | Age estimation using methylation-sensitive high-resolution melting (MS-HRM) in both healthy felines and those with chronic kidney disease

Domestic cat age was estimated with a mean absolute error (MAE) of 3.83 years in the model based on two gene regions and MS-HRM. Health conditions influenced the accuracy of the model. Specifically, the models built on cats with chronic kidney disease (CKD) had lower accuracy than those built on healthy cats. The snow leopard-specific model (i.e., the model that resets the model settings for snow leopards) had better accuracy (MAE = 2.10 years) than that obtained using the domestic cat model directly. This implies that the markers could be utilized across species, although changing the

model settings when targeting different species could lead to better estimation accuracy. The snow leopard-specific model also successfully distinguished between sexually immature and mature individuals.

3.2 | A cost-effective blood DNA methylation-based age estimation method in domestic cats, Tsushima leopard cats (Prionailurus bengalensis euptilurus), and Panthera species, using targeted bisulfite sequencing and machine learning models

The models used 8–23 selected CpG sites from 2–5 gene regions achieved satisfactory accuracy—the MAE of the best models was 1.97, 1.35, and 1.55 years in domestic cats, Tsushima leopard cats, and *Panthera* spp., respectively. The models in domestic cats and Tsushima leopard cats were applicable to all individuals regardless of health conditions, indicating the high applicability of the models to samples collected from diverse situations. The estimated epigenetic age was consistent with the age and age stage estimated with morphological methods in the wild-born individuals who had lived in captivity for long or relatively short periods.

3.3 | The first fecal DNA methylation-based age estimation in an endangered felid species—Tsushima leopard cats (Prionailurusbengalensis euptilurus) using methylation-sensitive high-resolution melting (MS-HRM)

Methylation data could only be measured from 27 out of 55 fecal samples. Samples defecated in water had low detectability in *ELOVL2* (p = 0.011). The best clock only contained one marker (*SLC12A5*) as an explanatory variable, and the MAE was 2.54 years.

4 | Discussion

This thesis successfully built epigenetic clocks with satisfactory accuracy (MAE = 1.35–1.97 years) at a low cost (about only $3–7 per sample) for seven felid species from blood samples. I also showed the possibility of using fecal samples to estimate the age of endangered Tsushima leopard cats (MAE = 2.54 years). To further improve the accuracy of fecal-based age estimation, improved experiments and a larger sample size are needed. For future studies, one suggestion for sampling, whatever the sample type, is to sample in multiples of one year or less to obtain an average predicted age as the final predicted age since variations were observed within the predicted age of samples collected in a short period of time. This, in turn, would provide a more accurate and precise average final predicted age.

Before applying to conservation, it is necessary to validate the models with wild individuals with known ages. In the future, it is expected to develop age-estimation models that are robust to various environments and common among closely related species, by selecting new genetic regions and adjusting model parameters. Through the models, it will be possible to estimate the age of a wide variety of wild individuals from zoo captive individuals. If the age of rescued individuals from the wild can be estimated, this could contribute to *ex-situ* conservation, for example, by improving the treatment and environment to suit the age of the individuals and giving suggestions on whether adding them to candidate lists for captive breeding pairs. If age pyramids can be estimated from feces alone in the wild, where direct observation is difficult, it will also be possible to predict future population sizes, providing important information for *in-situ* conservation.