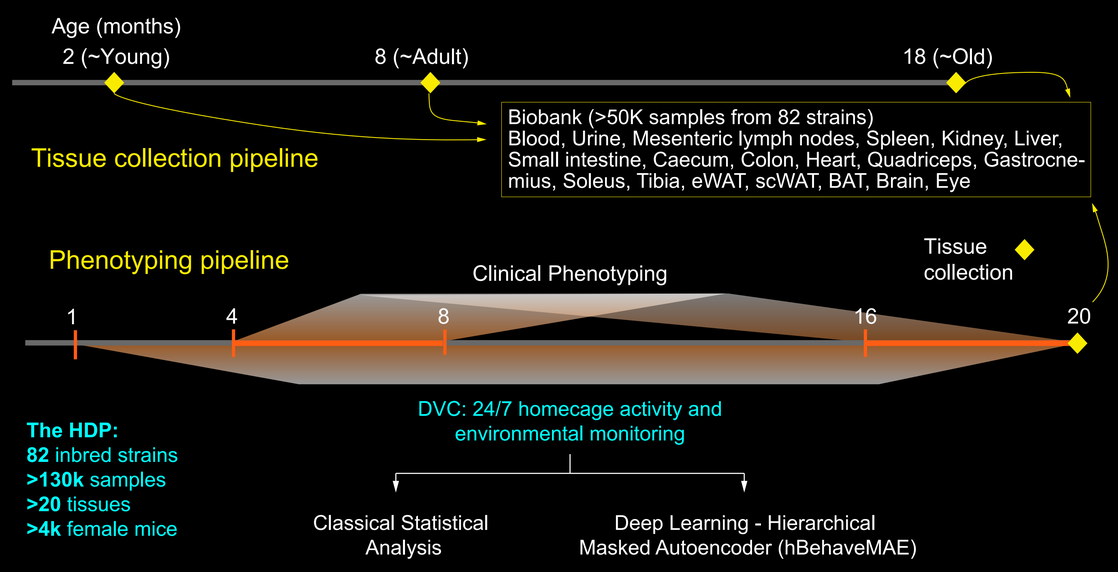
Deep phenotyping via hierarchical learning of mouse movement: decoding aging through activity

Gaurav Bhole[1], Giacomo von Alvensleben[1], Jon Lecumberri[1], Andy Bonnetto[2], Michał Grudzień[2], Alexander Mathis[2], Johan Auwerx[1]

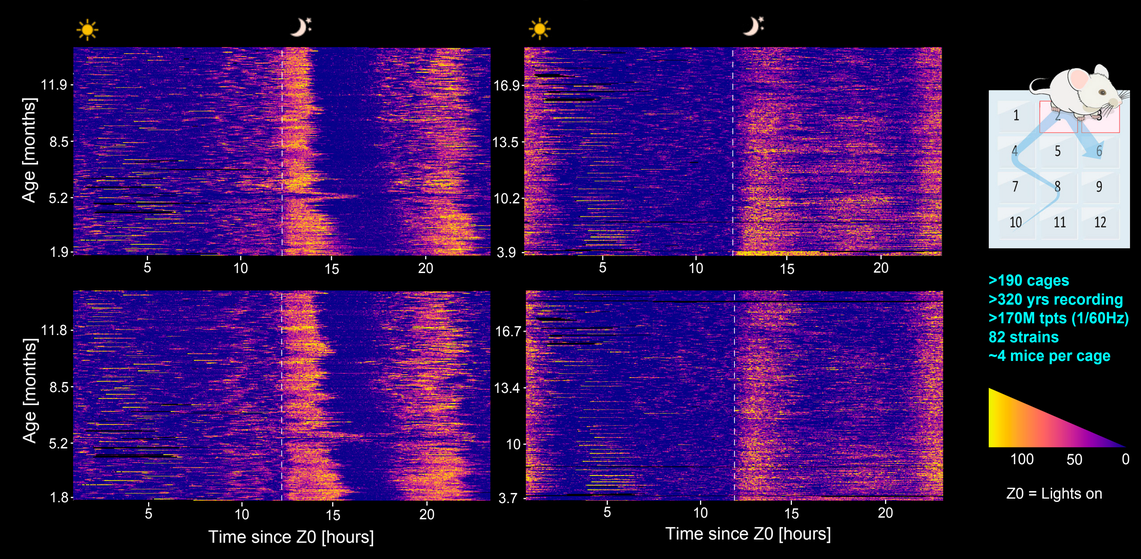
1 - Laboratory of Integrative Systems Physiology, 2 - Alexander Mathis Group of Computational Neuroscience and AI

1 The HDP: An integrated Platform for Healthspan Discovery



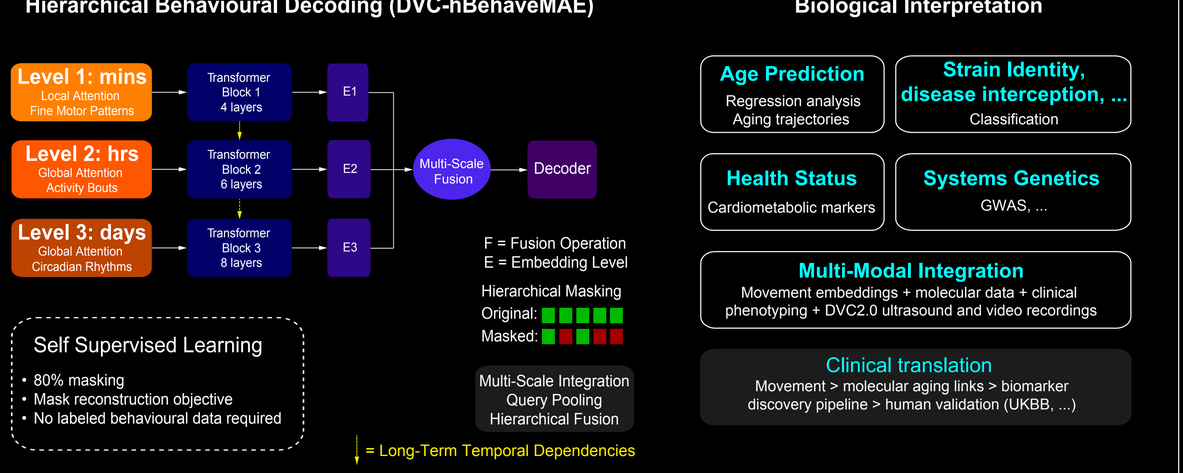
The Healthspan Diversity Panel represents an unprecedented longitudinal study of aging in laboratory mice, integrating continuous behavioural monitoring with comprehensive multi-modal phenotyping across the lifespan. Our platform combines 24/7 Digital Ventilated Cage (DVC) monitoring of over 4,000 female mice from 82 genetically diverse inbred strains, capturing lifelong activity patterns from 2 months to 18+ months of age. A large collection of neurobehavioural and cardiometabolic clinical traits with direct or similar correspondence to traits measured in human biobanks such as the UKBB, is measured at two age ranges corresponding to adult and old ages. This continuous behavioural surveillance is complemented with systematic tissue collection at key developmental timepoints, enabling deep molecular profiling of over 130,000 tissue samples including brain, skeletal muscle, and metabolic organs. The integration of these data with longitudinal behavioural and locomotion features, extracted with classical methods and a hierarchical masked autoencoder from the DVC recordings, creates a unified framework to decipher the genetic architecture of healthspan. This will allow us to identify novel biomarkers of successful aging enabling early interception and intervention.

2 The Digital Ventilated Cage (DVC) system: Patterns of activity are strain-specific



The DVC home caging system utilizes capacitive sensor technology to capture continuous mouse activity without disrupting natural behaviours. Each DVC board contains 12 capacitive electrodes placed beneath the cage, detecting electrical field perturbations from mouse movement at a native resolution of 4Hz, subsequently aggregated into 1-min intervals to get sensor relative activation time percentages. The system is coupled to a module continuously measuring environmental parameters that could affect mice behaviour, including environmental noise and human presence in proximity of the cages. This non-invasive approach applied to the HDP results in a total of over 170 million timepoints (>40 billion timepoints at native resolution), corresponding to over 320 years of cumulative recordings across the entire mice lifespan in over 190 cages. Measuring continuous group locomotion activity pattern enables population-scale behavioural phenotyping with standardized environmental conditions. The high longitudinal throughput of the DVC applied to a genetic reference population provide an unprecedented window to study natural behaviour during aging. Replication and differences in group activity patterns in independent cages within a strain and across strains respectively provide a strong indication of genetic determination of activity. Because of the very complex nature and high dimensionality of those data, we introduce here a deep-learning analysis framework complementing classical approaches.

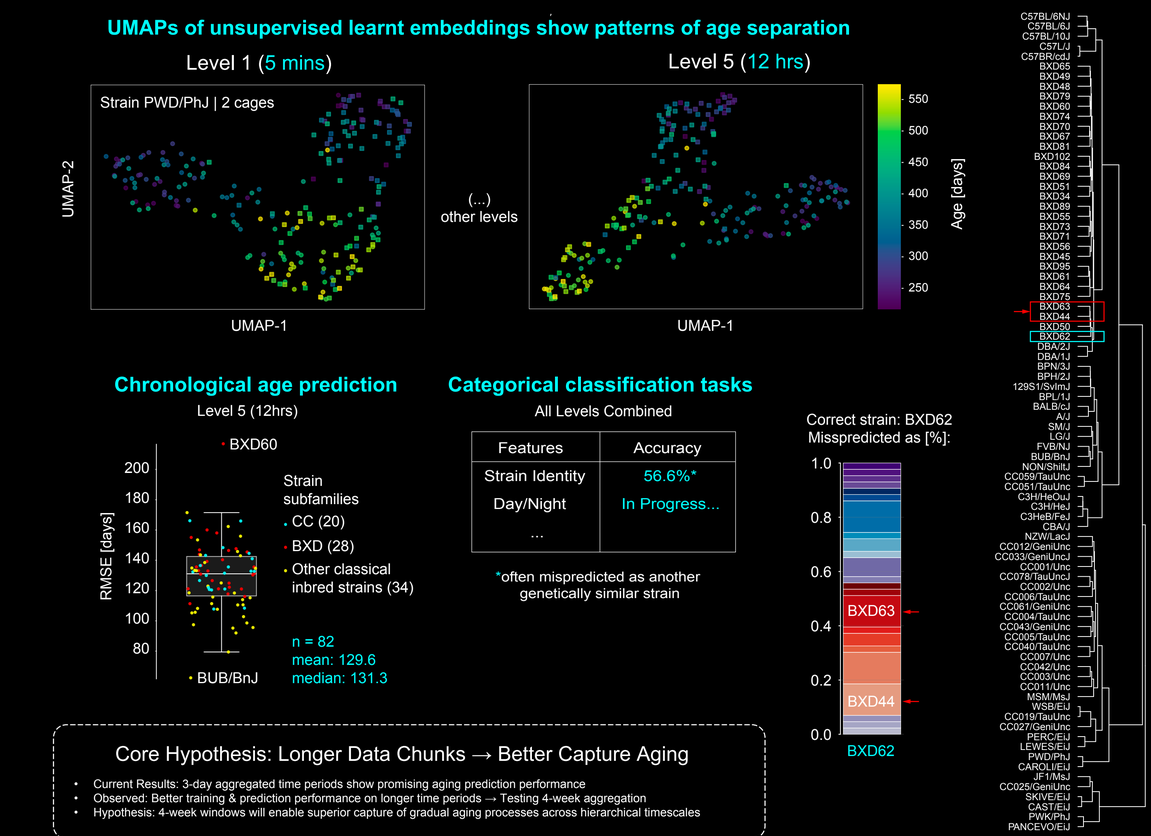
3 Discovering aging signatures through hBehaveMAE*: A hierarchical masked autoencoder



The hBehaveMAE* architecture processes continuous behavioural DVC time series data through a hierarchical masked autoencoder transformer framework (DVC-hBehaveMAE) designed to capture complex behavioural patterns across multiple temporal scales. DVC recordings from the 12 electrodes are first converted into spatio-temporal patches using learned linear projections with separable positional embeddings for time and spatial dimensions. The hierarchical encoder consists of multiple transformer blocks operating seuentially at different temporal resolutions: lower levels process minute-scale patterns using local attention to capture minute-scale behavioural dynamics, with each subsequent level building upon the previous layer's representations while employing global attention to model activity over longer time scales to represent long-term behavioural patterns. Query pooling attention mechanisms enable fusion across spatial and temporal dimensions at each hierarchical level. The self-supervised learning objective uses a block-based masking at an 80% ratio with hierarchical propagation, allowing the model to learn meaningful representations without requiring behavioural annotations. A single-layer transformer decoder reconstructs masked portions of the input sequence using L2 loss computed only on masked tokens. We hypothesized and show in box 4 that this architecture enables extraction of meaningful multiscale behavioural embeddings that allow prediction of key features such as age or strain identity. Ultimately, our goal is to integrate multiple data sources to extract interpretable latent features, enabling precise longitudinal genetic mapping of different aspects of healthspan in mice to be validated in humans with the other data layers collected in the HDP.

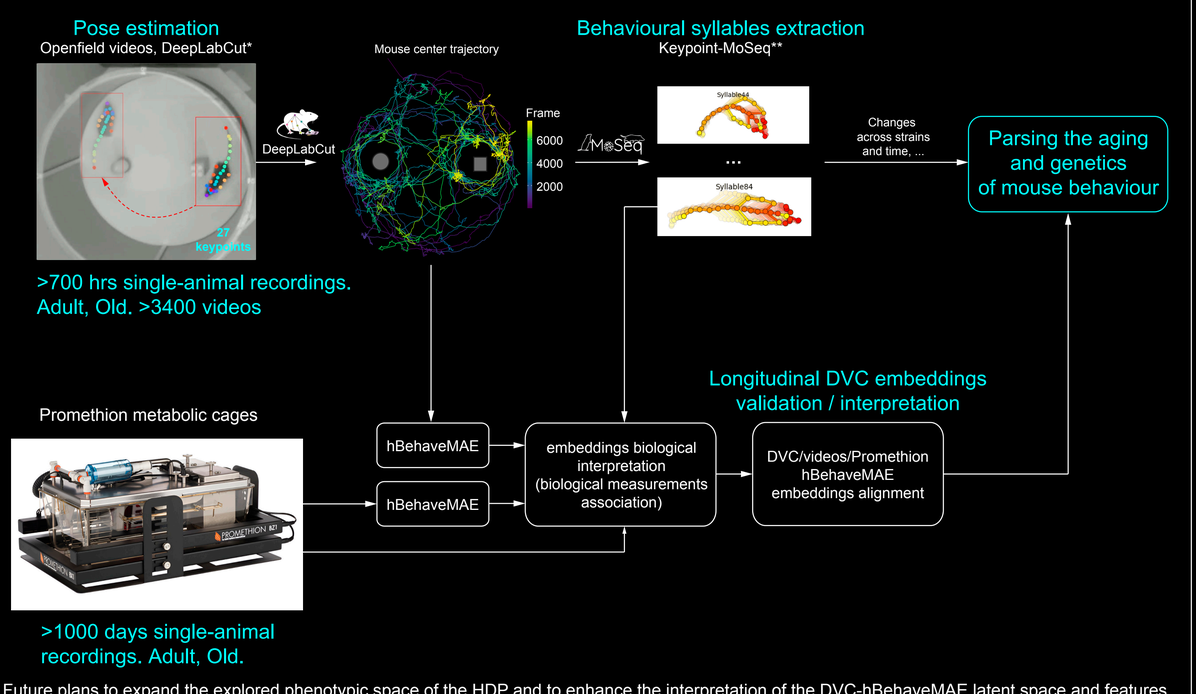
*hBehaveMAE: Stoffl, Lucas, et al. 'Elucidating the Hierarchical Nature of Behavior with Masked Autoencoders'. Computer Vision – ECCV 2024, Springer Nature Switzerland, 2025

4 HDP embeddings recapitulate phenotyping features



Our hierarchical masked autoencoder learns interpretable behavioural representations across multiple temporal scales. UMAP visualization of various levels reveal a clear age separation of input data chunks (3 days of recordings), providing a first indication that unsupervised learning from movement data might recapitulate key phenotypic features. To quantify this hypothesis, chronological age prediction achieves a mean absolute error of about 4 months (varying between 2 months and 6 months), while strain classification reaches 56.6% accuracy across 82 genetically diverse backgrounds. Inspection of misclassified strains revealed that errors often occur between genetically similar strains. This is depicted in the example above using a preliminary HDP genetic tree. To better represent biological features occurring on a long time scale, we are optimizing the DVC-hBehaveMAE framework to enable training and embedding extraction after training on much larger data chunks (>= 4 weeks vs 3 days). Preliminary inspections of test runs show superior performances in this case, although training performed over an aggregation of 3-days already seems to capture meaningful relationships through behavioural patterns and validates the biological relevance of learned representations.

5 Future Prospects



Future plans to expand the explored phenotypic space of the HDP and to enhance the interpretation of the DVC-hBehaveMAE latent space and features include leveraging over 700 hours of single-animal video recordings across adult and old mice for pose estimation using DeepLabCut*. To extract behavioural syllables and identify aging patterns, we will apply Keypoint-MoSeq**. hBehaveMAE will be used to derive latent features that will be interpreted by associative analysis with the extracted syllables. In a similar way, Promethion metabolic cages providing over 1000 days of continuous activity and metabolic monitoring data will be used to extract meaningful latent features. Combined with the use of the classical HDP phenotypes and the DVC2.0 system for DVC-hBehaveMAE embeddings interpretation (Box 3), we predict that this approach will allow for further expansion of our latents' biological validation through correlative analysis of cross-platform embeddings. This multi-modal approach will enable the study of aging and mouse behaviour through integrated longitudinal analysis of healthspan trajectories across complementary behavioural measurement systems for enhanced precision.

*DeepLabCut: Mathis, Alexander, et al. 'DeepLabCut: Markerless Pose Estimation of User-Defined Body Parts with Deep Learning'. Nature Neuroscience
**Keypoint-Moseq: Weinreb, Caleb, et al. 'Keypoint-MoSeq: Parsing Behavior by Linking Point Tracking to Pose Dynamics'. Nature Methods

6 Conclusion

- The HDP is a comprehensive multi-modal resource to model behavioural activity of mice
- The DVC data captures meaningful phenotypic and genetic features through activity recordings
- DVC-hBehaveMAE is a powerful way to hierarchically dissect complex longitudinal data

Laboratory of Integrative Systems Physiology

hBehaveMAE

Mathis Group of Computational Neuroscience & AI

Trans-bmic Population Analysis

École polytechnique fédérale de Lausanne, Switzerland

Mathis Group for Computational Neuroscience and AI