Universidade do Minho
Escola de Ciências

Diogo Gonçalves Barardo

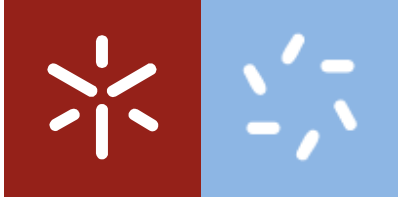
Creation of databases of ageing-related drugs and statistical analysis and applied machine learning for the prioritization of potential lifespan-extension drugs

Julho de 2016

Creation of databases of ageing-related drugs and statistical analysis and applied machine learning for the prioritization of potential lifespan-extension drugs

Diogo Gonçalves Barardo

UMinho | 2016



Universidade do Minho
Escola de Ciências

Diogo Gonçalves Barardo

Creation of databases of ageing-related drugs and statistical analysis and applied machine learning for the prioritization of potential lifespan-extension drugs

Dissertação de Mestrado
Mestrado em Biofísica e Bionanossistemas

Trabalho efetuado sob a orientação do
Doutor João Pedro de Magalhães
Professora Doutora Margarida Casal

Julho de 2016

DECLARAÇÃO

Nome: Diogo Gonçalves Barardo

Endereço electrónico: diogog.barardo@gmail.com **Telefone**: +351 913829131

Número do Bilhete de Identidade: 13615259

Título da dissertação: Creation of databases of ageing-related drugs and statistical analysis and applied machine learning for the prioritization of potential lifespan-extension drugs

Orientadores:

Doutor João Pedro de Magalhães

Professora Doutora Margarida Casal

Ano de conclusão: 2016

Mestrado em Biofísica e Bionanossistemas

Universidade do Minho, ____/____/_____

Assinatura: _____

ACKNOWLEDGEMENTS

The strict official structure of this dissertation does not contemplate a dedication section, if it did, undoubtedly, this thesis would have been dedicated to Doctor João Pedro de Magalhães.
In Pedro, I found an exigent but principled supervisor, a person truly passionate about his craft, a world-class biogerontologist, and, above all, a reliable friend. Thank you for trusting in my potential, and for giving me all the opportunities and intellectual freedom that allowed me to grow so much as a scientist. I am proud to be your supervisee and for holding the honor of being the main curator of DrugAge.
I sincerely hope that my present and future work are capable of reflecting the wisdom that you tried to bestow on me.

My parents went through enormous financial sacrifices and still supported me unconditionally. Love is truly a superior kind of force.
I am grateful to Professor Andreia Gomes for going above and beyond, in kindness and sheer effort, for her students. They say that imagination is unlimited, but I am not able to conceive a better leader to guide a new generation of scientists.
I would like to thank Professor Margarida Casal for never failing to be there when I needed.
Professor Alexey Moskalev and Professor Alex Zhavoronkov played a crucial role in shaping my scientific world view by showing me that scientific cooperation is alive and well. For this, you have my gratitude.
Even without being explicitly mentioned in the main text of this dissertation, Daniel Thornton, Doctor Danielle Newby and Professor Alex Freitas (my “unofficial co-supervisor”), all played fundamental roles in some aspects of the current thesis. Indeed, without them, the contributions of this work would be considerably less impressive. Thank you.
All the contributors and friends that made the DrugAge project a reality in such a short time frame have my full appreciation.
Lastly, but not least, I have to thank Harikrishnan Thoppil for the friendship and all the exciting scientific discussions.

RESUMO

Durante os últimos séculos, o sucesso da medicina moderna tem consistentemente aumentado a esperença média de vida da humanidade. Esta maior longevidade é acompanhado por uma mudança de paradigma: multimorbidade, causada pela acumulação de doenças relacionadas com o envelhecimento, é agora a nossa principal preocupação, ao invés das doenças fatais imediatas (por exemplo infeções) do passado. As populações envelhecidas presentemente observadas nos países desenvolvidos, já estão a ter repercussões negativas no ideal do estado social e é esperado que estas se alastrem para o resto do mundo. A solução científica para este problema assenta em desenvolver terapias anti-envelhecimento.

Nas décadas recentes, o conceito de envelhecimento como um processo biológico fixado foi desafiado e indubitavelmente refutado. Atualmente, conhecem-se mais de um milhar de genes que modificam a longevidade em organismos modelo, e simples modificações no estilo de vida como uma dieta de restrição calórica prolongam a esperança de vida em primatas não-humanos. Infelizmente, as descobertas até hoje realizadas estão ainda para ser traduzidas em terapias anti-envelhecimento com impacto em seres humanos. Neste trabalho nós oferecemos várias contribuções científicas para ajudar a mitigar a iminente crise da população envelhecida.

A nossa contribuição mais proeminente é a criação da base de dados DrugAge (http://genomics.senescence.info/drugs/). Este recurso sem paralelo congila sistematicamente informação relativa a ensaios de envelhecimento de drogas que aumentaram a longevidade em organismos modelo. DrugAge é grátis, está curada manualmente e é composta por 1316 entradas representando 418 substâncias diferentes provenientes de estudos conduzidos em 27 organismos modelo. Usámos a informação presente na DrugAge para: treinar um algoritmo para estimar o potencial anti-envelhecimento de novos compostos; realizar o enriquecimento funcional de DrugAge; comparar DrugAge com os genes anti-envelhecimento conhecidos; revelar que género não influencia a performance the compostos anti-envelhecimento em organismos modelo.

Um capítulo independente é dedicado a aplicar a reutilização de drogas para acelerar a descoberta de drogas anti-envelhecimento em humanos. Depois de fazer a correspondência entre um meta-repositório de interações droga-gene e os genes anti-envelhecimento de organismos modelo, encontrámos 16 compostos com um considerável potencial para afetar o processo de envelhecimento. Duas combinações de drogas são sugeridas para serem testadas em organismos modelo.

ABSTRACT

Over the last few centuries, the success of modern medicine has consistently increased the average life expectancy of mankind. This extended longevity came a paradigm-shift: multimorbidity is now our top concern, instead of the immediate fatal diseases (e.g. infections) of the past. The aged populations currently observed in developed countries, are already having negative recursions in the social state ideal and are expected to spread to the rest of the world. The scientific solution to this predicament lies in developing anti-aging therapies.

In the recent decades, the idea that aging is not a fixed biological process was challenged and thoroughly refuted. There are now more than a thousand different genes known to alter lifespan in model organisms, and simple lifestyle interventions like a caloric restriction diet prolong the lifespan of non-human primates. Unfortunately, the discoveries made so far are yet to be translated into meaningful human anti-aging therapies. In this work, we offer several scientific contributions to help mitigate the looming aging crisis.

Our most prominent contribution is the creation of the DrugAge database (http://genomics.senescence.info/drugs/). This unparalleled resource systematically compiles information regarding drug lifespan assays that increased the lifespan of model organisms. DrugAge is free, manually curated and is composed of 1316 entries featuring 418 different compounds from studies across 27 model organisms. We used the information provided on DrugAge to: train an algorithm for the prediction of the anti-aging potential of new compounds; conduct the functional enrichment of DrugAge; compare DrugAge with the known anti-aging genes; show that gender does not influence the performance of anti-aging compounds in model organisms.

A separate section is dedicated to applying drug repurposing to accelerate the discovery of anti-aging drugs in humans. After matching a meta-repository of drug-gene interactions with the known anti-aging genes in model organisms, we found 16 drugs with significant potential to affect the aging process. Two drug combinations are suggested to be tried in model organisms.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ABAT	4-aminobutyrate aminotransferase
AKT	AKT serine/threonine kinase
AUC	Area under the curve
CRISPR	Clustered regularly interspaced short palindromic repeats
DGIdb	Drug-Gene Interaction Database
Gmean	Geometric mean of recall and specificity
GO	Gene Ontology™
HAGR	Human Ageing Genomic Resources
HDAC	Histone deacetylase
IIS	Insulin and insulin-like growth factor
ITP	Interventions Testing Program
MOE	Molecular Operating Environment
OECD	Organization for Economic Cooperation and Development
OGDH	Oxoglutarate dehydrogenase
PDE4	Phosphodiesterase 4
R&D	Research and Development
SMILES	Simplified molecular-input line-entry system

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1. AGING

The aging phenomenon and its omnipresence in our life lead to us intuitively think that it is a "natural part of life" and not much can be done about it. While this might not be true for all animals[1]–[6], it surely seemed to be the case with humans. Therefore until recently Biogerontology did not enjoy mainstream popularity as an academic discipline. This situation changed dramatically when the severe underpinnings of an increasingly older world population started to be felt, and scientific progress showed us that the rate at which organisms age is everything but set in stone.

1.1.1. SOCIOECONOMIC PERSPECTIVE

In developed countries it is clear that humans are living longer[7] (falling mortality rates and increased longevity) while reproducing less[8] (falling fertility rate), resulting in population aging. Contrary to conventional believe the latter cause seems to be the main driver[9] of population aging. Forecasts indicate that population aging is expected to continue[10], [11] and extend globally[12].

Total fertility rate, calculated as the average number of children per woman implied by current women of all ages in a given year, is below the replacement rate of 2.1 for two-thirds of the countries (Figure 1.1).

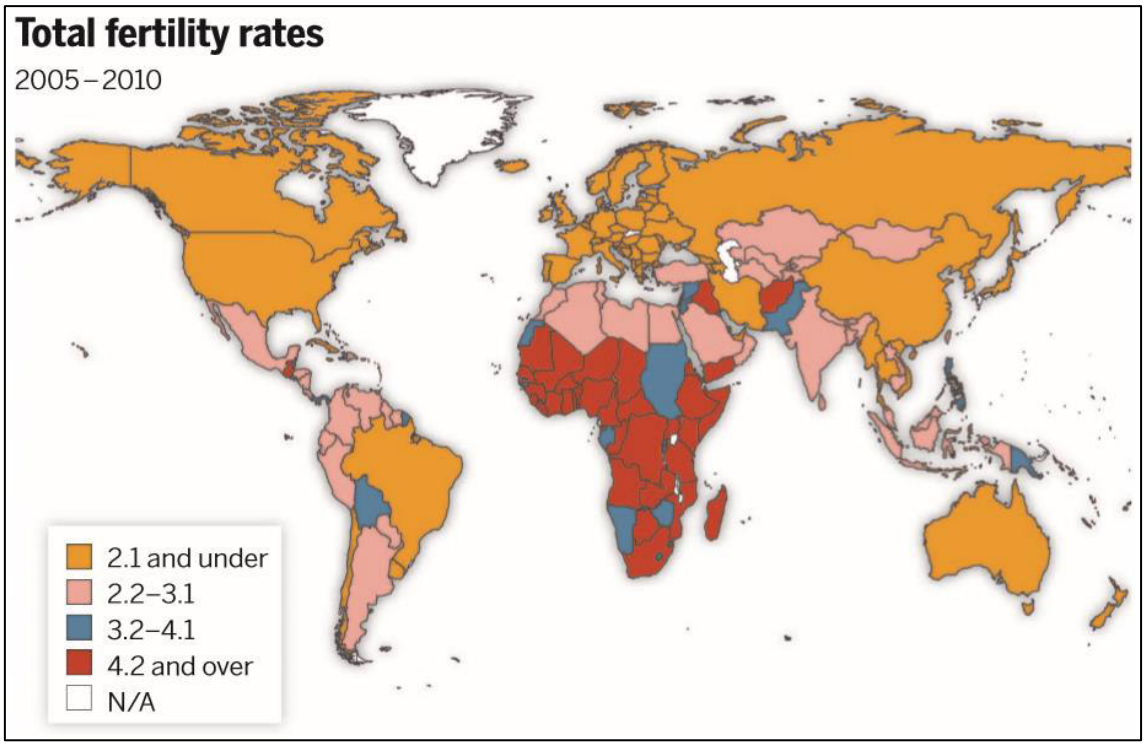


Figure 1.1: World map of total fertility rates (2005-2010). From[11].

An aged population increases the demand for public spending on pensions and healthcare[13], calling into question the very foundations of the welfare state and posing a risk of creating inter-generational conflict. Several policies and approaches have been aimed to address this "demographic deficit"[14] (Figure 1.2) such as encouraging increases in the fertility and immigration rates[15] and making the population work for longer[9].

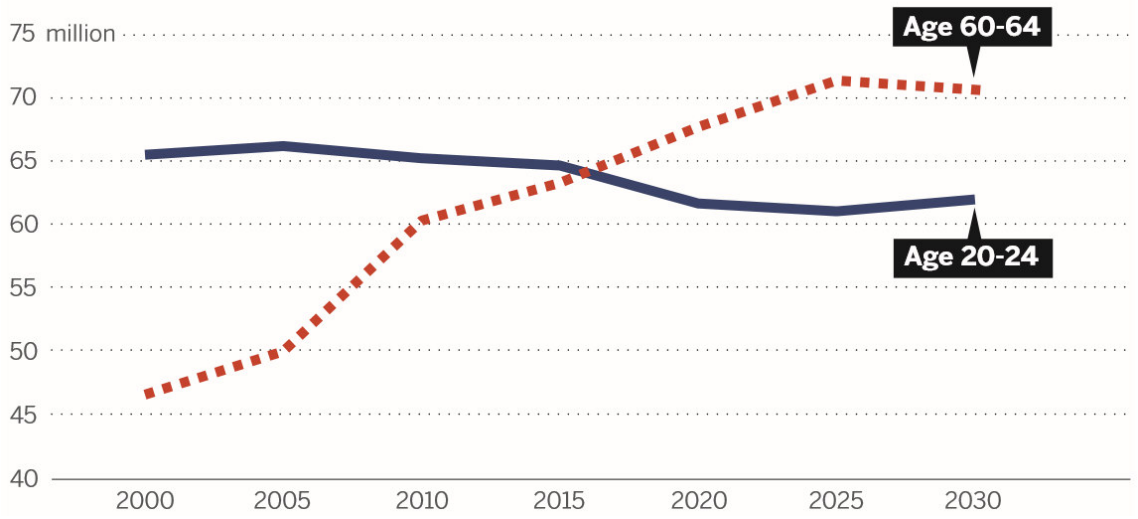


Figure 1.2: Organization for Economic Cooperation and Development (OECD) demographic deficit (2000-2030). Observed and projected size of the incoming (20–24) and outgoing (60–64) working-age cohorts in OECD countries From [11].

The rise in the total amount of people suffering from illnesses or even disability, caused by the increase in longevity[16], is so worrisome and prevalent that it has been termed “epidemic of frailty”[17]. The signature of this epidemic is the health status being compromised by complex chronic long-term diseases, instead of acute infections. The present work has implications for this so-called “chronic disease burden”[18], [19].

1.1.2. BIOLOGICAL PERSPECTIVE

As mention previously, there was a paradigm shift, from the scientific point-of-view, that propelled the study of the aging process - Gerontology – to the mainstream of scientific research[20]. This turn of event was the discovery that the aging rate is extremely malleable, and it can be exemplified by the now classic study headed by Cynthia Kenyon, that showed that a single genetic mutation doubled the average lifespan of *Caenorhabditis elegans*[21]. Since then, several genetic[22], pharmacological[23] and lifestyle[24] interventions (Figure 1.3) were discovered to influence lifespan drastically in a plethora of distinct model organisms. The volume of scientific research and interest was such that a new subfield of gerontology was born – Biogerontology - dedicated primarily to achieving healthy old age (extending healthspan) accompanied by improved longevity (increased lifespan).

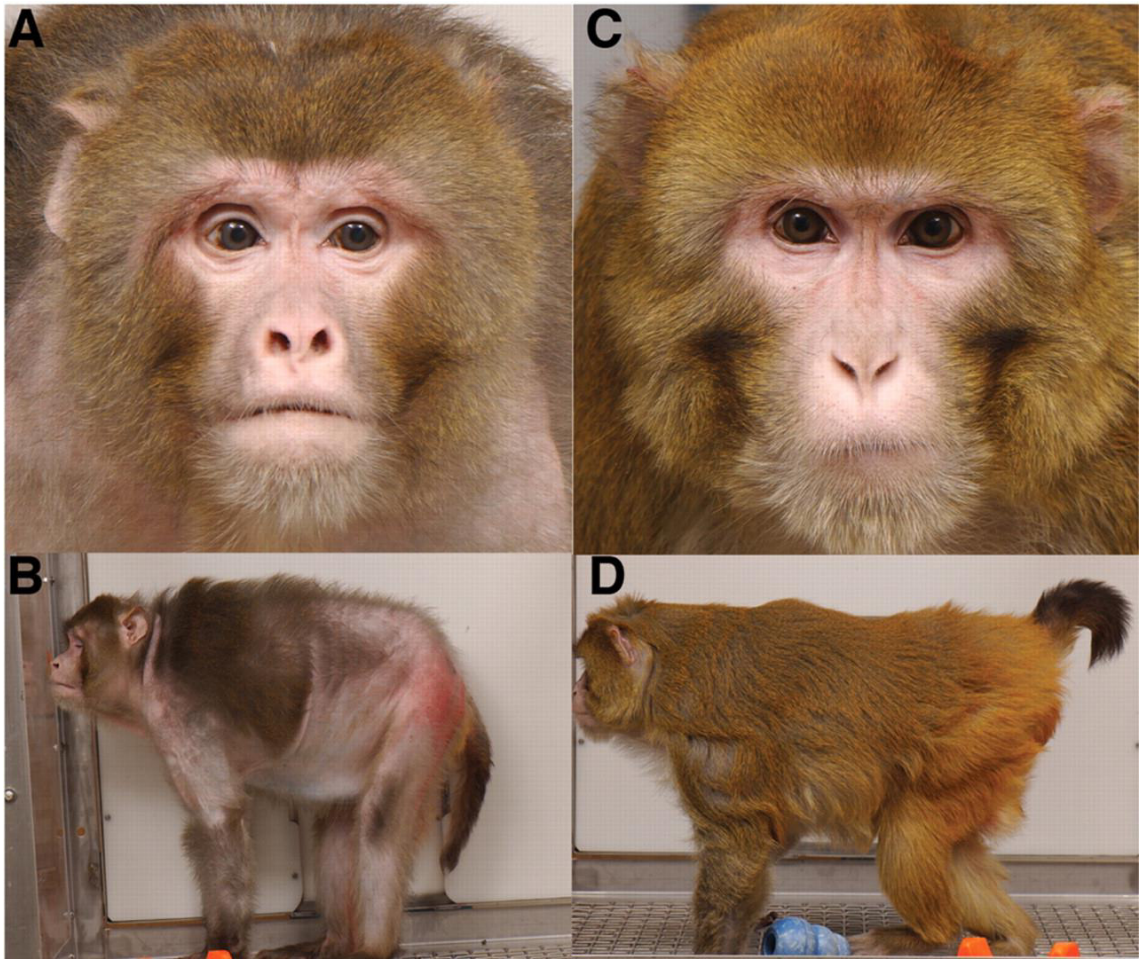


Figure 1.3: Rhesus monkey appearance in old age. **A** and **B** - Photographs of a typical control animal at 27.6 years of age (about the average life span). **C** and **D** - Photographs of an age-matched animal on caloric restriction. From [25].

Although it is easy to empirical distinguish between the phenotype of an adult and senior citizen, aging is among the most complex biological processes[26]. Broadly defined, aging is characterized by a progressive loss of physiological integrity, leading to impaired function and increased vulnerability to death. Worth rephrasing, namely because it goes unnoticed by the general public, is the fact that aging is an independent primary risk factor for major human pathologies including cancer, diabetes, cardiovascular diseases, and neurodegenerative disorders[27].

Biological theories of aging abound[28]–[35], but there is nil consensus over which one of them is superior and if even any one of them manages to encapsulate the entire aging-associated phenomena[36]. To briefly outline the biological basis of aging we decided to be agnostic in relation to biological theories of aging. Instead, we followed the categorization based on the hallmarks of aging[27] (but taking into consideration recent results not available to the authors at the time),

which helps to conceptualize its essence and underlying mechanisms in an evidence-based manner.

There are currently nine hallmarks (emphasis on mammalian aging) guiding the biogerontology field (Figure 1.4). The hallmarks are treasured from a pragmatic standpoint due to the criteria that must be fulfilled in order to be considered a "hallmark" (especially the second and third requisites): (i) a hallmark should manifest itself during normal aging; (ii) its experimental aggravation should accelerate aging; and (iii) its experimental amelioration should retard the normal aging process and, hence, increase healthspan.

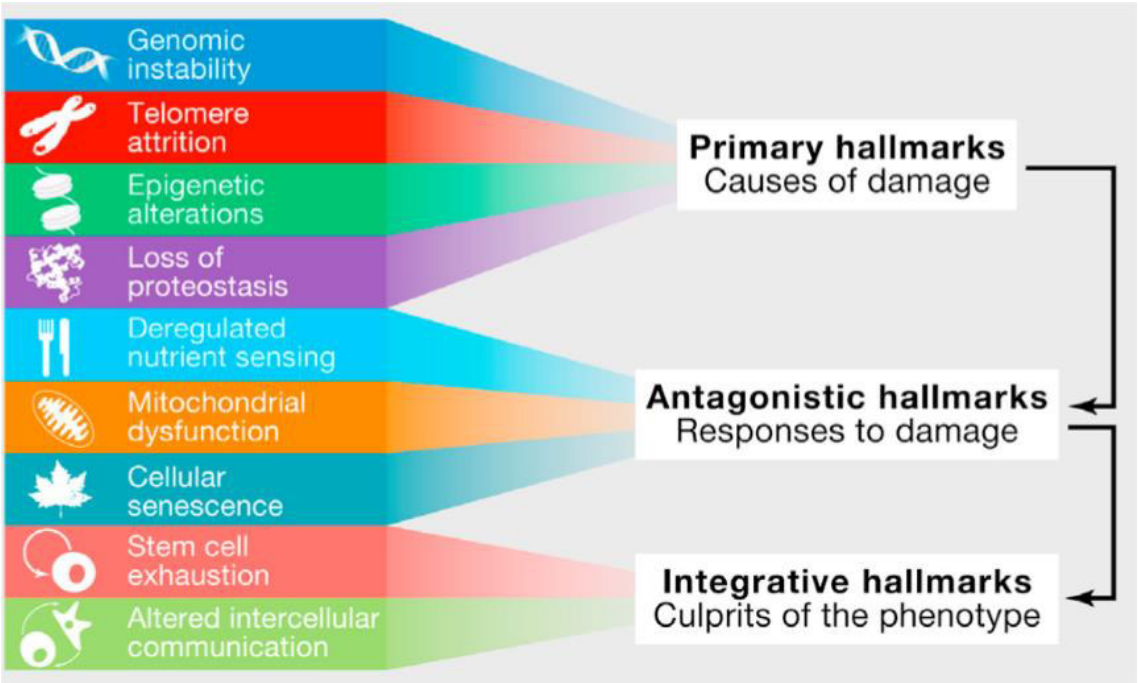


Figure 1.4: The hallmarks of aging and their functional interconnections. The proposed nine hallmarks of aging are grouped into three categories. **Top** - Those hallmarks considered to be the primary causes of cellular damage. **Middle** - Those deemed to be part of compensatory or antagonistic responses to the damage. These reactions initially mitigate the damage, but eventually, if chronic or exacerbated, they become deleterious themselves. **Bottom** - Integrative hallmarks that are the result of the previous two groups of hallmarks and are ultimately responsible for the functional decline associated with aging. From [27].

Genomic Instability

The very existence of complex networks of DNA repair mechanisms can only be explained in evolutionary terms if the integrity of DNA molecules is continuously challenged throughout life[37]. Unfortunately, the genetic lesions are minimized but not completely dealt with, resulting in the slow accumulation of genetic damage.

Premature aging diseases (also called progeroid syndromes), such as Werner syndrome and Bloom syndrome, partially resemble normal aging and are the consequence of increased DNA

damage[38]. Furthermore, experimental reinforcement or debilitation of DNA repair mechanisms delays or accelerates aging, respectively[39]. Of notice ii that in addition to nuclear DNA, mitochondrial DNA may contribute to aging, at least in accelerating it[40].

Another source of genomic damage is genomic instability caused by defects in the nuclear lamina. The nuclear lamina is a dynamical architecture that plays a role in genomic maintenance by providing a scaffold for tethering chromatin and protein complexes that regulate genomic stability. Not only does nuclear lamina suffers changes in its constitution with normal aging[41], [42], mutations in genes associated with it cause accelerated aging syndromes such as the Hutchinson-Gilford[43], [44] and the Néstor-Guillermo[45] progeria syndromes.

Telomere Attrition

Telomeres are the chromosomal regions corresponding to the terminal ends of linear DNA molecules. Most mammalian somatic cells do not express telomerase (the only DNA polymerase with the capacity to completely replicate telomeres) and telomeres exhaustion ultimately leads to their senescence and/or apoptosis[46], [47].

In humans, telomerase deficiency is associated with a higher risk of developing diseases that involve the loss of the regenerative capacity of different tissues, such as pulmonary fibrosis, dyskeratosis congenita and aplastic anemia[48].

In model organisms, the causality of telomere length in organismal aging is firmly established, genetically modified mice displaying shortened or lengthened telomeres exhibit decreased[49], [50] or increased lifespan[51], respectively.

Epigenetic Alterations

Similar to genetic instability, epigenetic instability (epigenetic alterations) is so robustly associated with aging that the most accurate biological clock is based exclusively on epigenetic marks[52].

Among histone modifications, the gene SIRT6 is the most important hallmark in mammals. Through histone H3K9 deacetylation it regulates genomic stability, NF-κB signaling and glucose homeostasis[53]–[55]. Mutant mice deficient or overexpressing *Sirt6*, display accelerated aging or longer lifespan, respectively.

Furthermore, gain- and loss-of-function studies have confirmed that transcriptional alterations of miRNAs[56] and chromatin alterations[57]–[59] causally modulate aging in invertebrates.

The attractiveness of the epigenome is that its age-related changes are not irreversible. This universe of opportunities for designing anti-aging therapies is already showing promising results[60].

Loss of Proteostasis

Even in the hypothetical condition that the genome was completely error-free, if the mechanisms providing quality-control and assurance to the proteome (proteostasis) lose their efficacy with time, aging would still occur, with the only difference being that the process would initiate downstream of the genome, at the proteome level. Indeed proteostasis is affected by aging[61] and chronic expression of aberrant proteins contributes to the development of some age-related pathologies, such as Alzheimer's disease, Parkinson's and cataracts[62].

Two types of biological processes are involved in proteostasis: chaperone-mediated protein folding and proteolytic systems; conceptually (in a simplified manner) the first works as a protein quality-check and repair point and the second as a group of aberrant protein clearance pathways.

Transgenic flies[63], worms[64] and mice[65] that overexpress chaperones live longer, whereas mutant mice deficient in a co-chaperone belonging to the heat-shock family exhibit accelerated aging phenotypes[66].

Regarding proteolytic systems: chemical inducers of autophagy (autophagy-lysosomal system) extend lifespan in yeast[67], flies[68], worms[69], [70] and mice[71], [72]; and increasing ubiquitin-proteosome activity is known to extend lifespan in yeast[73] and nematodes[74], [75].

Deregulated Nutrient Sensing

The fact that dietary restriction (Figure 1.3) increases lifespan and/or healthspan in all the eukaryote species investigated so far[25], [76], [77] is a statement of the importance of nutrient-sensing in the aging process.

The most researched nutrient-sensing pathway is the "insulin and insulin-like grow factor (IGF-1)" pathway, which participates in glucose-sensing. An initial observation of its role creates a paradox: on the one hand, the activity level of this pathway experiences a decline in aged humans, as well as mouse models of premature aging[78]; on the other hand further attenuation of its signaling intensity, through multiple genetic interventions, in worms, flies and mice robustly prolongs lifespan[79]–[81]. The solution lies in considering that these genetic manipulations extend longevity because insulin and IGF-1 pathway (IIS) signaling leads to lower rates of cell growth and metabolism, hence culminating in lower rates of cell damage; and assuming that, for the same reason, aged organisms naturally decrease IIS activity in an attempt to maximize survival. It would logically follow that minimizing IIS activity would maximize lifespan in the same proportion. However, this is contradicted by evidence. In reality, low but not extremely low levels of IIS signaling are optimal for an aged organism[82].

Another intensively studied nutrient-sensing system is the mammalian target-of-rapamycin (mTOR), and it senses amino-acid concentrations. Genetic downregulation of mTORC1 (one of the complexes of which mTOR kinase is part of) extends lifespan in several model organisms[83], [84]. What's more, rapamycin is considered the most robust lifespan-extending pharmacological intervention in mammals[85].

Mitochondrial Dysfunction

It is no surprise that the classical theory of aging – the mitochondrial theory of aging[86] – finds a culprit in mitochondria if we note that: life is based on energy flow, and the cellular "power plants" are the mitochondria; it is known that their efficacy diminishes with aging[87].

The mitochondrial theory of aging proposes that global cellular damage is the consequence of a positive feedback loop between mitochondrial dysfunction and production of reactive oxygen species. The suggested paradigm was recently superseded by the mitohormesis concept, supported by experimental evidence showing that increased production of reactive oxygen species prolongs lifespan[88]–[91] in model organism instead of accelerating aging and that antioxidants fail to extend lifespan[92], [93].

The mitohormesis paradigm states that mild exposure to toxic factors triggers defensive responses that end up overcompensating for the damage, resulting in a superior cellular fitness in the pos- *versus* pre-exposure states[94].

Cellular Senescence

The amount of senescent cells (cells exhibiting irreversible cell cycle arrest coupled to stereotyped phenotypic changes) increases with aging[95]. It is not known if this accumulation is due to an increased rate of generation of senescent cells and/or a decrease in their rate of clearance. Nonetheless, new evidence unambiguously establishes a causal link between the aging phenotype and senescence cell population size[96].

The belief in the potential of directly targeting cellular senescence as an anti-aging intervention is exemplified by the definition and exploration of a new functional class of drugs termed senolytics[97].

Stem Cells Exhaustion

Adult organisms retain several stem cell niches that are responsible for tissue regeneration throughout their lifetime.

Almost by definition, the regenerative potential of tissues diminishes with aging. Experimentally, it is found that functional attrition of stem cells is ubiquitous across cell populations[98]–[100]. Although the deficient proliferation of stem cells is detrimental to healthy aging, an excessive proliferation rate, which translates into niche depletion, leads to premature aging in flies[101].

Parabiosis experiments, in which and old and young mice share one circulatory system, have shown that the old mice experience a reversal in stem cell functional decline coupled with increased health- and lifespan[102], consistent with the view of stem cells exhaustion as an integrated hallmark. This experimental evidence suggests that stem cell rejuvenation therapies might be feasible and with benefic repercussions at the organism level[103].

Altered Intercellular Communications

Aging-associated changes are present not only at the individual cell level but also on how cells, tissues, and organs interact with one another. Altered intercellular communications are the most holistic hallmark of aging. An old cell might communicate differently with its neighbors and the aggregated effect of miscommunicating aged cells is the general deregulation of neurohormonal signaling characteristic of aged organisms.

Among the types of aging-associated changes in intercellular communications, one can count chronic inflammation[104] and immunosenescence[105], but from a therapeutic perspective the most interesting is, arguably, “contagious aging”.

Contagious aging is the spatial propagation of aging, in which an aged cell, tissue or organ, leads to aging-specific deterioration of another cell, tissue or organ. For example, senescent cells induce senescence in neighbor cells via chemical signaling[106]. Therapeutic significance does not lie in contagious aging *per se*, but instead in the experimental validation of its corollary consisting in the inversion of the concept; in other words, anti-aging interventions targeting a particular tissue can have the side-effect of slowing or halting the aging process in other tissues[51], [107].

1.1.3. ANTI-AGING THERAPIES

Lifespan-extending interventions can be one of three natures: pharmaceutical; genetic; or lifestyle modification.

The main lifestyles changes that are shown to increase healthspan or thought to extend lifespan are physical exercise and dietary modifications, namely caloric restriction (Figure 1.3).

Physical exercise does not extend lifespan, but it improves healthspan[108]. Exercising guidelines have long been a recommendation of health-concerned organizations from all over the world, and

rightly so[109]. Nonetheless, exercise alone cannot make up for the modern society lifestyle. The increasingly sedentary lifestyle negates or attenuates many of the health benefits conferred from physical exercising, independently of the physical activity level (exercise intensity)[110].

A calorie restricted diet, which consists in reducing the daily caloric intake by up to 40%, consistently extends healthspan and lifespan in many model organisms[111] and improves diseases risk factors in humans[112]. Unfortunately, its benefits of delaying the majority of age-associated diseases come at the price of disrupting the homeostatic state of the adult human body, for example, it leads to menstrual irregularities, loss of libido and slower wound healing[113]. The clear failure in translating the potential of caloric restriction to humans resulted in two refinements of the concept: fasting-mimicking diets and caloric restriction mimetics.

Fasting-mimicking diets aim to maximize the upsides and minimize the side-effects of purely restricting calories, by imposing their limitation in the time domain. They come in the form of: time-restricted feeding, which imposes a daily time window for the feeding period; intermittent fasting, which features ad libitum days followed by days without food or with a very low-calorie daily intake; or cycling relatively extended periods of normal dieting and calorie restricted dieting[114]. The initial results of fasting-mimicking diets do seem to improve the applicability of the caloric restriction concept to humans[115], although only studies using cohorts of small sizes are available. Another limiting factor in using fasting-mimicking diets as anti-aging interventions is that due to their multifactorial nature many variables still need to be optimized and tailored to special populations. As a case in point, old mice are negatively affected by a 4-day mimicking-fasting diet, but not by the same diet when this is restricted to just three days[116].

Caloric restriction mimetics[117] are small compounds that when ingested trigger phenotypic changes akin to the ones elicited by following a calorie restricted diet. The ultimate goal is to design a pill that would be taken daily by anyone and improve the healthspan and lifespan of the individual, akin to the canonical anti-aging miracle pill.

The mere existence of research focused on finding caloric restriction mimetics as a perfection of a calorie restricted diet brings to surface two crucial facts: that there is enough evidence suggesting that such compounds might actually exist (otherwise we would not see any scientific endeavor on this facet, especially since alternatives, like fasting-mimicking diets, are showing promising results); and that pharmacological interventions are more desired than their lifespan modification counterparts (in which poor compliance is a common issue).

More than a thousand genetic interventions are known to result in lifespan extension of model organisms[22]. However, this otherwise mighty potential is betrayed by the ethical and technical challenges associated with the application of gene therapy in humans. The ethical factor (notice that we are referring to intervention in adult humans and not germline editions) is a limitation secondary to the technical challenges. If there were no safety issues with gene therapy, we believe that gene therapy would become widely adopted and even ordinary.

In the technical front, the discovery of the CRISPR technology, the technology with the fastest adoption rate in history[118], is revolutionizing gene therapy. It does so because it is cheaper, highly-customizable (due its modular structure) and more efficient in gene editing (including knock-ins) than the competing techniques. As a potential anti-aging intervention, it could be used to target the genome and epigenome[119], and be administered by gene vectors or to stem cells that would then be transplanted into the patient[120].

The unprecedented rate of progress of the CRISPR technology has not yet arrived at the point in which the rate of off-target mutations is not larger than the rate of spontaneous mutation in the human genome. Until then, the usefulness of gene therapy for anti-aging interventions remains at a distance.

The last type of anti-aging intervention left to discuss is pharmacological interventions.

For virtually every single chronic disease for which an oral pharmaceutical treatment is viable, they are the option of choice. The hassle-free administration mode allows the patient to adhere to the doctor's orders, maximizing the success of the treatment.

For the first time in humanity's history there exists enough scientific knowledge and evidence to support an anti-aging clinical trial. This pioneer clinical trial is not the fruit of chance but a culmination of years of dedication in fighting aging.

Several years ago the National Institutes of Health recognized that aging could be targeted and established the National Institute of Aging Interventions Testing Program.(ITP) This program is conducted at multiple centers and tests potential anti-aging interventions in genetically heterogeneous (outbred) mice. After its successful results in ITP, metformin is now going to be subjected to a randomized, controlled clinical trial[121].

Even though the landscape of aging research provides reasons for being optimistic, it is worth noticing that metformin is just a single compound, and a single compound intervention is always theoretically suboptimal in manipulating a multifactorial process such as aging. Biogerontologists

are well aware of this short-coming and are searching and exploring other potential life-extending compounds[122].

1.2. MACHINE LEARNING

The human brain is a fascinating organ. Through millions of years of evolution, it became expert in pattern recognition, abstract thinking, and language[123]. Unfortunately, its organic nature sets rigid limitations in the volume of information that it can process[124] and makes it only adept to the simplistic modeling of the world[125].

The alternative that bypasses the constraint of an organic computational architecture is to use computers to store information and appropriately extract meaningful knowledge from it, or, in one word, to learn. Machine learning is the field of study that is centered in giving computers the ability to learn without being explicitly programmed by static instructions[126]. The learning capabilities of machine learning applications have benefited greatly from Moore's law and progresses in applied mathematics, statistics, and computer science. They are the only available avenue for inference in highly nonlinear systems of unprecedented scale exhibiting a rich landscape of interactions[127].

1.2.1. CLASSIFICATION TASK

The tasks that machine learning is applied to can be categorized accordingly to the type of output desired from the machine-learned system. The task faced in the current work is called classification, and it consists of the model outputting a class label given an example as input. In our case, we will be doing binary classification since there are only two possible classes that a given sample instance can be assign to: or it belongs to group X or it doesn't.

An additional way of categorizing machine learning tasks is based on the nature of the learning "signal" or "feedback" available to the learning system[128]. Categorized in this way our binary classification task belongs to the "supervised learning" realm. It does so because we are going to use labeled examples to teach the learning system a mapping of their features to their class memberships.

An interesting concept useful in a classification task, and that we are going to take advantage of is to program the learning system to output not only the class label that it classifies a given sample to belong to but the probability of it belonging to each of the available classes. For example, a system trained to predict the gender of a person by using as input the individuals' height, instead of outputting just "Male" or "Female", it would output "0.35 chance of Male and 0.65 of Female".

The last type of output is usually referred to as "class probabilities" and it reflects the degree of confidence with which the system classified each sample.

1.2.2. ALGORITHMS

A machine learning algorithm can be loosely defined as the set of formalized instructions that one takes to build a machine learning system.

There are thousands of algorithms available to choose from, so the choice should be tailored to the specific problem and objectives at hand. For our binary classification task, we seek an algorithm capable of state-of-art performance in datasets of similar nature to ours. Ideally, it should also naturally output class probabilities. As it turned out, our classification task has two additional special traits that need to be taken into consideration: the existence of class imbalance and the fact that the number of predictors far exceeds the number of observations available to train the algorithm.

Our dataset is class imbalanced because one of the classes as significantly more samples than the other. If a static rule that outputted the majority class label independently of the input sample were created, it would achieve an accuracy proportional to the class imbalance of the dataset without actually doing any learning. The disregard for class imbalance might lead to use algorithms that bias toward the majority class, in other words, they fail to predict the minority class. In our case, this consideration has additional weight since the minority class is the targeted class of interest.

When the number of predictors is much larger than the number of samples (also called "$p \gg n$") we are dealing with a high-dimensional dataset[129]. The danger is that there are more ways to separate data samples (subset the high-dimensional sample space) than there are samples. In this setting the algorithms overfit (find ways to explain the noise of the data), losing generalization capabilities.

Random Forest

Taking into consideration the characteristics of our dataset, we choose to base our classifier on the random forest algorithm[130].

It is called "forest" because its outputted prediction results from the aggregation (e.g. using the majority vote or the average of class probabilities) of the output of several independent instances of a decision tree algorithm (Figure 1.5). The "random" part originates from the fact that each decision tree only as access to a random fraction of the available data[130]. This source of

randomness gives “biodiversity” to the forest and therefore it is also the source of the robustness of the random forest algorithm.

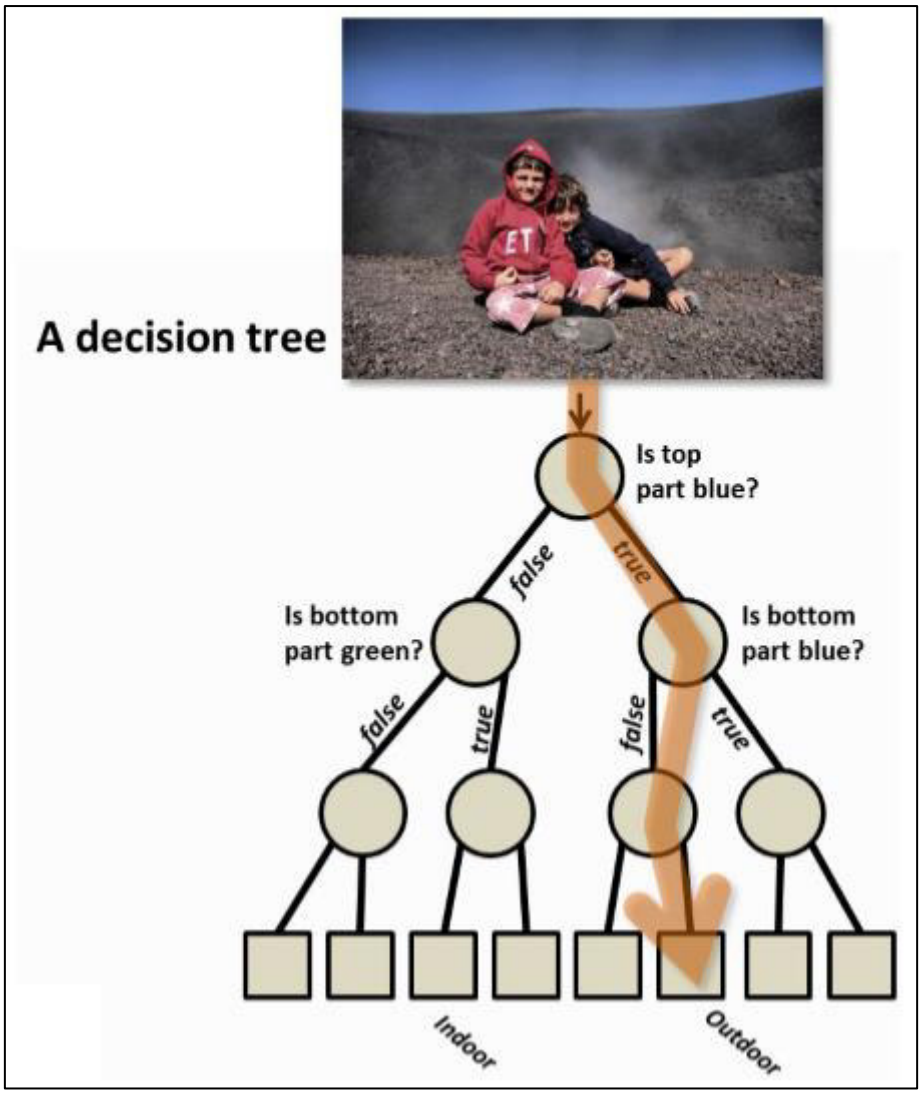


Figure 1.5: Decision tree. A decision tree is a hierarchical structure and the basic unit of the random forest algorithm. Each internal node stores a split (or test) function to be applied to the incoming data. Each leaf stores the final answer (predictor). Here we show an illustrative decision tree used to figure out whether a photo represents an indoor or outdoor scene. From [131].

Notice that because the decision trees are grown independently, with only the aggregation step requiring information from all of them, the random forest algorithm is renowned for its scaling properties.

Random forest can handle binary/categorical and continuous input features simultaneously since all types of inputs are easily partitioned by a decision tree (anything can be converted into a binary logic test).

Ensemble learners, that is methods that are based on the aggregation of many independent classifiers, and, more particularly, random forests hold the state of art results in high-dimensional datasets[132], [133].

Class imbalance deteriorates the performance of virtually all algorithms[134]. Nonetheless, in a large scale comparison of the most popular algorithms in ten imbalanced experimental datasets, random forest outperformed all the other models[135].

1.3. MOTIVATION AND FRAMEWORK

The contributions of the present dissertation are discussed in detail in chapters 2 and 3. Although they can be read in an almost independent fashion, both are guided by one unifying working framework.

We have commented upon the enormous burdens that aging poses at the population and individual levels. We have also enumerated its known underlying biological mechanisms and malleability. From the last two statements, we can logically derive the motivation of this work: the aging process itself should be the direct target of potential therapeutic interventions. We even go further: aging-ameliorating interventions should be a top priority of medical research given its unique potential in stopping the impending social disaster.

Allocation of funds to aging-specific research is the fairest among all possible choices of public spending in medical research if we consider that, not the majority, but every single member of society will directly benefit from it (we will all be old eventually). Furthermore, it is also the choice that maximizes return-on-investment, since age is the primary risk for nearly every major cause of mortality in developed nations[136]. Lastly, and more important, when people were surveyed about if they wished to live longer given that their healthspan would also increase, the result was a resounding “yes”[137].

The classical paradigm of biomedical research is focused on the treatment of individual diseases. This disease-centric tradition as made profound contributions to human health, helping people live longer than ever before. However, longer lives without access to successful therapies in preventing, ameliorating, or postponing the aging process, culminate, paradoxically, in a longer undesirable life stage that continuously accumulates multiple disabilities and morbidities. In lay terms: we are living longer, but also dragging out suffering.

The chosen working framework is grounded on a contrastingly different paradigm: that directly targeting and slowing the aging process delays the onset and progression of all aging-related

diseases. The clinical importance of this "longevity dividend" cannot be overstated, as even maximum success in curing a single chronic illness is insufficient to halt a multimorbidity state (and sometimes it even aggravates the other co-morbidities).

An alternative view of our working paradigm is that it is focused on prevention rather than treatment. This hypothesis considers that increasing healthspan (compressing morbidity) is the best choice to face the otherwise unmanageable chain of events that cause aging before it even starts.

We mentioned before that anti-aging genetic interventions have not reached technological maturity yet and explained why pharmacological interventions dominate their lifestyle modification analogs, therefore therapies of the pharmacological type have the largest likelihood of being adopted post-haste on a global scale.

1.4. OBJECTIVES

Biogerontology's focus is on efforts to understand, prevent, cure or minimize age-related impairments, and the contributions of the current work belong to its domain.

Working from the framework exposed in the previous subsection, the specific objectives of this thesis are:

Objective 1: Investigate if there are drugs significantly enriched in interactions with human homologs of experimentally verified lifespan-extending genes.

Objective 2: If the answer to Objective 1 is affirmative, prioritize the obtained drugs for further lifespan testing in model organisms.

Objective 3: Create a database dedicated to lifespan assays of drugs known to prolong lifespan in model organisms.

Objective 4: Conduct initial analyses based on the information provided by the database created in Objective 3.

Objective 5: Compile biological and chemical information on drugs that have their lifespan effect known in *C. elegans*, and train and optimize a classifier so that it can be used for readily assessing the anti-aging potential of an unexplored compound, in future works.

1.5. CONTRIBUTIONS

Motivated by the arguments above and set on the objectives enumerated in the previous sub-section, we contribute to our field of study in the following ways:

Contribution 1: We examined thousands of drugs present in a meta-database of drug-gene interactions and found that 16 of them have a significant number of meaningful connections with human homologs of model organisms' lifespan-extending genes.

Contribution 2: We analyzed the conceptual viability of the top candidate anti-aging drugs and ended up suggesting two distinct drug combinations that should be tested in model organisms.

Contribution 3: We curated almost all the existing literature regarding anti-aging drugs assayed in model organisms and present the resulting effort in the form of a new resource: "DrugAge: Database of Aging-Related Drugs". DrugAge is free to use, and it is open to the scrutiny of the scientific community at http://genomics.senescence.info/drugs/.

Contribution 4: We show evidence reinforcing the notion that the more complex an organism is, the smaller the relative magnitude of average and maximum lifespan extension achieved by pharmaceutical interventions.

Contribution 5: Through the use of two distinct statistical methodologies we clearly show that, in general, sexual dimorphism is not a concerning factor affecting the performance of anti-aging drugs in model organisms.

Contribution 6: We show that there is a statistically significant strong correlation between mean lifespan changes and maximum lifespan changes in drug lifespan assays.

Contribution 7: We contrasted, for the first time, anti-aging interventions of the pharmacological and genetic types by comparing the functional analysis and genetic overlap between DrugAge and GenAge.

Contribution 8: We successfully trained and optimized a classifier for gauging the anti-aging potentiality of untested chemical compounds.

CHAPTER 2: DRUG REPURPOSING FOR AGING

2.1. INTRODUCTION

Having previously mentioned the rationale for the need to tackle the aging process and to ideally do so through interventions of the pharmaceutical type, the challenge being considered now is: what is the best way to go about it.

The straightforward approach would be to design *de novo* drugs specifically targeting the aging process. Nonetheless, even assuming that we ignore the complexity and uncertainty wrapping the aging process, the drug research and development (R&D) process itself is full of obstacles.

Even with the rate of technological progress following Moore's law, bafflingly, every nine years the cost of drug development doubles (Figure 2.1)[138]. To make matters more worrisome, this increase in the cost of drug development is not compensated by the increase in the number of newly approved drugs entering the market. In fact, the latter rate as remained flat since the 1950s[139]. Such a depressing state of affairs is forecasted to lead to the stagnation of investment in drug R&D projects[138]. In actuality, the market sentiment regarding the potential financial returns of the pharmaceutical sector is so negative, that the only viable solution relies on the direct application of securitization techniques to improve its risk-to-reward profile[140], [141].

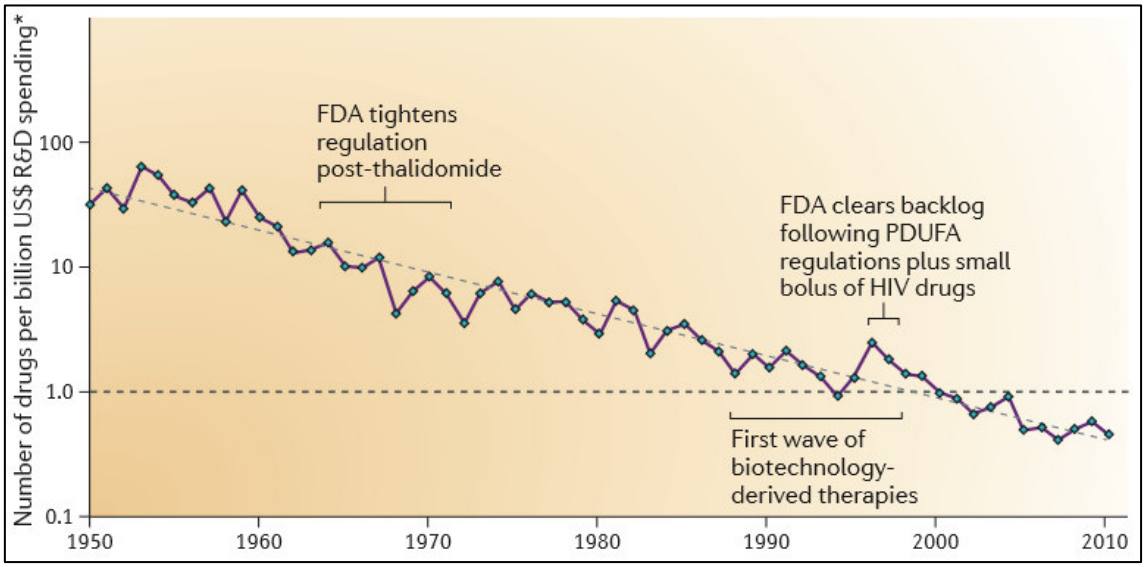


Figure 2.1: The number of new drugs approved by the US Food and Drug Administration (FDA) per billion US dollars (inflation-adjusted) spent on R&D has halved roughly every nine years. From [138].

From a technical perspective, the disappointing results of a candidate drug journey from bench to bedside mainly originate from issues with clinical side effects and tolerability[142], [143].

Drug repurposing allows bypassing the drug industry R&D bottlenecks by trying to re-apply compounds that were found to be safe and tolerable in human clinical trials to treat health conditions that are not their original intended target. Notice that it is of no importance if the drug being repurposed was successfully approved for its primary target. As long as it is deemed safe, a compound can be subjected to a new therapeutic hypothesis, and its efficacy promptly evaluated in a new clinical trial.

One case of successful drug repurposing is azidothymidine. This compound was deliberately developed as an anti-cancer agent and shown to inhibit oncogenic viruses (and tumor proliferation)[144]. Failure in treating cancer did not stop it from being approved as an anti-HIV therapy, with its previously known antiviral activity convincingly inhibiting HIV replication[145].

Two clear advantages are unique to the drug repurposing paradigm:

- it is a financial free lunch (the investment to develop, register and fully test the compounds has already been made) with potential applications only limited by the scientific community ingenuity;
- most of the 17 years period that it takes to translate research findings into clinical practice[146] can be skipped by immediately testing the hypothesis at the clinical trial stage.

Compelled by the advantages of the drug repurposing paradigm and inspired by its previous achievements, it is the goal of this section to apply a bioinformatics approach to gauge the lifespan-extension potential of already developed compounds. If compounds with considerable potential are to be found, they will be ranked (and therefore prioritized) in an effort of advancing drug repurposing in biogerontology.

2.2. METHODS

Any drug repurposing project starts by defining the library of compounds to be screened. We find an ideal repository of drugs in the Drug-Gene Interaction Database (DGIdb)[147]. DGIdb is the largest meta-database of drug-gene interactions, updated weekly, publicly available, and it considers information from ongoing clinical trials. Its goal is to "provide users with the most current knowledge of clinically actionable drug–gene interactions"[147].

Together with using DGIdb as the source of candidate lifespan-extending drugs and their respective drug-gene interactions, we also made use of the GenAge Database of Aging-Related Genes[22] to gather gene targets for anti-aging therapies. GenAge is also the perfect resource for our purpose

because it is the most complete database of genetic interventions that influence lifespan in model organisms.

Version 2.22 of DGIdb was downloaded and imported into RStudio[148] (version 0.99.879) running R[149] (version 3.2.3). The base file contained 3158 human genes establishing 29708 interactions with 11636 unique compounds.

Build 18 of GenAge was downloaded and imported into Microsoft Excel 2013. With this step, we obtained information regarding 2242 manipulated genes in model organisms. GenAge is a repository of genetic experiments that affect lifespan in model organisms, but we are only interested in the subset, of 1184 genes, that extend lifespan when intervened upon. Furthermore, it is crucial to take into consideration the type of genetic manipulation that when applied to each gene culminates in lifespan extension. This consideration subsets the life-extending genes into two categories:

- One group of lifespan-extending GenAge genes which increase longevity when they are knockdown or knockout (reduced levels of gene expression) and are called “anti-longevity” genes;
- Another group of genes which only prolongs lifespan when they are over-expressed or knock-in (increased gene activity) and are defined as “pro-longevity” genes.

After data cleaning, we found 988 and 175, anti-and pro-longevity lifespan-extending animal genes, respectively.

We used the dbOrtho web service of bioDbnet[150] (version 2.1) to convert the two subsets of GenAge lifespan-prolonging genes into human orthologues. The two resulting subsets of 995 and 198 unique human orthologues were then imported to RStudio. Due to the orthologue conversion 58 genes were common to both subsets of human genes and removed as their nature as anti- or pro-longevity is ambiguous. Of our 1077 total non-redundant human orthologues of GenAge lifespan-extending genes, 140 were pro-longevity genes and 937 were anti-longevity genes.

Of the 1077 non-redundant human orthologues, 357 (59 and 298, pro- and anti-longevity genes respectively) were present in DGIdb, interacting with 2064 distinct drugs.

Table 2.1: Categorization of DGIdb´s drug-gene interactions according to their effect on gene expression levels.

Anti	Neutral	Pro
antagonist	n/a	activator
antagonist, antibody	binder	agonist
antagonist, inhibitor	modulator	agonist, modulator
antagonist, inhibitor, competitive	antibody	agonist, partial agonist
antagonist, inhibitory allosteric modulator	other/unknown	chaperone
antagonist, multitarget	allosteric modulator	cofactor
Antisense	multitarget	inducer
antisense oligonucleotide	ligand	partial agonist
binder, antagonist	agonist, antagonist	partial agonist, agonist
Blocker	antagonist, partial agonist	potentiator
Cleavage	vaccine	product of
competitive, inhibitor	adduct	stimulator
Inhibitor	partial agonist, antagonist	
inhibitor, antagonist	immunotherapy	
inhibitor, competitive	antagonist, agonist	
inhibitory allosteric modulator	positive allosteric modulator	
inverse agonist	agonist, inhibitor	
multitarget, antagonist		
negative modulator		
partial antagonist		
Suppressor		

Previously we made the distinction between pro-longevity and anti-longevity genes. This distinction is relevant now that we have gathered all the gene interactions because it is time to select the ones that we can assume that would influence lifespan.

A drug that inhibits pro-longevity gene is not expected to extend lifespan because this interaction would only be meaningful (for our purposes) if it was an interaction that would increase the expression of the gene. Several other possible cases of drug-gene interaction could be considered with the variables being the nature of a longevity gene (anti or pro) and the type of interaction between a drug and that gene. In order to focus exclusively on the drug-gene interactions that prolong lifespan we assigned all sorts of drug-gene interactions present in DGIdb to the categories (

Table 2.11):

- "Anti" - types of drug-gene interactions that might extend lifespan if an anti-longevity gene is being considered;
- "Pro" - interactions that increase gene expression (or similar magnification of gene influence) and therefore should be taken into account when we are dealing with a pro-longevity gene;
- "Neutral" - interactions that couldn't be included in neither of the other categories because they are ambiguous (case-specific) or non-informative.

Some of the cases related to some interaction types were manually checked to assure the quality of our analysis. Of these 2064 drugs, only 489 had interactions in a meaningful way, that is, the interaction type was the required for a specific gene to exert its anti-aging effects.

Table 2.2: Summary of the 16 chemical compounds found to be significantly enriched.

Drug Primary Name	AdjPvalue	SuccSample	FailSample	SuccPop	FailPop
Dacinostat	7.07E-13	10	0	98	3050
Givinostat	7.07E-13	10	0	98	3050
Abexinostat	7.07E-13	10	0	98	3050
Belinostat	7.51E-12	10	1	98	3049
Vorinostat	7.51E-12	10	1	98	3049
Pivanex	7.01E-10	8	0	100	3050
Sodium phenylbutyrate	7.01E-10	8	0	100	3050
Panobinostat	2.42E-08	10	8	98	3042
Valproic acid	3.18E-07	10	12	98	3038
Romidepsin	2.05E-05	5	0	103	3050
CHR-3996	6.21E-04	4	0	104	3050
Roflumilast	6.21E-04	4	0	104	3050
Entinostat	3.01E-03	4	1	104	3049
Ipatasertib	1.85E-02	3	0	105	3050
MK2206	1.85E-02	3	0	105	3050
Ibudilast	1.98E-02	4	3	104	3047

Next, we measured the statistical enrichment of relevant lifespan-extending interaction for each drug through the one-sided Fisher's exact test, using the "exact2x2" R package (version

1.4.1)[151] , with null hypothesis being that a drug has the same proportion of interactions with lifespan-extending genes (of the type required to prolong lifespan) as the proportion present when the entire dataset (DGIdb) is considered. After correcting for multiple hypotheses using Holm's correction[152]($\alpha = 0.05$), we found 16 compounds significantly enriched in correctly oriented directed edges with human orthologues of lifespan-extending GenAge genes (Table 2.2).

2.3. RESULTS AND DISCUSSION

The results of the current methodology are part of a manuscript that is currently in the final phase of the peer review process for publication in the scientific journal "Human Molecular Genetics".
We follow with the analysis and discussion of the results from two perspectives: the drugs obtained, and the gerontologic information of the genes by which they exert their effects. It is our believe that the second adds valuable insights to the first, namely in developing a contextualization for the expected magnitude of their potential anti-aging clinical effects.

2.3.1. ENRICHED DRUGS

A first contemplation of the list of enriched drugs and all of the genes that they target reveals that all the implicit drug-gene interactions belong to the anti-longevity category, in other words, the candidate compounds shall delay aging solely by inhibiting the expression of anti-longevity genes. We do not have any explanation that justifies why pro-longevity drug-gene interactions were not enriched.
We can cluster our statistically enriched drugs into three broad classes accordingly to the lifespan-extending genes that they target (Table 2.3):

- histone deacetylases inhibitors (dacinostat, givinostat, abexinostat, belinostat, vorinostat, pivanex, sodium phenylbutyrate, panobinostat, valproic acid, romidepsin, CHR-3996, entinostat);
- Akt signaling inhibitors (ipatasertib and MK-2206);
- phosphodiesterase inhibitors (roflumilast and ibudilast).

Dacinostat[153] is still an experimental drug, and as yet to reach clinical trials.
Givinostat has reached the phase II of clinical trials a few times. Unfortunately, it causes thrombocytopenia[154], which excludes its application as a candidate anti-aging therapy.

Abexinostat is a pan-HDAC inhibitor that mostly targets HDAC1. Its common side effects, which include thrombocytopenia, neutropenia, fatigue and anemia[155], do not justify the trade-off between quality of life and potential lifespan-extending effects that would be subjacent to its usage.

Table 2.3: Enriched drugs and their respective gene targets.

	Dacinostat	Givinostat	Abexinostat	Belinostat	Vorinostat	Pivanex	Sodium phenylbutyrate	Panobinostat	Valproic acid	Romidepsin	CHR-3996	Roflumilast	Entinostat	Ipatasertib	MK2206	Ibudilast
ABAT									X							
AKT1														X	X	
AKT2														X	X	
AKT3														X	X	
HDAC1	X	X	X	X	X	X	X	X	X	X	X		X			
HDAC10	X	X	X	X	X			X								
HDAC2	X	X	X	X	X	X	X	X	X	X	X		X			
HDAC3	X	X	X	X	X	X	X	X	X	X	X		X			
HDAC4	X	X	X	X	X	X	X	X	X	X						
HDAC5	X	X	X	X	X	X	X	X	X							
HDAC6	X	X	X	X	X			X								
HDAC7	X	X	X	X	X	X	X	X	X							
HDAC8	X	X	X	X	X	X	X	X	X	X	X		X			
HDAC9	X	X	X	X	X	X	X	X	X							
OGDH									X							
PDE4A												X				X
PDE4B												X				X
PDE4C												X				X
PDE4D												X				X

As in the two previous drugs, Belinostat's clinical side effects (among them one can count anemia, thrombocytopenia, dyspnea, and neutropenia[156]) also exclude it as a desirable anti-aging intervention.

Vorinostat is another unattractive candidate anti-aging compound. It causes hematological toxicity, mainly in the form of thrombocytopenia[157].

Pivanex is relatively well-tolerated for a compound of the HADC-inhibiting class. Its only side effects are nausea and dysgeusia[158].

Sodium phenylbutyrate is used to treat urea cycle disorders. It is deemed very safe and tolerable, for example, in a phase II clinical study applying it to treat amyotrophic lateral sclerosis it was tolerated in all dosages tested (including at a dosage of more than double the suggested therapeutic dosage) and histone acetylation was decreased by half[159].

Panobinostat is a broad-spectrum HDAC-inhibitor. We disregard its feasibility as an anti-aging intervention because it is administered intravenously (which would raise compliance issues) and it causes the side effects typical of its functional drug class: thrombocytopenia, fatigue, and neutropenia[160].

Valproic acid is the outlier in our group of enriched HDAC inhibitors as it also inhibits the OGDH (oxoglutarate decarboxylase)[161] and ABAT (4-aminobutyrate aminotransferase)[162], [163] genes. It was first approved in 1967 as an anticonvulsive drug and therefore its long-term chronic effects are well-studied. Its chronic ingestion leads to significant weight gain and if it were to be applied as an anti-aging intervention, platelet, and hepatic functions would have to be closely monitored mainly for the risk of thrombocytopenia (which has an incidence of 12%)[164].

Romidepsin mostly inhibits HDAC1 and HDAC2. It severely compromises healthspan by causing extremely fatigue[165] along with the expected toxicities: nausea, vomiting, fatigue, and transient thrombocytopenia and granulocytopenia[166].

For the treatment of solid tumors, CHR-3996 display a favorable toxicological profile, but the same cannot be said for the treatment of aging, as it caused atrial defibrillation even in the lowest dosage in a phase I clinical trial[167].

Entinostat is a potent inhibitor of HADC1 and HADC3. Its adverse events reported in several clinical trials included anorexia, nausea, hypoalbuminemia, hypophosphatemia, fatigue, headache, diarrhea, neutropenia, thrombocytopenia and leukopenia. These are all reversible, but because an anti-aging pharmacological intervention is continuous in nature, its toxicological profile is unacceptable[168].

Ipatasertib is a highly selective pan-Akt inhibitor. It is currently undergoing phase II clinical trials for prostate (clinicaltrials.gov NTC01485861) and gastric cancer (clinicaltrials.gov NTC01896531). Initial studies seem to indicate that it is relatively well-tolerated for a cancer drug, but it causes diarrhea, fatigue and hyperglycemia[169].

MK-2206 is also a highly selective pan-Akt inhibitor and it was the first allosteric small molecule inhibitor of Atk to enter clinical development[170]. It has and still is being tested in multiple phases I and II clinical trials. Rashes tend to be the only reported common side effect and are observed in a dose-dependent manner[171].

Roflumilast is a highly selective phosphodiesterase-4 inhibitor approved for the treatment of chronic obstructive pulmonary disease. In contrast to the majority of drugs for treating chronic obstructive pulmonary disease, it is not inhaled but instead taken orally, a once a day. The only clinical side effects observed with its administration are diarrhea (in a few instances leading to hospitalization) and weight loss (an average of 2kg and more relevant in obese patients)[172], [173]. The intestinal distress might be attenuated by the titration of the dose up to the recommended dosage[174].

In Japan, ibudilast has been approved as a daily treatment of asthma for more than two decades. It is a non-selective phosphodiesterase inhibitor, and therefore it causes moderate gastrointestinal adverse effects. However, it should be noted that this is the only common side-effect and that it disappears after 2-4 days[175]. Highly encouraging and emphasizing its safety and tolerability is the fact that in a recent systematic study of drug repurposing for secondary progressive multiple sclerosis (a disease requiring long-term treatment, just like in anti-aging therapies) ibudilast was one of the few selected lead candidates for clinical evaluation[176].

2.3.2. BIOGERONTOLOGY OF THE TARGETED GENES

4-aminobutyrate aminotransferase

The 4-aminobutyrate gene is a newly aging-associated gene, and therefore its role in aging is still unknown. Its deletion in yeast extends mean chronological lifespan by 15-50%[177].

AKT serine/threonine kinases

The AKT serine/threonine kinase 1 (AKT1) gene isoform was studied in several human genome-wide association studies. In Caucasian populations the results are mixed, some studies find SNPs in AKT1 significantly associated with longevity[178]–[180], while others do not[181]–[183]. In a Han Chinese population, two AKT1 haplotypes were significantly represented in long-lived individuals[184].

Haploinsufficiency of Akt1 (the mouse homolog of AKT1) extends mean lifespan by 8% in males and 15% in females when compared to wild-type. The mutant mice had the same weight but less body fat. In *C. elegans*, the same studied reported that inhibition of akt1 (the worm homolog of

AKT1 and AKT3 isoforms) extends lifespan[185], which is consistent with previous literature that also used RNA interference to inhibit akt1[186].

Deletion of *Saccharomyces cerevisiae*'s SCH9 (the yeast orthologue of all the three AKT human isoforms) more than triples its mean chronological lifespan[187]–[189].

It must be mention that, contrary to what the deletion experiments in yeast might suggest, maximal reduction of AKT expression should not be a therapeutic goal as AKt1, AKt2 or AKt3 null mice are all viable but display severely compromised healthspan[190], [191].

Histone deacetylases

The human HDAC1 gene has a homolog in the gene Rpd3 of *D. melanogaster*. A decrease in Rpd3 gene expression mimics caloric restriction, and there is sexual dimorphism in the magnitude of lifespan extension that it induces. Males heterozygous for hypomorphic or null mutation of Rpd3 have a lifespan extension of 33% and 41-47%, respectively. While females heterozygous for a hypomorphic allele have a 52% increase in lifespan, and those carrying a null mutation, do not display median lifespan extension.[192]

In yeast, there are two homologs of human HDAC genes: HDA1 is a homolog of HDAC4, HDAC5, HDAC6, HDAC7 and HDAC9; and RPD3 is a homolog of HDAC2, HDAC3, and HDAC8. Only the HDAC1 is considered to be a homolog of both of the yeast genes.

Deletion of RPD3 extends replicative lifespan by 41%[193], [194] and, corroborating the results obtained in flies; there was no additive effect with caloric restriction. Deletion of HDA1 has no effect[194] or a very moderate one[67] of its own on longevity but acts synergistically with caloric restriction to increase life span.

Oxoglutarate dehydrogenase

The human oxoglutarate dehydrogenase gene (OGDH) is a homolog of the *C. elegans*, *S. cerevisiae,* and *S. pombe*; odgh-1, KGD1, and SPBC3H7.03c genes, respectively.

In worms, RNA interference of odgh-1 extended lifespan by 79%[195], a result congruent with the increase in lifespan observed when SPBC3H7.03c suffers deletion in fission yeast[196]. Curiously, in baker's yeast, KGD1 deletion halves the lifespan in two different studies[177], [197].

Phosphodiesterase 4 isoforms

All the four human phosphodiesterase isoforms (PDE4A, PDE4B, PDE4C, and PDE4D) have the dunce gene of *D. melanogaster* as a homolog. Cyclic adenosine monophosphate

phosphodiesterase-deficient dunce mutants enjoyed a maximum lifespan extension of about 70%[198].

2.3.3. RECOMMENDED DRUG COMBINATIONS

Aging is a multifactorial debilitation, even at the genetic level, so it makes sense that a pharmaceutical intervention with it as a therapeutic target should benefit from a composite set of drugs that individually target each of its contributing factors.

We can attempt to suggest combinations of this kind based on the targets of our enriched drugs (Table 2.3) and their side effects. These combinations can then be tested in model organisms and evaluated using lifespan and healthspan endpoints.

We shall give priority to tolerability and safety criterions for two reasons: the lifespan-extending effects of the obtained drugs are still unknown, and so they cannot be taken into consideration; and that for an intervention that it is likely going to be chronic, healthspan concerns acquire increased importance.

Of the histone deacetylases inhibitors cluster, sodium phenylbutyrate presents the most attractive toxicological profile, and it is our recommendation.

MK-2206 has, arguably, slightly less unpleasant common side effects (rashes) than ipatasertib (which might cause fatigue and hyperglycemia), and, therefore, it is the AKT-inhibitor that we prefer for a long-term intervention.

Both, roflumilast and ibudilast, have a safe history of treating long-term conditions. They also share similar toxicological profiles, which imply that either of them could be used for our purposes.

Of notice, is that two classes of enriched drugs have gastrointestinal distress as a side effect, so an addictive aggravation of the same is to be expected when applied to humans (and this will probably elude lifespan assays in lower animals).

In sum, our recommended drug cocktail, targeting lifespan extension without hindering healthspan, is composed of sodium phenylbutyrate and MK-2006 plus either roflumilast or ibudilast. Ideally, both versions would be experimentally tested.

CHAPTER 3: DRUGAGE

3.1. INTRODUCTION

Common sense and the evidence presented in the introductory chapter support the notion that pharmaceutical interventions are the preferred type of interventions to ameliorate the aging phenotype, which, by its turn, would make it the kind with the potentially widest and fastest socioeconomic impact.

In order to develop an evidence-based anti-aging intervention we naturally started by searching for a tool that systematizes the results from drug lifespan assays performed so far. For our surprise no such resource existed, and therefore creating one became our priority.

In the rest of this chapter, we will enumerate the steps taken to create the "DrugAge: Database of Ageing-Related Drugs". DrugAge is a free online resource (available at http://genomics.senescence.info/drugs/) driven by feedback from the scientific community. The underlying data is the result of an ongoing manual curation effort, with all entries referencing to their corresponding PubMed record. The interface is intuitive, fast and responsive. The integration of summary tables, clear graphic displays, and annotations using third-party databases allows the user to develop additional insights without ever having to leave DrugAge. The database is free to download on the website, and we encourage feedback and further data submission.

We will also analyze and discuss:

- the statistical analysis of several of the insights that are, for the first time, revealed thanks to the wealth of data available in DrugAge;
- its statistical and functional enrichment analysis;
- the genetic overlap between DrugAge and GenAge (a database of ageing-related genes);
- and build a machine learning classifier that estimates the probability of a new compound resulting in lifespan extension (namely in *C. elegans*).

3.2. STATE OF THE ART

A biogerontologist looking for genomic information has at his disposal the "Human Ageing Genomic Resources"[199]. This online resource integrates information into three freely available databases: "GenAge" - a curated database of candidate human ageing-related genes and genes associated with longevity and/or aging in model organisms.; "GenDR"[200] - a curated database of genes related to dietary restriction in model organisms either from genetic manipulation experiments or

gene expression profiling; and "LongevityMap"[201] – a database of human genetic variants associated with longevity.

This wealth of aging-related genetic information is in absolute contrast with the lack of resources regarding pharmaceutical assays and longevity, which would leave our not so hypothetical biogerontologist to his means, destined to personally mine and curate the hundreds of drug lifespan-assays available in the scientific literature.

One of the first aging-related databases is the Aging Genes and Interventions Database[202]. However, this resource is largely outdated, which reminds us that a resource that it is not regularly maintained and updated to reflect the latest scientific findings is bound to become completely useless, and in a relatively short time frame. This crucial insight will be kept with us for the rest of this chapter.

AgeFactDB[203] is a recent meta-database that collects aging factors and their lifespan data focusing primarily on the integration of existing aging databases (with the pharmaceutical ones being all outdated or abandoned). New data is added to AgeFactDB through automated data-mining of research paper abstracts and homology analysis, which although valuable, lacks the confidence associated with experimentally validated data.

We can, therefore, feel the urgent need for aging-related databases dedicated to pharmaceutical interventions, which are actively and consistently maintained to a high standard.

3.2.1. GEROPROTECTORS.ORG

The youngest aging-related database is "Geroprotectors.org"[23]. It takes special consideration not only because the author of this dissertation is one of its co-authors, but mainly due to having been developed at the same time as DrugAge and representing an improvement over the aforementioned state of the art in many ways.

Geroprotectors.org is a manually curated database of geroprotectors freely available online (http://geroprotectors.org/). A "geroprotector" is any intervention that aims to increase longevity, or that reduces, delays or impedes the onset of age-related pathologies by hampering aging-associated processes, repairing damage or modulating stress resistance. This broad definition implies that, for example, magnesium is considered a geroprotector based only on correlational evidence, between its dietary intake and mortality in adults at high risk of cardiovascular disease, originated from an epidemiological study[204]. The infeasibility of conducting lifespan assays in humans means that correlation analyses and interventions that ameliorate biomarkers of aging are

preferred, although one cannot scientifically prove that interventions using evidence of this nature actually extends human lifespan.

Although the first database of geroprotectors is of utmost importance for any physician and biogerontologist, its focus is distinct from DrugAge. Geroprotectors.org features a wider scope, more relevant from a translational/clinical perspective, rather than a pure science paradigm, for example, it considers epidemiological studies, clinical trials, drug human approval status, side effects and toxicological information. DrugAge, as we will soon see, will focus strictly on compiling drug lifespan assays in model organisms and in doing so with the ideal goal of exhaustively collecting all the literature available so far. The attempt to fulfill this audacious goal will culminate in DrugAge featuring more than double the number of lifespan assays relative to Geroprotectors.org. It must also be mention that DrugAge is going to be integrated into the Human Ageing Genomic Resources (HAGR), creating a one-stop online resource for biogerontologists' pharmacological and genomic needs.

A wise scientist will take advantage of the complementarity between the two databases. An ideal pharmaceutical intervention would maximize the translational potential of a compound based on Geroprotectors.org, with its credibility and expected magnitude of longevity effects based on experimental evidence, carried in model organisms, as indicated on DrugAge (and the rest of HAGR). Only the full information should be considered a robust indicator of a compound's feasibility to meaningfully extend lifespan and healthspan in humans.

3.3. TOOLS

With the current reproducibility crisis overshadowing more than 70% of the scientific results[205], it is essential to take deliberate measures to lessen its disgraceful effects.

Almost all the software used in this thesis is freeware, and the full source code is indiscriminately available upon request (requests should be addressed to diogog.barardo@gmail.com). We shall be revealing our software choices along the way with the exception of the ClueGO software, which demands a longer explanation.

3.3.1. CLUEGO

As announced before, we are going to be conducting functional enrichment analyses. A functional enrichment analysis consists in two steps.

First, we enumerate the biological entities present in our sample and map them to a curated bio-ontology. A bio-ontology is a schematic that incorporates biological entities into one or several concepts (e.g. pathways) and depicts the relation between these concepts[206]. The second step is the statistical analysis of the obtained concepts, against a reference background, to check the rarity/significance of each of them.

By far, the largest bio-ontology is Gene Ontology™ (GO)[207]. This ontology is updated daily by the Gene Ontology Consortium, which is supported by an open bioinformatics community. Virtually all the statistical enrichment analysis software packages are built upon this ontology. The next question is: "Which software to choose?". We continue with the sharing of our experience in this regard, which culminates in ClueGO being our tool of choice. The goal is to expose our rationale to the scrutiny of the reader.

Our first choice was to use PANTHER tools[208], which is an online web resource that allows, among other analyses, to do functional enrichment analysis. It is actively maintained (an important key attribute taking into reflection the high frequency of GO updates) and offers additional information based on protein evolutionary clustering (PANTHER™ is an acronym for "Protein Analysis Through Evolutionary Relationships"[209]).

Displaying only results with P<0.05; click here to display all results

	Homo sapiens (REF)	Client Text Box Input (▽ Hierarchy NEW! ⓘ)				
PANTHER GO-Slim Biological Process	#	#	expected	Fold Enrichment	+/-	P value
nitric oxide biosynthetic process	6	14	.65	21.40	+	3.48E-12
↳nitrogen compound metabolic process	1099	190	119.81	1.59	+	1.42E-07
↳metabolic process	8247	1427	899.03	1.59	+	2.93E-108
neuromuscular synaptic transmission	6	10	.65	15.29	+	4.74E-07
↳synaptic transmission	331	132	36.08	3.66	+	2.66E-33
↳cell-cell signaling	633	234	69.01	3.39	+	5.28E-55
↳cell communication	3006	686	327.69	2.09	+	9.69E-80
↳cellular process	6708	1207	731.26	1.65	+	6.97E-92
ferredoxin metabolic process	8	13	.87	14.91	+	2.59E-09
↳protein metabolic process	2692	397	293.46	1.35	+	7.36E-08
↳primary metabolic process	6825	1195	744.01	1.61	+	2.71E-82

Figure 3.1: Actual screenshot showing some of the output of PHANTER tools' statistical overrepresentation test analysis of data coming from an alpha version of DrugAge.

Figure 3.1 depicts an example of results of an overrepresentation statistical test using PHANTER's "Gene List Analysis" (http://www.pantherdb.org/). The more indentation a GO term has, the higher it is located on the hierarchical level; for instance: "protein metabolic process" is a subcategory of "primary metabolic process". We can observe that "cellular process" is a GO term

tremendously enriched, and that as it is decomposed into more specific/lower level GO terms, the magnitude of the statistical significance decreases. The decline of significance accompanying further decomposition is what is expected to happen since fewer elements in a subset imply that the p-value range loses granularity. Even in this typical case, we already can appreciate that the enriched results should be corrected for a hierarchical structure. Otherwise, the vast majority of the strongest enriched terms would be too general to allow any biological insight.

The need for weighting the structure of the ontology in the enrichment test is even more pronounced in the case of the decomposition of "metabolic process" into "nitric oxide biosynthetic process". The top level, "metabolic process", is more enriched than the second, "nitrogen compound metabolic process", in a typical manner. However, confusion settles in when we consider the p-value of the lowest level. It is clearly more significant than its parent, tempting us to immediately rationalize that its signal is so strong that it offsets the hierarchical structure bias. Such justification would be correct if its significance would be the highest of the decomposition, but this is not the case: "metabolic process" still is the term manifesting the highest statistical significance. We are now left in a grey zone: the lowest level is more meaningful from an annotation standpoint, and considerable enriched (enough to partial offset the hierarchical structure bias), and simultaneously, the maximum enrichment is exhibited by the most general term, which is uninformative, but shouldn't be ignored for ad-hoc reasons. The only resource that we found that empower us to solve dilemmas of this kind while still using the latest version of GO is ClueGO.

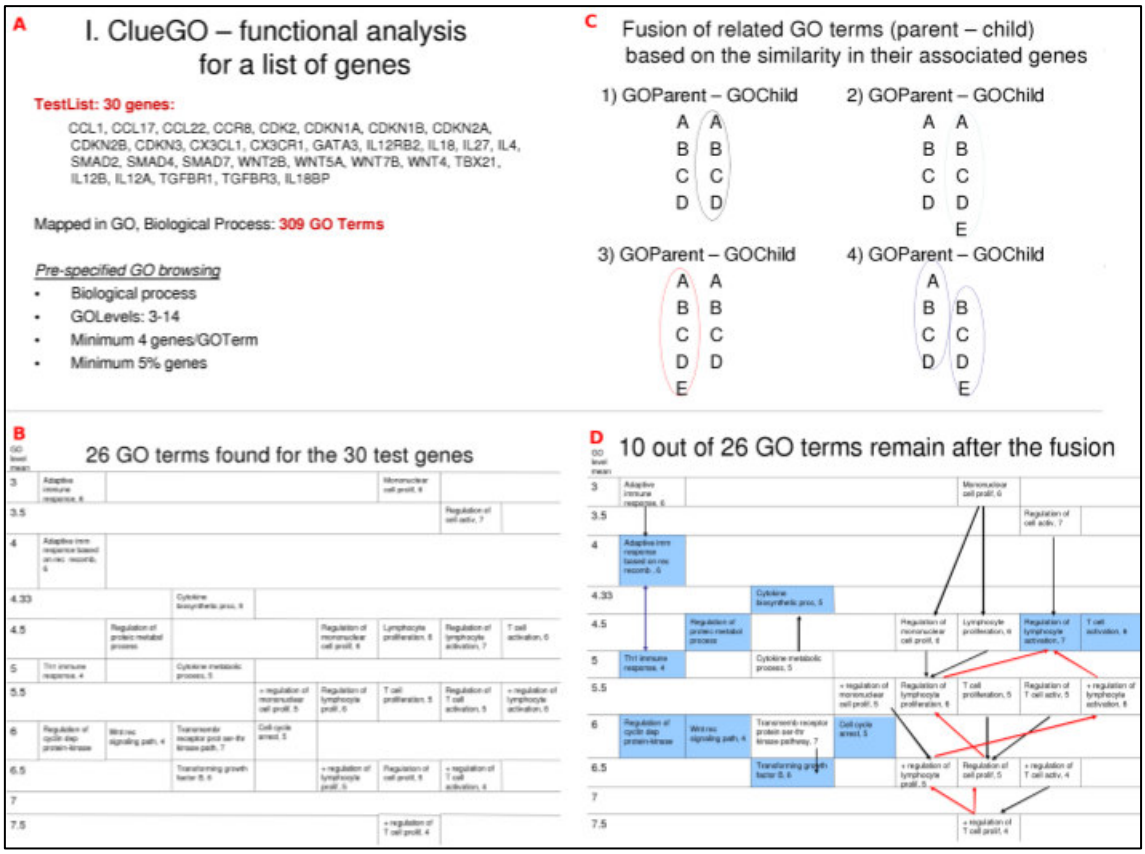


Figure 3.2: GO term fusion example. **A** - The test gene list and the total number of corresponding GO terms. Selection criteria applied in ClueGO. **B** - GO terms selected under the specified selection criteria. **C** - Fusion criteria. It applies to GO terms in parent-child relation, with similar associated genes. The most representative parent or child term is preserved. The terms are organized based on their mean level specificity (up: general terms with low specificity, down: specific terms). **D** - GO terms maintained after fusion (colored in blue). Colored arrows show the fusion criteria used (concordance with step C).

ClueGo[210] is a free plug-in for the freeware Cytoscape[211], developed to easily annotate a list of genes backed by several distinct ontologies. In our case, we were particularly interested in the "GO Term Fusion" feature. This option takes advantage of the stratified nature of GO and uses it to remove redundancy through the fusion of related GO terms that have similar genes, namely: the terms in parent-child relationship that share similar genes (identical, or with one gene difference) are assessed, and the most representative parent or child term is retained (Figure 3.2). Praxis show us that GO terms fusion is sufficient to account for the hierarchical noise.

3.4. CURATION

The literature mining for DrugAge, which is focus entirely on lifespan-extending compounds, came from three sources: freely submitted data by the scientific community and mining of pre-existing aging-related databases and PubMed.

Among the pre-existing databases that were thoroughly analyzed were: AgeFactDB; Geroprotectors.org; and the Aging Genes and Interventions Database.

We adopted the following search term in the manual mining of PubMed literature: *"increases lifespan" OR "lifespan increase" OR "lifespan extension" OR "prolongs lifespan" OR "antiaging agent" OR "extends lifespan"*; and restricted the query to literature published before the year 2016. All literature mined was manually curated and subjected to quality-control. Exclusion of research papers or some of their assays (partial exclusion) were based on the following criteria:

- assays without a control group or statistical analysis of the results (exceptions were left to the curator's (the author of this thesis) discretion in order to include some research papers that contained information deemed essential for DrugAge, e.g. a large-scale assay;
- assays that used disease or short-lived strains (relative to the reference strain);
- lifespan assays based on abnormal diets, e.g. high-fat diets;
- experiments in which animals were kept in non-standard environmental conditions, e.g. high-temperature;
- experiments that used paper-specific mutant strains, e.g. strains resulting from the knockout of one very particular gene;
- research articles that were not available in the English language;
- additionally, in the case of data submitted by the community, only data originated from research papers indexed by PubMed were considered.

In the end, data belonging to 325 distinct research articles passed the quality-control stage and were therefore included. Efforts were also made to include data that contradicted the purported lifespan-extending results already included in DrugAge, with the aim of reflecting the potential controversy associated with each compound (although we do not claim an almost exhaustive compilation of contrary evidence, as this is not easily defined by search terms).

DrugAge offers unprecedented scope, comprising 1316 lifespan assays using 418 distinct lifespan-extending compounds on 27 unique model organisms, including 71 individual strains (Figure 3.3). When used in reference to DrugAge data the terms "drug" or "compound" (as we will be doing in the next sections) are slightly imprecise. DrugAge includes well-defined chemical compounds such

as resveratrol, but also complex substances like apple flesh. The reader is advised to keep such consideration in mind.

3.5. WEB INTERFACE

DrugAge is freely available at http://genomics.senescence.info/drugs/index.php. From the DrugAge home page, some basic search functionality allows the main dataset to be quickly accessed and filtered according to keywords e.g. the experimental drug or organism.

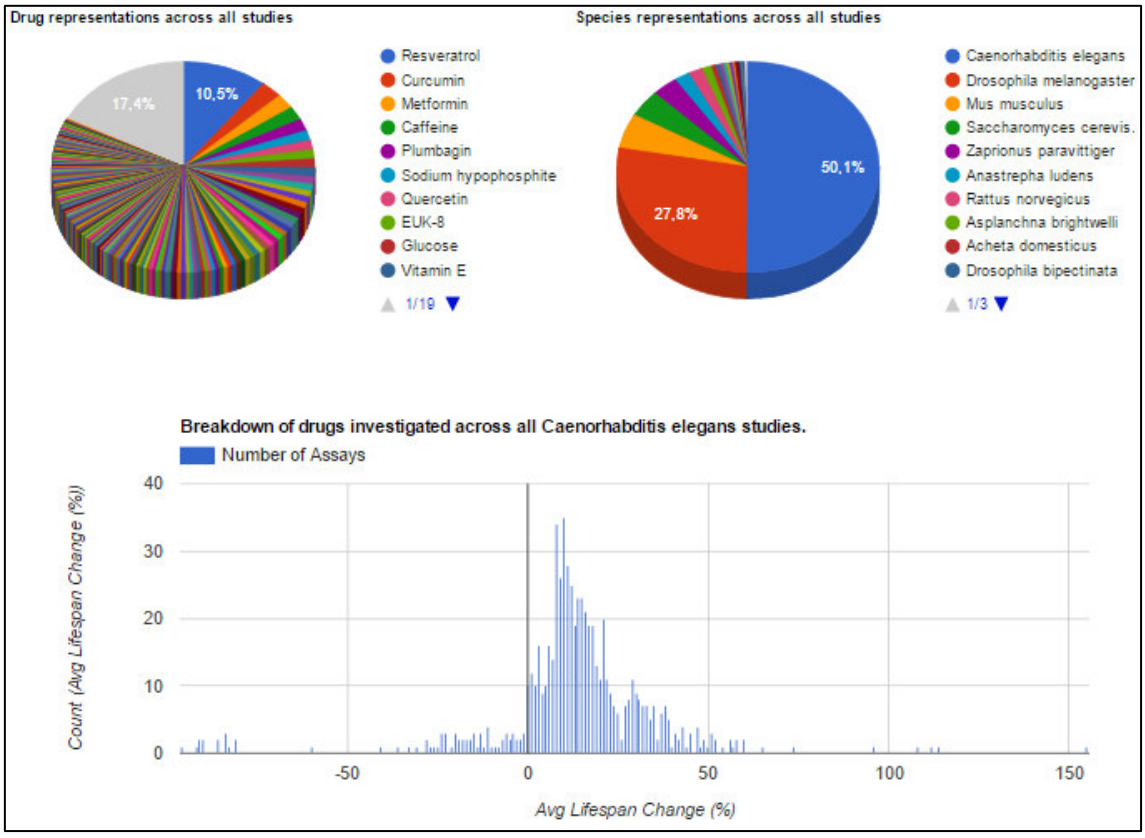


Figure 3.3: Partial screenshot of the "DrugAge Summary" tab (full page available at http://genomics.senescence.info/drugs/visualisation.php).

The user-friendly data browsing interface provides the basic means to search across all DrugAge fields and to retrieve drugs based on a range of lifespan effect values. Besides the key experimental parameters recorded for each assay, each entry is properly referenced with links to the original PubMed record and to any associated GenAge genes that the drug is known to interact with (Figure 3.4), and a considerable amount of entries even include crucial additional information in the form of curator's (this thesis' author) annotations (under the "Notes" field).

Figure 3.4: Screenshot of a DrugAge browsing experience.

The hyperlinks on the data browser, as depicted in Figure 3.4, also accept specific drugs and organisms to be retrieved separately and in more depth, permitting a more detailed investigation of the results associated with a specific drug or organism. These pages allow the data to be visualized and explored on a per drug or organism basis, such as how species impact on lifespan effect for a given drug.

The contribution of the current thesis' author in regards to the development of the web interface was limited to providing feedback as an alpha- and beta-tester (he played no part in the programming of this resource). This subsection is present in this dissertation only for the sake of completeness, as we feel that otherwise the full relevance of DrugAge as a scientific resource (and therefore the practical value of our contributions) would not be explicitly demonstrated. We invite the reader to experience DrugAge personally to discover several unmentioned capabilities, which include integration of DGIdb and STITCH drug targets.

We must quickly mention that in the web interface and for the rest of this section, we will loosely use the field/term "average lifespan" to encompass average and median lifespan changes.

3.6. STATISTICAL ANALYSES

Under this subsection, we will present exploratory and statistical analyses of the wealth of information compiled on DrugAge. When these are based on only a subset of DrugAge, for example,

a limited number of species, it is because we excluded the rest of strata due to insufficient sample size.

Our first inquiry was to observe the shape of the distribution of average and maximum lifespan changes across distinct model organisms. We found that the magnitude of average life expectancy changes per assay is highly species-dependent, and it appears to be inversely correlated with organism complexity (Figure 3.5).

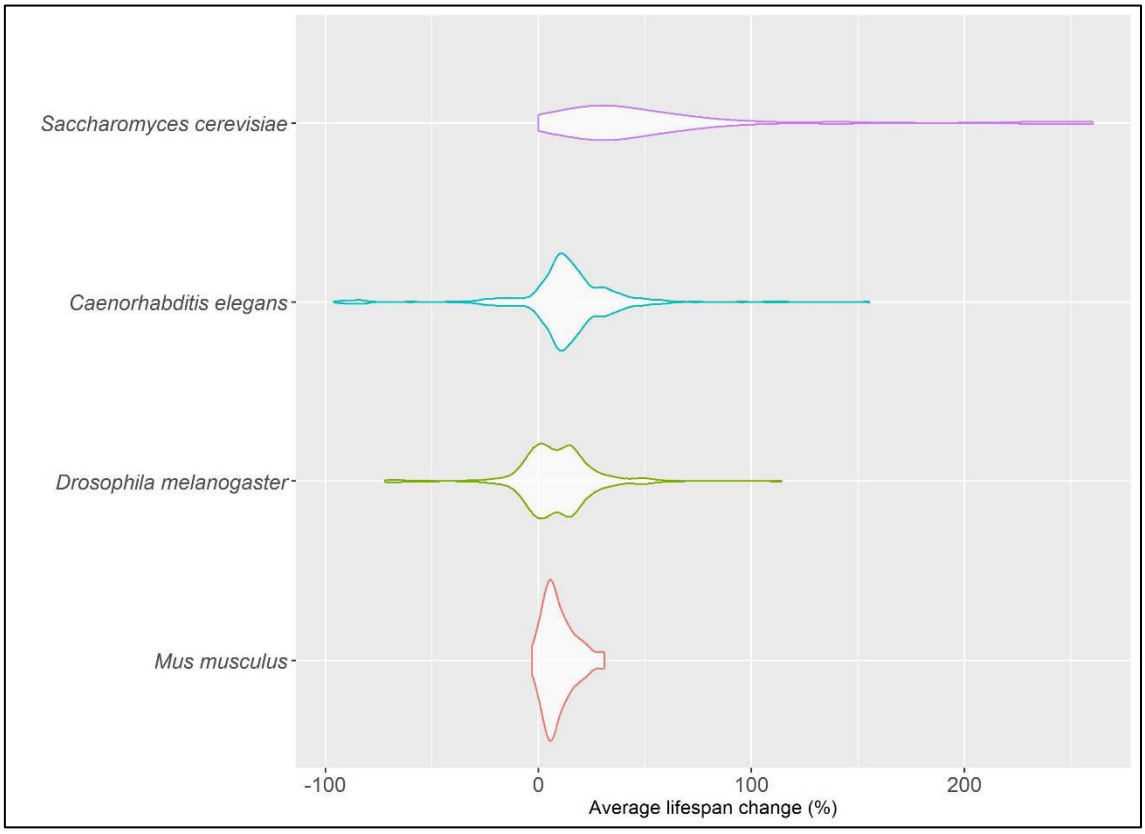


Figure 3.5: Violin plot based on the average lifespan changes reported on 42 *S. cerevisiae*, 653 *C. elegans*, 357 *D. melanogaster* and 63 *M. musculus* lifespan assays.

The extent of maximum lifespan changes are more modest and not as dissimilar across species (Figure 3.6). We could only make use of data referent to three species because this metric is widely underreported.

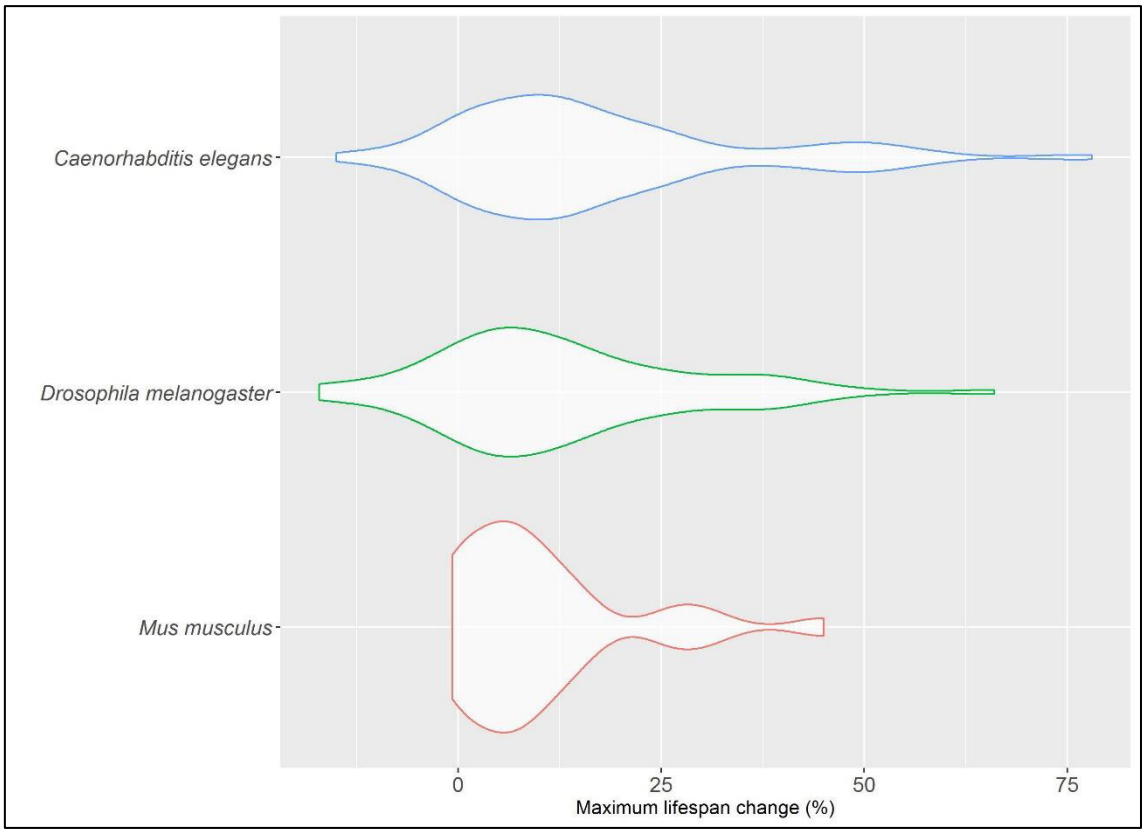


Figure 3.6: Violin plot of the maximum lifespan changes obtained in 140 *C. elegans*, 74 *D. melanogaster*, and 27 *M. musculus* lifespan assays

In the entire DrugAge, only 326 assays simultaneously reported average and maximum lifespan changes. We noted an extremely significant Pearson's product-moment correlation (two-tailed test done in R) of around 0.85 (95 percent confidence interval ranged from 0.82 to 0.88), between these metrics (Figure 3.7).

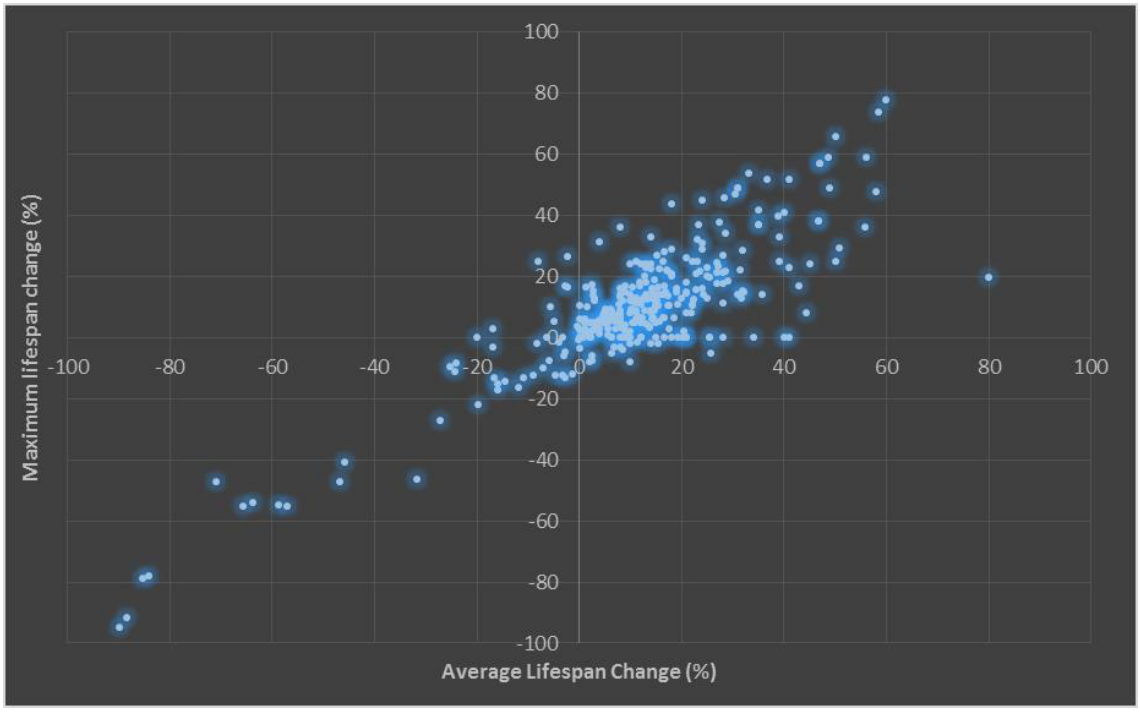


Figure 3.7: Scatter plot of average lifespan change (horizontal axis) and maximum lifespan change (vertical axis) from assays that measured both. Linear correlation p-value < 2.2e-16 (the assumed alternative hypothesis is that true correlation is not equal to 0).

We also focused on assays for males and females, more concretely we used two different statistical analysis to determine if gender is a key factor influencing the magnitude of lifespan changes in model organisms.

First, we noticed that two of the species, *Mus musculus,* and *Drosophila melanogaster*, have a sufficient amount of assays reporting gender-specific average lifespan effects to allow to test if exists sexual dimorphism in the general response to a pharmaceutical intervention aimed to extend average life expectancy. Using the "dgof" R package, we obtained a p-value of the two-sided Kolmogorov-Smirnov test, with the null hypothesis being that the data of both genders originated from the same distribution, of 0.9948 and 0.4794, for mice and fly, respectively. Therefore, we conclude that sexual dimorphism does not seem to be a significant factor regarding pharmaceutical interventions to extend lifespan in *Mus musculus* or *Drosophila melanogaster* (Figure 3.8), as the distributions of the magnitude of average life expectancy changes are not statically different across genders of the two species in consideration.

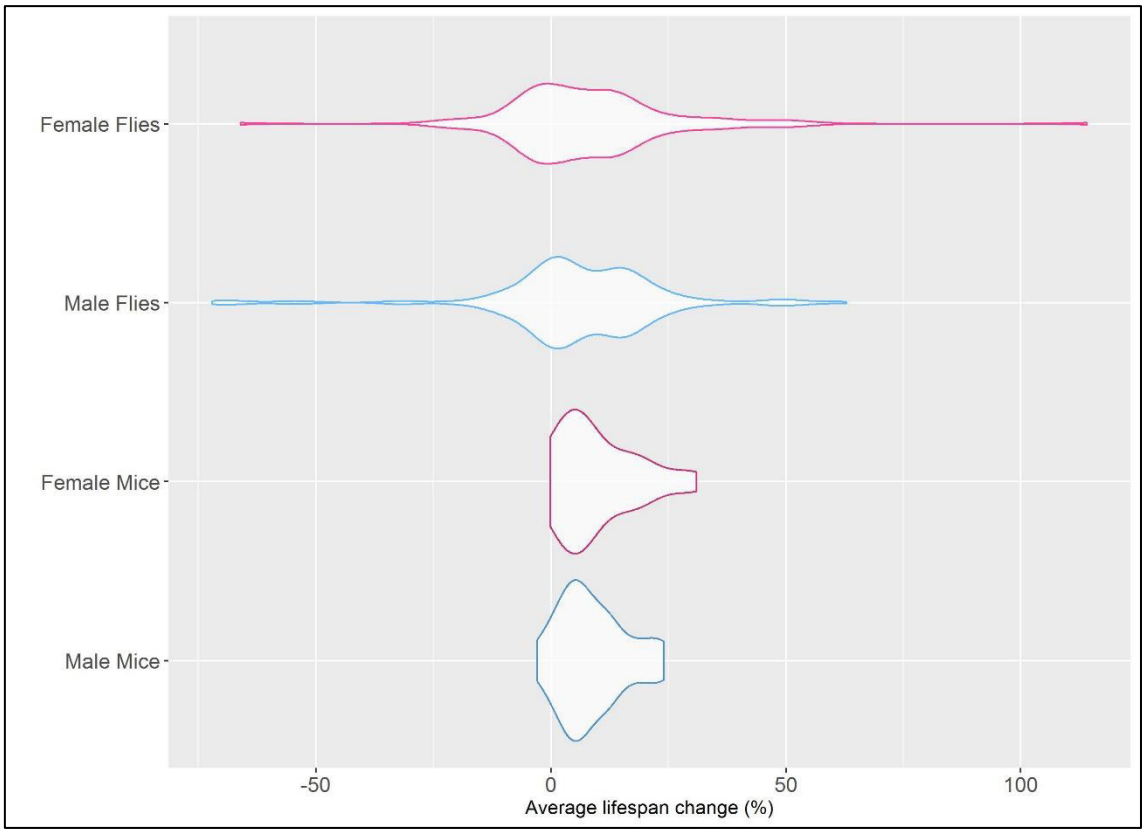


Figure 3.8: Violin plot of the average lifespan change in lifespan assays that displays gender-specific statistics for mice (p-value>0.99) and flies (p-value>0.47). It aggregates information from 133 and 164 lifespan assays using female and male flies, respectively; and from 25 and 30 lifespan assays in female and male mice.

Secondly, we also undertook an alternative analysis of sexual dimorphism by considering data from all species in DrugAge and manually curating it so that only lifespan assays that were conducted in the same experimental conditions and measured gender-specific endpoints were contemplated. This method as the advantage of generating enough data for also assessing the role of gender on the magnitude of maximum lifespan changes. Through the application of the same workflow as before on curated data obtained from 11 species of model organisms, we obtained statistical significant (p-value$< 2.2 \times 10^{-16}$) Pearson's product-moment correlations of approximately 0.88 (95 percent confidence interval ranging from 0.83 to 0.91) and 0.90 (95 percent confidence interval from 0.84 to 0.94) for the magnitude of average and maximum lifespan changes across genders, respectively (Figure 3.9).

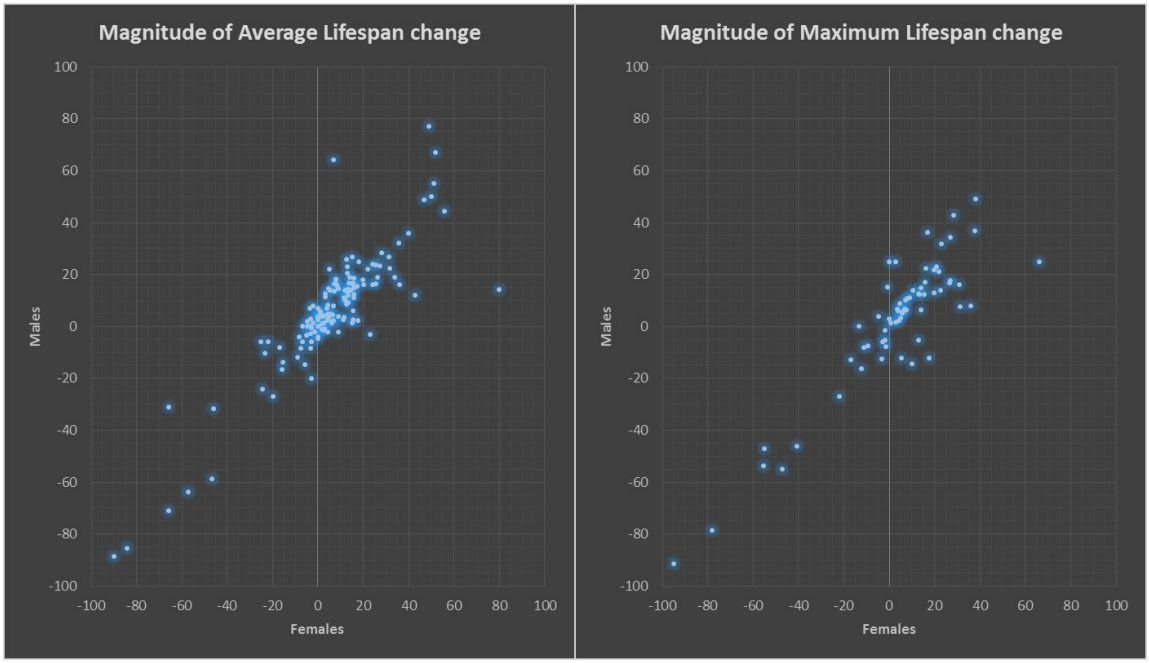


Figure 3.9: Side-by-side display of average (n=140) and maximum (n=65) lifespan changes among gender-paired lifespan assays.

3.7. ENRICHMENT FUNCTIONAL ANALYSES OF DRUGAGE

In order to discover enriched Gene Ontology terms in targets of compounds in DrugAge, we used DGIdb 3 to obtain a list of gene interacting partners for all drugs in DrugAge. To optimize the number of hits, we endeavored in manual curation, to ensure that the compound name field in DrugAge matched PubChem synonyms. From the set of 418 compounds/substances present in DrugAge, 90 were found to have a corresponding record in DGIdb, which translated into 411 distinct interacting genes composing the genes list that is going to be our input for the statistical functional enrichment analysis.

Statistical functional enrichment analyses require defining a background dataset, which was taken as the 3090 distinct Ensembl Gene Ids found in the DGIdb.

The DrugAge interactors list was passed as input to Cytoscape (version 3.3.0) plug-in ClueGO (version 2.2.5) to compute functionally enriched GO Terms. ClueGO used information from all four available GO ontologies (08.04.2016) and InterPro (10.04.2016) and matched 3050 entries from our background list. We are aware that ClueGO offers other ontologies, including Reactome[212] and WikiPathways[213], nonetheless, internal experimentation taught us that the extra annotations obtained by selecting these databases (even when we made use of ClueGO's clustering based on GO functional groups) is not enough to compensate for the loss of interpretability (the other

databases do not have a hierarchical schema, and therefore the GO term fusion algorithm cannot be applied).

The parameters chosen for the ClueGO analysis were "Use GO Term Fusion" and the entire "GO Tree Interval". We imposed no threshold on the "GO Term/Pathway Selection". There were 3044 annotated DrugAge interacting genes. The statistical enrichment was evaluated through a right-sided mid-P-value hypergeometric test with Bonferroni step-down (Holm-Bonferroni method), corrected against the reference set which was the list of DGIdb genes. We decided on this background so that our enrichment results reflect DrugAge and not DGIdb, in other words, we eliminated the bias that results from having to use DGIdb as an intermediary step in our workflow. At the significance threshold chosen (corrected mid-p-value < 0.01), we detected 182 enriched GO terms (Table 3.1) out of a universe of 4830 fused GO terms.

Table 3.1: Top-15 enriched GO terms and their statistical significance (corrected mid-p-value).

GO Term	Term statistical significance
regulation of blood circulation	1.6E-14
glutathione derivative biosynthetic process	3.4E-14
blood circulation	1.7E-13
regulation of system process	2.3E-13
calcium channel complex	3E-13
calcium ion transmembrane transport	1.3E-12
heart contraction	4.1E-12
glutathione metabolic process	4.7E-12
voltage-gated calcium channel complex	5.4E-12
calcium ion transport	1E-11
regulation of heart contraction	3.7E-11
inorganic ion transmembrane transport	4.7E-11
response to xenobiotic stimulus	6.6E-11
cellular detoxification	6.8E-11
xenobiotic metabolic process	7.7E-11

3.8. COMPARATIVE ANALYSIS *VERSUS* GENAGE

We also wanted to explore the overlap between DrugAge interacting genes and genes previously associated with aging. The rationale behind such analysis is to determine if the known genetic and pharmaceutical anti-aging interventions in model organisms, represented by GenAge and DrugAge, respectively, are acting upon the same genetic targets.

The overlap analysis was based on the 1124 human orthologues of genes that extend lifespan in model organisms according to the GenAge database.

Of the 1124 genes, 287 (25.5%) were known to interact with drugs in DGIdb. Of these, 65 (29.3%) were also DrugAge interacting genes (Figure 3.10).

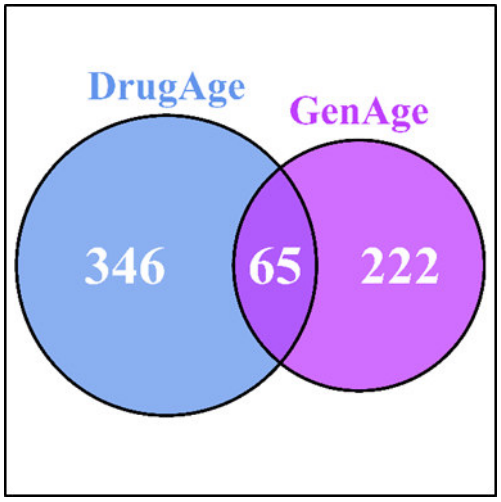


Figure 3.10: Venn diagram displaying the number of unique and shared DGIdb genes between DrugAge-interacting genes and human orthologues of GenAge lifespan-extending genes. The overlap has 27 more genes than expected by chance.

The statistical overlap between human genes mapped to DrugAge and in GenAge was evaluated by Blaker's Exact Test using the "exact2x2" R package[151]. The background chosen was the set of 3090 DGIdb genes. Genes counted as "DrugAge" in Figure 3.10 are all 411 DrugAge-interacting genes in DGIdb (originated from the previous subsection). We identified a statistically significant (p-value$= 6.024 \times 10^{-6}$) degree of overlap between the genes arising from the two databases.

3.9. APPLIED MACHINE LEARNING

This subsection is dedicated to the training and optimization of random forests for the classification of new compounds according to their estimated likelihood of having anti-aging effects. The output of our algorithm is going to be class probabilities, and we intend to make use of these as a proxy for anti-aging potential.

A toy example showcasing the usefulness of class probabilities can go as follow: if "compound A" is classified as 0.65 likely of being anti-aging (and therefore 0.35 likely of not having anti-aging effects) and "compound B" is assigned a 0.90 probability of being an anti-aging substance (and hence 0.10 probability of not causing anti-aging outcomes), everything being equal, compound B

should be prioritized over compound A for lifespan experiments in model organisms (especially *Caenorhabditis elegans*, for reasons that will become clear later in this subsection).

To train the classifier, we need a dataset of labeled samples and features that vary among these samples. When we were curating the literature for DrugAge, we noticed that the majority of lifespan assays belonged to *C. elegans* and concomitantly started to search the literature for drugs that failed to extend lifespan (this includes chemical compounds that have no effect on or decrease lifespan) in worms. Our decision to compile a dataset that is based specifically in *C. elegans* is grounded on two considerations:

- *C. elegans* has by far the largest representation on DrugAge (Figure 3.3) and it is also the only species that we found a considerable amount of compounds that failed to prolong lifespan;
- had we considered the entire DrugAge data; we would have to face the challenge of how to select and organize biological features across species, for example, information and chemical reactions in a biological pathway might flow differently, contrarily or even be inexistent in some model organisms.

The raw dataset mined from the literature contained 283 anti-aging chemicals and 1272 that did not cause lifespan prolongation. The list of compounds was then sent to an external collaborator that curated the dataset for redundancy and annotated it for chemical features. The next paragraph enumerates the steps taken in the chemoinformatics workflow only succinctly as it is not the area of expertise of this thesis' author.

The simplified molecular-input line-entry system (SMILES) codes of the drugs were extracted (and checked for duplicates) from PubChem[214], [215] and ChemSpider[216], both of which are free online chemical repositories. Then the Molecular Operating Environment[217] (MOE), version 2013v0.8, was used to desalt the chemicals and minimize the structure for the posterior calculation of molecular descriptors. Another check for duplicates and errors was done using the desalted minimized SMILES outputted by MOE. Chemical descriptors for the compounds were calculated from two software MOE and Advanced Chemical Development and any descriptor with zero variance or with more than 98% constant values were removed. The chemical curation was then considered complete. A dataset comprised of 229 DrugAge drugs that prolong lifespan in *C. elegans* (from here forward denominated as "positive set") and 1163 compounds that didn't reveal anti-aging effects in worms (henceforth referred to as "negative set") for which chemical features were found was returned to us.

In 11 of the 268 chemical descriptors, some drugs (never more than twelve per chemical descriptor) had missing values (there were 98 such values in total). Instead of discarding these drugs, we decided to proceed with the imputation of these values. To this end, we selected the "missForest" R package (version 1.4)[218]. MissForest is a state of the art truly nonparametric imputation method that makes no assumptions regarding the data, requires no tuning, avoids the need to holdout data for an imputation testing set, is computationally efficient and is capable of handling high-order non-linear interactions among variables even in high-dimensional datasets. All these impressive characteristics are accomplished in a quite uncomplicated way, as the method consists simply in a well-defined algorithm that iteratively repeats two steps: it fits a random forest to the observable data, and then it predicts the missing values. The performance of the imputation is assessed, for our continuous chemical descriptors, by comparing the absolute difference between true imputation error and out-of-bag imputation error estimate in all simulation runs (for greater detail, please see the original paper[218]).

We took care to remove the class membership variable (a categorical feature that reveals whether a given drug belongs to the positive or the negative set) of our dataset of drugs from the matrix that missForest used as input because we do not want to create an optimistically biased classifier due to some information regarding the true class label of the drugs in the testing sets being somehow present in the imputed value. When missForest reached the stoppage criterion, the normalized mean root square error[219] was less than 0.069.

The biological features of our dataset were generated in the following way:

1. We acquired genes that interact with the drugs in the dataset by mining all DGIdb drug-gene interactions and the drug-protein interactions in the STITCH 4.0 database[220] with a confidence score larger than 0.450 (which the authors of STITCH define as confidence of medium strength) and limited to only the top-100 protein-drug interactions with the highest STITCH confidence score for each compound. This threshold was set in order to restrict the influence of the 479 proteins, out of the possible 11352, that were judged outliers for having interactions with more than 100 distinct compounds;
2. We annotated the mined biological drug-interacting entities to GO terms using ClueGO, with the same parameters that were used in the enrichment analysis of DrugAge except that the background was the built-in *Homo sapiens* reference set. We decided on using human-centered annotations because not only is the GO of *C. elegans* very poor in comparison, but also due to our ultimate objective being the prediction of drugs anti-aging

potential in human beings. By mapping drugs to genes/proteins to GO terms, we are indirectly bridging drugs and their respective GO terms. Using the GO Term Fusion allows diminishing the dimensionality of our final dataset as only non-redundant GO terms are going to be considered biological features.

The 10757 outputted GO terms were interpreted as categorical biological features. We were unable to find biological features for all the drugs present in our dataset. Drugs for which no biological descriptor could be found were eliminated, with our final dataset consisting of 190 and 783 compounds belonging to the positive and negative set, respectively.

Our dataset is imbalanced, but we expect that the actual probability distribution of lifespan-extending compounds in the known chemical universe to be even more imbalanced. Supporting evidence is found in the largest high-throughput assay for lifespan-extending drugs to date (the best proxy for the underlying probability distribution currently available). The authors screened 88000 compounds in *C. elegans* and found that only 115 extended lifespan[221]. This rarity implies that the magnitude of class imbalance properly approximates the real class membership probability distribution.

The algorithm optimization and training were carried using the "mlr" R package (developer version 2.9), which is a general machine learning interface that works as a wrapper for a plethora of learning algorithms available in distinct R packages. We are going to train random forests that the mlr package imports from the "ranger" R package[222].

We optimized the random forest algorithm for four versions of our dataset: one featuring only chemical features; another featuring only biological features; a third version featuring both, chemical and biological, types of features; and a fourth version using only chemical features but applied to the (larger) dataset of drugs that were present before we discarded the compounds for which no biological features could be found.

The parameters of the random forest algorithm that we tuned were the number of trees in the forest and the number of variables to possibly split at in each node (routinely called "mtry" in the machine learning field). It is ordinary to tune the number of trees when optimizing a random forest algorithm, but tuning the mtry parameter is rare, as a case in point, just one of the articles considered in a recent review of the application of the random forest algorithm in the life sciences proceeded to do so[223]. We followed the formula to obtain possible parameter values from the cited research article[224], which means that three values for the mtry parameter are going to be evaluated: the square root of the number of features in the dataset and half and double this value.

In setting the number of trees parameter in a random forest there are no consensus or rules-of-thumb to guide us, therefore, after reading the literature that applied random forests to similar classification tasks, we decided to cover a relatively wide range of typical values: $\{100,300,500,700,900\}$.

We used a nested cross-validation with grid search for tuning. Grid search consists in evaluating the performance of every possible configuration of parameters. The nested cross-validation had ten folds in each inner loop and ten folds in the outer loop.

Another crucial aspect in any optimization is deciding on the metric to be optimized. Because there is no metric that is superior to all the others; we opted to optimize the algorithms separately for three distinct metrics: area under the curve (AUC); F1 measure; and the geometric mean of recall and specificity (Gmean). The increased computational cost is more than compensated by having three metrics, each capturing distinct components of the performance, guiding our choice of parameters.

Recapitulating, each of the random forests, applied to the four different datasets, for all the combinations of values in the optimization grid, was also optimized for each of the three performance metrics: AUC, F1 and Gmean.

We compared the performance of each tested configuration of the random forest algorithm by calculating the median rank of the ranked median inner loop out-of-sample performances across all 10-folds of the outer loop of the nested cross-validation.

The summary of the performances obtained for the dataset with only the chemical features and for the dataset in which both types of features were used uncontroversially indicate the optimal values for the parameters of the random forest algorithm, as there is an agreement for all the different metrics utilized as optimization endpoints (Table 3.2). Coincidently, for both datasets, the combination of the largest values for each parameter is the one displaying the most robust performance.

Table 3.2: Performance summary of optimizations run on the dataset with both types of features and on the one that used just chemical descriptors. The numbers in a colored background are the relative median rank performances, lower values (greener background) are better. Values in colorless background correspond to values taken for the number of trees and mtry parameters. In **bold and colored blue** is the performance of the parameters configuration that we considered best, for each dataset.

Biological and chemical features				Just the chemical features			
AUC	*52*	*105*	*210*	AUC	*8*	*16*	*33*
100	11,5	12	12,5	*100*	9	9	7
300	11,5	8	7	*300*	11,5	8,5	6
500	9	5,5	6	*500*	6,5	8	8
700	8	4,5	6,5	*700*	11,5	8	5,5
900	7	5,5	**4,5**	*900*	7	8,5	**5,5**
Gmean	*52*	*105*	*210*	Gmean	*8*	*16*	*33*
100	9	7	2,5	*100*	9	4	5
300	7,5	4	2	*300*	4,5	7	3
500	9	5	2	*500*	4	5	7,5
700	9	5	3	*700*	8	3	4
900	8	5,5	**2**	*900*	4,5	8,5	**3**
F1	*52*	*105*	*210*	F1	*8*	*16*	*33*
100	7,5	3,5	3	*100*	8,5	4,5	5,5
300	8,5	4	1	*300*	6,5	7,5	3,5
500	7	6	3,5	*500*	4	4,5	8
700	7	5,5	4	*700*	9	2,5	5
900	7	6	**3**	*900*	4	7,5	**2,5**

The situation is not as clear for the dataset that used all drugs (before deleting drugs for which no biological features could be obtained) and for the dataset that used only features of the biological type (Table 3.3). So much so, that we decided on testing a pair of combinations of values for each dataset.

Table 3.3: Performance summary of optimizations run on the dataset version with both types of features and on the one that used just chemical descriptors. The numbers in a colored background are the relative median rank performances, lower values (greener background) are better. Values in colorless background correspond to values taken for the number of trees and mtry parameters. In **bold and colored blue** are the performances of the parameters configurations that we considered superior, for each dataset.

Just chemical features using drugs available before deletion				Just the biological features			
AUC	*8*	*16*	*33*	AUC	*52*	*104*	*207*
100	11,5	**4**	6	*100*	6,5	8	11,5
300	14	6,5	5	*300*	**3**	8	11
500	8	9,5	9	*500*	5,5	6,5	11,5
700	10	**5**	7	*700*	6	7	11,5
900	12	5,5	5,5	*900*	5,5	7	**12**
Gmean	*8*	*16*	*33*	Gmean	*52*	*104*	*207*
100	9	**2,5**	4,5	*100*	8	3,5	3,5
300	5,5	8,5	2	*300*	**8,5**	5	4
500	3,5	5	8,5	*500*	7,5	5	3
700	8,5	**2**	4,5	*700*	8	4	3
900	6	9	2	*900*	7,5	4	**2,5**
F1	*8*	*16*	*33*	F1	*52*	*104*	*207*
100	9	**3**	6	*100*	8	6	2
300	5,5	9	3	*300*	**8**	4	2,5
500	2	5,5	8	*500*	7,5	4	3
700	8,5	**2,5**	6	*700*	7	4	3
900	5,5	9	2,5	*900*	7,5	3	**2,5**

The performances of the optimal random forests for each dataset version were contrasted through the median AUC in the outer 10-fold cross validation loop (Table 3.4). We opted for optimizing the AUC because it is the more commonly reported performance metric in the literature[132].

Table 3.4: Summary of the performance of random forests optimized for each version of our dataset. In **bolded blue** is our best model.

Dataset	**Number of trees**	**mtry**	**median AUC**
Biological and chemical features	**900**	**210**	**0.8**
Chemical features before deletion	100	16	0.781
	700	16	0.774
Just the biological features	300	52	0.716
	900	207	0.707
Just the chemical features	900	33	0.675

3.10. DISCUSSION

The foremost contribution of this section (and thesis) is the creation of DrugAge, a database of compounds and drugs experimentally proven to extend average and/or maximum lifespan in model organisms. Only compounds, drugs, and substances experimentally shown to extend lifespan in a statistically significant manner in at least one experiment were included. Conflicting and negative results were then added to provide a balanced literature survey of the effects of each compound or drug.

DrugAge is a freely-available manually curated database offering several tools in an easy browsing experience. Its sheer scope and meticulous curation are unmatched, establishing it as a crucial resource for the scientific community, which was in great need of a database of anti-aging drugs.

The practical impact of DrugAge data was demonstrated in the rest of this section, with all the analyses and methodologies using DrugAge as an information source.

Common sense in the gerontology field tells us that the biological complexity of an animal is inversely proportional to the relative malleability of its aging phenotype. Actual evidence supporting this assumption, in the context of anti-aging pharmaceutical interventions, is finally revealed by us (Figure 3.5). This discrepancy is particularly evident when one compares yeast and mice, the two extremes of biological complexity in the considered species.

Likewise considered common knowledge is the fact that the magnitude of maximum lifespan changes is generally more modest than changes in average lifespan. Once again with found scientific evidence corroborating empirical knowledge (Figure 3.6). Of unclear importance, we also noticed that the distribution of maximum lifespan changes is not as dissimilar across species as the distribution of the magnitudes of average/median lifespan effects.

We shall now take the opportunity to bring attention to a reality: the maximum lifespan effects are severely under-reported. There is hardly any excuse for this state of affairs since measuring maximum lifespan is free (in a financial sense, as no additional investment has to be made[225]) and a panoply of freeware that is already going to be used in the statistical testing of median/average lifespan changes has tools dedicated specifically to this end.

Based on 326 lifespan assays simultaneously reported average and maximum lifespan changes and we calculated a highly significant strong linear correlation of around 0.85 (Figure 3.7). The fact that, in general, pharmaceutical interventions that prolong average life expectancy extend maximum lifespan in a proportional manner is fascinating.

Let us recall that some anti-aging intervention, such as physical exercise, extend lifespan but have no repercussions on maximum lifespan[226]. This anti-aging profile is attributed (empirically) to a reduced rate of incidence of certain mid-life diseases. On the contrary, interventions that increase maximum lifespan are ascribed to a slowing down of the "fundamental process of aging".

Our results indicate that: or drugs exert their anti-aging effects by simultaneously reducing the rate of developing mid-life diseases and targeting the fundamental aging process, or that such conceptualizations are irrelevant/artificial in the generality of anti-aging interventions of the pharmaceutical type. A practical implication of the high linear correlation between average and maximum lifespan changes is that for a high-throughput screen searching for anti-aging chemical compounds, either one of the lifespans can be the sole endpoint used without increasing the chance of false negatives.

We proceeded to investigate the possible causal influence of gender on the magnitude of the anti-aging effects of pharmaceutical interventions. We did so by using two distinct methodologies.

The first species-centered analysis is especially relevant for biologists that work with *M. musculus* and *D. melanogaster*. We show that for each of these two species (the only species that we deemed to contain enough data points to permit meaningful conclusions) the distribution of average/median lifespan changes is not statistically different across genders (Figure 3.8).

The second analysis aims for a more general view and considers lifespan results from all DrugAge species as long as they have a paired assay conducted in the opposite gender (in the exactly same experimental conditions). In other words, we sacrifice the accuracy gained by controlling for species, in the hopes of increasing our dataset size (and therefore inference power). We were successful in our endeavor, as data originated from 11 different species of model organisms formed a larger dataset that even allowed us to test the potential role of gender in maximum lifespan changes.

We obtained exceptionally significant linear correlations of approximately 0.88 and 0.90, for the magnitude of average and maximum lifespan changes across genders, respectively (Figure 3.9).

The agreement between the results of the two analyses that discredits the existence of sexual dimorphism in the response to anti-aging drugs must not lead to hastily conclude that the same applies to humans. It is crucial to remember that humans are unique in this aspect, with one of the genders (females) known to have a ubiquitous survival advantage[227].

This set of holistic analyses of anti-aging interventions are a new contribution to our field of study, and they would not have been possible without DrugAge.

The functional statistical enrichment analysis of DrugAge drugs actually consisted in gauging the functional enrichment of GO terms of DGIdb genes that interact with DrugAge drugs relative to all the genes in the DGIdb database. The enrichment was carried in the ClueGO software because it deals with possible interpretability issues by fusing redundant GO terms. At the chosen significance threshold (multiple hypotheses corrected mid-p-value < 0.01), we detected 182 enriched GO terms from a total of 4830 fused GO terms. As a reference to the reader, the GO term "aging" is associated with a corrected mid-p-value of 0.0056.

Exhaustive inspection of the set of statistically enriched terms allowed us to extract biologically meaningful patterns (Table 3.5).

Table 3.5: Conceptual clustering of enriched GO terms and their statistical significance. *p-value is actually multiple hypotheses corrected mid-p-value.

Categories	Sub-categories	GO term	p-value
Hormesis	Positive ROS regulation	positive regulation of reactive oxygen species biosynthetic process	5.10E-04
		positive regulation of reactive oxygen species metabolic process	2.30E-03
	General ROS terms	reactive oxygen species biosynthetic process	5.90E-05
		regulation of reactive oxygen species biosynthetic process	2.50E-04
		reactive oxygen species metabolic process	2.20E-03
		regulation of reactive oxygen species metabolic process	4.80E-03
	Xenobiotics	response to xenobiotic stimulus	6.60E-11
		xenobiotic metabolic process	7.70E-11
Glutathione		glutathione derivative biosynthetic process	3.40E-14
		glutathione metabolic process	4.70E-12
		glutathione peroxidase activity	2.90E-08
		glutathione binding	6.30E-04
Histone deacetylation		histone deacetylase complex	2.20E-03
		NAD-dependent histone deacetylase activity	5.30E-03
		histone H3 deacetylation	5.30E-03
Sulfur compoudns		sulfur compound metabolic process	7.90E-06
		sulfur compound biosynthetic process	2.10E-04

The hormesis paradigm is severely enriched in the GO terms resulting from the functional enrichment of DrugAge, with both xenobiotic and reactive oxygen species related terms. More closely it can be observed that all the reactive oxygen species terms significantly enriched belong to the positive regulation type, that is, their increase/activation. In other words, all of the opposite terms implying a decrease in free radicals abundance, remain insignificant (although being present in the universe of possible fused GO terms). Such observation might suggest that for lifespan extension, reactive oxygen species should be analyzed from the hormesis paradigm[94], which is contrary to the classical view of reactive oxygen species as purely nocive[86].

In line with previous works showing that glutathione levels decrease in aged humans[228] and manipulations that increase them extend lifespan in flies[229], glutathione-associated terms are another strongly enriched process.

Two other processes are supported by several enriched terms and are known to be implicated in aging: histone deacetylation (namely histone H3 deacetylation[230]) and sulfur compound[231] involving-processes.

To examine the possible overlap among genetic and pharmaceutical types of anti-aging interventions, we compared the set of human homologs of GenAge lifespan-prolonging genes present in DGIdb with DrugAge-interacting DGIdb genes (Figure 3.10). Blaker's test assigned a p-value of 6.024×10^{-6} for the statistical significance of the overlap between the two lists of genes. The results imply that there is a statistically significant number of anti-aging genes that were the common target of pharmaceutical and genetic manipulations. While highly statistically significant, the overlap is modest in relation to the total number of genes target by either type of interventions. We speculate that experts on one kind of interventions should benefit from seeking inspiration from colleagues specialized in the other type of manipulations. Interdisciplinarity holds untapped potential regarding candidate anti-aging therapies.

Setting ourselves for screening large libraries of chemical compounds in the search for candidate anti-aging chemicals, we trained and optimized the random forest algorithm in a dataset of drugs that are known to successfully or unsuccessfully prolong lifespan in *C. elegans*. We characterized each drug by using molecular descriptors and biological features (GO terms annotated in drug-interacting genes). The random forest trained on a dataset containing both, biological and chemical, types of features obtains a median AUC in a 10-fold cross validation of 0.8 (the highest that we were able to achieve) which is sufficient for our purposes (Table 3.4). This trained model is the one that we are deploying for predicting the class membership of compounds in large-scale drug screens for new candidate anti-aging compounds.

We must also comment that taking advantage of both kinds of features was preferable to using just one, even if fewer data samples are available. Such fact stresses the superiority of interdisciplinary research among chemo- and bioinformaticians.

We are currently finishing a manuscript focused on introducing the DrugAge database to the scientific community. The majority of the analyses examined in this section are also included to showcase the practical utility of DrugAge. The author of this dissertation is a co-first author in said

manuscript (the other co-first author is the person responsible for developing the DrugAge web interface).

CHAPTER 4: CONCLUSIONS AND FUTURES PERSPECTIVES

4.1. CONCLUSIONS

We have made the case for tackling the imminent social crisis, caused by an aged population, through the superior means of translational biogerontology. We demonstrated that pharmaceutical interventions are the preferred choice to undertake this endeavor and the only viable option with potentially real application in the near-term future. Although the theoretical advantages of this paradigm are evident, much is yet to be translated to praxis. Our contributions shorten this gap.

In chapter 2, we crossed information regarding life-extending genes with the drugs targeting them in an effort to use drug repurposing to delay the intrinsic mechanisms of aging. More concretely, we mapped the human homologs of lifespan-extending genes of model organisms, which are curated by GenAge, to the compounds of a meta-database of drug-gene interactions, DGIdb. Sixteen compounds, which interact with nineteen anti-longevity human genes, were determined statistically enriched. Two anti-aging drug combinations were then suggested after considering the enriched drugs clinical profile and potential redundancy in their targeted genes.

In Chapter 3, we developed DrugAge and applied the information wherein contained in several methodologies.

DrugAge is a new database of life-extending compounds and drugs in model organisms that accurately reflects the current knowledge of pharmacological manipulations of aging. Exploratory analyses of DrugAge data suggests that maximum and average lifespan changes are linearly correlated in a strong fashion and that gender does not affect them.

We gained some biological insights from the functional enrichment of DrugAge, namely that anti-aging drugs should be analyzed through the lens of the hormesis paradigm in detriment of free-radical damage theories.

Assuming that GenAge and DrugAge are the best proxies for the scientific knowledge regarding genetic and pharmaceutical anti-aging therapies in model organisms, respectively, we strove to understand the relation between these two categories of potential anti-aging treatments. The study of the possible overlap between human orthologues of GenAge lifespan-prolonging genes present in DGIdb and DrugAge-interacting DGIdb genes concluded that there is statistically significant overlap, nonetheless very modest in magnitude when this is measured in the relative number of overlapping genes.

In preparation of further advancing the paradigm that biogerontology is the best way to face the impending social crisis, we trained and optimized a random forest algorithm for high-throughput *in*

silico screen of drug libraries with the aim of discovering and prioritize new candidate anti-aging chemical compounds. The optimal configuration of the algorithm performed at more than sufficient level for the intended task and therefore we are fully ready to move to the screening stage (work that we are undertaking at the moment).

In sum, we hope to have accelerated the present and future development of biogerontology as The pure science on which translational medicine stands.

4.2. FUTURES PERSPECTIVES

Regarding section 2, much still needs to be done. We hope to follow on it by confirming the lifespan-extending potential of the enriched drugs and suggested drug combinations through lifespan assays in model organisms.

Since we already made full use of the currently available knowledge regarding lifespan-extending genes, the only way to improve our bioinformatics approach would be to use more chemicals, in other words, use several repositories as the sources of drug-gene interactions. We encourage our colleagues to advance our methodology, as the drug repurposing paradigm as the potential to accelerate the bench-to-bedside transition by more than a decade.

Concerning DrugAge, we intend to keep it a first-line resource for biogerontologists. The author, as DrugAge's official curator, vows to keep it up-to-date and is already curating information for the next update.

We envision and encourage cheminformaticians, bioinformaticians, biologists, clinicians and even the nutraceutical industry to use DrugAge to create a world with less suffering.

Any application of machine learning can always be improved upon, and the author has intentions to do so in the immediate future. Four straightforward refinements are: to screen and optimize a vast library of distinct machine learning algorithms; to create a methodology that weights drugs in proportion to the magnitude of their lifespan prolongation effects; to tune a given algorithm parameters using meta-algorithms[232] instead of grid search; to set the classification task in the Bayesian paradigm, including setting a prior for the class probability membership based on a meta-review of high-throughput screens for anti-aging compounds.

REFERENCES

[1] G. . Cailliet, a. . Andrews, E. . Burton, D. . Watters, D. . Kline, and L. . Ferry-Graham, “Age determination and validation studies of marine fishes: do deep-dwellers live longer?,” *Exp. Gerontol.*, vol. 36, no. 4–6, pp. 739–764, 2001.

[2] J. D. Congdon, R. D. Nagle, O. M. Kinney, and R. C. Van Loben Sels, “Hypotheses of aging in a long-lived vertebrate, Blanding’s turtle (Emydoidea blandingii),” in *Experimental Gerontology*, 2001, vol. 36, no. 4–6, pp. 813–827.

[3] J. D. Congdon, R. D. Nagle, O. M. Kinney, R. C. Van Loben Sels, T. Quinter, and D. W. Tinkle, “Testing hypotheses of aging in long-lived painted turtles (Chrysemys picta),” in *Experimental Gerontology*, 2003, vol. 38, no. 7, pp. 765–772.

[4] T. A. Ebert and J. R. Southon, “Red sea urchins (Strongylocentrotus franciscanus) can live over 100 years: Confirmation with A-bomb 14carbon,” *Fish. Bull.*, vol. 101, no. 4, pp. 915–922, 2003.

[5] J. K. Miller, “Escaping senescence: Demographic data from the three-toed box turtle (Terrapene carolina triunguis),” *Exp. Gerontol.*, vol. 36, no. 4–6, pp. 829–832, 2001.

[6] A. G. Bodnar and J. A. Coffman, “Maintenance of somatic tissue regeneration with age in short- and long-lived species of sea urchins,” *Aging Cell*, vol. 2, pp. 1–10, 2016.

[7] A. Zhavoronkov and M. Litovchenko, “Biomedical progress rates as new parameters for models of economic growth in developed countries,” *Int. J. Environ. Res. Public Health*, vol. 10, no. 11, pp. 5936–5952, 2013.

[8] G. Nargund, “Declining birth rate in Developed Countries : A radical policy re-think is required,” *F, V&V ObGyn*, vol. 1, no. 3, pp. 191–193, 2009.

[9] D. E. Bloom, D. Canning, and G. Fink, “Implications of population ageing for economic growth,” *Oxford Rev. Econ. Policy*, vol. 26, no. 4, pp. 583–612, 2010.

[10] G. K. Vincent and V. a Velkoff, “The Next Four Decades The Older Population in the United States: 2010 to 2050,” *U.S. Census Bur.*, vol. P25–1138, no. May, 2010.

[11] S. Harper, “Economic and Social Implications of Aging Societies,” *Science (80-.).*, vol. 346, no. 587, pp. 73–92, 2014.

[12] G. A. Petsko, “A seat at the table,” *Genome Biol.*, vol. 9, no. 12, p. 113, 2008.

[13] K. Howse, “Healthy ageing: the role of health care services.,” *Perspect. Public Health*, vol. 132, no. 4, pp. 171–7, 2012.

[14] R. Lee and A. Mason, “Fertility, Human Capital, and Economic Growth over the Demographic Transition,” *Eur. J. Popul. / Rev. Eur. Démographie*, vol. 26, no. 2, pp. 159–182, 2010.

[15] C. Dustmann, T. Frattini, and C. Halls, “Assessing the fiscal costs and benefits of A8 migration to the UK,” *Fisc. Stud.*, vol. 31, no. 1, pp. 1–41, 2010.

[16] J. W. Vaupel, “Biodemography of Human Ageing,” *Nature*, vol. 464, no. 7288, pp. 536–42, 2010.

[17] B. Rechel, E. Grundy, J. M. Robine, J. Cylus, J. P. MacKenbach, C. Knai, and M. McKee, “Ageing in the European Union,” *Lancet*, vol. 381, no. 9874, pp. 1312–1322, 2013.

[18] E. Nolte and C. M. McKee, “Measuring the health of nations: Updating an earlier analysis,” *Health Aff.*, vol. 27, no. 1, pp. 58–71, 2008.

[19] D. Stuckler, “Population causes and consequences of leading chronic diseases: A comparative analysis of prevailing explanations,” *Milbank Q.*, vol. 86, no. 2, pp. 273–326, 2008.

[20] W. Luyten, P. Antal, B. P. Braeckman, J. Bundy, F. Cirulli, C. Fang-Yen, G. Fuellen, A. Leroi, Q. Liu, P. Martorell, A. Metspalu, M. Perola, M. Ristow, N. Saul, L. Schoofs, K. Siems, L. Temmerman, T. Smets, A. Wolk, and S. I. S. Rattan, “Ageing with elegans: a research proposal to map healthspan pathways.,” *Biogerontology*, 2016.

[21] C. Kenyon, J. Chang, E. Gensch, A. Rudner, and R. Tabtiang, “A C. elegans mutant that lives twice as long as wild type.,” *Nature*, vol. 366, no. 6454, pp. 461–4, 1993.

[22] R. Tacutu, T. Craig, A. Budovsky, D. Wuttke, G. Lehmann, D. Taranukha, J. Costa, V. E. Fraifeld, and J. P. De Magalh??es, “Human Ageing Genomic Resources: Integrated databases and tools for the biology and genetics of ageing,” *Nucleic Acids Res.*, vol. 41, no. D1, 2013.

[23] A. Moskalev, E. Chernyagina, J. P. de Magalh??es, D. Barardo, H. Thoppil, M. Shaposhnikov, A. Budovsky, V. E. Fraifeld, A. Garazha, V. Tsvetkov, E. Bronovitsky, V. Bogomolov, A. Scerbacov, O. Kuryan, R. Gurinovich, L. C. Jellen, B. Kennedy, P. Mamoshina, E. Dobrovolskaya, A. Aliper, D. Kaminsky, and A. Zhavoronkov, “Geroprotectors.org: A new, structured and curated database of current therapeutic interventions in aging and age-related disease,” *Aging (Albany. NY).*, vol. 7, no. 9, pp. 616–628, 2015.

[24] S. Anton and C. Leeuwenburgh, “Fasting or caloric restriction for healthy aging,” *Experimental Gerontology*, vol. 48, no. 10, pp. 1003–1005, 2013.

[25] R. J. Colman, R. M. Anderson, S. C. Johnson, E. K. Kastman, K. J. Kosmatka, T. M. Beasley, D. B. Allison, C. Cruzen, H. a Simmons, J. W. Kemnitz, and R. Weindruch, “Caloric

restriction delays disease onset and mortality in rhesus monkeys.," *Science*, vol. 325, no. 5937, pp. 201–204, 2009.

[26] L. Hayflick, "Biological aging is no longer an unsolved problem," in *Annals of the New York Academy of Sciences*, 2007, vol. 1100, pp. 1–13.

[27] C. López-Otín, M. A. Blasco, L. Partridge, M. Serrano, and G. Kroemer, "The hallmarks of aging.," *Cell*, vol. 153, no. 6, pp. 1194–217, Jun. 2013.

[28] T. Fulop, J. M. Witkowski, G. Pawelec, C. Alan, and A. Larbi, "On the immunological theory of aging," in *Aging: Facts and Theories*, vol. 39, 2014, pp. 163–176.

[29] A. A. Freitas and J. P. De Magalhães, "A review and appraisal of the DNA damage theory of ageing," *Mutat. Res. - Rev. Mutat. Res.*, vol. 728, no. 1–2, pp. 12–22, 2011.

[30] L. Hayflick and P. S. Moorhead, "The serial cultivation of human diploid cell strains," *Exp. Cell Res.*, vol. 25, no. 3, pp. 585–621, 1961.

[31] I. Afanas'ev, "Signaling and Damaging Functions of Free Radicals in Aging-Free Radical Theory, Hormesis, and TOR.," *Aging Dis.*, vol. 1, no. 2, pp. 75–88, 2010.

[32] J. Bjorksten and H. Tenhu, "The crosslinking theory of aging–added evidence.," *Exp. Gerontol.*, vol. 25, no. April 1989, pp. 91–95, 1990.

[33] C. D. Rollo, "Aging and the Mammalian regulatory triumvirate.," *Aging Dis.*, vol. 1, no. 2, pp. 105–38, 2010.

[34] D. van Heemst, "Insulin, IGF-1 and longevity," *Aging Dis.*, vol. 1, no. 2, pp. 147–57, 2010.

[35] M. Davidovic, G. Sevo, P. Svorcan, D. P. Milosevic, N. Despotovic, and P. Erceg, "Old age as a privilege of the 'selfish ones'.," *Aging Dis.*, vol. 1, no. 2, pp. 139–46, 2010.

[36] K. Jin, "Modern Biological Theories of Aging.," *Aging Dis.*, vol. 1, no. 2, pp. 72–74, 2010.

[37] C. J. Lord and A. Ashworth, "The DNA damage response and cancer therapy," *Nature*, vol. 481, no. 7381, pp. 287–294, 2012.

[38] C. R. Burtner and B. K. Kennedy, "Progeria syndromes and ageing: what is the connection?," *Nat. Rev. Mol. Cell Biol.*, vol. 11, no. 8, pp. 567–578, 2010.

[39] D. J. Baker, M. M. Dawlaty, T. Wijshake, K. B. Jeganathan, L. Malureanu, J. H. van Ree, R. Crespo-Diaz, S. Reyes, L. Seaburg, V. Shapiro, A. Behfar, A. Terzic, B. van de Sluis, and J. M. van Deursen, "Increased expression of BubR1 protects against aneuploidy and cancer and extends healthy lifespan.," *Nat. Cell Biol.*, vol. 15, no. 1, pp. 96–102, 2013.

[40] M. Vermulst, J. Wanagat, G. C. Kujoth, J. H. Bielas, P. S. Rabinovitch, T. A. Prolla, and L. A. Loeb, "DNA deletions and clonal mutations drive premature aging in mitochondrial mutator mice.," *Nat. Genet.*, vol. 40, no. 4, pp. 392–394, 2008.

[41] C. D. Ragnauth, D. T. Warren, Y. Liu, R. McNair, T. Tajsic, N. Figg, R. Shroff, J. Skepper, and C. M. Shanahan, "Prelamin a acts to accelerate smooth muscle cell senescence and is a novel biomarker of human vascular aging," *Circulation*, vol. 121, no. 20, pp. 2200–2210, 2010.

[42] P. Scaffidi and T. Misteli, "Lamin A-dependent nuclear defects in human aging.," *Science*, vol. 312, no. 5776, pp. 1059–63, 2006.

[43] A. De Sandre-Giovannoli, R. Bernard, P. Cau, C. Navarro, J. Amiel, I. Boccaccio, S. Lyonnet, C. L. Stewart, A. Munnich, M. Le Merrer, and N. Lévy, "Lamin a truncation in Hutchinson-Gilford progeria.," *Science*, vol. 300, p. 2055, 2003.

[44] M. Eriksson, W. T. Brown, L. B. Gordon, M. W. Glynn, J. Singer, L. Scott, M. R. Erdos, C. M. Robbins, T. Y. Moses, P. Berglund, A. Dutra, E. Pak, S. Durkin, A. B. Csoka, M. Boehnke, T. W. Glover, and F. S. Collins, "Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome.," *Nature*, vol. 423, no. 6937, pp. 293–298, 2003.

[45] R. Cabanillas, J. Cadiñanos, J. A. F. Villameytide, M. Pérez, J. Longo, J. M. Richard, R. Álvarez, N. S. Durán, R. Illán, D. J. González, and C. López-Otín, "Néstor-Guillermo progeria syndrome: A novel premature aging condition with early onset and chronic development caused by BANF1 mutations," *Am. J. Med. Genet. Part A*, vol. 155, no. 11, pp. 2617–2625, 2011.

[46] M. Fumagalli, F. Rossiello, M. Clerici, S. Barozzi, D. Cittaro, J. M. Kaplunov, G. Bucci, M. Dobreva, V. Matti, C. M. Beausejour, U. Herbig, M. P. Longhese, and F. d'Adda di Fagagna, "Telomeric DNA damage is irreparable and causes persistent DNA-damage-response activation.," *Nat. Cell Biol.*, vol. 14, no. 4, pp. 355–65, 2012.

[47] G. Hewitt, D. Jurk, F. D. M. Marques, C. Correia-Melo, T. Hardy, A. Gackowska, R. Anderson, M. Taschuk, J. Mann, and J. F. Passos, "Telomeres are favoured targets of a persistent DNA damage response in ageing and stress-induced senescence.," *Nat. Commun.*, vol. 3, p. 708, 2012.

[48] M. Armanios and E. H. Blackburn, "The telomere syndromes.," *Nat. Rev. Genet.*, vol. 13, no. 10, pp. 693–704, 2012.

[49] M. Armanios, J. K. Alder, E. M. Parry, B. Karim, M. A. Strong, and C. W. Greider, "Short Telomeres are Sufficient to Cause the Degenerative Defects Associated with Aging," *Am. J. Hum. Genet.*, vol. 85, no. 6, pp. 823–832, 2009.

[50] K. L. Rudolph, S. Chang, H. W. Lee, M. Blasco, G. J. Gottlieb, C. Greider, and R. A. DePinho, "Longevity, stress response, and cancer in aging telomerase-deficient mice.," *Cell*, vol. 96, no. 5, pp. 701–712, 1999.

[51] A. Tomás-Loba, I. Flores, P. J. Fernández-Marcos, M. L. Cayuela, A. Maraver, A. Tejera, C. Borr??s, A. Matheu, P. Klatt, J. M. Flores, J. Vi??a, M. Serrano, and M. A. Blasco, "Telomerase Reverse Transcriptase Delays Aging in Cancer-Resistant Mice," *Cell*, vol. 135, no. 4, pp. 609–622, 2008.

[52] S. Horvath, "DNA methylation age of human tissues and cell types.," *Genome Biol.*, vol. 14, no. 10, p. R115, 2013.

[53] L. Zhong, A. D'Urso, D. Toiber, C. Sebastian, R. E. Henry, D. D. Vadysirisack, A. Guimaraes, B. Marinelli, J. D. Wikstrom, T. Nir, C. B. Clish, B. Vaitheesvaran, O. Iliopoulos, I. Kurland, Y. Dor, R. Weissleder, O. S. Shirihai, L. W. Ellisen, J. M. Espinosa, and R. Mostoslavsky, "The Histone Deacetylase Sirt6 Regulates Glucose Homeostasis via Hif1alpha," *Cell*, vol. 140, no. 2, pp. 280–293, 2010.

[54] T. L. a Kawahara, E. Michishita, A. S. Adler, M. Damian, E. Berber, M. Lin, R. a McCord, K. C. L. Ongaigui, L. D. Boxer, H. Y. Chang, and K. F. Chua, "SIRT6 links histone H3 lysine 9 deacetylation to NF-kappaB-dependent gene expression and organismal life span.," *Cell*, vol. 136, no. 1, pp. 62–74, 2009.

[55] Y. Kanfi, V. Peshti, R. Gil, S. Naiman, L. Nahum, E. Levin, N. Kronfeld-Schor, and H. Y. Cohen, "SIRT6 protects against pathological damage caused by diet-induced obesity," *Aging Cell*, vol. 9, no. 2, pp. 162–173, 2010.

[56] K. Larson, S. J. Yan, A. Tsurumi, J. Liu, J. Zhou, K. Gaur, D. Guo, T. H. Eickbush, and W. X. Li, "Heterochromatin formation promotes longevity and represses ribosomal RNA synthesis," *PLoS Genet.*, vol. 8, no. 1, 2012.

[57] T. Smith-Vikos and F. J. Slack, "MicroRNAs and their roles in aging.," *J. Cell Sci.*, vol. 125, no. Pt 1, pp. 7–17, 2012.

[58] Y. Shen, J. Wollam, D. Magner, O. Karalay, and A. Antebi, "A Steroid Receptor-MicroRNA Switch Regulates Life Span in Response to Signals from the Gonad," *Science (80-.).*, vol. 338, no. 6113, pp. 1472–1476, 2012.

[59] N. Liu, M. Landreh, K. Cao, M. Abe, G.-J. Hendriks, J. R. Kennerdell, Y. Zhu, L.-S. Wang, and N. M. Bonini, "The microRNA miR-34 modulates ageing and neurodegeneration in Drosophila," *Nature*, vol. 482, no. 7386, pp. 519–523, 2012.

[60] S. Peleg, F. Sananbenesi, A. Zovoilis, S. Burkhardt, S. Bahari-Javan, R. C. Agis-Balboa, P. Cota, J. L. Wittnam, A. Gogol-Doering, L. Opitz, G. Salinas-Riester, M. Dettenhofer, H. Kang, L. Farinelli, W. Chen, and A. Fischer, "Altered histone acetylation is associated with age-dependent memory impairment in mice," *Science (80-.).*, vol. 328, no. 5979, pp. 753–756, 2010.

[61] H. Koga, S. Kaushik, and a M. Cuervo, "Protein Homeostasis and Aging: the importance of exquisite quality control," *Ageing Res. Rev.*, vol. 10, no. 2, pp. 205–215, 2012.

[62] E. T. Powers, R. I. Morimoto, A. Dillin, J. W. Kelly, and W. E. Balch, "Biological and chemical approaches to diseases of proteostasis deficiency.," *Annu. Rev. Biochem.*, vol. 78, no. August, pp. 959–991, 2009.

[63] G. Morrow, M. Samson, S. Michaud, and R. M. Tanguay, "Overexpression of the small mitochondrial Hsp22 extends Drosophila life span and increases resistance to oxidative stress.," *FASEB J.*, vol. 18, no. 3, pp. 598–599, 2004.

[64] G. a Walker and G. J. Lithgow, "Lifespan extension in C. elegans by a molecular chaperone dependent upon insulin-like signals.," *Aging Cell*, vol. 2, no. 2, pp. 131–139, 2003.

[65] W. R. Swindell, M. M. Masternak, J. J. Kopchick, C. A. Conover, A. Bartke, and R. A. Miller, "Endocrine regulation of heat shock protein mRNA levels in long-lived dwarf mice," *Mech. Ageing Dev.*, vol. 130, no. 6, pp. 393–400, 2009.

[66] J.-N. Min, R. a Whaley, N. E. Sharpless, P. Lockyer, A. L. Portbury, and C. Patterson, "CHIP deficiency decreases longevity, with accelerated aging phenotypes accompanied by altered protein quality control.," *Mol. Cell. Biol.*, vol. 28, no. 12, pp. 4018–25, 2008.

[67] T. Eisenberg, H. Knauer, A. Schauer, S. Büttner, C. Ruckenstuhl, D. Carmona-Gutierrez, J. Ring, S. Schroeder, C. Magnes, L. Antonacci, H. Fussi, L. Deszcz, R. Hartl, E. Schraml, A. Criollo, E. Megalou, D. Weiskopf, P. Laun, G. Heeren, M. Breitenbach, B. Grubeck-Loebenstein, E. Herker, B. Fahrenkrog, K.-U. Fröhlich, F. Sinner, N. Tavernarakis, N. Minois, G. Kroemer, and F. Madeo, "Induction of autophagy by spermidine promotes longevity.," *Nat. Cell Biol.*, vol. 11, no. 11, pp. 1305–14, 2009.

[68] I. Bjedov, J. M. Toivonen, F. Kerr, C. Slack, J. Jacobson, A. Foley, and L. Partridge, "Mechanisms of Life Span Extension by Rapamycin in the Fruit Fly Drosophila melanogaster," *Cell Metab.*, vol. 11, no. 1, pp. 35–46, 2010.

[69] E. J. O'Rourke, P. Kuballa, R. Xavier, and G. Ruvkun, "Omega-6 Polyunsaturated fatty acids extend life span through the activation of autophagy," *Genes Dev.*, vol. 27, no. 4, pp. 429–440, 2013.

[70] D. C. Rubinsztein, G. Mariño, and G. Kroemer, "Autophagy and aging," *Cell*, vol. 146, no. 5. pp. 682–695, 2011.

[71] M. Matsumoto, S. Kurihara, R. Kibe, H. Ashida, and Y. Benno, "Longevity in mice is promoted by probiotic-induced suppression of colonic senescence dependent on upregulation of gut bacterial polyamine production," *PLoS One*, vol. 6, no. 8, 2011.

[72] K. Soda, Y. Dobashi, Y. Kano, S. Tsujinaka, and F. Konishi, "Polyamine-rich food decreases age-associated pathology and mortality in aged mice," *Exp. Gerontol.*, vol. 44, no. 11, pp. 727–732, 2009.

[73] U. Kruegel, B. Robison, T. Dange, G. Kahlert, J. R. Delaney, S. Kotireddy, M. Tsuchiya, S. Tsuchiyama, C. J. Murakami, J. Schleit, G. Sutphin, D. Carr, K. Tar, G. Dittmar, M. Kaeberlein, B. K. Kennedy, and M. Schmidt, "Elevated proteasome capacity extends replicative lifespan in saccharomyces cerevisiae," *PLoS Genet.*, vol. 7, no. 9, 2011.

[74] G. Liu, J. Rogers, C. T. Murphy, and C. Rongo, "EGF signalling activates the ubiquitin proteasome system to modulate C. elegans lifespan.," *EMBO J.*, vol. 30, no. 15, pp. 2990–3003, 2011.

[75] D. Vilchez, I. Morantte, Z. Liu, P. M. Douglas, C. Merkwirth, A. P. C. Rodrigues, G. Manning, and A. Dillin, "RPN-6 determines C. elegans longevity under proteotoxic stress conditions," *Nature*, vol. 489, no. 7415, pp. 263–268, 2012.

[76] J. a. Mattison, G. S. Roth, T. M. Beasley, E. M. Tilmont, A. M. Handy, R. L. Herbert, D. L. Longo, D. B. Allison, J. E. Young, M. Bryant, D. Barnard, W. F. Ward, W. Qi, D. K. Ingram, and R. de Cabo, "Impact of caloric restriction on health and survival in rhesus monkeys from the NIA study," *Nature*, vol. 489, no. 7415, pp. 318–321, 2012.

[77] L. Fontana, L. Partridge, and V. D. Longo, "Extending healthy life span–from yeast to humans.," *Science*, vol. 328, no. 5976, pp. 321–6, Apr. 2010.

[78] B. Schumacher, I. Van Der Pluijm, M. J. Moorhouse, T. Kosteas, A. R. Robinson, Y. Suh, T. M. Breit, H. Van Steeg, L. J. Niedernhofer, W. Van Ijcken, A. Bartke, S. R. Spindler, J. H. J. Hoeijmakers, G. T. J. Van Der Hors, and G. A. Garinis, "Delayed and accelerated aging share common longevity assurance mechanisms," *PLoS Genet.*, vol. 4, no. 8, 2008.

[79] L. Fontana, L. Partridge, and V. D. Longo, "Extending healthy life span - From yeast to humans," *Science (80-.).*, vol. 328, pp. 321–326, 2010.

[80] L. C. Foukas, B. Bilanges, L. Bettedi, W. Pearce, K. Ali, S. Sancho, D. J. Withers, and B. Vanhaesebroeck, "Long-term p110alpha PI3K inactivation exerts a beneficial effect on metabolism," *EMBO Mol. Med.*, vol. 5, no. 4, pp. 563–571, 2013.

[81] A. Ortega-Molina, A. Efeyan, E. Lopez-Guadamillas, M. Mu??oz-Martin, G. G??mez-L??pez, M. Ca??amero, F. Mulero, J. Pastor, S. Martinez, E. Romanos, M. Mar Gonzalez-Barroso, E. Rial, A. M. Valverde, J. R. Bischoff, and M. Serrano, "Pten positively regulates brown adipose function, energy expenditure, and longevity," *Cell Metab.*, vol. 15, no. 3, pp. 382–394, 2012.

[82] G. Mariño, A. P. Ugalde, A. F. Fernández, F. G. Osorio, A. Fueyo, J. M. P. Freije, and C. López-Otín, "Insulin-like growth factor 1 treatment extends longevity in a mouse model of human premature aging by restoring somatotroph axis function.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 37, pp. 16268–73, 2010.

[83] S. C. Johnson, P. S. Rabinovitch, and M. Kaeberlein, "mTOR is a key modulator of ageing and age-related disease.," *Nature*, vol. 493, no. 7432, pp. 338–45, 2013.

[84] D. W. Lamming, L. Ye, P. Katajisto, M. D. Goncalves, M. Saitoh, D. M. Stevens, J. G. Davis, A. B. Salmon, A. Richardson, R. S. Ahima, D. A. Guertin, D. M. Sabatini, and J. A. Baur, "Rapamycin-induced insulin resistance is mediated by mTORC2 loss and uncoupled from longevity.," *Science*, vol. 335, no. 6076, pp. 1638–43, 2012.

[85] D. E. Harrison, R. Strong, Z. D. Sharp, J. F. Nelson, C. M. Astle, K. Flurkey, N. L. Nadon, J. E. Wilkinson, K. Frenkel, C. S. Carter, M. Pahor, M. a Javors, E. Fernandez, and R. a Miller, "Rapamycin fed late in life extends lifespan in genetically heterogeneous mice.," *Nature*, vol. 460, no. 7253, pp. 392–395, 2009.

[86] D. Harman, "The biologic clock: the mitochondria?," *J. Am. Geriatr. Soc.*, vol. 20, no. 4, pp. 145–7, May 1972.

[87] D. R. Green, L. Galluzzi, and G. Kroemer, "Mitochondria and the autophagy-inflammation-cell death axis in organismal aging.," *Science*, vol. 333, no. 6046, pp. 1109–12, 2011.

[88] A. Mesquita, M. Weinberger, A. Silva, B. Sampaio-Marques, B. Almeida, C. Leão, V. Costa, F. Rodrigues, W. C. Burhans, and P. Ludovico, "Caloric restriction or catalase inactivation extends yeast chronological lifespan by inducing H2O2 and superoxide dismutase activity.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 34, pp. 15123–8, 2010.

[89] J. M. Van Raamsdonk and S. Hekimi, "Deletion of the mitochondrial superoxide dismutase sod-2 extends lifespan in Caenorhabditis elegans," *PLoS Genet.*, vol. 5, no. 2, 2009.

[90] S. Schmeisser, K. Zarse, and M. Ristow, "Lonidamine extends lifespan of adult Caenorhabditis elegans by increasing the formation of mitochondrial reactive oxygen species," *Horm. Metab. Res.*, vol. 43, no. 10, pp. 687–692, 2011.

[91] S. Schmeisser, K. Schmeisser, S. Weimer, M. Groth, S. Priebe, E. Fazius, D. Kuhlow, D. Pick, J. W. Einax, R. Guthke, M. Platzer, K. Zarse, and M. Ristow, "Mitochondrial hormesis links low-dose arsenite exposure to lifespan extension," *Aging Cell*, vol. 12, no. 3, pp. 508–517, 2013.

[92] R. Doonan, J. J. McElwee, F. Matthijssens, G. A. Walker, K. Houthoofd, P. Back, A. Matscheski, J. R. Vanfleteren, and D. Gems, "Against the oxidative damage theory of aging: Superoxide dismutases protect against oxidative stress but have little or no effect on life span in Caenorhabditis elegans," *Genes Dev.*, vol. 22, no. 23, pp. 3236–3241, 2008.

[93] V. I. Pérez, H. Van Remmen, A. Bokov, C. J. Epstein, J. Vijg, and A. Richardson, "The overexpression of major antioxidant enzymes does not extend the lifespan of mice," *Aging Cell*, vol. 8, no. 1, pp. 73–75, 2009.

[94] M. Ristow and K. Schmeisser, "Mitohormesis: Promoting Health and Lifespan by Increased Levels of Reactive Oxygen Species (ROS).," *Dose. Response.*, vol. 12, no. 2, pp. 288–341, 2014.

[95] C. Wang, D. Jurk, M. Maddick, G. Nelson, C. Martin-ruiz, and T. Von Zglinicki, "DNA damage response and cellular senescence in tissues of aging mice," *Aging Cell*, vol. 8, no. 3, pp. 311–323, 2009.

[96] D. J. Baker, B. G. Childs, M. Durik, M. E. Wijers, C. J. Sieben, J. Zhong, R. A. Saltness, K. B. Jeganathan, G. C. Verzosa, A. Pezeshki, K. Khazaie, J. D. Miller, and J. M. van Deursen, "Naturally occurring p16(Ink4a)-positive cells shorten healthy lifespan.," *Nature*, vol. 530, no. 7589, pp. 184–9, 2016.

[97] Y. Zhu, T. Tchkonia, T. Pirtskhalava, A. C. Gower, H. Ding, N. Giorgadze, A. K. Palmer, Y. Ikeno, G. B. Hubbard, M. Lenburg, S. P. O'Hara, N. F. Larusso, J. D. Miller, C. M. Roos, G. C. Verzosa, N. K. Lebrasseur, J. D. Wren, J. N. Farr, S. Khosla, M. B. Stout, S. J. Mcgowan, H. Fuhrmann-Stroissnigg, A. U. Gurkar, J. Zhao, D. Colangelo, A. Dorronsoro, Y. Y. Ling, A. S. Barghouthy, D. C. Navarro, T. Sano, P. D. Robbins, L. J. Niedernhofer, and J. L. Kirkland, "The Achilles' heel of senescent cells: From transcriptome to senolytic drugs," *Aging Cell*, 2015.

[98] A. V Molofsky, S. G. Slutsky, N. M. Joseph, S. He, R. Pardal, J. Krishnamurthy, N. E. Sharpless, and S. J. Morrison, "Increasing p16INK4a expression decreases forebrain progenitors and neurogenesis during ageing," *Nature*, vol. 443, no. 7110, pp. 448–452, 2006.

[99] R. Gruber, H. Koch, B. A. Doll, F. Tegtmeier, T. A. Einhorn, and J. O. Hollinger, "Fracture healing in the elderly patient," *Experimental Gerontology*, vol. 41, no. 11. pp. 1080–1093, 2006.

[100] I. M. Conboy and T. A. Rando, "Heterochronic parabiosis for the study of the effects of aging on stem cells and their niches," *Cell Cycle*, vol. 11, no. 12. pp. 2260–2267, 2012.

[101] M. Rera, S. Bahadorani, J. Cho, C. L. Koehler, M. Ulgherait, J. H. Hur, W. S. Ansari, T. Lo, D. L. Jones, and D. W. Walker, "Modulation of longevity and tissue homeostasis by the Drosophila PGC-1 homolog," *Cell Metab.*, vol. 14, no. 5, pp. 623–634, 2011.

[102] M. Scudellari, "Ageing research: Blood to blood," *Nature*, vol. 517, pp. 426–429, 2015.

[103] T. A. Rando and T. Wyss-Coray, "Stem cells as vehicles for youthful regeneration of aged tissues," *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, vol. 69. pp. S39–S42, 2014.

[104] C. Franceschi and J. Campisi, "Chronic inflammation (Inflammaging) and its potential contribution to age-associated diseases," *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, vol. 69. pp. S4–S9, 2014.

[105] G. A. Poland, I. G. Ovsyannikova, R. B. Kennedy, N. D. Lambert, and J. L. Kirkland, "A systems biology approach to the effect of aging, immunosenescence and vaccine response," *Current Opinion in Immunology*, vol. 29, no. 1. pp. 62–68, 2014.

[106] G. Nelson, J. Wordsworth, C. Wang, D. Jurk, C. Lawless, C. Martin-Ruiz, and T. von Zglinicki, "A senescent cell bystander effect: Senescence-induced senescence," *Aging Cell*, vol. 11, no. 2. pp. 345–349, 2012.

[107] J. Durieux, S. Wolff, and A. Dillin, "The cell-non-autonomous nature of electron transport chain-mediated longevity," *Cell*, vol. 144, no. 1, pp. 79–91, 2011.

[108] R. Garcia-Valles, M. C. Gomez-Cabrera, L. Rodriguez-Mañas, F. J. Garcia-Garcia, A. Diaz, I. Noguera, G. Olaso-Gonzalez, and J. Viña, "Life-long spontaneous exercise does not prolong lifespan but improves health span in mice.," *Longev. Heal.*, vol. 2, no. 1, p. 14, 2013.

[109] M. Reiner, C. Niermann, D. Jekauc, and A. Woll, "Long-term health benefits of physical activity–a systematic review of longitudinal studies.," *BMC Public Health*, vol. 13, no. 1, p. 813, 2013.

[110] A. Biswas, P. I. Oh, G. E. Faulkner, R. R. Bajaj, M. A. Silver, M. S. Mitchell, and D. A. Alter, "Sedentary Time and Its Association With Risk for Disease Incidence, Mortality, and Hospitalization in AdultsA Systematic Review and Meta-analysisSedentary Time and Disease Incidence, Mortality, and Hospitalization," *Ann. Intern. Med.*, vol. 162, no. 2, pp. 123–132, 2015.

[111] L. Fontana and L. Partridge, "Promoting health and longevity through diet: From model organisms to humans," *Cell*, vol. 161, no. 1, pp. 106–118, 2015.

[112] E. Ravussin, L. M. Redman, J. Rochon, S. K. Das, L. Fontana, W. E. Kraus, S. Romashkan, D. A. Williamson, S. N. Meydani, D. T. Villareal, S. R. Smith, R. I. Stein, T. M. Scott, T. M. Stewart, E. Saltzman, S. Klein, M. Bhapkar, C. K. Martin, C. H. Gilhooly, J. O. Holloszy, E. C. Hadley, and S. B. Roberts, "A 2-year randomized controlled trial of human caloric restriction: Feasibility and effects on predictors of health span and longevity," *Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.*, vol. 70, no. 9, pp. 1097–1104, 2015.

[113] A. J. Dirks and C. Leeuwenburgh, "Caloric restriction in humans: Potential pitfalls and health concerns," *Mechanisms of Ageing and Development*, vol. 127, no. 1. pp. 1–7, 2006.

[114] V. D. Longo and S. Panda, "Fasting, Circadian Rhythms, and Time-Restricted Feeding in Healthy Lifespan," *Cell Metab.*, vol. 23, no. 6, pp. 1048–1059, 2016.

[115] K. a Varady, S. Bhutani, M. C. Klempel, C. M. Kroeger, J. F. Trepanowski, J. M. Haus, K. K. Hoddy, and Y. Calvo, "Alternate day fasting for weight loss in normal weight and overweight subjects: A randomized controlled trial," *Nutr. J.*, vol. 12, no. 1, p. 146, 2013.

[116] S. Brandhorst, I. Y. Choi, M. Wei, C. W. Cheng, S. Sedrakyan, G. Navarrete, L. Dubeau, L. P. Yap, R. Park, M. Vinciguerra, S. Di Biase, H. Mirzaei, M. G. Mirisola, P. Childress, L. Ji, S. Groshen, F. Penna, P. Odetti, L. Perin, P. S. Conti, Y. Ikeno, B. K. Kennedy, P. Cohen, T. E. Morgan, T. B. Dorff, and V. D. Longo, "A Periodic Diet that Mimics Fasting Promotes Multi-System Regeneration, Enhanced Cognitive Performance, and Healthspan," *Cell Metab.*, vol. 22, no. 1, pp. 86–99, 2015.

[117] S. Calvert, R. Tacutu, S. Sharifi, R. Teixeira, P. Ghosh, and J. P. de Magalhães, "A network pharmacology approach reveals new candidate caloric restriction mimetics in C. elegans," *Aging Cell*, vol. 15, no. 2, pp. 256–266, 2016.

[118] E. Pennisi, "The CRISPR Craze," *Science (80-.).*, vol. 341, no. 6148, pp. 833–836, 2013.

[119] I. B. Hilton, A. M. D'Ippolito, C. M. Vockley, P. I. Thakore, G. E. Crawford, T. E. Reddy, and C. A. Gersbach, "Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers," *Nat. Biotechnol.*, vol. 33, no. 5, pp. 510–7, 2015.

[120] A. G. Bassuk, A. Zheng, Y. Li, S. H. Tsang, and V. B. Mahajan, "Precision Medicine: Genetic Repair of Retinitis Pigmentosa in Patient-Derived Stem Cells.," *Sci. Rep.*, vol. 6, p. 19969, 2016.

[121] N. Barzilai, J. P. Crandall, S. B. Kritchevsky, M. A. Espeland, K. Alagiakrishnan, S. Sankaralingam, M. Ghosh, L. Mereu, P. Senior, C. Algire, O. Moiseeva, X. Deschênes-Simard, L.

Amrein, L. Petruccelli, E. Birman, B. Viollet, G. Ferbeyre, M. N. Pollak, J. S. Allard, E. J. Perez, K. Fukui, P. Carpenter, D. K. Ingram, R. de Cabo, G. Anfossi, I. Russo, K. Bonomo, M. Trovati, V. N. Anisimov, A. Bartke, V. N. Anisimov, L. M. Berstein, P. A. Egormin, T. S. Piskunova, I. G. Popovich, M. A. Zabezhinski, M. L. Tyndyk, M. V. Yurova, I. G. Kovalenko, T. E. Poroshina, A. V. Semenchenko, V. N. Anisimov, P. A. Egormin, T. S. Piskunova, I. G. Popovich, M. L. Tyndyk, M. N. Yurova, M. A. Zabezhinski, I. V. Anikin, A. S. Karkach, A. A. Romanyukha, V. N. Anisimov, L. M. Berstein, I. G. Popovich, M. A. Zabezhinski, P. A. Egormin, T. S. Piskunova, A. V. Semenchenko, M. L. Tyndyk, M. N. Yurova, I. G. Kovalenko, T. E. Poroshina, V. R. Aroda, S. Edelstein, R. Goldberg, W. C. Knowler, S. Marcovina, T. Orchard, G. A. Bray, D. Schade, M. Temprosa, N. H. White, et al., C. A. Bannister, S. E. Holden, S. Jenkins-Jones, C. L. Morgan, J. P. Halcox, G. Schernthaner, J. Mukherjee, C. J. Currie, A. Bartke, J. C. Wright, J. A. Mattison, D. K. Ingram, R. A. Miller, G. S. Roth, N. Barzilai, D. M. Huffman, R. H. Muzumdar, A. Bartke, C. Batandier, B. Guigas, D. Detaille, M. Y. El-Mir, E. Fontaine, M. Rigoulet, X. M. Leverve, H. R. Bridges, A. J. Jones, M. N. Pollak, J. Hirst, F. Cabreiro, C. Au, K. Y. Leung, N. Vergara-Irigaray, H. M. Cochemé, T. Noori, D. Weinkove, E. Schuster, N. D. Greene, D. Gems, C. Cheng, C. H. Lin, Y. W. Tsai, C. J. Tsai, P. H. Chou, T. H. Lan, K. Cho, J. Y. Chung, S. K. Cho, H. W. Shin, I. J. Jang, J. W. Park, K. S. Yu, J. Y. Cho, R. de Cabo, D. Carmona-Gutierrez, M. Bernier, M. N. Hall, F. Madeo, W. De Haes, L. Frooninckx, R. Van Assche, A. Smolders, G. Depuydt, J. Billen, B. P. Braeckman, L. Schoofs, L. Temmerman, A. DeCensi, M. Puntoni, A. Guerrieri-Gonzaga, M. Cazzaniga, D. Serrano, M. Lazzeroni, A. Vingiani, O. Gentilini, M. Petrera, G. Viale, et al., D. P. P. R. Group, F. A. Duca, C. D. Côté, B. A. Rasmussen, M. Zadeh-Tahmasebi, G. A. Rutter, B. M. Filippi, T. K. Lam, L. Fontana, L. Partridge, L. Fontana, L. Partridge, V. D. Longo, M. Foretz, S. Hébrard, J. Leclerc, E. Zarrinpashneh, M. Soty, G. Mithieux, K. Sakamoto, F. Andreelli, B. Viollet, M. Foretz, B. Guigas, L. Bertrand, M. Pollak, B. Viollet, S. Gandini, M. Puntoni, B. M. Heckman-Stoddard, B. K. Dunn, L. Ford, A. DeCensi, E. Szabo, R. Goldberg, M. Temprosa, J. Otvos, J. Brunzell, S. Marcovina, K. Mather, R. Arakaki, K. Watson, E. Horton, E. Barrett-Connor, R. Goldberg, M. Temprosa, V. R. Aroda, E. L. Barrett-Connor, M. Budoff, J. Crandall, D. Dabelea, E. S. Horton, K. J. Mather, T. J. Orchard, et al., M. Guo, J. Mi, Q. M. Jiang, J. M. Xu, Y. Y. Tang, G. Tian, B. Wang, S. Haffner, M. Temprosa, J. Crandall, S. Fowler, R. Goldberg, E. Horton, S. Marcovina, K. Mather, T. Orchard, R. Ratner, E. Barrett-Connor, D. P. P. R. Group, D. E. Harrison, R. Strong, Z. D. Sharp, J. F. Nelson, C. M. Astle, K. Flurkey, N. L. Nadon, J. E. Wilkinson, K. Frenkel, C. S. Carter, et al., D. E. Harrison, R. Strong, D. B. Allison, B. N. Ames, C. M. Astle, H. Atamna, E. Fernandez, K. Flurkey, M. A. Javors, N. L. Nadon, et al., K. S. Jadhav, C.

M. Dungan, D. L. Williamson, S. C. Johnson, P. S. Rabinovitch, M. Kaeberlein, E. Karnevi, K. Said, R. Andersson, A. H. Rosendahl, E. Kickstein, S. Krauss, P. Thornhill, D. Rutschow, R. Zeller, J. Sharkey, R. Williamson, M. Fuchs, A. Köhler, H. Glossmann, et al., W. C. Knowler, E. Barrett-Connor, S. E. Fowler, R. F. Hamman, J. M. Lachin, E. A. Walker, D. M. Nathan, D. P. P. R. Group, W. C. Knowler, S. L. Eldelstein, R. B. Goldberg, R. T. Ackermann, J. P. Crandall, J. C. Florez, S. E. Fowler, W. H. Herman, E. S. Horton, S. E. Kahn, D. P. P. R. Group, et al., A. Kooy, J. de Jager, P. Lehert, D. Bets, M. G. Wulffelé, A. J. Donker, C. D. Stehouwer, S. Kordes, M. N. Pollak, A. H. Zwinderman, R. A. Mathôt, M. J. Weterman, A. Beeker, C. J. Punt, D. J. Richel, J. W. Wilmink, J. D. Lalau, A. Vermersch, L. Hary, M. Andrejak, F. Isnard, J. Quichaud, G. W. Landman, N. Kleefstra, K. J. van Hateren, K. H. Groenier, R. O. Gans, H. J. Bilo, M. S. Lee, C. C. Hsu, M. L. Wahlqvist, H. N. Tsai, Y. H. Chang, Y. C. Huang, C. P. Lexis, I. C. van der Horst, E. Lipsic, W. G. Wieringa, R. A. de Boer, A. F. van den Heuvel, H. W. van der Werf, R. A. Schurer, G. Pundziute, E. S. Tan, G.-I. Investigators, et al., G. Libby, L. A. Donnelly, P. T. Donnan, D. R. Alessi, A. D. Morris, J. M. Evans, F. Lien, A. Berthier, E. Bouchaert, C. Gheeraert, J. Alexandre, G. Porez, J. Prawitt, H. Dehondt, M. Ploton, S. Colin, et al., B. Liu, Z. Fan, S. M. Edgerton, X. Yang, S. E. Lind, A. D. Thor, V. D. Longo, A. Antebi, A. Bartke, N. Barzilai, H. M. Brown-Borg, C. Caruso, T. J. Curiel, R. de Cabo, C. Franceschi, D. Gems, et al., J. Lu, J. Shi, M. Li, B. Gui, R. Fu, G. Yao, Z. Duan, Z. Lv, Y. Yang, Z. Chen, et al., J. A. Luchsinger, T. Perez, H. Chang, P. Mehta, J. Steffener, G. Pradabhan, M. Ichise, J. Manly, D. P. Devanand, E. Bagiella, T. C. Ma, J. L. Buescher, B. Oatis, J. A. Funk, A. J. Nash, R. L. Carrier, K. R. Hoyt, A. Martin-Montalvo, E. M. Mercken, S. J. Mitchell, H. H. Palacios, P. L. Mote, M. Scheibye-Knudsen, A. P. Gomes, T. M. Ward, R. K. Minor, M. J. Blouin, et al., R. A. Miller, D. E. Harrison, C. M. Astle, R. A. Floyd, K. Flurkey, K. L. Hensley, M. A. Javors, C. Leeuwenburgh, J. F. Nelson, E. Ongini, et al., R. A. Miller, D. E. Harrison, C. M. Astle, J. A. Baur, A. R. Boyd, R. de Cabo, E. Fernandez, K. Flurkey, M. A. Javors, J. F. Nelson, et al., O. Moiseeva, X. Deschênes-Simard, E. St-Germain, S. Igelmann, G. Huot, A. E. Cadar, V. Bourdeau, M. N. Pollak, G. Ferbeyre, M. Monami, C. Colombi, D. Balzi, I. Dicembrini, S. Giannini, C. Melani, V. Vitale, D. Romano, A. Barchielli, N. Marchionni, et al., E. M. Moore, A. G. Mander, D. Ames, M. A. Kotowicz, R. P. Carne, H. Brodaty, M. Woodward, K. Boundy, K. A. Ellis, A. I. Bush, A. Investigators, et al., N. L. Nadon, R. Strong, R. A. Miller, J. Nelson, M. Javors, Z. D. Sharp, J. M. Peralba, D. E. Harrison, V. Nair, S. Sreevalsan, R. Basha, M. Abdelrahim, A. Abudayyeh, A. R. Hoffman, S. Safe, T. P. Ng, L. Feng, K. B. Yap, T. S. Lee, C. H. Tan, B. Winblad, B. Onken, M. Driscoll, B. I. Pérez-Revuelta, M. M. Hettich, A. Ciociaro, C. Rotermund, P. J. Kahle, S. Krauss, D. A. Di Monte, D. Preiss, S. M. Lloyd, I. Ford,

J. J. McMurray, R. R. Holman, P. Welsh, M. Fisher, C. J. Packard, N. Sattar, B. J. Quinn, M. Dallos, H. Kitagawa, A. B. Kunnumakkara, R. M. Memmott, M. C. Hollander, J. J. Gills, P. A. Dennis, Y. Saisho, B. Salani, S. Maffioli, M. Hamoudane, A. Parodi, S. Ravera, M. Passalacqua, A. Alama, M. Nhiri, R. Cordera, D. Maggi, C. Slack, A. Foley, L. Partridge, D. L. Smith, C. F. Elam, J. A. Mattison, M. A. Lane, G. S. Roth, D. K. Ingram, D. B. Allison, Y. M. Song, Y. H. Lee, J. W. Kim, D. S. Ham, E. S. Kang, B. S. Cha, H. C. Lee, B. W. Lee, R. J. Stevens, R. Ali, C. R. Bankhead, M. A. Bethel, B. J. Cairns, R. P. Camisasca, F. L. Crowe, A. J. Farmer, S. Harrison, J. A. Hirst, et al., R. Strong, R. A. Miller, C. M. Astle, R. A. Floyd, K. Flurkey, K. L. Hensley, M. A. Javors, C. Leeuwenburgh, J. F. Nelson, E. Ongini, et al., M. E. Tinetti, G. McAvay, M. Trentalange, A. B. Cohen, H. G. Allore, L. Tosca, C. Ramé, C. Chabrolle, S. Tesseraud, J. Dupont, C. H. Tseng, U. P. D. S. (UKPDS) Group, H. J. Whittington, A. R. Hall, C. P. McLaughlin, D. J. Hausenloy, D. M. Yellon, M. M. Mocanu, J. E. Wilkinson, L. Burmeister, S. V. Brooks, C. C. Chan, S. Friedline, D. E. Harrison, J. F. Hejtmancik, N. Nadon, R. Strong, L. K. Wood, et al., Z. Xie, K. Lau, B. Eby, P. Lozano, C. He, B. Pennington, H. Li, S. Rathi, Y. Dong, R. Tian, et al., Z. Zheng, H. Chen, J. Li, T. Li, B. Zheng, Y. Zheng, H. Jin, Y. He, Q. Gu, and X. Xu, "Metformin as a Tool to Target Aging," *Cell Metab.*, vol. 23, no. 6, pp. 1060–1065, 2016.

[122] S. Kumar and D. B. Lombard, "Finding Ponce de Leon's Pill: Challenges in Screening for Anti-Aging Molecules," *F1000Research*, vol. 5, no. 0, p. 406, 2016.

[123] J. B. Tenenbaum, C. Kemp, T. L. Griffiths, and N. D. Goodman, "How to grow a mind: statistics, structure, and abstraction.," *Science*, vol. 331, no. 6022, pp. 1279–1285, 2011.

[124] N. Cowan, "The Magical Mystery Four: How Is Working Memory Capacity Limited, and Why?," *Curr. Dir. Psychol. Sci.*, vol. 19, no. 1, pp. 51–57, 2010.

[125] M. K. Transtrum, B. B. Machta, K. S. Brown, B. C. Daniels, C. R. Myers, and J. P. Sethna, "Perspective: Sloppiness and emergent theories in physics, biology, and beyond," *J. Chem. Phys.*, vol. 143, no. 1, 2015.

[126] P. W. and Simon, "Too Big to Ignore : The Business Case for Big Data," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2013.

[127] A. R. Joyce and B. O. Palsson, "The model organism as a system: integrating 'omics' data sets.," *Nat. Rev. Mol. Cell Biol.*, vol. 7, no. 3, pp. 198–210, 2006.

[128] S. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach, 3rd edition," *Prentice Hall*, pp. 1 – 1132, 2009.

[129] M. Slawski, M. Daumer, and A.-L. Boulesteix, "CMA: a comprehensive Bioconductor package for supervised classification with high dimensional data.," *BMC Bioinformatics*, vol. 9, p. 439, 2008.

[130] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[131] A. Criminisi, "Decision Forests: A Unified Framework for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning," *Found. Trends® Comput. Graph. Vis.*, vol. 7, no. 2–3, pp. 81–227, 2011.

[132] M. Khalilia, S. Chakraborty, and M. Popescu, "Predicting disease risks from highly imbalanced data using random forest," *BMC Med Inf. Decis Mak*, vol. 11, p. 51, 2011.

[133] W. Z. W. Zhang, F. Z. F. Zeng, X. W. X. Wu, X. Z. X. Zhang, and R. J. R. Jiang, "A Comparative Study of Ensemble Learning Approaches in the Classification of Breast Cancer Metastasis," *2009 Int. Jt. Conf. Bioinformatics, Syst. Biol. Intell. Comput.*, pp. 268–271, 2009.

[134] R. Blagus and L. Lusa, "Class prediction for high-dimensional class-imbalanced data.," *BMC Bioinformatics*, vol. 11, no. 1, p. 523, 2010.

[135] T. M. Khoshgoftaar, M. Golawala, and J. Van Hulse, "An Empirical Study of Learning from Imbalanced Data Using Random Forest," *19th IEEE Int. Conf. Tools with Artif. Intell. 2007)*, vol. 2, pp. 310–317, 2007.

[136] M. Kaeberlei, P. S. Rabinovitch, and G. M. Martin, "Healthy aging: The ultimate preventative medicine," *Science (80-.).*, vol. 350, no. 6265, pp. 1191–1193, 2015.

[137] Y. Donner, K. Fortney, S. R. G. Calimport, K. Pfleger, M. Shah, and J. Betts-LaCroix, "Great Desire for Extended Life and Health amongst the American Public," *Front. Genet.*, vol. 6, no. 21, pp. 3903–3904, 2016.

[138] J. W. Scannell, A. Blanckley, H. Boldon, and B. Warrington, "Diagnosing the decline in pharmaceutical R&D efficiency.," *Nat. Rev. Drug Discov.*, vol. 11, no. 3, pp. 191–200, 2012.

[139] B. Munos, "Lessons from 60 years of pharmaceutical innovation.," *Nat. Rev. Drug Discov.*, vol. 8, no. 12, pp. 959–968, 2009.

[140] J.-M. Fernandez, R. M. Stein, and A. W. Lo, "Commercializing biomedical research through securitization techniques," *Nat. Biotechnol.*, vol. 30, no. 10, pp. 964–975, 2012.

[141] D. E. Fagnan, J. M. Fernandez, A. W. Lo, and R. M. Stein, "Can financial engineering cure cancer?," in *American Economic Review*, 2013, vol. 103, no. 3, pp. 406–411.

[142] J. A. DiMasi, L. Feldman, A. Seckler, and A. Wilson, "Trends in risks associated with new drug development: success rates for investigational drugs.," *Clin. Pharmacol. Ther.*, vol. 87, no. 3, pp. 272–7, 2010.

[143] J. Arrowsmith and P. Miller, "Trial Watch: Phase II and Phase III attrition rates 2011–2012," *Nat. Rev. Drug Discov.*, vol. 12, no. 8, pp. 569–569, 2013.

[144] T. S. Lin and W. H. Prusoff, "Synthesis and biological activity of several amino analogues of thymidine.," *J. Med. Chem.*, vol. 21, no. 1, pp. 109–112, 1978.

[145] R. Yarchoan and S. Broder, "Development of antiretroviral therapy for the acquired immunodeficiency syndrome and related disorders. A progress report," *N Engl J Med*, vol. 316, no. 9, pp. 557–564, 1987.

[146] Z. S. Morris, S. Wooding, and J. Grant, "The answer is 17 years, what is the question: understanding time lags in translational research.," *J. R. Soc. Med.*, vol. 104, no. 12, pp. 510–20, 2011.

[147] A. H. Wagner, A. C. Coffman, B. J. Ainscough, N. C. Spies, Z. L. Skidmore, K. M. Campbell, K. Krysiak, D. Pan, J. F. McMichael, J. M. Eldred, J. R. Walker, R. K. Wilson, E. R. Mardis, M. Griffith, and O. L. Griffith, "DGIdb 2.0: mining clinically relevant drug-gene interactions," *Nucleic Acids Res*, vol. 44, no. D1, pp. D1036–44, 2016.

[148] RStudio, "RStudio: Integrated development environment for R," *The Journal of Wildlife Management*, vol. 75, no. 8. 2012.

[149] R Developement Core Team, "R: A Language and Environment for Statistical Computing," *R Foundation for Statistical Computing*, vol. 1. p. 409, 2015.

[150] U. Mudunuri, A. Che, M. Yi, and R. M. Stephens, "bioDBnet: The biological database network," *Bioinformatics*, vol. 25, no. 4, pp. 555–556, 2009.

[151] M. P. Fay, "Confidence intervals that match fisher's exact or blaker's exact tests," *Biostatistics*, vol. 11, no. 2, pp. 373–374, 2010.

[152] S. Holm, "A simple sequentially rejective multiple test procedure.," *Scand. J. Stat.*, vol. 6, no. 2, pp. 65–70, 1979.

[153] P. Atadja, L. Gao, P. Kwon, N. Trogani, H. Walker, M. Hsu, L. Yeleswarapu, N. Chandramouli, L. Perez, R. Versace, A. Wu, L. Sambucetti, P. Lassota, D. Cohen, K. Bair, A. Wood, and S. Remiszewski, "Selective Growth Inhibition of Tumor Cells by a Novel Histone Deacetylase Inhibitor, NVP-LAQ824," *Cancer Res.*, vol. 64, no. 2, pp. 689–695, 2004.

[154] M. Galli, S. Salmoiraghi, J. Golay, A. Gozzini, C. Crippa, N. Pescosta, and A. Rambaldi, "A phase II multiple dose clinical trial of histone deacetylase inhibitor ITF2357 in patients with relapsed or progressive multiple myeloma," *Ann. Hematol.*, vol. 89, no. 2, pp. 185–190, 2010.

[155] A. M. V. Evens J. M.Harb, W. A.Gordon, L. I.Langdon, R.Grant, B.Sprague, J.Plasencia, C.Sirisawad, M.Yue, J.Luan, Y.Siek, A.Zhou, L.Balasubramanian, S.Bartlett, N. L., "A phase II multicenter study of the histone deacetylase inhibitor (HDACi) abexinostat (PCI-24781) in relapsed/ refractory follicular lymphoma (FL) and mantle cell lymphoma (MCL) [Abstract 55]," *54th American Society of Hematology Annual Meeting and Exposition*. 2012.

[156] O. A. O'Connor, S. Horwitz, T. Masszi, A. Van Hoof, P. Brown, J. Doorduijn, G. Hess, W. Jurczak, P. Knoblauch, S. Chawla, G. Bhat, M. R. Choi, J. Walewski, K. Savage, F. Foss, L. F. Allen, and A. Shustov, "Belinostat in patients with relapsed or refractory peripheral T-cell lymphoma: Results of the pivotal phase II BELIEF (CLN-19) study," *J. Clin. Oncol.*, vol. 33, no. 23, pp. 2492–2499, 2015.

[157] E. Galanis, K. A. Jaeckle, M. J. Maurer, J. M. Reid, M. M. Ames, J. S. Hardwick, J. F. Reilly, A. Loboda, M. Nebozhyn, V. R. Fantin, V. M. Richon, B. Scheithauer, C. Giannini, P. J. Flynn, D. F. Moore, J. Zwiebel, and J. C. Buckner, "Phase II trial of vorinostat in recurrent glioblastoma multiforme: a north central cancer treatment group study.," *J. Clin. Oncol.*, vol. 27, no. 12, pp. 2052–8, 2009.

[158] T. Reid, F. Valone, W. Lipera, D. Irwin, W. Paroly, R. Natale, S. Sreedharan, H. Keer, B. Lum, F. Scappaticci, and A. Bhatnagar, "Phase II trial of the histone deacetylase inhibitor pivaloyloxymethyl butyrate (Pivanex, AN-9) in advanced non-small cell lung cancer," *Lung Cancer*, vol. 45, pp. 381–386, 2004.

[159] M. E. Cudkowicz, P. L. Andres, S. a Macdonald, R. S. Bedlack, R. Choudry, R. H. Brown, H. Zhang, D. a Schoenfeld, J. Shefner, S. Matson, W. R. Matson, and R. J. Ferrante, "Phase 2 study of sodium phenylbutyrate in ALS.," *Amyotroph. Lateral Scler.*, vol. 10, no. 2, pp. 99–106, 2009.

[160] D. J. DeAngelo, a Spencer, K. N. Bhalla, H. M. Prince, T. Fischer, T. Kindler, F. J. Giles, J. W. Scott, K. Parker, a Liu, M. Woo, P. Atadja, K. K. Mishra, and O. G. Ottmann, "Phase Ia/II, two-arm, open-label, dose-escalation study of oral panobinostat administered via two dosing schedules in patients with advanced hematologic malignancies.," *Leukemia*, vol. 27, no. 8, pp. 1628–36, 2013.

[161] C. U. Johannessen and S. I. Johannessen, "Valproate : Past , Present , and Future Mechanisms of Action," vol. 9, no. 2, pp. 199–216, 2003.

[162] W. Löscher, "Anticonvulsant and biochemical effects of inhibitors of gaba aminotransferase and valproic acid during subchronic treatment in mice," *Biochem. Pharmacol.*, vol. 31, no. 5, pp. 837–842, 1982.

[163] J. H. Ha, D. U. Lee, J. T. Lee, J. S. Kim, C. S. Yong, J. A. Kim, J. S. Ha, and K. Huh, "4-Hydroxybenzaldehyde from Gastrodia elata B1. is active in the antioxidation and GABAergic neuromodulation of the rat brain," *J. Ethnopharmacol.*, vol. 73, no. 1–2, pp. 329–333, 2000.

[164] Y. Sher, A. C. Miller Cramer, A. Ament, S. Lolak, and J. R. Maldonado, "Valproic Acid as a Treatment for Hyperactive Delirium: Rationale and Literature Review," *Psychosomatics*, 2015.

[165] M. Haigentz, M. Kim, C. Sarta, J. Lin, R. S. Keresztes, B. Culliney, A. G. Gaba, R. V. Smith, G. I. Shapiro, L. R. Chirieac, J. M. Mariadason, T. J. Belbin, J. M. Greally, J. J. Wright, and R. I. Haddad, "Phase II trial of the histone deacetylase inhibitor romidepsin in patients with recurrent/metastatic head and neck cancer," *Oral Oncol.*, vol. 48, no. 12, pp. 1281–1288, 2012.

[166] R. L. Piekarz, R. Frye, M. Turner, J. J. Wright, S. L. Allen, M. H. Kirschbaum, J. Zain, H. M. Prince, J. P. Leonard, L. J. Geskin, C. Reeder, D. Joske, W. D. Figg, E. R. Gardner, S. M. Steinberg, E. S. Jaffe, M. Stetler-Stevenson, S. Lade, A. T. Fojo, and S. E. Bates, "Phase II multi-institutional trial of the histone deacetylase inhibitor romidepsin as monotherapy for patients with cutaneous T-cell lymphoma.," *J. Clin. Oncol.*, vol. 27, no. 32, pp. 5410–7, 2009.

[167] U. Banerji, L. Van Doorn, D. Papadatos-Pastos, R. Kristeleit, P. Debnam, M. Tall, A. Stewart, F. Raynaud, M. D. Garrett, M. Toal, L. Hooftman, J. S. De Bono, J. Verweij, and F. a L. M. Eskens, "A phase I pharmacokinetic and pharmacodynamic study of CHR-3996, an oral class I selective histone deacetylase inhibitor in refractory solid tumors," *Clin. Cancer Res.*, vol. 18, no. 7, pp. 2687–2694, 2012.

[168] J. Knipstein and L. Gore, "Entinostat for treatment of solid tumors and hematologic malignancies.," *Expert Opin. Investig. Drugs*, vol. 20, no. 10, pp. 1455–67, 2011.

[169] S. J. Isakoff, J. c. Bendell, A. Cervantes, J.-C. Soria, L. R. Molife, and S. M. et al. Sanabria-Bohorquez, "Phase Ib dose-escalation study of an Akt inhibitor ipatasertib (Ipat) in combination with docetaxel (Doc) or paclitaxel (Pac) in patients (pts) with metastatic breast cancer (MBC)," *Cancer Res.*, vol. 75, no. 9 Supplement, pp. P6–12–02–P6–12–02, 2015.

[170] Y. Cheng, X. Ren, Y. Zhang, R. Patel, A. Sharma, H. Wu, G. P. Robertson, L. Yan, E. Rubin, and J. M. Yang, "eEF-2 kinase dictates cross-talk between autophagy and apoptosis induced by Akt

inhibition, thereby modulating cytotoxicity of novel Akt inhibitor MK-2206," *Cancer Res.*, vol. 71, no. 7, pp. 2654–2663, 2011.

[171] Y. Oki, M. Fanale, J. Romaguera, L. Fayad, N. Fowler, A. Copeland, F. Samaniego, L. W. Kwak, S. Neelapu, M. Wang, L. Feng, and A. Younes, "Phase II study of an AKT inhibitor MK2206 in patients with relapsed or refractory lymphoma," *Br. J. Haematol.*, vol. 171, no. 4, pp. 463–470, 2015.

[172] Y. Oba and N. A. Lone, "Efficacy and safety of roflumilast in patients with chronic obstructive pulmonary disease: A systematic review and meta-analysis," *Ther. Adv. Respir. Dis.*, vol. 7, no. 1, pp. 13–24, 2013.

[173] J.-H. Yan, W.-J. Gu, and L. Pan, "Efficacy and safety of roflumilast in patients with stable chronic obstructive pulmonary disease: a meta-analysis.," *Pulm. Pharmacol. Ther.*, vol. 27, no. 1, pp. 83–9, 2014.

[174] D. P. Tashkin, "Roflumilast : the new orally active, selective phophodiesterase-4 inhibitor, for the treatment of COPD.," *Expert Opin. Pharmacother.*, vol. 15, no. 1, pp. 85–96, 2014.

[175] P. Rolan, M. R. Hutchinson, and K. W. Johnson, "Ibudilast: a review of its pharmacology, efficacy and safety in respiratory and neurological disease," *Expert Opin Pharmacother*, vol. 10, no. 17, pp. 2897–2904, 2009.

[176] H. M. Vesterinen, P. Connick, C. M. J. Irvine, E. S. Sena, K. J. Egan, G. G. Carmichael, A. Tariq, S. Pavitt, J. Chataway, M. R. Macleod, and S. Chandran, "Drug repurposing: A systematic approach to evaluate candidate oral neuroprotective interventions for secondary progressive multiple sclerosis," *PLoS One*, vol. 10, no. 4, pp. 1–18, 2015.

[177] G. T. Laschober, D. Ruli, E. Hofer, C. Muck, D. Carmona-Gutierrez, J. Ring, E. Hutter, C. Ruckenstuhl, L. Micutkova, R. Brunauer, A. Jamnig, D. Trimmel, D. Herndler-Brandstetter, S. Brunner, C. Zenzmaier, N. Sampson, M. Breitenbach, K. U. Fr??hlich, B. Grubeck-Loebenstein, P. Berger, M. Wieser, R. Grillari-Voglauer, G. G. Thallinger, J. Grillari, Z. Trajanoski, F. Madeo, G. Lepperdinger, and P. Jansen-D??rr, "Identification of evolutionarily conserved genetic regulators of cellular aging," *Aging Cell*, vol. 9, no. 6, pp. 1084–1097, 2010.

[178] J. Deelen, H. W. Uh, R. Monajemi, D. Van Heemst, P. E. Thijssen, S. B??hringer, E. B. Van Den Akker, A. J. M. De Craen, F. Rivadeneira, A. G. Uitterlinden, R. G. J. Westendorp, J. J. Goeman, P. E. Slagboom, J. J. Houwing-Duistermaat, and M. Beekman, "Gene set analysis of GWAS data for human longevity highlights the relevance of the insulin/IGF-1 signaling and telomere maintenance pathways," *Age (Omaha).*, vol. 35, no. 1, pp. 235–249, 2013.

[179] W. M. Passtoors, M. Beekman, J. Deelen, R. van der Breggen, A. B. Maier, B. Guigas, E. Derhovanessian, D. van Heemst, A. J. M. de Craen, D. A. Gunn, G. Pawelec, and P. E. Slagboom, "Gene expression analysis of mTOR pathway: Association with human longevity," *Aging Cell*, vol. 12, no. 1, pp. 24–31, 2013.

[180] L. Pawlikowska, D. Hu, S. Huntsman, A. Sung, C. Chu, J. Chen, A. H. Joyner, N. J. Schork, W. C. Hsueh, A. P. Reiner, B. M. Psaty, G. Atzmon, N. Barzilai, S. R. Cummings, W. S. Browner, P. Y. Kwok, and E. Ziv, "Association of common genetic variation in the insulin/IGF1 signaling pathway with human longevity," *Aging Cell*, vol. 8, no. 4, pp. 460–472, 2009.

[181] M. Nygaard, M. Soerensen, F. Flachsbart, J. Mengel-From, Q. Tan, S. Schreiber, A. Nebel, K. Christensen, and L. Christiansen, "AKT1 fails to replicate as a longevity-associated gene in Danish and German nonagenarians and centenarians.," *Eur. J. Hum. Genet.*, vol. 21, no. 5, pp. 574–7, 2013.

[182] F. Di Cianni, D. Campa, F. Tallaro, C. Rizzato, F. De Rango, R. Barale, G. Passarino, F. Canzian, F. Gemignani, A. Montesanto, S. Landi, and G. Rose, "MAP3K7 and GSTZ1 are associated with human longevity: A two-stage case-control study using a multilocus genotyping," *Age (Omaha).*, vol. 35, no. 4, pp. 1357–1366, 2013.

[183] J. Deelen, M. Beekman, H. W. Uh, Q. Helmer, M. Kuningas, L. Christiansen, D. Kremer, R. van der Breggen, H. E. D. Suchiman, N. Lakenberg, E. B. van den Akker, W. M. Passtoors, H. Tiemeier, D. van Heemst, A. J. de Craen, F. Rivadeneira, E. J. de Geus, M. Perola, F. J. van der Ouderaa, D. A. Gunn, D. I. Boomsma, A. G. Uitterlinden, K. Christensen, C. M. van Duijn, B. T. Heijmans, J. J. Houwing-Duistermaat, R. G. J. Westendorp, and P. E. Slagboom, "Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited," *Aging Cell*, vol. 10, no. 4, pp. 686–698, 2011.

[184] N. Li, H. Luo, X. Liu, S. Ma, H. Lin, and R. Chen, "Association study of polymorphisms in FOXO3, AKT1 and IGF-2R genes with human longevity in a Han Chinese population," *Oncotarget*, vol. 7, no. 1, 2015.

[185] A. Nojima, M. Yamashita, Y. Yoshida, I. Shimizu, H. Ichimiya, N. Kamimura, Y. Kobayashi, S. Ohta, N. Ishii, and T. Minamino, "Haploinsufficiency of Akt1 Prolongs the Lifespan of Mice," *PLoS One*, vol. 8, no. 7, 2013.

[186] B. Hamilton, Y. Dong, M. Shindo, W. Liu, I. Odell, G. Ruvkun, and S. S. Lee, "A systematic RNAi screen for longevity genes in C. elegans," *Genes and Development*, vol. 19, no. 13. pp. 1544–1555, 2005.

[187] M. Weinberger, L. Feng, A. Paul, D. L. Smith, R. D. Hontz, J. S. Smith, M. Vujcic, K. K. Singh, J. A. Huberman, and W. C. Burhans, "DNA replication stress is a determinant of chronological lifespan in budding yeast," *PLoS One*, vol. 2, no. 8, 2007.

[188] V. D. Longo, "The Ras and Sch9 pathways regulate stress resistance and longevity," in *Experimental Gerontology*, 2003, vol. 38, no. 7, pp. 807–811.

[189] P. Fabrizio, F. Pozza, S. D. Pletcher, C. M. Gendron, and V. D. Longo, "Regulation of longevity and stress resistance by Sch9 in yeast.," *Science*, vol. 292, no. 5515, pp. 288–290, 2001.

[190] W. S. Chen, P. Z. Xu, K. Gottlob, M. L. Chen, K. Sokol, T. Shiyanova, I. Roninson, W. Weng, R. Suzuki, K. Tobe, T. Kadowaki, and N. Hay, "Growth retardation and increased apoptosis in mice with homozygous disruption of the akt1 gene," *Genes Dev.*, vol. 15, no. 17, pp. 2203–2208, 2001.

[191] B. Dummler, O. Tschopp, D. Hynx, Z.-Z. Yang, S. Dirnhofer, and B. A. Hemmings, "Life with a single isoform of Akt: mice lacking Akt2 and Akt3 are viable but display impaired glucose homeostasis and growth deficiencies.," *Mol. Cell. Biol.*, vol. 26, no. 21, pp. 8042–51, 2006.

[192] B. Rogina, S. L. Helfand, and S. Frankel, "Longevity regulation by Drosophila Rpd3 deacetylase and caloric restriction," *Science (80-.).*, vol. 298, no. 5599, p. 1745, 2002.

[193] S. Kim, A. Benguria, C. Y. Lai, and S. M. Jazwinski, "Modulation of life-span by histone deacetylase genes in Saccharomyces cerevisiae.," *Mol. Biol. Cell*, vol. 10, no. 10, pp. 3125–36, 1999.

[194] J. C. Jiang, J. Wawryn, H. M. C. Shantha Kumara, and S. M. Jazwinski, "Distinct roles of processes modulated by histone deacetylases Rpd3p, Hda1p, and Sir2p in life extension by caloric restriction in yeast," *Exp. Gerontol.*, vol. 37, no. 8–9, pp. 1023–1030, 2002.

[195] Y. Kim and H. Sun, "Functional genomic approach to identify novel genes involved in the regulation of oxidative stress resistance and animal lifespan," *Aging Cell*, vol. 6, no. 4, pp. 489–503, 2007.

[196] C. Rallis, L. López-Maury, T. Georgescu, V. Pancaldi, and J. Bähler, "Systematic screen for mutants resistant to TORC1 inhibition in fission yeast reveals genes involved in cellular ageing and growth.," *Biol. Open*, vol. 3, no. 2, pp. 161–71, 2014.

[197] P. Fabrizio, S. Hoon, M. Shamalnasab, A. Galbani, M. Wei, G. Giaever, C. Nislow, and V. D. Longo, "Genome-wide screen in Saccharomyces cerevisiae identifies vacuolar protein sorting, autophagy, biosynthetic, and tRNA methylation genes involved in life span regulation," *PLoS Genet.*, vol. 6, no. 7, pp. 1–14, 2010.

[198] J. J. Tong, S. E. Schriner, D. McCleary, B. J. Day, and D. C. Wallace, "Life extension through neurofibromin mitochondrial regulation and antioxidant therapy for neurofibromatosis-1 in Drosophila melanogaster.," *Nat. Genet.*, vol. 39, no. 4, pp. 476–485, 2007.

[199] R. Tacutu, T. Craig, A. Budovsky, D. Wuttke, G. Lehmann, D. Taranukha, J. Costa, V. E. Fraifeld, and J. P. de Magalhães, "Human Ageing Genomic Resources: integrated databases and tools for the biology and genetics of ageing.," *Nucleic Acids Res.*, vol. 41, no. Database issue, pp. D1027–33, 2013.

[200] D. Wuttke, R. Connor, C. Vora, T. Craig, Y. Li, S. Wood, O. Vasieva, R. Shmookler Reis, F. Tang, and J. P. de Magalhães, "Dissecting the Gene Network of Dietary Restriction to Identify Evolutionarily Conserved Pathways and New Functional Genes," *PLoS Genet.*, vol. 8, no. 8, 2012.

[201] A. Budovsky, T. Craig, J. Wang, R. Tacutu, A. Csordas, J. Lourenço, V. E. Fraifeld, and J. P. De Magalhães, "LongevityMap: A database of human genetic variants associated with longevity," *Trends in Genetics*, vol. 29, no. 10. pp. 559–560, 2013.

[202] M. Kaeberlein, B. Jegalian, and M. McVey, "AGEID: A database of aging genes and interventions," *Mechanisms of Ageing and Development*, vol. 123, no. 8. pp. 1115–1119, 2002.

[203] R. Hühne, T. Thalheim, and J. Sühnel, "AgeFactDB - The JenAge Ageing Factor Database - Towards data integration in ageing research," *Nucleic Acids Res.*, vol. 42, no. D1, 2014.

[204] M. Guasch-Ferré, M. Bulló, R. Estruch, D. Corella, M. A. Martínez-González, E. Ros, M. Covas, F. Arós, E. Gómez-Gracia, M. Fiol, J. Lapetra, M. Á. Muñoz, L. Serra-Majem, N. Babio, X. Pintó, R. M. Lamuela-Raventós, V. Ruiz-Gutiérrez, and J. Salas-Salvadó, "Dietary magnesium intake is inversely associated with mortality in adults at high cardiovascular disease risk.," *J. Nutr.*, vol. 144, no. 1, pp. 55–60, 2014.

[205] R. Surveyed, "Is there a reproducibility crisis?," *Nature*, vol. 533, pp. 452–454, 2016.

[206] J. B. L. Bard and S. Y. Rhee, "Ontologies in biology: design, applications and future challenges.," *Nat. Rev. Genet.*, vol. 5, no. 3, pp. 213–222, 2004.

[207] The Gene Ontology Consortium, "Gene Ontology Consortium: going forward," *Nucleic Acids Res.*, vol. 43, no. D1, pp. D1049–D1056, 2015.

[208] P. D. Thomas, A. Kejariwal, N. Guo, H. Mi, M. J. Campbell, A. Muruganujan, and B. Lazareva-Ulitsky, "Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools," *Nucleic Acids Res.*, vol. 34, no. WEB. SERV. ISS., 2006.

[209] H. Mi, S. Poudel, A. Muruganujan, J. T. Casagrande, and P. D. Thomas, "PANTHER version 10: expanded protein families and functions, and analysis tools.," *Nucleic Acids Res.*, vol. 44, no. D1, pp. D336–42, 2015.

[210] G. Bindea, B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W. H. Fridman, F. Pagès, Z. Trajanoski, and J. Galon, "ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks," *Bioinformatics*, vol. 25, no. 8, pp. 1091–1093, 2009.

[211] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker, "Cytoscape: A software Environment for integrated models of biomolecular interaction networks," *Genome Res.*, vol. 13, no. 11, pp. 2498–2504, 2003.

[212] D. Croft, A. F. Mundo, R. Haw, M. Milacic, J. Weiser, G. Wu, M. Caudy, P. Garapati, M. Gillespie, M. R. Kamdar, B. Jassal, S. Jupe, L. Matthews, B. May, S. Palatnik, K. Rothfels, V. Shamovsky, H. Song, M. Williams, E. Birney, H. Hermjakob, L. Stein, and P. D'Eustachio, "The Reactome pathway knowledgebase," *Nucleic Acids Res.*, vol. 42, no. D1, 2014.

[213] A. R. Pico, T. Kelder, M. P. Van Iersel, K. Hanspers, B. R. Conklin, and C. Evelo, "WikiPathways: Pathway Editing for the People," *PLoS Biology*, vol. 6, no. 7. pp. 1403–1407, 2008.

[214] Y. Wang, T. Suzek, J. Zhang, J. Wang, S. He, T. Cheng, B. A. Shoemaker, A. Gindulyte, and S. H. Bryant, "PubChem BioAssay: 2014 update," *Nucleic Acids Res.*, vol. 42, no. D1, 2014.

[215] S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang, and S. H. Bryant, "PubChem Substance and Compound databases.," *Nucleic Acids Res.*, vol. 44, no. D1, pp. D1202–13, 2016.

[216] A. J. Williams, "Chemspider: A Platform for Crowdsourced Collaboration to Curate Data Derived From Public Compound Databases," in *Collaborative Computational Technologies for Biomedical Research*, 2011, pp. 363–386.

[217] ChemicalComputingGroupInc., "Molecular Operating Environment (MOE)," *Sci. Comput. Instrum.*, vol. 22, no. 1, p. 32, 2004.

[218] D. J. Stekhoven and P. Buehlmann, "Missforest-Non-parametric missing value imputation for mixed-type data," *Bioinformatics*, vol. 28, no. 1, pp. 112–118, 2012.

[219] S. Oba, M. Sato, I. Takemasa, M. Monden, K. Matsubara, and S. Ishii, "A Bayesian missing value estimation method for gene expression profile data," *Bioinformatics*, vol. 19, no. 16, pp. 2088–2096, 2003.

[220] M. Kuhn, D. Szklarczyk, S. Pletscher-Frankild, T. H. Blicher, C. Von Mering, L. J. Jensen, and P. Bork, "STITCH 4: Integration of protein-chemical interactions with user data," *Nucleic Acids Res.*, vol. 42, no. D1, 2014.

[221] M. Petrascheck, X. Ye, and L. B. Buck, "An antidepressant that extends lifespan in adult Caenorhabditis elegans.," *Nature*, vol. 450, no. 7169, pp. 553–556, 2007.

[222] M. N. Wright and A. Ziegler, "ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R," *arXiv Prepr.*, 2015.

[223] W. G. Touw, J. R. Bayjanov, L. Overmars, L. Backus, J. Boekhorst, M. Wels, and A. F. T. Sacha van Hijum, "Data mining in the life science swith random forest: A walk in the park or lost in the jungle?," *Brief. Bioinform.*, vol. 14, no. 3, pp. 315–326, 2013.

[224] A. Statnikov, L. Wang, and C. Aliferis, "A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification," *BMC Bioinformatics*, vol. 9, no. 1, p. 319, 2008.

[225] S.-K. Jung, B. Aleman-Meza, C. Riepe, and W. Zhong, "QuantWorm: a comprehensive software package for Caenorhabditis elegans phenotypic assays.," *PLoS One*, vol. 9, no. 1, p. e84830, Jan. 2014.

[226] J. O. Holloszy, "Exercise increases average longevity of female rats despite increased food intake and no growth retardation.," *J. Gerontol.*, vol. 48, no. 3, pp. B97–100, 1993.

[227] S. N. Austad and K. E. Fischer, "Sex Differences in Lifespan," *Cell Metab.*, vol. 23, no. 6, pp. 1022–1033, 2016.

[228] R. V. Sekhar, S. G. Patel, A. P. Guthikonda, M. Reid, A. Balasubramanyam, G. E. Taffet, and F. Jahoor, "Deficient synthesis of glutathione underlies oxidative stress in aging and can be corrected by dietary cysteine and glycine supplementation," *Am. J. Clin. Nutr.*, vol. 94, no. 3, pp. 847–853, 2011.

[229] W. C. Orr, S. N. Radyuk, L. Prabhudesai, D. Toroser, J. J. Benes, J. M. Luchak, R. J. Mockett, I. Rebrin, J. G. Hubbard, and R. S. Sohal, "Overexpression of glutamate-cysteine ligase extends life span in Drosophila melanogaster," *J. Biol. Chem.*, vol. 280, no. 45, pp. 37331–37338, 2005.

[230] H.-H. Xu, T. Su, and Y. Xue, "Histone H3 N-terminal acetylation sites especially K14 are important for rDNA silencing and aging.," *Sci. Rep.*, vol. 6, p. 21900, 2016.

[231] H. S. Shim and V. D. Longo, "A protein restriction-dependent sulfur code for longevity," *Cell*, vol. 160, no. 1–2, pp. 15–17, 2015.

[232] J. Snoek, H. Larochelle, and R. P. Adams, “Practical Bayesian Optimization of Machine Learning Algorithms,” *Adv. Neural Inf. Process. Syst.*, vol. 25, pp. 2960–2968, 2012.