

## Review

# Optimizing Translational Research for Exceptional Health and Life Span: A Systematic Narrative of Studies to Identify Translatable Therapeutic Target(s) for Exceptional Health Span in Humans

Nalini Raghavachari, PhD,\*<sup>✉</sup> Beth Wilmot, PhD, and Chhanda Dutta, PhD

Division of Geriatrics & Clinical Gerontology, National Institute on Aging, NIH, Bethesda, Maryland, USA.

\*Address correspondence to: Nalini Raghavachari, PhD, Division of Geriatrics & Clinical Gerontology, National Institute on Aging, NIH, Gateway Building, Suite 3C307, 7201 Wisconsin Avenue, Bethesda, MD 20892-9205, USA. E-mail: [nalini.raghavachari@nih.gov](mailto:nalini.raghavachari@nih.gov)

Received: January 26, 2022; Editorial Decision Date: March 6, 2022

**Decision Editor:** Lewis A. Lipsitz, MD, FGSA

## Abstract

**Background:** Exceptional longevity as manifested by the lower incidence and delayed onset of age-related disabilities/diseases that include cardiovascular disease, Alzheimer's disease, cancer is believed to be influenced by inherent protective molecular factors in exceptionally long-lived individuals. Unraveling these protective factors could lead to the discovery of therapeutic target(s) and interventions to promote healthy aging.

**Methods:** In this context, the National Institute on Aging has established a collection of translational longevity research projects (ie, the Long-Life Family Study, the Longevity Consortium, Longevity Genomics, and the Integrative Longevity Omics) which are generating large omics data sets spanning the human genome to phenome and have embarked on cross-species multiomic data analyses integrating human and nonhuman species that display wide variation in their life spans.

**Results:** It is expected that these studies will discover key signaling pathways that influence exceptional health span and identify therapeutic targets for translation to enhance health and life span. Other efforts related to translational longevity research include the “Comprehensive Evaluation of Aging-Related Clinical Outcomes and Geroproteins study,” which focuses on potential effects in humans of polypeptides/proteins whose circulating levels change with age, and for which experimental evidence indicates reversal or acceleration of aging changes. The “Predictive Human Mechanistic Markers Network” is devoted to the development of predictive markers of aging, for target engagement when testing novel interventions for healthy aging.

**Conclusion:** We describe here the significance, the unique study design, categories of data sets, analytical strategies, and a data portal to facilitate open science and sharing of resources from these longevity studies to identify and validate potential therapeutic targets for healthy aging.

**Keywords:** Longevity, Healthy Aging, Protective factors, Translation, multi-omics

Aging, manifested by a gradual decline of normal physiological functions over time, increases susceptibility to major age-related chronic diseases that include atherosclerosis, most cancers, mild cognitive impairment, dementias, Parkinson's disease, Alzheimer's disease, and other neurodegenerative diseases, type 2 diabetes, renal dysfunction, osteoarthritis, blindness, frailty, chronic obstructive pulmonary disease, sarcopenia, and many others (1–4). There are, however, individuals who survive well beyond average life expectancy and compress the period of time that they experience age-related disease(s) and/or disability toward the end of their very relatively long lives

(3,5). Epidemiological and population-based longevity studies across the world have shown that exceptional longevity and health span are highly correlated. Long-lived individuals exhibit reduced morbidity and delayed onset of age-related disabilities even though they may be carriers of risk factors for certain diseases, suggesting that they may possess genetically mediated factors that modulate fundamental biological processes underlying aging and protect them from developing many age-related diseases. Exceptional longevity represents an extreme phenotype and the ability to achieve exceptional health span is likely to be influenced by differing domain-specific factors that

affect preservation of performance in individual physiologic systems (eg, respiratory, cardiovascular, immune) or functional domains (eg, mobility, cognition) as well as biological processes that span those systems (eg, energetics, genetic instability, cell senescence, and stem cell exhaustion). While the traditional approach to discovering treatments for aging-related diseases and syndromes has focused on the underlying causes and mechanisms, there are valuable insights to be gained from also discovering protective genetic and lifestyle factors in individuals who show relatively slow rates of aging changes, markedly delay, or escape aging-related diseases and disability, and subsequently achieve extreme old ages. Understanding the functional effects of protective human genetic variants and cellular factors associated with exceptional longevity could lead to the identification of novel targets for interventions to mimic their favorable effects. Hence, studies on exceptionally long-lived individuals may be extremely valuable for identifying cellular factors and biological processes that protect individuals from onset of age-related diseases.

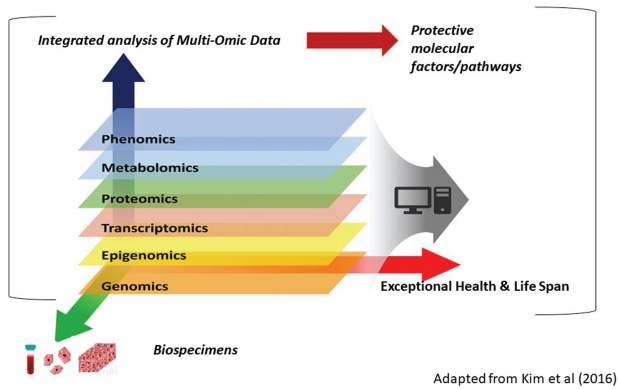
In support of the above premise, the National Institute on Aging (NIA) has facilitated research seeking to identify longevity-associated genetic variants (LAVs), both common and rare, via the funding of genome wide association studies (GWAS) and sequencing studies focusing on exceptionally long-lived individuals. Relevant GWAS and sequencing studies pursued so far have suggested that identifying genome-wide significant LAVs is not trivial and requires analysis of large sample-sized human cohorts. GWAS analyses of exceptional longevity and healthy aging in humans have identified variants of the apolipoprotein E (ApoE) and the gene FOXO3, encoding the transcription factor forkhead box (FOXO3A) genes as modestly associated with longevity and health span. However, there is substantial variation in the impact of these variants upon survival according to ethnicity, sex, and age of survival (6–12). Given that exceptional longevity is likely a complex genetic trait, it is not surprising that polygenic risk scores (13), candidate genes (14), and agnostically constructed genetic signatures (5) suggest that single protective gene variants have weak-to-modest effects upon survival to extreme age. However, the right combination of such variants may have strong effects. It is very likely that specific combinations of genetic variants and lifestyle choices according to sex, ethnicity, and rarity of survival (eg, to 100, 105, or 110+ years) explain why centenarians (1 per 5 000), semisupercentenarians (1 per 250 000, ages 105–109 years), and supercentenarians (1 per 5 million) are so rare. There is also evidence that the more extreme the age, the more phenotypically homogeneous the sample of centenarians is in terms of age of onset of aging-related diseases and disability, and the more powerful the genetic influence is upon survival (5).

In parallel, complementary hypothesis-driven candidate gene studies involving model organisms have identified several genes associated with longevity or decreased mortality rates. Candidate gene approaches considered in human association studies have been motivated in part by orthologs of the genes harboring those variants in model organisms (15). For example, variants in or near many lipoprotein metabolism genes may affect longevity-related phenotypes and include the ApoE, FOXO, insulin/insulin like growth factor (IGF1) signaling pathway genes, lamin A/C (LMNA), RNA editing genes, and heat shock factor (HSF)227 genes (16–18). This suggests that model organism studies, as well as studies that focus on age-related diseases and factors, like lipoproteins, can inform studies of human longevity and the identification and validation of LAVs. However, it should be noted that many genes shown to influence life span in model organisms have human orthologs that have not been shown to either influence human longevity or harbor variants

associated with human longevity-related phenotypes. In some cases, many genes found to be associated with life span in model organism studies simply do not have good and reliably determined orthologs in humans. This also raises the need for more sophisticated tools that can reliably identify orthologs in humans. Furthermore, it is often underappreciated that even if a gene or set of genes are conserved across species, the elements regulating those genes might not be, suggesting that differences in the roles or impact those genes might have in different species could be pronounced (19).

It is now widely recognized that use of GWAS for human longevity translational research is limited by the sample sizes of relevant cohorts (20–22), and there is a need to exploit complementary assays and approaches to derive functional context of LAVs. Another limitation in the development of pharmacologic interventions based on GWAS findings is the need to explore the network of biological pathway(s) in which the identified variants participate. The translation of GWAS findings has also been challenging as most single-nucleotide polymorphisms (SNPs) fall within noncoding regions, with lack of knowledge on functional or mechanistic action on the variant of interest at the physiologic level. Thus, it appears that strategies based on genetics alone present substantial limitations as many gene products have multiple targets that interact with each other and with multiple genes. Herein, a multidimensional view of the genome has evolved to be beneficial to evaluate the functional effects of the variants of interest. The relationships of genetic variants to expression profiles of RNA, proteins, and metabolites that influence longevity and health are also complex. Many biological processes are highly dynamic, and their regulation and functionality involve a multitude of interactions between the intertwined genome, epigenome, transcriptome, proteome, and metabolome (23). The advent of genomic technologies has made it possible to collect multiple highly entangled layers of cellular information starting from genome to phenome. These layers are highly intertwined and regulate biological processes involving a multitude of interactions across layers, particularly in a complex trait like exceptional longevity. Integrative data analyses of these layers could reveal mechanisms that remain concealed in a single-layered approach. Thus, there is a need for multiomics/integrative approaches (ie, omics profiles) to yield better predictors of healthy aging phenotypes, identify protective molecular factors/biological pathways influencing exceptional health and life span, and to guide the assessment of drugs targeted at specific molecules. Figure 1 illustrates a simple model of integrated multiomic analysis.

At present most of the life and health span studies across populations are conducted in humans due to the possibility of collection of extensive and immensely valuable physiologic, clinical, and pharmacologic data. However, the modest difference in life span observed in multiple human populations along with other stochastic factors, environmental exposures termed as “exposomes,” and other genetic variants within the cohort can limit the detection of biological factors associated with exceptional life span. Added to this is the inaccessibility of tissues from humans and the need to wait to harvest samples postmortem to understand tissue-specific biology. As a complementary approach, comparative biology studies may be more powerful through analyses of nonhuman species that show wide variation in life span across species starting from 2 years as in rodents to 210 years as in the bowhead whale, and even within species as in primates (24–29). This potentially greater signal strength may improve insights into crucial factors influencing life span and health span without requiring the large sample sizes of human GWAS. In addition, tissue-specific information difficult to procure in



**Figure 1.** Integrated multiomic approach for exceptional health and life span. Multidimensional approaches to systems understanding of exceptional life span. Given the multifaceted nature of the healthy aging process, multidimensional approaches are required for the systems understanding of the healthy aging process to identify protective molecular factors and biological pathways governing exceptional life span from blood, cells, and tissues of human and nonhuman species. The various molecular layers encompassing genomic, epigenomic, transcriptomic, proteomic, metabolomic, and phenomic data, derived from high-throughput omics technologies are essential for an in-depth analysis by novel integrative approaches across species to identify protective molecular factors and biological pathways that influence exceptional health and life span.

human studies can be obtained in comparative biology studies. Thus, integrating human and nonhuman data may be a more promising approach for identifying the protective genomic factors and associated biological pathway(s) contributing to exceptional health and life span and in turn, the druggable therapeutic target(s) to facilitate healthy aging.

Translational genomics has gained a lot of momentum in the fields of cancer and cardiovascular diseases yet is in infancy within the healthy aging field. Identification and validation of druggable target(s) is the first step in the drug development process. Most importantly, drug target(s) guided by genomics findings have been shown to be more successful in drug discovery (30–32). Once a therapeutic target has been developed, drug discovery relies on extensive screening of chemical libraries to detect compounds with activity against the target. With the availability of extensive drug-target interaction data in public databases such as PubChem and ChemBL, chemoinformatic approaches can be exploited to assess the potential activity of drugs.

The major purpose of the manuscript is to provide an overview of each of the exceptional longevity projects, categories of data sets generated by these longevity studies, and possible analytical tools to identify and validate potential therapeutic targets as shown in Table 1. It is believed that through public sharing of the rich resources of data, biospecimens, analytical tools, and other resources could facilitate and advance the ongoing research on characterization and translation of factors contributing to exceptional life and health span.

### Description of NIA Human Longevity Translational Projects and Related Studies

The NIA is currently supporting a collection of human longevity translational research projects which include the Long-Life Family Study (LLFS), Longevity Consortium (LC), Integrative Longevity

Omics (ILO), and Longevity Genomics (LG) to identify LAVs (both common and rare) as well as omics profiles associated with exceptional healthy aging and longevity. While these longevity projects converge on the major goal of identifying protective factors and translatable targets for healthy aging, interestingly, each study is fairly unique in its study design and analytical strategies. Moreover, these large-scale projects have begun to generate a wealth of multiomic data from exceptionally long-lived individuals and from species with varied life spans which in the near future will provide opportunities for the development and application of novel analytical approaches and tools for data integration and ultimately advance translational research on human longevity. Other NIA-supported research projects pertinent to the identification and translation of protective factors contributing to human health span include the Comprehensive Evaluation of Aging-Related Clinical Outcomes and Geroproteins (CARGO) project and the Predictive Human Mechanistic Markers Network. A description of these different research efforts follows.

### Long-Life Family Study

LLFS (<https://longlifefamilystudy.wustl.edu>), a family-based study, was initiated in 2005 with the major goal of identifying familial rare variants that provide strong protection against age-related disabilities and diseases and contribute to exceptional longevity and exceptionally healthy aging phenotypes. This contrasts with most longitudinal studies where the focus is on the identification of factors contributing to adverse health outcomes. LLFS as described by Wojczynski and others (33) is a multicenter, binational (the United States and Denmark), longitudinal study comprising 539 families with 4 953 participants with ~55% male and female and ~97% Whites, including 2 generations (extremely long-lived parents, their siblings and offspring, and spouses of the offspring) and is focused on determining the degree and patterns of familial transmission and aggregation of exceptionally healthy survival and longevity. LLFS cohort was created applying the “family longevity selection score,” or FLoSS (2), that selects for families with sibships that were rare by virtue of the number of siblings achieving rare percentiles of survival available for study. A FLoSS score of 7 or greater was chosen as the eligibility criteria for the proband generation with a minimum family size of 3 who are willing and able to provide informed consent and participate in the interview and in-home examination including providing a blood sample for serum and DNA extraction. The threshold of 7 was chosen based on the observation that only 0.2% of the FLoSS sibships of the Framingham Heart Study (FHS) meet this threshold, in contrast to over 30% of the LLFS screening families. The second phase (2011–2013) activities of LLFS included the collection of additional phenotypic information via telephone follow-up and development of assays for biospecimens (1). The longitudinal phase of the study began in 2014, with the initiation of a second home visit to all surviving LLFS participants for repeat collection of phenotype data and biospecimens.

Analyses of the LLFS families so far have shown that they are, on average, much healthier than their age-/sex-matched peers and their phenotypes are highly heritable cross-sectionally and longitudinally (6) with different families showing different healthy aging characteristics and profiles in key pathways of healthy aging (cognitive, metabolic, inflammatory, etc.) (7). The prevalence of many common diseases and their risk factors are substantially lower in LLFS than in reference population cohorts such as the FHS and CHS (33). Comprehensive linkage analysis of the LLFS sample identified extremely strong genetic linkage peaks for

**Table 1.** Summary of EL Projects

Study	Research Resources	Planned Analyses	Integration With Other Studies
Long-Life Family Study (LLFS) Long-lived families	WGS Methylomics Transcriptomics (longitudinal) Proteomics Metabolomics (longitudinal) Known serum biomarkers Phenotypic data + cognitive functions	Linkage analyses to identify rare variants Integrated multiomic analyses Identification of mechanistic pathway(s)	LC, LG, and ILO
Longevity Consortium (LC) Human centenarians	Metabolomics	Integrated analysis for longevity omics profiles	LLFS, ILO, and LG
Long-lived individuals (MrOS, SOF, CHS)	Proteomics Chemoinformatics Phenotypic data + cognitive functions	Identification of mechanistic pathway(s) Druggability of longevity pathways	
Cell lines of birds, rodents, and primates	Metabolomics	Identification of human orthologs	
Slow-aging mice—plasma and tissues	Proteomics		
Longevity Genomics (LG)	Publically available data PrediXcan/MetaXcan  Chemoinformatics In vitro models/iPSCs/organoids	Identification of LAV and LAG Identification of eQTLs Mendelian randomization DrugTarget identification Validation of potential targets	LC, LLFS, and ILO
Integrative Longevity Omics (ILO) Centenarians/controls	Transcriptomics Methylomics Proteomics Metabolomics Microbiomics iPSC—centenarians	Integrated omic analyses Identification of mechanistic pathway(s)	LLFS, LC, and LG
Nonhuman species ~100 species of mammal - multiple tissues/cell lines eg, primates, birds, rodents, bowhead whale	Transcriptomics Proteomics Metabolomics WGS	Human orthologs for genes/proteins Life-span-related mechanisms	

*Notes:* eQTL = expression quantitative trait locus; iPSC = induced pluripotent stem cell; LAG = longevity-associated gene; LAV = longevity-associated variant; MrOS = Osteoporotic Fractures in Men; SOF = Study of Osteoporotic Fractures; WGS = whole-genome sequencing.

cross-sectional as well as longitudinal trajectory rates of change for a wide variety of healthy aging domains such as exceptional cognitive performance and lack of Alzheimer’s disease (AD) (8). Preliminary evidence from the Danish Health Registry suggests that, at least in Denmark, the protection persists into the third generation, with significantly lower rates of medical conditions across the disease spectrum (34). The study is now generating whole-genome sequencing (WGS) data on this unique cohort to identify the rare protective variants driving these strong linkage peaks. In summary, LLFS is poised to conduct extensive transcriptomics, methylomics, metabolomics, and proteomics on these selected high-linkage pedigrees, and begin to move from GWAS to integrated multiomic analyses to unravel the biological genes associated with exceptional longevity. LLFS is also in the process of enrolling approximately 800 grandchildren of the proband generation to examine if the grandchildren likely carry more copies of the rare protective alleles prevalent in these families.

### Longevity Consortium

The LC (<https://www.longevityconsortium.org>) began as a planning grant in 2000 to bring together researchers interested in longevity and aging from a variety of disciplines, including laboratory scientists who study longevity in model systems; epidemiologists who have enrolled and followed large cohorts of older people; experts in genomic methods; biostatisticians interested in the analysis of the associations between genetic data and complex traits; demographers who study patterns of mortality and life expectancy; and gerontologists and geriatricians to accomplish the goal of identifying genes that affect longevity and aging in humans (35). The discussions from these multidiscipline scientists led to the establishment of the LC in the next phase of the grant in 2010 to accomplish the initial broad goals of identifying common genetic variants associated with survival in exceptionally long-lived individuals, conducting studies on long-lived mice to identify genes associated with long life span, and studying the associations between telomere length and human

longevity. Specifically, the consortium of scientists focused their initial efforts on identifying SNPs in insulin signaling pathways in association with exceptional survival; correlating telomere length with exceptional survival and the effects of growth hormone and insulin signaling in life-span determination in mouse models and identifying genetic variants associated with cellular stress resistance and longevity in wild-type mice.

In the current phase of the grant that started in 2019, the consortium is pursuing more refined studies focusing on collecting comprehensive phenotypic and blood-based multiomic data from cohort studies with participants who survived to within the oldest 2 percentile of the population, along with a specific centenarian enrollment initiative. In addition, cross-species integrative multiomic analyses on fibroblasts from species of varied life span will also be performed. Most importantly, integrative multiomic data analyses in human and nonhuman species will be conducted as a comparative biology initiative. In order to accomplish the major objective of identifying factors that influence human longevity and health span, the LC is enrolling a new cohort of centenarians, and it is leveraging unique and very large human study cohorts, state-of-the-art molecular assays, and cross-species studies designed to exploit evolutionary orthologous relationships between genes (36). These approaches as well as the employment of novel data analytic methods, systems biology techniques, and chemoinformatic analyses may facilitate data integration and more coherent views of the major molecular pathways and processes influencing human longevity amenable to modulation.

In brief, the LC has proposed to collect multiomic data from large human cohorts that include centenarians, healthy long-lived individuals, longitudinal and pedigree data whenever possible from these cohorts to identify genetic variants (including rare variants) that are associated with human longevity and health span. These variants will also be tested for their presence in individuals with and without specific diseases and for sex-specific differences. In addition to the identification of variants, LC will be evaluating important cellular and molecular phenotypes associated with human longevity and health span using leukocyte gene expression as well as serum proteome and metabolome assays performed on deeply phenotyped centenarians as well as individuals that fall within the oldest 2 percentiles of the population who are/were enrolled in the Osteoporotic Fractures in Men (MrOS) Study, the corresponding study in women known as the Study of Osteoporotic Fractures (SOF), CHS, and the Dynamics of Health and Body Composition Study (Health ABC). Relevant analyses will leverage methods accommodating heterogeneity (eg, sex differences), causal Mendelian randomization (MR) methods, and very large unique LC human cohorts along with public domain data sets.

In the comparative biology component of LC, multiomic data will be generated and analyzed from cells derived from long-lived and short-lived species of birds, rodents, and primates, as well as mice treated with drugs that extend longevity. Data mining from these nonhuman species is expected to discover factors such as proteins, metabolites, gene expression profiles, and tissue-specific molecular pathways associated with exceptional longevity. Where appropriate, the human orthologs of these factors will be tested for association with human longevity and aging using the LC cohorts. In addition, findings from the human studies will motivate studies of orthologous factors in mice and cultured cells. Systems biological approaches will integrate results from various analyses for LAVs, transcriptomic, metabolomics, and proteomic profiles from human and nonhuman species to identify molecular pathways and regulatory networks potentially causally related to healthy human aging and longevity.

As part of the cross-species, cross-program initiative of the LC, proteomic and metabolomic data sets generated by the intervention testing program on mice are being used to develop novel methods and piloting analytical approaches for integrative analyses of omic data sets to understand the aging process. The developed methods could potentially be applied for integrating human cohort data sets. Eventually, from a translational standpoint, chemoinformatic strategies will be employed to assess the druggability (37) of the various genetic and molecular factors found to be associated with human aging and longevity. Several state-of-the-art drug discovery databases will be used to identify approved drugs that could be repurposed for longevity research and identify novel compounds or strategies for developing a novel compound that targets relevant factors associated with exceptional longevity.

In all these analyses, state-of-the-art molecular assays such as high-throughput genotyping, DNA sequencing, transcriptomics, proteomics, and metabolomics, cross-species studies to exploit evolutionary orthologous relationships between genes, both to identify conserved mechanisms affecting longevity and to provide mouse model systems through which pharmacological intervention studies can be pursued. This will be followed by transformative data analytic methods, systems biology approaches, and chemoinformatic analyses to facilitate the integration of the data and results to identify major molecular pathways and processes influencing human longevity.

### ILO Study

There are 2 phases of the ILO (<https://longevityomics.org>) human and comparative species projects. In the first phase, the New England Centenarian Study (NECS) and the Einstein Centenarian and Offspring studies established a standardized deep phenotypic data and biological sample collection (fasting and immediately frozen blood and fecal samples) protocol which they are carrying out as they enroll 700 participants from each study site (total  $n = 1\,400$ ). The protocol is nearly identical to that used by the LC's centenarian project and very similar to the protocol established by the LLFS. ILO also aims to generate induced pluripotent stem cell (iPSC) lines from 15 centenarians, ages >103 years old. iPSCs will be differentiated into hepatocyte, sarcomere (muscle), and neuron cell lines. These human study efforts are paralleled by comparative transcriptomic, proteomic, metabolomic, and microbiomic studies of nonhuman mammalian species of widely different life spans.

In its second phase, ILO will generate transcriptomic, methylomic, metabolomic, proteomic, and microbiomic data from the biological samples collected from centenarians and centenarian offspring. There is also a possibility in future for generating WGS and WES data sets from the centenarian cohort of the ILO. In about a third of the centenarian cohort, earlier samples collected from participants when they were enrolled in the NECS, and the Einstein study will facilitate generation of longitudinal data. Methods for multiomic data integration developed during the first phase will be used to discover molecular profiles that associate with EL and healthy aging phenotypes, including delay of or escape from AD. Integration of molecular profiles derived from functional studies of resiliency in iPSC-derived neurons, sarcomeres, and hepatocytes along with molecular profiles associated with increased life span from multiple species will point to mechanisms leading to the discovery of candidate small molecule and compound therapeutics to enhance health span. iPSC-derived batches of specific cell types will be particularly unique and valuable shared resources to address the challenge of studying cells of various tissue types in very old participants.



In parallel to the human cohorts, studies on long- or short-lived variants within nonhuman laboratory organisms and other nonhuman species comprising 60 divergent species with substantial variation in life span will be conducted. Multiomic data (genomic, transcriptomic, proteomic, and metabolomic) generated from 5 different tissues from these species will be integrated with the human cohort data sets as part of the comparative biology project to understand the molecular mechanisms that influence the rate of aging, resistance to age-related diseases, and to unravel mechanisms responsible for the more than 10-fold differences in life span between species of mammals. These proposed studies are some of the first to champion the notion that the “triangulation” of disparate scientific studies and discoveries, that is, the attempt to unify results from different study designs based on their biological coherence, is the optimal way to advance identification of longevity-related conserved genes, pathways, and targets for longevity-enhancing, geroprotective drugs of relevance to humans (38).

### Longevity Genomics

The LG (<https://www.longevitygenomics.org>) was established in 2015 to develop translational strategies to promote healthy human aging based on findings from genomic studies of aging. Three types of findings have been used to develop translational strategies: (i) genetic variants associated with human longevity (LAV) and life span from GWAS, (ii) expression profiles associated with human longevity (LAG), normal aging, or healthy aging, and (iii) candidate genes and physiological factors with evidence for a role in human aging. The study team has developed analytical tools to data-mine publicly available data sets from different human cohorts (ie, UK Biobank, CHARGE, FHS) to identify LAVs and LAGs. The overall approach is to follow up LAVs identified from GWAS of longevity to identify targets for translation. For every LAV, there can be many potential LAGs in the LAV association region that might mediate the variant's association with longevity. By integrating multiple lines of evidence, including expression quantitative trait locus (eQTL), PrediXcan studies, positional overlap, and chromatin interaction studies, LAVs will be linked to candidate LAGs. MR approaches have been adapted to understand the causal associations of the variant/gene and to evaluate the potential impact of modulating tissue-specific LAG expression on subclinical risk and disease processes related to longevity. To identify genetic variants associated with gene expression, an eQTL analysis has been performed by the study team using the Genotype-Tissue Expression (GTEx) database (<https://gtexportal.org/home/>). The GTEx project consists of eQTL data on nonimmortalized cells from primary tissues, and the wide diversity of tissue types collected, including brain, heart, blood, thyroid, muscle, lung, skin, and adipose tissue. The investigators have also made extensive use of PrediXcan in this project to predict tissue-specific gene expression with SNP genotypes measured in human cohorts such as Health ABC, MrOS, and SOF. The study team has also been evaluating cheminformatic strategies using web-based drug discovery databases to discover longevity-associated drugs (37,39).

With the approaches described above, the study so far has identified 3 novel genes. (i) CEACAM19—a cell adhesion molecule that is widely expressed across tissues and is associated with multiple age-related traits, including chronic inflammation, AD, and blood lipids; (ii) ADAMTS7—a thrombospondin type 1 motif 7 associated with cardiovascular diseases, diastolic blood pressure, and red cell distribution width; and (iii) BLOC1S1 associated with atrial fibrillation, hearing loss, and longevity due to its role in regulating

mitochondrial functions. Scientists working outside of the LG team are currently following up with translational pilot studies on these 3 genes in animal and in vitro models to delineate the biological processes underlying their role in healthy aging.

The NIA also supports other clinical research, which is integral to the discovery and testing of therapeutic targets for healthy aging and is complementary to the ongoing translational genomics research by the Human Longevity Translational Projects.

### Comprehensive Evaluation of Aging-Related Clinical Outcomes and Geroproteins

The CARGO project was established in 2017 by an NIA initiative that focused on the clinical translation of endogenous proteins and polypeptides found to reverse or accelerate aging changes in laboratory animal models. For example, this included proteins whose circulating levels are higher in young animals than in old animals and may reverse certain age-related changes. Conversely, there are proteins and polypeptides whose circulating concentrations increase with age and induce changes associated with aging when administered to young animals. The CARGO study is utilizing the extensive phenotypic data and stored biospecimens available from longitudinal cohorts such as CHS, Health ABC, and the Baltimore Longitudinal Study of Aging to examine the effects of specific pro- and antiaging polypeptides and proteins (eg, follistatin, eotaxin, growth differentiation factor 11, myostatin, oxytocin) on a range of health outcomes, including physical performance problems/disability and risk for dementia. Analyses of data and biospecimens collected in epidemiologic and clinical studies can provide information on the relationships between aging outcomes and the levels of these proteins and polypeptides, identifying possible therapeutic targets and therapies to influence aging-related outcomes based on modulating the levels of such proteins and polypeptides.

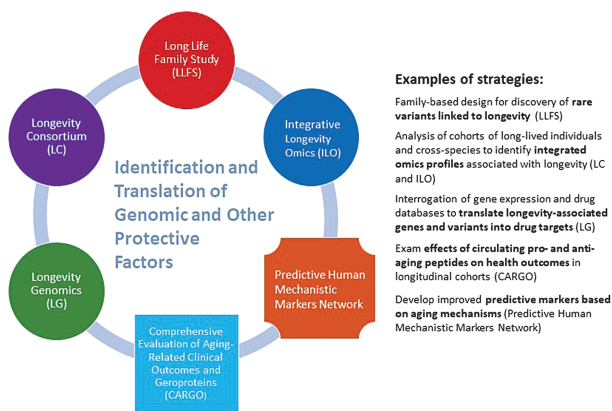
### Predictive Human Mechanistic Markers Network

The prospect for developing novel therapies based on factors contributing to health span and longevity would be enhanced with the availability of improved predictive markers which reflect the activity of fundamental mechanisms of aging such as stem cell exhaustion, macromolecular damage, cell senescence, mitochondrial function, and epigenetic regulation, to name a few. The availability of reliable and validated “mechanism-proximal” markers could substantially improve our ability to identify promising targets for interventions to extend health span, help to distinguish factors influencing specific conditions from those influencing multiple aging outcomes, and serve as markers of target engagement in geroscience-based drug development studies. Among the issues which need to be addressed in the development of such markers are those related to the proper collection and storage conditions of biospecimens, technical performance of the assays, including scale-up methodologies and high-throughput optimization. There is also a need to improve our understanding of the relationships between marker levels in one tissue (eg, serum) and levels in other organs/tissues (skeletal muscle, adipose tissue). Equally important for the validation of such human mechanistic predictive markers is the development of statistical methodology and application of appropriate inferential methods to evaluate the relationships of mechanistic markers to aging-related outcomes, and to clinical, functional, and physiological predictors of these outcomes. With these goals in mind, the Predictive Human Mechanistic Markers Network (<https://predictivebiomarkers.org>) was established through an NIA initiative and launched in 2019. This

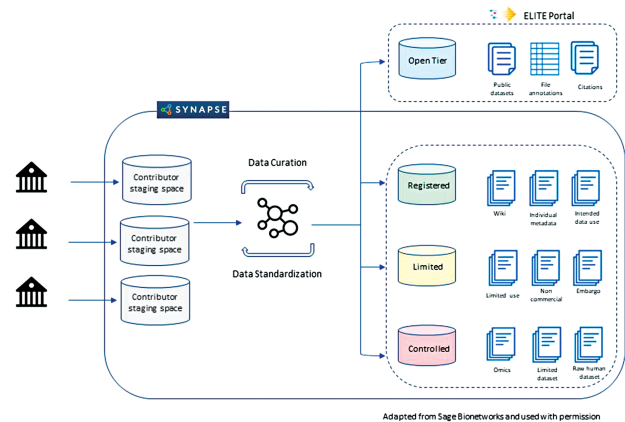
network consists of 5 separate projects focusing on the development of senescence-derived biomarkers of aging, validation and optimization of epigenetic clocks, markers reflecting viral burden and systemic inflammation, assessment of nuclear morphology in different cell types, omics profiles, and blood-based markers of mitochondrial bioenergetics. The network investigators are also conducting collaborative analysis to assess different markers within the same group of individuals to identify a potential combination of markers that have greater predictive value for specific health outcomes than the use of any single type of marker.

### Facilitating Open Science Via the NIA Exceptional Longevity Translational rEsources Portal

Research progress made by the collection of human longevity translational research projects has contributed to the availability of complex multidimensional biological data and unique research resources (eg, biospecimens from long-lived individuals and from cross-species comparisons, analytical tools, drug screening databases/tools, cutting edge algorithms) which can be leveraged by the broader scientific community to advance and expand multidisciplinary research on exceptional health spans and longevity. Moreover, the anticipated large volumes of omics data (eg, genomics, transcriptomics, methylomics, proteomics, metabolomics) to be generated by these studies will create greater opportunities for data mining/integrative analyses across different data types to identify omics profiles associated with exceptional longevity, including possible protection against age-related dementias/AD. Based on the highly successful model of NIA's the Accelerating Medicines Partnership Program for Alzheimer's Disease Knowledge Portal (40) to facilitate open science, NIA created the Exceptional Longevity Translational rEsources (ELITE) portal (an NIH-designated data repository; <https://adknowledgeportal.synapse.org/ELITE>) following FAIR principles and using the same bioinformatics platform as NIA's existing AD Knowledge Portal to consolidate and make readily available the valuable research resources generated by ongoing human longevity translational and related projects as illustrated in Figure 2 to a wide variety of researchers, including the AD community. Notably, the data governance policies and procedures for the ELITE portal will enable sharing of age-specific data from long-lived individuals in a secure manner (Figure 3). Adhering to NIH policies



**Figure 2.** Integration of different strategies to accelerate human longevity translational research.



Adapted from Sage Bionetworks and used with permission

**Figure 3.** Exceptional Longevity Translational rEsources (ELITE) portal infrastructure. Synapse platform infrastructure for data ingress and secure data storage for each individual project. Data are accessed through a 4-tier system from open access for summary data through 3 tiers of increasing stringency. Access to individual data for all ages is constrained to the top tier. Figure created by Sage Bionetworks and used with permission.

on data/resource sharing, all generated data and information about unique biological resources, including the EL-iPSC-derived hepatocytes and neurons, from the above-described projects will be shared with the scientific community using a tiered access model to preserve data privacy and security. There are no plans for ELITE to include a separate biorepository component, but it will direct investigators to the university-based biorepositories to obtain access to resources such as the centenarian iPSC-derived hepatocytes and neurons. The ELITE Portal includes a collaborative research workspace to facilitate combining multidimensional data, cloud computing, and software tools in a secure environment to allow investigators to access, deposit, and analyze different data types from multiple sources.

In the forthcoming years, the NIA will continue to enhance the capability of the portal infrastructure to facilitate data sharing and interface with other relevant data sets. Future portal development will include enhanced visualization functionality. The size and complexity of multiscale characterizations of biological data create challenges in effective data analysis and presentation. Graphical tools for organizing, integrating, and visualizing data and results will facilitate viewing and interpreting data and results (41). Innovations to visualize analyses for integrating across layers of diverse data types and the interconnected nature of high-dimensional data sets will be incorporated into the portal or as links to developed tools.

### Activities to Facilitate Scientific Synergies Across the Human Longevity Translational Research Projects

It is expected that improved interactions and scientific collaboration across EL study teams by leveraging analytical strategies, data, and other resources will enable and speed up the process of target identification and translational strategies to enhance health span. The NIA also supports other clinical research, which is integral to the discovery and testing of therapeutic targets for healthy aging and is complementary to the ongoing translational genomics research by the Human Longevity Translational Projects. In this respect, the study teams and NIA have initiated a number of integrative activities, as shown below.

Examples of ongoing and planned areas of enhanced interactions and synergy across these projects include joint activities to enable education, resources sharing, discussion of data harmonization strategies, validation and optimization of Inter-Lab analyses, and cross-platform comparisons for multiomic data generation. Most importantly, the NIA has organized biannual joint meetings of the exceptional longevity study teams to facilitate discussions and brainstorm ideas for the collective development of research strategies to accomplish the major goals of the human longevity translational research projects. In addition, NIA has initiated discussions of the study teams with National Center for Advancing Translational Science and NIH Intramural Center for Alzheimer's and Related Dementias to make use of the available research resources and work in partnership for the translation of protective factors associated with exceptional health and life span. In future, the joint meetings would be extended to other longevity researchers outside of the abovementioned studies to facilitate collaborations and sharing of data and resources.

## Future Perspectives

All evidence to date suggests healthy aging and longevity traits will be a polygenic combination of many variants with small clinical effects as well as many rare variants with larger effects associated with small population attributable risk/protection. Because biology works as systems and networks of genes and variants, not individual variants in isolation, integrative analysis approaches are believed to be powerful for discovery as well as more complete in explaining the biology of complex traits such as healthy aging. Application of systems biology and network analysis scans across multiple omics data sets across different populations, and multiple species will discover novel variants and new biology for healthy aging and longevity. Using these novel analytical methods to characterize the molecular landscape of divergent human and nonhuman species exhibiting substantial variation in life and health span leverages tools and data from the public domain allowing comprehensive analyses and triangulation of disparate scientific studies and discoveries. This appears to be an optimal way to accomplish the goals of developing therapeutics for enhancing health span. The aging research field is uniquely positioned in this era of genomic medicine with a wealth of resources such as large data sets from multiple aging studies, novel analytical tools for data mining, cloud infrastructure for hosting and analyzing big data, drug discovery databases, gene editing technologies like CRISPR/CAS9, optimized iPSC methodologies, and in vitro model systems including tissue chips to discover genetic and omic factors related to longevity and exceptional health span and functional genomic studies for translation. In particular, the NIA has now catalyzed and triggered a concerted effort across exceptional longevity studies that converge on the major goal of advancing the discovery of longevity-related variants, conserved genes, pathways, and therapeutic targets for longevity-enhancing, geroprotective drugs of relevance to humans.

## Funding

Funding was provided by the Division of Geriatrics and Clinical Gerontology, National Institute on Aging, and National Institutes of Health.

## Conflict of Interest

None declared.

## Acknowledgments

The authors wish to acknowledge Dr. Evan Hadley for his guidance, encouragement, and critical review of the manuscript. The authors sincerely thank the principal investigators of the human longevity translational research studies, CARGO, and Predictive Markers Network for their careful review, input, and recommendations in the writing of the manuscript. The authors also would like to thank the NIA staff Dr. Max Guo, Dr. Suzana Petanceska, and Ms. Kathleen Mercure for their contribution in the preparation of this manuscript.

## Disclaimer

The views and thoughts expressed in the text belong solely to the authors and these views are not expressed on behalf of the government.

## References

1. Stout MB, Justice JN, Nicklas BJ, Kirkland JL. Physiological aging: links among adipose tissue dysfunction, diabetes, and frailty. *Physiology (Bethesda)*. 2017;32(1):9–19. doi:10.1152/physiol.00012.2016
2. Lipsky MS, King M. Biological theories of aging. *Dis Mon*. 2015;61(11):460–466. doi:10.1016/j.disamonth.2015.09.005
3. Ismail K, Nussbaum L, Sebastiani P, et al. Compression of morbidity is observed across cohorts with exceptional longevity. *J Am Geriatr Soc*. 2016;64(8):1583–1591. doi:10.1111/jgs.14222
4. Burch JB, Augustine AD, Frieden LA, et al. Advances in geroscience: impact on healthspan and chronic disease. *J Gerontol A Biol Sci Med Sci*. 2014;69(suppl 1):S1–S3. doi:10.1093/gerona/glu041
5. Sebastiani P, Solovieff N, Dewan AT, et al. Genetic signatures of exceptional longevity in humans. *PLoS One*. 2012;7(1):e29848. doi:10.1371/journal.pone.0029848
6. Sebastiani P, Gurinovich A, Nygaard M, et al. APOE alleles and extreme human longevity. *J Gerontol A Biol Sci Med Sci*. 2019;74(1):44–51. doi:10.1093/gerona/gly174
7. Sebastiani P, Bae H, Sun FX, et al. Meta-analysis of genetic variants associated with human exceptional longevity. *Aging (Albany NY)*. 2013;5(9):653–661. doi:10.18632/aging.100594
8. Broer L, Buchman AS, Deelen J, et al. GWAS of longevity in CHARGE consortium confirms APOE and FOXO3 candidacy. *J Gerontol A Biol Sci Med Sci*. 2015;70(1):110–118. doi:10.1093/gerona/glu166
9. Bae H, Gurinovich A, Malovini A, et al. Effects of FOXO3 polymorphisms on survival to extreme longevity in four centenarian studies. *J Gerontol A Biol Sci Med Sci*. 2018;73(11):1439–1447. doi:10.1093/gerona/glx124
10. Chen R, Morris BJ, Donlon TA, et al. FOXO3 longevity genotype mitigates the increased mortality risk in men with a cardiometabolic disease. *Aging (Albany NY)*. 2020;12(23):23509–23524. doi:10.18632/aging.202175
11. Donlon TA, Davy PMC, Willcox BJ. Analysis of FOXO3 gene polymorphisms associated with human longevity. *Methods Mol Biol*. 2019;1890:251–258. doi:10.1007/978-1-4939-8900-3\_21
12. Davy PMC, Allsopp RC, Donlon TA, Morris BJ, Willcox DC, Willcox BJ. FOXO3 and exceptional longevity: insights from hydra to humans. *Curr Top Dev Biol*. 2018;127:193–212. doi:10.1016/bs.ctdb.2017.10.001
13. Tesi N, van der Lee SJ, Hulsman M, et al. Polygenic risk score of longevity predicts longer survival across an age continuum. *J Gerontol A Biol Sci Med Sci*. 2021;76(5):750–759. doi:10.1093/gerona/glaa289
14. Ryu S, Han J, Norden-Krichmar TM, et al. Genetic signature of human longevity in PKC and NF- $\kappa$ B signaling. *Aging Cell*. 2021;20(7):e13362. doi:10.1111/acel.13362
15. Ma S, Gladyshev VN. Molecular signatures of longevity: insights from cross-species comparative studies. *Semin Cell Dev Biol*. 2017;70:190–203. doi:10.1016/j.semcdb.2017.08.007
16. Jeck WR, Siebold AP, Sharpless NE. Review: a meta-analysis of GWAS and age-associated diseases. *Aging Cell*. 2012;11(5):727–731. doi:10.1111/j.1474-9726.2012.00871.x
17. Deelen J, Evans DS, Arking DE, et al. A meta-analysis of genome-wide association studies identifies multiple longevity genes. *Nat Commun*. 2019;10(1):3669. doi:10.1038/s41467-019-11558-2



18. Conneely KN, Capell BC, Erdos MR, et al. Human longevity and common variations in the LMNA gene: a meta-analysis. *Aging Cell*. 2012;11(3):475–481. doi:10.1111/j.1474-9726.2012.00808.x
19. King MC, Wilson AC. Evolution at two levels in humans and chimpanzees. *Science*. 1975;188(4184):107–116. doi:10.1126/science.1090005
20. Barrett JH, Taylor JC, Illes MM. Statistical perspectives for genome-wide association studies (GWAS). *Methods Mol Biol*. 2014;1168:47–61. doi:10.1007/978-1-4939-0847-9\_4
21. Song K, Mosteller M, Lawson M, Nelson MR. Practical limitations to estimating heritability in pharmacogenetic studies. *Pharmacogenomics*. 2013;14(8):851–852. doi:10.2217/pgs.13.86
22. Ball RD. Designing a GWAS: power, sample size, and data structure. *Methods Mol Biol*. 2013;1019:37–98. doi:10.1007/978-1-62703-447-0\_3
23. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. *Genome Biol*. 2017;18(1):83. doi:10.1186/s13059-017-1215-1
24. Lewis KN, Rubinstein ND, Buffenstein R. A window into extreme longevity; the circulating metabolomic signature of the naked mole-rat, a mammal that shows negligible senescence. *Geroscience*. 2018;40(2):105–121. doi:10.1007/s11357-018-0014-2
25. Ball HC, Levari-Shariati S, Cooper LN, Aliani M. Comparative metabolomics of aging in a long-lived bat: insights into the physiology of extreme longevity. *PLoS One*. 2018;13(5):e0196154. doi:10.1371/journal.pone.0196154
26. Lewis KN, Soifer I, Melamud E, et al. Unraveling the message: insights into comparative genomics of the naked mole-rat. *Mamm Genome*. 2016;27(7–8):259–278. doi:10.1007/s00335-016-9648-5
27. Deweerdt S. Comparative biology: looking for a master switch. *Nature*. 2012;492(7427):S10–S11. doi:10.1038/492S10a
28. Austad SN. Diverse aging rates in metazoans: targets for functional genomics. *Mech Ageing Dev*. 2005;126(1):43–49. doi:10.1016/j.mad.2004.09.022
29. Wilkinson GS, South JM. Life history, ecology and longevity in bats. *Aging Cell*. 2002;1(2):124–131. doi:10.1046/j.1474-9728.2002.00020.x
30. Nelson MR, Tipney H, Painter JL, et al. The support of human genetic evidence for approved drug indications. *Nat Genet*. 2015;47(8):856–860. doi:10.1038/ng.3314
31. Hurler MR, Nelson MR, Agarwal P, Cardon LR. Impact of genetically supported target selection on R&D productivity. *Nat Rev Drug Discov*. 2016;15(9):596–597. doi:10.1038/nrd.2016.187
32. King EA, Davis JW, Degner JF. Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLoS Genet*. 2019;15(12):e1008489. doi:10.1371/journal.pgen.1008489
33. Wojczynski MK, Lin SJ, Sebastiani P, et al. NIA long life family study: objectives, design, and heritability of cross sectional and longitudinal phenotypes. *J Gerontol A Biol Sci Med Sci*. 2022;77(4):717–727. doi:10.1093/gerona/glab333
34. Christensen K, Wojczynski MK, Pedersen JK, et al. Mechanisms underlying familial aggregation of exceptional health and survival: a three-generation cohort study. *Aging Cell*. 2020;19(10):e13228. doi:10.1111/accel.13228
35. Kahn A; Longevity Consortium. The Longevity Consortium: harnessing diverse approaches to understand the genetic basis of human longevity and healthy aging. An introduction to a series of articles. *Ageing Res Rev*. 2011;10(2):179–180. doi:10.1016/j.arr.2010.04.008
36. Hadley E, Evans DS, Perls TT. The Longevity Consortium: multi-omics integrative approach to discovering healthy aging and longevity determinants. *Innovation in Aging*. 2019;3:S208–S208. doi:10.1093/geroni/igz038.757
37. McCarrison J, Girke T, Goetz LH, Miller RA, Schork NJ. Genetic support for longevity-enhancing drug targets: issues, preliminary data, and future directions. *J Gerontol A Biol Sci Med Sci*. 2019;74(suppl\_1):S61–S71. doi:10.1093/gerona/glz206
38. Munafò MR, Davey Smith G. Robust research needs many lines of evidence. *Nature*. 2018;553(7689):399–401. doi:10.1038/d41586-018-01023-3
39. Duan Y, Evans DS, Miller RA, Schork NJ, Cummings SR, Girke T. signatureSearch: environment for gene expression signature searching and functional interpretation. *Nucleic Acids Res*. 2020;48(21):e124. doi:10.1093/nar/gkaa878
40. Greenwood AK, Montgomery KS, Kauer N, et al. The AD knowledge portal: a repository for multi-omic data on Alzheimer's disease and aging. *Curr Protoc Hum Genet*. 2020;108(1):e105. doi:10.1002/cphg.105
41. Subramanian I, Verma S, Kumar S, Jere A, Anamika K. Multi-omics data integration, interpretation, and its application. *Bioinform Biol Insights*. 2020;14:1177932219899051. doi:10.1177/1177932219899051