Topic: Prediction of longevity through mathematical modeling using complex heterogeneous dataset from *Drosophila melanogaster*

Submitter: Kuhn Estelle, Lemler Sarah, Michael Rera

Laboratories: MaIAGE (INRAE), Fédération de mathématiques (CentraleSupélec), Centre de Recherche Interdisciplinaire (INSERM, Université de Paris)

Ageing is generally seen as a continuous and progressive alteration of an organism. Although a broad range of ageing modes can be observed amongst living organisms, commonly observed age-related features named hallmarks of ageing have been described. Their incidence increases with age, while improving healthspan generally delays their apparition, accelerating their apparition strongly decreases healthspan. The past 30 years have seen the identification of the first genetic components of ageing, mainly thanks to model organisms such as *Drosophila melanogaster* [1,2]. Yet, little is known about the drivers of the processes involved and important questions remain. Do all the hallmarks of ageing have a common cause? Are there key genes controlling the whole ageing phenomenon? Is there a universal ageing process? And can the aging process be slowed, stopped or even reversed?

Recently, a new tool has been developed that enables *in vivo* identification of physiologically old individuals characterized by dramatically higher risk of death than the rest of a given population [3,4]. Thanks to this innovative approach, a change in paradigm occurred regarding the way we study ageing, shifting from population-based approaches to ones centered on individuals. Based on a simple and low-cost feeding assay using a non-toxic food dye that is normally not absorbed by the intestine, it identifies individuals about to die of natural causes by showing a dramatic increase in their digestive tract's permeability, turning them blue. Their extended blue coloration is at the origin of the name of this new phenotype: Smurf. It has been shown that Smurfness is an age-dependent phenotype and a strong predictor of impending death [4]. Not only is this phenotype a better predictor of mortality than chronological age, but it is equally a strong predictor of so-called "hallmarks of ageing" such as impaired insulin signalling, loss of energy stores and decreased spontaneous activity [4]. Since then, we could extend this discovery to several model organisms frequently used in ageing studies [5]. Thus, Smurfness - this peculiar intestinal permeability phenotype - is broadly conserved through evolutionarily distant organisms for its age-dependent occurrence and more importantly as a strong predictor of incoming 'natural' death. Thanks to these results, we know that a 2-phase model of ageing is molecularly and physiologically relevant. Therefore life can be described as two consecutive phases separated by the Smurf transition [1]. Besides, a rich dataset is available coming from an experiment where more than 27000 female flies of 117 DGRP (Drosophila Genetic Reference Panel) lines were followed throughout their whole life: i) genotypic data, namely SNPs, for the 117 lines; ii) Smurf transition aggregated data; iii) individual survival data.

Building upon these early new biological findings and the rich dataset including dependent, heterogeneous and high-dimensional variables, the objective is to develop a mathematical model to improve our understanding of genomic determinants of longevity and better predict ageing. The aim is to jointly model survival data and smurf transition time using survival models and mixed models to take into account heterogeneity. The resulting model will also allow the selection of relevant SNPs that are important for longevity determination based on high dimensional variable selection tools [6]. Adapted estimation procedures will be developed involving potentially stochastic algorithms [7]. The predictive performance on the built model will be validated on an additional available dataset.

Profile of the candidate:

The applicant will have a strong background in applied mathematics, in particular in modeling, probability and statistics. He/she will also have interest in working at the interface with biology through collaborations.

References

[1] Tricoire, H. & Rera, M. A New, Discontinuous 2 Phases of Aging Model: Lessons from Drosophila melanogaster. PLoS One 10, e0141920 (2015).

[2] Rera, M., Clark, R. I. & Walker, D. W. Intestinal barrier dysfunction links metabolic and inflammatory markers of aging to death in Drosophila. Proc Natl Acad Sci U A 109, 21528–33 (2012).

[3] Mackay, T. F. et al. The Drosophila melanogaster Genetic Reference Panel. Nature 482, 173–8 (2012).

[4] Ivanov, D. K. et al. Longevity GWAS Using the Drosophila Genetic Reference Panel. J Gerontol Biol Sci Med 70, 1470–8 (2015).

[5] Dambroise, E. et al. Two phases of aging separated by the Smurf transition as a public path to death. Sci. Rep. 6, (2016).

[6] Guilloux A., Lemler S., Taupin M.-L., Adaptive kernel estimation of the baseline function in the Cox model, with high-dimensional covariates, Journal of Multivariate Analysis, 148, 141-159, (2016)

[7] Kuhn, E., Matias, C., Rebafka, T. Properties of the stochastic approximation EM algorithm with mini-batch sampling. Statistics and Computing, *30*(6), 1725-1739, (2020).