Plan Overview

A Data Management Plan created using DMPonline

Title: Developmental Organisation of Molecular Ageing Across Human Tissues

Creator: Tamsin Ederies

Affiliation: University of Birmingham

Funder: Biotechnology and Biological Sciences Research Council (BBSRC)

Template: UoB short template

Project abstract:

Human ageing is associated with molecular and physiological decline. However adult tissues do not age in the same way or at the same rate. While most studies interpret ageing as a largely tissue-specific process driven by accumulated damage, an alternative possibility is that ageing is partly constrained by developmental programmes established during embryogenesis. This project tests whether embryonic germ layer origin contributes to the organisation of transcriptomic ageing across human tissues. Uniformly processed GTEx bulk RNA-seq data from 30 tissues was downloaded from recount3. Tissues were classified as ectodermal, mesodermal, or endodermal according to embryonic origin, with mixed-lineage organs assigned on the basis of their epithelial component. Tissue-specific age effects were estimated using limma–voom, and these were then integrated within each germ layer using random-effects meta-analysis to quantify both the magnitude and cross-tissue consistency of ageing effects. Tissue-level similarity was subsequently assessed in a shared gene-effect space using distance-based analyses, and developmental relatedness was tested using a developmental ontology and Mantel analysis. Finally, weighted gene co-expression network analysis was used to determine whether ageing is organised into coordinated transcriptional programmes within each lineage.

The results support an intermediate model of ageing. Age-associated genes were largely lineage-specific, with limited overlap between germ layers, arguing against a single universal transcriptional ageing programme. However, tissues from the same germ layer were significantly more similar to one another than tissues from different lineages, although this effect was small and did not produce discrete clustering. The strongest developmental signal emerged at the level of network organisation, where each germ layer showed distinct programme architectures and biological enrichments. Mesodermal ageing was more fragmented and polycentric, whereas endodermal and ectodermal ageing was more compact and integrated. Overall, the findings suggest that developmental origin does not determine ageing trajectories, but does leave a measurable imprint on how ageing is transcriptionally organised across tissues.

ID: 204103

Start date: 30-09-2025

End date: 30-09-2028

Last modified: 12-05-2026

Copyright information:

Developmental Organisation of Molecular Ageing Across Human Tissues

Data description

What types of data will be used or created?

This project carries out secondary analysis of existing publicly available human molecular datasets and generates derived analytical outputs. The main data currently being used is bulk RNA-seq gene count matrices and associated sample metadata from the GTEx project. This is accessed through the recount3 resource. These data includes tissue identity, age in binned categories, sex, and processed gene expression counts. In later stages of the project, comparable publicly available epigenetic datasets may also be incorporated, most likely DNA methylation data and associated sample metadata.

In addition to the source datasets, the project generates several forms of derived data. These include filtered expression matrices, metadata tables, tissue-level differential expression outputs, meta-analysis summary tables, heterogeneity statistics, tissue-by-gene age-effect matrices, developmental distance matrices, WGCNA module membership tables, eigengene association tables, enrichment analysis outputs, and publication/thesis figures. The project also generates analysis scripts, intermediate R objects, and summary spreadsheets used for interpretation and visualisation.

How will the data be structured and documented?

The data is structured in a project directory with separate folders for raw input data, cleaned metadata, intermediate analysis objects, final result tables, scripts, figures, and thesis/manuscript documents. Files are named systematically by analysis stage, germ layer, tissue, and date where appropriate. Tabular outputs are stored in standard text-based formats such as CSV/TSV, scripts are stored as R files, and narrative documents are stored as Word documents or PDFs or .txt files. Readme files and script annotations are used to document file contents, variable meanings, processing steps, and the relationship between intermediate and final outputs.

Data storage and archiving

How will your data be stored and backed up?

Data is stored on secure University of Warwick systems and backed up using OneDrive. Working files, analysis scripts, result tables, and draft documents are organised in clearly structured directories. Versioned copies of key scripts, figures, and chapter drafts are retained to reduce the risk of accidental loss or overwriting. This is done using Git. No data is stored in git, only scripts are pushed.

Is any of the data of (ethically or commercially) sensitive nature? If so, how do you ensure the data are protected accordingly?

The project does not involve the collection of new participant data. It uses existing anonymised datasets that were generated and released by third-party resources under their own ethics guidelines. The main ethical consideration relates to responsible handling of human-derived molecular data and associated metadata. No attempt will be made to re-identify individuals, and the data will be used only in line with the terms and conditions of the original resource. No commercially sensitive data is involved.

Where will your data be archived in the long term?

In the long term, the data generated by the project will be archived through a combination of institutional storage and appropriate research data archiving routes. Final curated outputs needed to support the thesis and any publications (such as processed summary tables, code, and documentation) will be retained in long-term University-supported storage. Large raw source datasets obtained from public resources will generally not be re-archived by the project, since these are already maintained by the original providers.

Data sharing

Which data will you share, and under which conditions? How will you make the data available to others?

The main data that will be shared is derived data and supporting materials needed to understand and reproduce the findings. This is likely going to include analysis scripts, metadata documentation, processed summary tables, differential expression summaries, meta-analysis outputs, module-level summaries, enrichment results, and figure source data. Data sharing will be limited to materials that can be shared lawfully and responsibly under the terms of the original source datasets. Raw GTEx or other third-party source data will not be redistributed and users will instead be directed to the original repositories.