Title: Human Ageing Genomic Resources: updates on key databases in ageing research

Authors: João Pedro de Magalhães¹*, Zoya Abidi², Gabriel Arantes dos Santos¹, Roberto A. Avelar², Diogo Barardo³, Kasit Chatsirisupachai², Peter Clark¹, Evandro A. De-Souza⁴, Emily J. Johnson⁵, Inês Lopes², Guy Novoa²,⁵, Ludovic Senez¹, Angelo Talay¹, Daniel Thornton²,³,⁵, Paul Ka Po To¹

Affiliations:
1 Genomics of Ageing and Rejuvenation Lab, Institute of Inflammation and Ageing, University of Birmingham, B15 2WB, UK
2 Integrative Genomics of Ageing Group, Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool L7 8TX, UK
3 NOVOS Labs, 100 Park Avenue, 16th Fl, New York, NY 10017, USA
4 Department of Biochemistry and Tissue Biology, Institute of Biology, University of Campinas, Campinas 13083-970, SP, Brazil
5 Computational Biology Facility, Liverpool Shared Research Facilities, Faculty of Health and Life Sciences, University of Liverpool, Liverpool L69 7ZB, UK
$ Current affiliation: Department of Macromolecular Structures, Centro Nacional de Biotecnología - CSIC, 28049 Cantoblanco, Madrid, Spain
£ Current affiliation: Sygnature Discovery Ltd, Alderley Park, UK
*Correspondence. Telephone: +44 121 3713643; E-mail: jp@senescence.info

Abstract

Ageing is a complex and multifactorial process. For two decades, the Human Ageing Genomic Resources (HAGR) have aided researchers in the study of various aspects of ageing and its manipulation. Here we present the key features and recent enhancements of these resources, focusing on its six main databases. One database, GenAge, focuses on genes related to ageing, featuring 307 genes linked to human ageing and 2205 genes associated with longevity and ageing in model organisms. AnAge focuses on ageing, longevity, and life-history across animal species, containing data on 4645 species. DrugAge includes information about 1097 longevity drugs and compounds in model organisms such as mice, rats, flies, worms, and yeast. GenDR provides a list of 214 genes associated with the life-extending benefits of dietary restriction in model organisms. CellAge contains a catalogue of 866 genes associated with cellular senescence. The LongevityMap serves as a repository for genetic variants associated with human longevity, encompassing 3144 variants pertaining to 884 genes. Additionally, HAGR provides various tools as well as gene expression signatures of ageing, dietary restriction, and replicative senescence based on meta-analyses. Our databases are integrated, regularly updated, and manually curated by experts. HAGR is freely available online (https://genomics.senescence.info/).

Keywords: biogerontology, genetics, geroscience, longevity
Introduction

Ageing is one of the most complex biological processes, yet despite extensive studies, its underlying mechanisms remain to be elucidated (1,2). Furthermore, ageing is a major risk factor for mortality and several diseases, leading researchers from various fields to study this process (3). The Human Ageing Genomic Resources (HAGR) is an intuitive and powerful collection of online tools and databases that have greatly assisted scientists in addressing this complex problem.

HAGR first became publicly available online in 2004 and has been growing dynamically since, in parallel with the major growth and development of ageing-related research (4). Our resources include six main databases related to different aspects of ageing research (Figure 1), alongside other supplementary projects and general information on ageing biology. Given the impact of genetics on ageing, HAGR places a strong emphasis on genetic and genomics.

In this paper, we provide an overview of the current version of HAGR, summarizing each of its databases and tools and highlighting updates made since its previous release (5). Our goal is to provide an up-to-date guide to one of the leading and most frequently accessed online platforms used in biogerontology. HAGR is freely available at https://genomics.senescence.info/, with no registration required.

Online Resources and Databases

GenAge: The database of ageing-related genes

GenAge (https://genomics.senescence.info/genes/) is our benchmark database focused on genes associated with the ageing process, the so-called “gerontome”. It is in turn divided into two core databases: i) GenAge – Human; and ii) GenAge – Model Organisms. Both of these have been manually curated from the scientific literature. Detailed information about GenAge is available in earlier publications (6), what follows is a brief description.

GenAge – Human – is a repository of genes associated with human ageing, or at least genes that may significantly impact the human ageing phenotype and processes. It is important to note that our focus within HAGR is on biological ageing, not just age-related diseases. Because longevity is influenced by multiple factors beyond ageing, such as accidents, we focus on genes that potentially regulate the ageing process as a whole – or that at least influence various aspects of the ageing phenotype – rather than solely those affecting longevity (refer to the LongevityMap below for longevity-related genes). Each human gene entry was thus selected following a review of the literature and considering that genes can be associated with ageing based on different types
of studies and evidence. These studies include human genetic association studies, genetic manipulations in model organisms, and in vitro studies. Consequently, genes are classified into nine categories corresponding to their level of evidence (ranging from "indirect/inconclusive evidence linking the gene product to ageing" to "evidence directly linking the gene product to ageing in humans").

Genes commonly differentially expressed during mammalian ageing are also available to researchers in GenAge. A recent meta-analysis by our lab revealed global and tissue-specific gene expression changes during human ageing, with significant overlaps with both GenAge Human and with the LongevityMap (7). More specifically, we identified 449 upregulated and 162 downregulated genes with age across all tissues. This is a substantially larger mammalian ageing signature than previously, which consisted of 56 upregulated and 17 downregulated genes (8), possibly due to the larger number of studies now available.

GenAge – Model Organisms – comprises a list of genes associated with longevity or ageing in model organisms based on genetic manipulation experiments curated from the literature. To be clear, only genes that, when genetically manipulated, have a significant impact on ageing and/or longevity are included. Genes are classified into three categories: “pro-longevity”, “anti-longevity”, and “necessary for fitness”. Genes with conflicting results or insufficient data are unannotated and categorized as “unclear”. Information from model organisms is also leveraged to infer possible genes associated with human ageing in the aforementioned human dataset. The species distribution in GenAge is presented in Table 1 and includes the major traditional biomedical models such as mice, flies, worms, and yeast.

While the number of human ageing genes has only modestly changed over the years (5), existing entries are further curated, resulting in the addition of dozens of new bibliographic references. Observations concerning many genes are also regularly updated to reflect new findings. In regard to model organisms, in addition to including almost 100 new genes since the last update, we also added another five additional species, most notably S. pombe, an essential unicellular organism in ageing research – now featuring 61 entries (9,10).

**AnAge : The database of animal ageing and longevity**

AnAge (https://genomics.senescence.info/species/) is a curated database of ageing, longevity and life-history traits. Its primary aim is to support studies involving comparative ageing biology while also being of value to fields such as ecology and conservation.
For a comprehensive description of AnAge, please refer to an earlier publication (11). In brief, this database presents entries with a multitude of data, including data on maximum lifespan, metabolism, taxonomy, and additional life-history data, alongside relevant ageing phenotypes and observations. AnAge incorporates over 1400 articles and also highlights a list of 9 species exhibiting negligible senescence.

Although primarily focused on animal biology, AnAge also includes entries about plants (n=5) and fungi (n=4). Regarding animals, nine phyla are represented, with approximately 98% of the entries in Chordata, divided into 14 classes. The most well-represented classes are Aves, Mammalia, Teleostei, and Reptilia, respectively (Table 2). AnAge is of great value for ageing research as offers information on the wide range of lifespans across taxa (Figure 2), facilitating a variety of comparative studies in biogerontology.

AnAge is one of the oldest and most frequently used HAGR resources and is arguably the benchmark animal longevity database worldwide due to its constant updating and manual curation. The current version comprises 4671 entries, encompassing 4645 species and 26 taxa.

**DrugAge: The database of anti-ageing drugs**

Among HAGR’s recent additions, DrugAge (https://genomics.senescence.info/drugs/), consists of a manually curated compilation of drugs and compounds that extend longevity in model animals (12). Some compounds are listed multiple times because they have been tested across various species and doses, enabling more comprehensive assessments of their impact on longevity, as shown previously using DrugAge (13). Given our focus on ageing, compounds from studies involving disease-prone animals or harmful conditions are not included.

Presently, DrugAge features 1097 drugs or compounds evaluated in over 3200 experiments across 37 species, supported by a total of 656 references. Figure 3 illustrates the growth trajectory of HAGR databases over the years, where we can see a major improvement over the previous versions that is more marked for DrugAge. Indeed, longevity pharmacology has been exploding, and the growth in longevity drugs has outpaced the growth of longevity genes (14).

The recent surge of interest in anti-ageing drugs within health research and the pharmaceutical/biotechnology sectors underscores the significance of a scientifically reliable resource like DrugAge (14,15). Despite its relatively recent development, our database has established itself as a leading information source in geroscience and is, after GenAge and AnAge, the most widely accessed database within HAGR (Figure 4).
GenDR: The database of dietary restriction genes

Dietary restriction (DR) is a widely studied anti-ageing intervention, yet its underlying mechanisms remains poorly understood (16-18). Recognizing the genetic component of ageing, our GenDR database (https://genomics.senescence.info/diet/) compiles genes associated with DR to aid research and advance our understanding of the genetic and molecular mechanisms of DR-induced life-extension.

Further details about GenDR are available in a previous publication (19), but briefly this database includes two datasets: i) genes inferred from genetic manipulation experiments in model organisms to regulate the life-extending benefits of DR; and ii) mammalian genes whose expression is robustly altered due to DR derived from a meta-analysis (20). In total, GenDR comprises five model organisms and entails 214 genes inferred from genetic manipulations, alongside 172 genes derived from gene expression changes. To our knowledge, it remains the only database of genetic alterations associated with DR.

CellAge: The database of cell senescence genes

Cellular senescence is triggered by various stressors like telomere attrition during replication, the dysregulation of onco- and tumour-suppressor genes, and cellular or DNA damage from sources such as hydrogen peroxide and irradiation (21,22). Senescent cells undergo proliferative arrest and secrete a mix of proinflammatory factors known as the senescence-associated secretory phenotype (SASP) (23,24). As senescent cells continually produce proinflammatory factors they may contribute to inflammageing and hinder tissue repair and renewal (25). Cellular senescence has been linked to various ageing-related diseases including cancer, Alzheimer’s disease, osteoarthritis, and diabetes (26-29).

CellAge was compiled from a systematic search of the literature, and genes were included based on specific criteria, as described (30). Briefly, the CellAge database consists of genes inferred from genetic manipulations in vitro that induce (n=370, 42.7%) or inhibit (n=475, 54.8%) replicative (n=153), stress-induced (n=185), and oncogene-induced (n=238) cellular senescence (https://genomics.senescence.info/cells/). Some genes are involved in multiple classes of senescence. There are 21 genes that have an unclear effect on cellular senescence, inducing or inhibiting this process depending on experimental context. Additionally, there are 360 genes in CellAge where the mechanism by which they influence the senescence program is unclear.

The current version of the CellAge database contains 866 genes (31), a considerable increase from the 279 genes in the first build (30). Furthermore, we previously used ‘replicative
senescence’ as the default annotation for CellAge genes when the literature did not directly specify how the gene was influencing the senescence phenotype. We have now added a fourth annotation, ‘unclear,’ alongside the ‘replicative,’ ‘stress-induced,’ and ‘oncogene-induced’ tags, in order to better represent our knowledge of how these genes contribute towards cellular senescence. Previous entries have been updated to reflect this new annotation where applicable.

Furthermore, CellAge includes a list of 1259 genes differentially expressed during replicative senescence (525 and 734 over- and underexpressed, respectively) derived from a meta-analysis of senescent cells compared to proliferating counterparts (32).

**LongevityMap: The database of genetic association studies of longevity**

While human longevity derives from a complex interplay of factors, the heritability of human longevity has been estimated to be ~25% (33). The LongevityMap (https://genomics.senescence.info/longevity/) was developed to assist in cataloguing the increasing volume of data arising from genetic-variant studies of human longevity (34).

Succinctly, all entries within the LongevityMap were curated from the literature, excluding studies in cohorts of unhealthy individuals at baseline. Details on study design are provided for each entry, including population details, sample sizes, and indications of statistical significance, alongside negative results. In total, our database encompasses over 500 entries, comprising 884 genes and 3144 variants. The list is derived from 270 individual studies and presents 275 statistically significant results.

**Additional datasets, tools, and features**

In addition to our core databases, HAGR also integrates other relevant resources on ageing. Of note, the Digital Ageing Atlas (https://ageing-map.org/) serves as a platform for age-related changes that integrates age-related molecular, physiological, psychological, and pathological data to deliver an interactive database that centralizes human ageing-related changes (35). Despite its status as an external database to HAGR, we link to this portal from HAGR to provide additional ageing-related context to genes in HAGR. Likewise, our genomics resources include genome and transcriptome sequencing of the naked-mole rat (the longest-lived rodent) and the bowhead whale (the longest-lived mammal) (36,37). Moreover, HAGR encompasses projects on the relationship between ageing, cancer, and evolution. Beyond genes related to age-related diseases and transcriptional signatures of ageing, these projects host datasets that can be explored and downloaded. In addition, two bioinformatics tools for ageing research are featured in HAGR:
Ageing Research Computational Tools (ARCT), a Perl toolkit, and an SPSS script to determine the demographic rate of aging for a given population, as described previously (38).

Furthermore, HAGR provides information and news about ageing biology. Our website includes links to social media (Twitter: https://twitter.com/AgingBiology and Facebook: https://www.facebook.com/BiologyAgingNews) with the latest news in the biology and genetics of ageing and includes an educational resource on ageing (https://www.senescence.info/). Lastly, we maintain WhosAge (http://whoswho.senescence.info/), a compilation featuring 340 individuals and 65 companies contributing to biogerontology.

Usage examples in ageing and longevity

With over 1,000 citations, HAGR has been invaluable to multiple studies in various and diverse research topics in the biology of ageing. Of note in recent years, AnAge was used in a study to identify species with remarkable longevity with evolutionary implications for lifespan (39). AnAge has also facilitated data gathering on species’ maximum lifespan and relative age comparisons (40,41), body size values in a study exploring evolutionary pathways to SARS-CoV-2 resistance (42), and age of female sexual maturity in a major research endeavour on somatic mutation rates across mammals (43). HAGR’s data contributed to accurate predictions of basal metabolic rate and organ weights (44).

To highlight other recent examples of the use of HAGR, Townes et al. used GenAge in their work to identify new potential genes associated with longevity in C. elegans and S. cerevisiae (45). Moreover, GenAge and DrugAge facilitated the annotation and curation of ageing-related genes/proteins which lead to the identification of potential drug targets (46). In collaboration with other scientists, our research group has used GenDR to aid in the application of machine learning methods to identify DR-associated features (47). A recent study used DrugAge as the primary source of information where they identified lifespan-extending compounds in diverse model organisms, providing novel insights on lifespan extension (48). A co-regulated network of senescence genes in human tissues was created using CellAge’s features (49). The CellAge database was also utilized in a study on microglial senescence, where human senescence signatures and senescence-associated genes were retrieved (50), and in another research endeavour where it assisted in identifying genes differentially expressed in aged stem cells (51). Further, the association of host genes with ageing across various eukaryotic hosts was investigated using the CellAge and GenAge databases (52). Podder et al. used, among other databases, the Digital Ageing Atlas to discover longevity genes associated with nutrient sensing.
A new variant in HLA-DQB1 gene was associated with longevity and lipid homeostasis in a Chinese population study that applied LongevityMap to variant selection (54). Finally, a Cardoso et al. used several HAGR databases to generate biomarker panels for human frailty (55). These instances underscore the diverse ways in which HAGR databases can support ageing research.

**Discussion**

Ageing is a complex process that arises from the interplay of various molecular pathways and the environment; therefore, the study of its multiple components is pivotal for understanding ageing and developing interventions. HAGR was conceived to facilitate such comprehensive analyses by offering comprehensive, consistent, and accurate data.

Compared to previous versions of HAGR, our resources have seen significant updates and growth. As illustrated in Figure 4A, since its inception in 2004, HAGR has consistently expanded in terms of users and citations, emphasizing its importance within the scientific community. Among our primary resources, AnAge and GenAge continue to attract the most visitors (Figure 4B).

In addition to HAGR, other websites and databases also provide valuable resources for studying ageing. One noteworthy example is the Ageing Atlas, an online resource that employs diverse data to explore the ageing process in a multidimensional way (56). Another database, AgeFactDB, compiles ageing-related factors, incorporating our databases in its analyses (57). Nonetheless, HAGR stands out as a leading resource in biogerontology due to its integrated features, offering comprehensive tools, datasets, and insights into ageing and longevity. We anticipate that HAGR, together with other tools, will continue to advance the study of ageing biology and contribute to our overarching goal: developing a paradigm that explains ageing and improves human health and longevity.

**Acknowledgements:** We are grateful to past and present lab members and to the numerous individuals who have contributed our databases, with special recognition to Steve Austad for advice on AnAge. HAGR is developed active user input, and are thankful to the many users who have provided feedback and suggestions that have helped us improve our resources.

**Conflict of interest:** JPM is CSO of YouthBio Therapeutics, an advisor/consultant for the Longevity Vision Fund, 199 Biotechnologies, and NOVOS, and the founder of Magellan Science Ltd, a company providing consulting services in longevity science.
**Funding:** This work was supported by a grant from the Biotechnology and Biological Sciences Research Council (BB/R014949/1).
Table 1. Species in the GenAge database

<table>
<thead>
<tr>
<th>Database</th>
<th>Species</th>
<th>Number of genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human genes</td>
<td><em>Homo sapiens</em></td>
<td>307</td>
</tr>
<tr>
<td>Model organism genes</td>
<td><em>Mus musculus</em></td>
<td>136</td>
</tr>
<tr>
<td></td>
<td><em>Caenorhabditis elegans</em></td>
<td>889</td>
</tr>
<tr>
<td></td>
<td><em>Drosophila melanogaster</em></td>
<td>202</td>
</tr>
<tr>
<td></td>
<td><em>Saccharomyces cerevisiae</em></td>
<td>911</td>
</tr>
<tr>
<td></td>
<td><em>Caenorhabditis briggsae</em></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><em>Danio rerio</em></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><em>Mesocricetus auratus</em></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><em>Podospora anserina</em></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td><em>Schizosaccharomyces pombe</em></td>
<td>61</td>
</tr>
<tr>
<td><strong>Total for model organisms</strong></td>
<td></td>
<td><strong>2205</strong></td>
</tr>
</tbody>
</table>
Table 2. Taxonomy of species in the AnAge database

<table>
<thead>
<tr>
<th>Kingdom Animalia</th>
<th>Phylum</th>
<th>Number of species</th>
<th>Average longevity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Annelida</td>
<td>3</td>
<td>283.3</td>
</tr>
<tr>
<td></td>
<td>Arthropoda</td>
<td>16</td>
<td>11.6</td>
</tr>
<tr>
<td></td>
<td>Chordata</td>
<td>4557</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Cnidaria</td>
<td>2</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Echinodermata</td>
<td>2</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>Mollusca</td>
<td>49</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Nematoda</td>
<td>2</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>Platyhelminthes</td>
<td>3</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>Porifera</td>
<td>2</td>
<td>8275</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phylum Chordata</th>
<th>Class</th>
<th>Number of species</th>
<th>Average longevity (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Amphibia</td>
<td>181</td>
<td>14.9</td>
</tr>
<tr>
<td></td>
<td>Actinopterygii</td>
<td>4</td>
<td>27.4</td>
</tr>
<tr>
<td></td>
<td>Asciidiacea</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Aves</td>
<td>1513</td>
<td>17.8</td>
</tr>
<tr>
<td></td>
<td>Cephalaspidomorpha</td>
<td>16</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td>Chondrichthytes</td>
<td>117</td>
<td>25.5</td>
</tr>
<tr>
<td></td>
<td>Chondrostei</td>
<td>14</td>
<td>71.2</td>
</tr>
<tr>
<td></td>
<td>Cladistei</td>
<td>1</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Coelacanthis</td>
<td>1</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Dipnoi</td>
<td>3</td>
<td>37.1</td>
</tr>
<tr>
<td></td>
<td>Holostei</td>
<td>4</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Mammalia</td>
<td>1349</td>
<td>19.8</td>
</tr>
<tr>
<td></td>
<td>Reptilia</td>
<td>547</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Teleostei</td>
<td>806</td>
<td>16.9</td>
</tr>
</tbody>
</table>
Figures

**Figure 1:** Overview of the main databases in the Human Ageing Genomic Resources (HAGR).
**Figure 2:** Distribution of maximum lifespan across selected orders using data from the AnAge database. Six mammalian orders, two bird, two reptile, two amphibian, and one fish order (each with over 40 species in the AnAge database) are included. Kernel density estimation was used for the distribution with default parameters. In the maximum lifespan distribution graph, the x-axis represents the maximum lifespan in years (log scale) while the y-axis indicates the number of species at each lifespan.
Figure 3: Time-series growth of the Human Ageing Genomic Resources (HAGR) databases. The figure illustrates the number of entries across different HAGR databases over the years. Gene expression signatures are not included.
Figure 4: Human Ageing Genomic Resources (HAGR) visitors, citations, and database statistics. (A) HAGR unique visitors per month and citations per year of HAGR’s papers. (B) Usage by percentage of the different HAGR databases in 2022.
References


