CrossMark

OPINION ARTICLE

Ageing with elegans: a research proposal to map healthspan pathways

**Walter Luyten · Peter Antal · Bart P. Braeckman · Jake Bundy ·
Francesca Cirulli · Christopher Fang-Yen · Georg Fuellen ·
Armand Leroi · Qingfei Liu · Patricia Martorell · Andres Metspalu ·
Markus Perola · Michael Ristow · Nadine Saul · Liliane Schoofs ·
Karsten Siems · Liesbet Temmerman · Tina Smets ·
Alicja Wolk · Suresh I. S. Rattan**

Abstract Human longevity continues to increase world-wide, often accompanied by decreasing birth rates. As a larger fraction of the population thus gets older, the number of people suffering from disease or disability increases dramatically, presenting a major societal challenge. Healthy ageing has therefore been selected by EU policy makers as an important priority (http://www.healthyageing.eu/european-policies-and-initiatives); it benefits not only the elderly but also their direct environment and broader society, as well as the economy. The theme of healthy ageing figures prominently in the Horizon 2020 programme (https://ec.europa.eu/programmes/horizon2020/en/h2020-section/health-demographic-change-and-wellbeing), which has launched several research and innovation actions (RIA), like "Understanding health, ageing and disease: determinants, risk factors and pathways" in the work programme on "Personalising healthcare" (https://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/693-phc-01-2014.html). Here we present our research proposal entitled "ageing with elegans" (AwE) (http://www.h2020awe.eu/),

W. Luyten (✉)
Department of Pharmaceutical and Pharmacological
Sciences, KU Leuven, Naamsestraat 59, Box 2469,
3000 Leuven, Belgium
e-mail: walter.luyten@med.kuleuven.be

P. Antal
Department of Measurement and Information Systems,
Budapest University of Technology and Economics,
Abiomics Europe Ltd., Budapest 1118, Hungary
e-mail: antal@mit.bme.hu

B. P. Braeckman
Biology Department, Ghent University, Proeftuinstraat
86N1, 9000 Ghent, Belgium
e-mail: bart.braeckman@ugent.be

J. Bundy
Department of Surgery and Cancer, Imperial College
London, South Kensington Campus Exhibition road,
SW7 2AZ London, UK
e-mail: j.bundy@imperial.ac.uk

F. Cirulli
Section of Behavioral Neuroscience, Department of Cell
Biology and Neurosciences, Istituto Superiore di Sanità,
00161 Rome, Italy
e-mail: francesca.cirulli@iss.it

C. Fang-Yen
Department of Bioengineering, School of Engineering and
Applied Science, University of Pennsylvania, 210 S 33rd
St, 240 Skirkanich Hall, Philadelphia, PA 19104, USA
e-mail: fangyen@seas.upenn.edu

G. Fuellen
Institute for Biostatistics and Informatics in Medicine and
Ageing Research, Rostock University Medical Center,
18057 Rostock, Germany
e-mail: fuellen@uni-rostock.de

funded by this RIA, which aims for better understanding of the factors causing health and disease in ageing, and to develop evidence-based prevention, diagnostic, therapeutic and other strategies. The aim of this article, authored by the principal investigators of the 17 collaborating teams, is to describe briefly the rationale, aims, strategies and work packages of AwE for the purposes of sharing our ideas and plans with the biogerontological community in order to invite scientific feedback, suggestions, and criticism.

Health and healthspan

Health is often described either as the absence of disease or as a vague concept of well-being, without clear objective measures. Although various parameters of frailty have been proposed (Fulop et al. 2010; Hubbard and Woodhouse 2010; Montesanto et al. 2010; Mitnitski et al. 2013), direct measures of health largely remain undefined. In the ongoing discourse about health, some attempts have been made to define health, healthspan and the quality of life. For example, taking into consideration the functionality of the living system as a crucial phenotype and the concept of activities of daily living (ADL), which is a well-established term (Engberg et al. 2008), health could be defined as a state of adequate physical and mental independence in ADL (Rattan 2013). Similarly, health has been defined as the ability to adapt and self-manage (Huber et al. 2011).

Most research on genetic and environmental factors in ageing has focused on lifespan rather than healthspan perhaps because the former is a "hard" measure, whereas there is still discussion on the best way to measure healthspan. For a long time there was also an often implicit assumption that lifespan and healthspan strongly correlate, but the two can clearly diverge (Avanesian et al. 2010; Garcia-Valles et al. 2013; Stroustrup et al. 2013; Bansal et al. 2015). In practical terms, one could consider healthspan as the life period when one is functionally independent and free from serious disease. Our consortium embraces this definition, for which detailed measuring parameters have recently been developed (e.g. the HLY indicator: http://www.eurohex.eu/index.php?option=ehleisproject). Furthermore, this definition can be used to develop animal models, and such healthspan model can then be used for an initial assessment of preventive or therapeutic measures, as well as for risk identification and the initial evaluation of potential biomarkers. Such an approach can also be useful to study interactions between genetic and various environmental factors.

Healthspan is affected by genetic as well as environmental factors, and their interactions.

A. Leroi
Department of Life Sciences, Imperial College London, South Kensington Campus Exhibition road, SW7 2AZ London, UK
e-mail: a.leroi@imperial.ac.uk

Q. Liu
Tsinghua University, B 407, Medical Science Building, Haidian District, 100084 Beijing, China
e-mail: liuqf@biomed.tsinghua.edu.cn

P. Martorell
Cell Biology Laboratory, Food Biotechnology Department, Biópolis SL, Paterna, 46980 Valencia, Spain
e-mail: patricia.martorell@biopolis.es

A. Metspalu
Estonian Genome Center & Institute of Molecular and Cell Biology, University of Tartu, Riia Street 23B, 51010 Tartu, Estonia
e-mail: andres.metspalu@ut.ee

M. Perola
Department of Health, National Institute for Health and Welfare, Helsinki, Finland
e-mail: markus.perola@thl.fi

M. Perola
Institute for Molecular Medicine Finland, University of Helsinki, Mannerheimintie 166, 00271 Helsinki, Finland

M. Perola
The Estonian Genome Center, University of Tartu, 50090 Tartu, Estonia

M. Ristow
Energy Metabolism Laboratory, Swiss Federal Institute of Technology (ETH) Zürich, Schorenstrasse 16, 8603 Schwerzenbach, Switzerland
e-mail: michael-ristow@ethz.ch

Information about the former is rapidly increasing thanks to exome and whole genome sequencing of individuals with unusually long lifespan. In (super)-centenarians healthspan appears to be relatively increased, with compressed morbidity at the end of life (Andersen et al. 2012). This is not due to the absence of risk alleles of common complex diseases but presumably the presence of protective genetic factors (Brooks-Wilson 2013). Although ever-decreasing DNA sequencing costs will most likely lead to a further explosion of sequence data, the functional significance of many candidate mutations and polymorphisms will not be easy to evaluate. On the other hand, many environmental factors or treatments are claimed to favour healthy ageing. However, obtaining objective evidence is difficult owing to the long treatment and observation periods required, the often limited mechanistic understanding and, in the case of humans, the genetic heterogeneity of the treated or exposed population. Furthermore, the complex interactions between genetic and environmental factors add to the difficulty of assessing the role of individual ones. Healthspan is also determined by stochastic processes. This is illustrated in cohorts of nematodes with identical genomic and environmental factors which, nevertheless, show clear individual differences in ageing and lifespan (Herndon et al. 2002; Suda 2014). However, since we cannot influence these stochastic effects, this study focuses on genetic and environmental parameters.

Therefore, we need a model system permitting rapid evaluation of candidate genetic and environmental factors, and for studying their interactions. We propose to develop *Caenorhabditis (C.) elegans,* which has already been well validated as a model system for lifespan (Lapierre and Hansen 2012), into a healthspan model. Healthspan would be assessed by monitoring the decline of functions where ageing is the primary risk factor, like cognitive (chemotaxis, learning) and motor processes (loss of muscle function, pharyngeal pumping rate and defecation rate), as well as stress resistance and - response.

The *C. elegans* model

Since Nobel Laureate Sydney Brenner established *C. elegans* as an animal model in 1974, many breakthrough discoveries in ageing research have been accomplished using this small nematode (Brenner 1974). The seminal discovery that it is possible to intervene in the ageing process by carefully manipulating genetic targets, such as those from the conserved insulin/IGF-1-like signalling (IIS) pathway, holds the promise of altering human life- and healthspan, at least to some extent (Friedman and Johnson 1988; Kenyon et al. 1993; Bartke 2008). A case in point is the finding that single nucleotide polymorphisms in the human FOXO3, orthologue of the *C. elegans daf-16* transcription factor in the IIS

N. Saul
Molecular Genetics Group, Faculty of Life Sciences, Institute of Biology, Humboldt-Universität zu Berlin, Chaussestr. 117, 10115 Berlin, Germany
e-mail: nadines1976@aol.com

L. Schoofs
Laboratory of Genomics and Proteomics, Department of Biology, KU Leuven, Naamsestraat 59, 3000 Leuven, Belgium
e-mail: liliane.schoofs@bio.kuleuven.be

K. Siems
Analyticon Discovery GmbH, Hermannswerder Haus 17, 14473 Potsdam, Germany
e-mail: k.siems@ac-discovery.com

L. Temmerman
Laboratory of Molecular and Functional Neurobiology, Department of Biology, KU Leuven, Naamsestraat 59, 3000 Leuven, Belgium
e-mail: liesbet.temmerman@bio.kuleuven.be

T. Smets
Cartagenia, Part of Agilent Technologies, Inc., Technologielaan 3, 3001 Louvain, Belgium
e-mail: tina.smets@agilent.com

A. Wolk
Division of Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, 17177 Stockholm, Sweden
e-mail: alicja.wolk@ki.se

S. I. S. Rattan
Laboratory of Cellular Ageing, Department of Molecular Biology and Genetics, Aarhus University, 8000 Aarhus-C, Denmark
e-mail: rattan@mbg.au.dk

pathway, are associated with an extension of human lifespan (Anselmi et al. 2009; Willcox et al. 2008). Other key regulatory pathways of energy metabolism with an evolutionary conserved role in ageing, like TOR and AMPK signalling, have also been characterized first in *C. elegans* and were validated in mammals afterwards (Kenyon 2010). Similarly, key cellular processes capable of modulating the rate of ageing, such as translation and autophagy, were first characterised as such in *C. elegans* and later verified in higher organisms (Hansen et al. 2007; Meléndez et al. 2003). This strong similarity at the cellular level is reflected at the organismal level. *C. elegans* displays many phenotypic traits of ageing that are surprisingly similar to those found in humans, including, but not limited to, sarcopenia, impaired locomotion, physical decline in neuron integrity and synapse outgrowth. *C. elegans* is also a proven asset to uncover the mechanism of action for treatments with health-promoting effects. A recent success story to illustrate this point is the discovery of the molecular processes underlying the healthspan-promoting effects of the antihyperglycemic drug metformin in *C. elegans* (Cabreiro et al. 2013; De Haes et al. 2014). Metformin is currently being considered for testing in a clinical trial to combat ageing in humans (Hayden 2015).

Specific objectives of AwE

A major goal of AwE is to create a panel of *C. elegans* reporter strains that each have a defect in different pathways affecting healthspan. The discovery of these pathways will be based on: (1) mining the biomedical literature for genes suggested to be involved in healthspan; (2) genetic screens of *C. elegans* for a healthspan phenotype, in which the screens will be based on RNAi or on the introduction of polymorphisms favouring healthy ageing as discovered in cohorts of healthy aged persons; and (3) testing of pure compounds or botanical preparations reported to slow ageing.

One of the aims of AwE is to develop an automated, high-throughput system for measuring *C. elegans* healthspan. We have developed a scalable microplate-based imaging and analysis platform to observe nematodes over multiple weeks in tens of thousands of worms at 8-µm resolution (Churgin and Fang-Yen 2015; Churgin et al. 2016). This system is based on the

WorMotel, a microfabricated platform we designed for imaging and cultivation of 240 individual *C. elegans* confined to separate wells. The animals are tracked by machine vision using a camera-based imaging system. Using a robotic plate handler and multiple imaging systems, up to 400 plates containing a total of 96,000 worms can be tracked throughout their lifespans. The automated method generates a record of spontaneous and stimulus-evoked activity and time of death, from which healthspan and lifespan curves are constructed. Data from the WorMotel are consistent with results from the manual method of survival curve acquisition for several mutants in both standard and stressful environments. These tools allow us to analyse measures of (healthy) ageing based on motor function, and greatly increase throughput of lifespan and healthspan assays.

The following specific aims will be targeted:

1. Perform a large-scale RNAi-screen in *C. elegans* to discover genes affecting healthspan.
2. Based on approximately 2500 whole genome sequences, identify genetic variants overrepresented in cohorts of persons with unusually long healthspan (as distinct from lifespan), taking into account sex differences. From these variants, the most robust (occurring in several ageing cohorts) and promising ones (e.g. coding polymorphisms, polymorphisms contributing the highest relative risk) will be selected for further study in *C. elegans*.
3. Introduce these human genetic variants into the orthologues of *C. elegans* and determine their effect on healthspan. Analyse *C. elegans* variants with longer (at least by 10 %) healthspan using proteomics, metabolomics and transcriptomics.
4. Select compounds (including nutraceuticals and endocrine disruptors), botanical preparations and diets with literature evidence for healthspan effects as well as a broader collection of natural products or plant extracts, test them in *C. elegans* and (for those showing effects) sort them into groups depending on the pathway(s) they affect. These compounds will not necessarily be suitable for drug development, but can be used as tools in molecular, cellular and (acute toxicity permitting) even animal experiments.
5. For botanical preparations improving healthspan in *C. elegans* via a novel mechanism (as assessed

by the aforesaid panel of *C. elegans* healthspan pathway mutants), purify the active components using bioassay-guided purification, and elucidate their mechanism of action in *C. elegans* using a combination of proteomics, metabolomics and transcriptomics.

6. Infer healthspan pathways by combining genetic evidence from healthy aged human cohorts and the RNAi screen in *C. elegans*, as well as the omics evidence and tool compounds, with public domain background knowledge using bioinformatics and artificial intelligence tools.

7. Create a panel of *C. elegans* reporter strains, each covering a particular healthspan pathway.

8. Based on the knowledge of healthspan pathways, develop human cell assays able to detect healthspan effects. Test the most promising compounds on human cell lines and in mice for healthspan effects.

Genetic factors and ageing

Genetic factors affecting ageing have been identified from spontaneous mutants showing an ageing phenotype (e.g. premature ageing syndromes) and from induced mutations in model organisms like *Drosophila* and *C. elegans*.

Spontaneous mutations can cause very interesting (including ageing) phenotypes, but we have no control over them, and in humans mutations affecting ageing lead mostly to accelerated ageing (Glueck et al. 1976), whereas we are interested in the opposite. It is known that genetic factors play a role also in increased longevity in humans, and their effect is more pronounced with very high age. However, few of those genes are known at present. Genome-wide association studies have looked at genetic markers (polymorphisms) that are linked to genes relevant for a particular condition. Such studies have revealed some regions containing candidate genes relevant for longevity, but in most cases further research is needed to demonstrate exactly which genes are involved. Thanks to the recent lowering of the cost of DNA sequencing, whole genomes can now be sequenced, and have yielded interesting insights in genetic factors determining extreme old age in (super)centenarians (Shadyab and LaCroix 2015). It is still a matter of debate to what extent the same genes are also relevant for

healthspan. Therefore, we want to discover such genes by studying polymorphisms overrepresented in groups of humans who stay healthy to an unusually old age. Even if successful, however, we cannot easily evaluate the physiological effect of these polymorphisms by introducing them into animals.

Induced mutations offer the researcher more control, and are possible even in some mammals (e.g. transgenic mice), but they are still cumbersome and (in addition the ageing phenotype) take long to study. Making one genetic change in mice can take a long time; making directed genetic changes in other mammals is even more difficult (although CRISPR/Cas9 technology may improve this). Assessing the effect on ageing of such a genetic change in mice can then take another 2-3 years.

Methods like saturation mutagenesis or genome-wide RNAi knockdown permit, in principle, the systematic interrogation of the entire genome for any phenotype. In practice, however, they are mostly used in invertebrate model organisms, which have the additional advantage of a short lifespan, so that ageing phenotypes can be studied in weeks rather than years. It is therefore not surprising that most genes affecting ageing have been found first in *Drosophila* or *C. elegans*. Importantly, subsequent studies in mice have generally confirmed the relevance of the corresponding genes (orthologues) in mammals (Bitto et al. 2015). These genetic studies, however, have focussed mostly on lifespan. Recently, measures have been developed for healthspan in *C. elegans* (Keith et al. 2014), but they require a great deal of manual work and are not scalable to a large number of such experiments. Therefore, we developed a platform in which healthspan can be measured in a highly parallel and automated fashion (Churgin and Fang-Yen 2015), permitting for the first time a systematic search for all the genes affecting healthspan in *C. elegans*. It is expected that these genes will also be relevant for humans.

Environmental factors and ageing

Like for genetic factors, most work has focused on lifespan. In contrast, however, to the genome (which is finite and can now be determined completely), there are many kinds of environmental factors (physical, chemical, biological, ecological, social, economic, etc.), and we typically cannot exhaustively and

accurately measure them. Moreover, whereas we have a good idea how genetic factors influence biology, such mechanistic insight is often lacking for environmental factors. In addition, whereas we can change any nucleotide in a genome at will in a number of organisms and study the effect of that change, it is difficult to modify a single environmental factor while keeping all other ones constant. Finally, the effect of an environmental factor does not only depend on the intensity and duration of the exposure, but also on the developmental stage of the organism that was exposed. Organisms may have particularly vulnerable periods, and effects of environmental exposure may not be immediately apparent, but only manifest themselves at later stages of development (e.g. in old age), or even in subsequent generations.

Given these myriad of possibilities, we need a system where we can quickly and inexpensively test candidate environmental factors for effects on healthspan, with the option of exposing at different developmental stages, and looking for both short- and long-term effects. Ideally, the test system should also permit the elucidation of the mechanism by which the environmental factor exerts its effect.

Interactions between environmental and genetic factors on ageing

Although interactions between known genetic and environmental factors (e.g. via epigenetics) have been documented on lifespan (Guha et al. 2013), the number of combinatorial possibilities has precluded an exhaustive exploration. For unknown genetic and/or environmental factors, systematic testing of interaction effects is obviously impossible. Even a very fast and inexpensive testing system for healthspan (which does not exist at present) would not permit a brute force solution by evaluating all possible combinations of known and suspected (let alone candidate) factors.

A much smarter approach would be to map the pathways affecting healthspan, so that genetic and environmental factors can be quickly categorised based on the pathways they affect. Reducing the vast number of factors to classes of factors with similar effects would greatly reduce the combinatorial problem. In principle, one need only study one representative from each class. Moreover, if one knows how the pathways affect healthspan, and how the factors

perturb these pathways, then this mechanistic understanding would permit predicting the effect on healthspan of a factor or combination of factors. Such predictions would greatly help in understanding which factors pose a risk for healthspan, and which ones could improve it. The former would be helpful for prevention, the latter for developing therapies.

Understanding healthspan pathways would also permit developing diagnostics to detect persons with pathway alterations (due to genetic or environmental factors) posing a healthspan risk, and discovering biomarkers that help (1) to develop therapies causing desirable changes in healthspan pathways or (2) to detect persons with particular alterations in healthspan pathways that stand to benefit from therapies directed at those alterations. AwE will not only map healthspan pathways, but also deliver a test system specific for different healthspan pathways as well as candidate biomarkers.

Positioning of AwE

A major goal of AwE is technology transfer. The Technology Readiness Level (TRL) estimates the maturity of specific technology elements within a research or development program (http://www.inno vationseeds.eu/Virtual_Library/Knowledge/TLR_Scale. kl). AwE starts at TRL2 (technology concept formulated) and aims to arrive at TRL4 (technology validated in lab) by the end, so that industrial partners can take over at that stage. One industrial partner (Biopolis S.L.) in the consortium already uses *C. elegans*, for the evaluation of natural products and nutrition on ageing. They will provide the necessary input to assure valorisation of project results. Some components of the project (like the WorMotel) will reach TRL7 level, as a prototype will be demonstrated in the operational environment of Biopolis.

Thanks to the high speed and low cost of tests based on *C. elegans*, many factors can be easily tested in parallel using different dosages or exposure windows, and even the systematic testing of combinations of factors is within reach. Moreover, using genetic methods or omics studies (transcriptomics, proteomics, metabolomics), the mechanisms by which factors cause changes in healthspan can be found out in *C. elegans*. By a systematic genetic screen of the entire *C. elegans* genome, we will map the genes

affecting healthspan. To verify that these genes are relevant also in humans, we will check whether variations in the corresponding human genes are over- or underrepresented in humans who lived to an old age in good health (compared to matched controls of the same age who are not in good health). We will also study exposure of *C. elegans* to the richest source of environmental chemicals that humans have been systematically exposed to: edible and medicinal plants. Since these are consumed deliberately but voluntarily by humans, and these plants typically have a good safety record, this chemical exposure poses no ethical problems. Some of these plants are even reported to improve healthspan in humans, and if we can demonstrate a similar effect in *C. elegans*, it is attractive to identify the active components and elucidate their mechanism of action.

The combination of genes and compounds affecting healthspan, with omics information on changes at the RNA, protein and metabolite level, will permit us to map the pathways that affect healthspan. By making genetic changes in each of these pathways, we can create a panel of *C. elegans* strains that respond more strongly or less strongly to compounds affecting that pathway. This set of strains will be tremendously useful for rapidly testing which pathways are affected by an environmental factor, and we can predict which environmental factors are likely to act antagonistically, additively or synergistically. The most promising factors emerging from the *C. elegans* tests will then be tested in human cells and mice.

We have chosen to focus in our project on factors that increase (rather than decrease) healthspan not because the latter are of lesser interest. Rather, we believe that since it is easier to disrupt than to ameliorate a complex system, many factors leading to a decrease in healthspan will be non-specific and will lead us astray in our search for healthspan pathways. However, this does not preclude the use of our *C. elegans* system to evaluate toxic factors for a possible adverse effect on healthspan that may be predictive for humans.

The project is highly interdisciplinary in the technologies it uses, for example molecular biology (including various omics methods), biochemistry, natural product chemistry, pharmacology, genetics, epidemiology, bioinformatics, biostatistics, artificial intelligence, behavioural sciences and laboratory automation. It spans levels from the molecular and

cellular to the organismal. The project is also trans-disciplinary in that scientists from many different fields collaborate to achieve a common scientific goal: a better understanding of healthy ageing.

Our project will benefit from international efforts to sequence genomes of humans with interesting ageing phenotypes. Specifically, the genomes of patients with ageing disorders as well as individuals who remain healthy until a very advanced age can provide useful information. Many EU member states have ongoing studies of cohorts of very old persons, whose life history and phenotype are well documented. Three consortium partners coordinate such studies (see below). We will also capitalise on the SurvCurv database developed by the European Bioinformatics Institute (https://www.ebi.ac.uk/thornton-srv/databases/SurvCurv/) as well as on the results of FP7 projects like IDEAL (Integrated research on Developmental determinants of Aging and Longevity), MARK-AGE (http://www.mark-age.eu/) to find biomarkers of ageing, and SHARE (http://www.share-project.org/home0.html): Survey of Health, Ageing and Retirement in Europe. For measuring healthspan in humans, the HLY indicator developed by the FP7 EHLEIS project (http://www.eurohex.eu/index.php?option=ehleisproject) will be used.

Our project directly capitalises on several cohort studies:

1. The FINRISK study (http://www.p3g.org/resources/brif/national-finrisk-study), a population-based prospective study, with surveys carried out every 5 years since 1972, where the number of participants with DNA samples is currently about 38,000. Participants are clinically examined in a standardized manner for mainly cardiovascular (but also other) phenotypes for half a day when they enter the study. All participants are followed up through various comprehensive validated registries (for example incident case follow-up for cardiovascular disease and diabetes) and from (very well validated and comprehensive) national electronic medical records. In addition, there are records of medication, mortality and hospital discharge diagnoses. Altogether in FINRISK there are currently genome-wide association study (GWAS) data on 20,000 whole-genome sequences, and NMR metabolomes for about 8000 individuals, and in 2016 these will increase. In 2007, a deep phenotyping of 518 persons was performed,

from whom three different single nucleotide polymorphism chips (more than a million genotypes) are available, all imputed to the 1000 Genome database. All of them have also an NMR metabolome, mass spectrometry metabolome and transcriptomic array on leukocyte RNA; Illumina 450 epigenetic chip analysis has been performed in 2014. These people were re-invited and re-sampled for clinical examination in April 2014.

2. The Central Sweden Cohort of over 100,000 participants consists of two population-based longitudinal cohorts: (1) the Swedish Mammography Cohort (SMC) of over 60,000 women born between 1914-1948, followed-up since 1987 (http://ki.se/en/imm/the-swedish-mammography-cohort-smc) and (2) the Cohort of Swedish Men (COSM) of nearly 50,000 men born between 1918-1952, followed-up since 1997 (http://ki.se/en/imm/cosm-a-cohort-of-50000-swedish-men). Some women reached 100 years in 2014 (as compared to an average survival of 84 years among Swedish women). A broad spectrum of phenotypic information from repeated questionnaires (whole cohort) and from clinical examinations (subcohort of 10,000 participants) is included in the database. There is detailed information available about health status of the participants (diagnosed diseases, surgeries, hospital discharge diagnoses, medications and cause of death) via annual linkages to the validated comprehensive Swedish health registries. The number of participants with available DNA is currently about 45,000. Among them, about 8000 have been genotyped employing the OncoArray that includes 534,000 single nucleotide polymorphisms.

3. The Estonian Biobank (http://www.geenivaramu.ee/en/access-biobank) is a population-based biobank. The cohort size is currently 51,535 gene donors ($\geq$18 years of age), and closely reflects the age, sex and geographical distribution of the Estonian population. All subjects have been recruited randomly by general practitioners and physicians in hospitals. They have filled out a Computer Assisted Personal Interview (CAPI) within a 1-2 h appointment at a doctor's office. The CAPI includes personal data (place of birth, place(s) of living, nationality etc.), genealogical data (family history of medical conditions spanning four generations), educational and occupational history, lifestyle data (physical

activity, dietary habits based on a food frequency questionnaire, smoking, alcohol consumption, women's health, quality of life). In addition, the Estonian Biobank has 40,000 participants with MCTQ (chronotype) data (https://www.bioinfo.mpg.de/mctq/core_work_life/core/introduction.jsp), and 15,000 with both MCTQ and genome-wide microarray data; 3000 participants have filled out the NEO-PI-R questionnaire (P.T. Costa Jr., R.R. McCrae. Revised NEO personality inventory and five-factor inventory professional manual Psychological Assessment Resources, Odessa, FL (1992)), and GWAS data are available on 2700 of them. Medical history and current health status are recorded according to ICD-10 codes, medication according to ATC codes (http://www.whocc.no/filearchive/publications/1_2013guidelines.pdf). Extra data are collected from patients with psychiatric disorders (MINI (Sheehan et al. 1998) and SSP interviews (Baron et al. 1981)). Anthropometric measurements, blood pressure (sitting position at the end of the interview), and resting heart rate are measured during the visit. GWAS analyses have been performed on 15,000 subjects genotyped with different Illumina microarrays. RNA samples from 2100 subjects (isolated from venous blood) are available for gene expression studies along with 45 biomarkers from serum and plasma. Metabolite analysis (NMR and mass spectrometry) from the same set of individuals (12,000 in total) are available. Full-genome sequencing and diagnostic exome sequencing is ongoing and there are almost 2500 full genomes and more than 2500 exomes available. Follow-up studies have started and 2000 gene donors have been contacted again and new samples and information obtained. In addition, the project database is linked with the national registries and hospital databases in order to get the up-to-date phenotypic information (endpoints and nonfatal events), which allows to draw the disease trajectories together with the prescriptions drugs obtained for the entire life of all gene donors in the Estonian biobank (Leitsalu et al. 2015).

The research part of the work programme will largely be based on laboratory experiments on a *C. elegans* model system, but will also make use of epidemiological evidence (cohort studies in persons

remaining healthy up to an unusually high age), as well as in silico analyses (of the genetic data and for pathway mapping). To demonstrate the validity of the *C. elegans* platform, treatments or genetic changes known to affect healthy ageing in mammals will be tested early in the project. Conversely, the new treatments or polymorphisms that we find to have the strongest effects in *C. elegans* will be tested in human cells and mice.

Integration of consortium data and public information, to infer healthspan-related pathways and biomarkers, and to suggest bioactive compounds, will be achieved by employing chemo- and bioinformatics methods rooted in network science, probabilistic graphical models, kernel methods and semantic technologies. Such fusion of chemical, pharmaceutical, clinical and multi-omics data, including genetic data, RNAi, gene and protein expression, and metabolomics will be performed at multiple levels, including causal inference based on interventional data. We take gene/protein/metabolite interaction networks as reference, and attach genetic/RNAi and metabolic information as appropriate, following the network-based link-score highlighting idea we pioneered (Warsow et al. 2010). We also wish to apply existing pipelines such as described by Ernst et al. (2015) for the cross-species analysis of differential data. Scoring of interactions by incorporating rich background knowledge allows feature reduction for machine learning of healthspan determinants and compound synergy. Additionally, the use of already established interaction networks supports mechanistic understanding, and even automated hypothesis and explanation generation using semantic knowledge bases. At the level of similarities and raw data, we will apply kernel methods for large-scale fusion in predictive tasks and Bayesian approaches over probabilistic graphical models (Arany et al. 2013; Juhasz et al. 2014) to infer robust biomarkers and causal relations.

In Horizon 2020, gender is a cross-cutting issue and must be integrated in the research and innovation agenda. Since many aspects of ageing differ between men and women, all genetic analyses in humans will be performed separately for the two sexes. In addition, lifestyle factors that may affect healthspan will be analysed for sex-specific biological factors and gender-related social factors. In *C. elegans*, sex (let alone gender) differences are much harder to explore since this nematode is a hermaphrodite. Rare males can be found, but because there are so few of them, experimental populations are very hard to set up. Sex and gender aspects will, however, be taken into account in the studies on mice. This will allow to bridge research performed in simpler model systems with data collected in the human population.

Conclusions and future outlook

AwE will advance beyond the state of the art in several ways. Whereas most studies so far have focused on lifespan, we will study healthspan. Most individuals are more interested in remaining in good health as they age, rather than simply living longer. This is also more attractive from a societal viewpoint since healthy elderly can contribute more to society.

Although *C. elegans* has been well validated as a useful model for ageing (indeed, some of our insights in the biology of ageing are based on this model organism), its use for healthspan studies has been limited. AwE will establish *C. elegans* as a **model system for healthspan**, with the goal to establish factors affecting healthspan and for the study of interactions between such factors.

The WorMotel **platform** that AwE will use for rapid testing of healthspan in *C. elegans* (Churgin and Fang-Yen 2015) is novel and permits for the first time a genome-wide RNAi screen for this phenotype, as well as high-throughput screens for compounds or other factors affecting healthspan. But since the platform is based on video-image analysis, any phenotype that can be deduced in this way is screenable. Moreover, it permits individual worms to be followed over many weeks, which is impossible with current competitor instruments. It also is much more compact even though it has a much higher throughput.

The **pathways** involved in healthspan are largely unknown, although it is conceivable that there is extensive overlap with lifespan pathways. We will systematically map those pathways genetically and create a panel of mutant *C. elegans* strains with defects in each pathway for the rapid triage of treatments claiming to promote healthy ageing.

A number of **botanical preparations** can delay or even reverse ageing in animal models, but the mechanism often remains unclear. We will identify the active compounds in botanical preparations traditionally used to combat ageing. Such activity-guided

purification is not feasible with mammalian models due to the time and cost involved, but is made possible by our *C. elegans* healthspan model, especially when combined with automated testing using the WorMotel.

The **mechanism of action** of selected compounds and their combinations, for which we can demonstrate strong healthspan effects in *C. elegans,* will be elucidated by proteomics, transcriptomics and metabolomics studies. The genetic and omics data will be used in combination with relevant public background information to infer biochemical and signal transduction pathways using novel algorithms based on bioinformatics and artificial intelligence.

The combined analysis of ageing pathways in *C. elegans*, mice and humans will yield many **potential biomarkers** at the DNA level. Moreover, the omics analyses in *C. elegans* will suggest biomarkers at the RNA, protein or metabolite level. In combination with insight into the mechanism of action of healthspan-enhancing compounds, these biomarkers should help the clinical development (as companion diagnostics) of those bioactive compounds, as well as the identification of persons that can benefit from them.

It is known that exposure early in life to certain **risk factors** can have health effects (including ageing) that become only apparent in the long term. Such effects are difficult to study in mammals, but are within reach of a *C. elegans* model. We will make use of the advantages of *C. elegans* for studying such effects on healthspan.

AwE will provide a better understanding of the **genetic and environmental factors**, as well as their **combinations**, that affect healthspan. The *C. elegans* model will provide evidence for bioactive compounds affecting healthspan, for future strategies to promote healthy ageing and clinical interventions. Likewise, the project will reveal a number of pathways involved in healthspan, which will not only improve our understanding but will permit sorting preventive and therapeutic measures into groups that may be combined to generate additive or even synergistic effects on healthspan. Moreover, each pathway will suggest biomarkers for healthspan that are not merely correlative but causative. Individualization of dietary interventions to prolong healthspan will benefit from such markers (Fuellen et al. 2015).

AwE has important **healthcare** implications. Age is an important risk factor for many diseases, including neurodegenerative diseases like Alzheimer's and

Parkinson's disease, cardiovascular disease, osteoporosis, arthritis, cancer, etc. A lot of research has focussed on combating these individual diseases, but if we could affect the underlying ageing process, reducing the decline in physiological function and compressing morbidity, the propensity for all these diseases may diminish (Finkel 2005; Seals and Melov 2014; Seals et al. 2015; Stevenson et al. 2015). Increasing healthspan is therefore likely to stave off a number of degenerative diseases typical of old age. This is supported by the observation that (super)centenarians do not typically lack a number of "disease genes", but appear to have protective genes that buffer the negative effects of the "disease genes" (Stevenson et al. 2015).

The need for healthy ageing is enormous, and has been called a "keystone for a sustainable Europe" (http://ec.europa.eu/health/archive/ph_information/indicators/docs/healthy_ageing_en.pdf). This is not surprising since the healthcare costs of the ageing European population are set to increase, and the burden of care for this segment of the population will fall on the shoulders of an ever-shrinking active population. This unfavourable demographic evolution could be countered by extending healthspan.

Finally, *C. elegans* is becoming **a fast and inexpensive in vivo tool for the food, cosmetics and pharmaceutical industries**, especially for the screening of compounds providing health benefits. Moreover, results obtained are generally consistent with higher animal models, which enable subsequent preclinical and clinical trials to be more oriented. Additionally, there is no ethical problem in the use of *C. elegans*, as this nematode is not regarded as an animal in the EU regulations (Directive 2010/63/EU). As our WorMotel and associated panel of strains gains acceptance in the pharma, food and cosmetics sector, this platform will permit companies in these sectors to develop new products for improving healthspan with much greater safety and confidence that they will achieve the desired effect in humans. AwE will also spur innovation in the diagnostic sector, with the generation of new healthspan biomarkers and novel toxicology assays for healthspan effects.

References

Andersen SL, Sebastiani P, Dworkis DA, Feldman L, Perls TT (2012) Health span approximates life span among many supercentenarians: compression of morbidity at the approximate limit of life span. J Gerontol A Biol Sci Med Sci 67:395–405

Anselmi CV, Malovini A, Roncarati R, Novelli V, Villa F, Condorelli G, Bellazzi R, Puca AA (2009) Association of the FOXO3A locus with extreme longevity in a southern Italian centenarian study. Rejuvenation Res 12:95–104

Arany Ádám, Bolgár Bence, Balogh Balázs, Antal Peter, Mátyus Péter (2013) Multi-aspect candidates for repositioning: data fusion methods using heterogeneous information sources. Curr Med Chem 20(1):95–107

Avanesian A, Khodayari B, Felgner JS, Jafari M (2010) Lamotrigine extends lifespan but compromises health span in Drosophila melanogaster. Biogerontology 11:45–52

Bansal A, Zhu LJ, Yen K, Tissenbaum HA (2015) Uncoupling lifespan and healthspan in Caenorhabditis elegans longevity mutants. Proc Natl Acad Sci USA 12(3):E277–E286

Baron M, Asnis L, Gruen R (1981) The schedule for schizotypal personalities (SSP): a diagnostic interview for schizotypal features. Psychiatr Res 4:213–228

Bartke A (2008) impact of reduced insulin-like growth factor-1/insulin signalling on aging in mammals: novel findings. Aging Cell 7:285–290

Bitto, A., A. M. Wang, C. F. Bennett and M. Kaeberlein (2015) Biochemical genetic pathways that modulate aging in multiple species. Cold Spring Harb Perspect Med 5: a025114

Brenner S (1974) The genetics of Caenorhabditis elegans. Genetics 77:71–94

Brooks-Wilson AR (2013) Genetics of healthy aging and longevity. Hum Genet 132:1238–1323

Cabreiro F, Au C, Leung K-Y, Vergara-Irigaray N, Cochemé HM, Noori T, Gems D (2013) Metformin retards aging in C. elegans by altering microbial folate and methionine metabolism. Cell 153(1):228–239

Churgin MA, Fang-Yen C (2015) An imaging system for C. elegans behavior. Methods Mol Biol 1327:199–207

Churgin MA, Jung S, Yu CC, Chen X, Raizen DM, Fang-Yen C (2016) High-throughput longitudinal imaging of healthspan and longevity in Caenorhabditis elegans (unpublished manuscript)

De Haes W, Frooninckx L, Van Assche R, Smolders A, Depuydt G, Billen J, Temmerman L (2014) Metformin promotes lifespan through mitohormesis via the peroxiredoxin PRDX-2. Proc Natl Acad Sci USA 111(24):E2501–E2509

Engberg H, Christensen K, Andersen-Ranberg K, Vaupel JW, Jeune B (2008) Improving activities of daily living in Danish centenarians–but only in women: a comparative study of two birth cohorts born in 1895 and 1905. J Gerontol A Biol Sci Med Sci 63:1186–1192

Ernst M, Abu Dawud R, Kurtz A, Schotta G, Taher L, Fuellen G (2015) Comparative computational analysis of pluripotency in human and mouse stem cells. Sci Rep 5:7927

Finkel T (2005) Radical medicine: treating ageing to cure disease. Nat Rev Mol Cell Biol 6:971–976

Friedman DB, Johnson TE (1988) A mutation in the age-1 gene in Caenorhabditis elegans lengthens life and reduces hermaphrodite fertility. Genetics 118:75–86

Fuellen G, Schofield P, Flatt T, Schulz R-J, Boege F, Kraft K, Simm A (2015) Living long and well: prospects for a personalized approach to the medicine of ageing. Gerontology

Fulop T, Larbi A, Witkowski JM, McElhaney J, Loeb M, Mitnitski A, Pawelec G (2010) Aging, frailty and age-related diseases. Biogerontology 11:547–563

Garcia-Valles R, Gomez-Cabrera MC, Rodriguez-Manas L, Garcia-Garcia FJ, Diaz A, Noguera I, Olaso-Gonzalez G, Vina J (2013) Life-long spontaneous exercise does not prolong lifespan but improves health span in mice. Longev Healthspan 2:14

Glueck CJ, Gartside P, Fallat RW, Sielski J, Steiner PM (1976) Longevity syndromes: familial hypobeta and familial hyperalpha lipoproteinemia. J Lab Clin Med 88:941–957

Guha S, Cao M, Kane RM, Savino AM, Zou S, Dong Y (2013) The longevity effect of cranberry extract in Caenorhabditis elegans is modulated by daf-16 and osr-1. Age (Dordr) 35:1559–1574

Hansen M, Taubert S, Crawford D, Libina N, Lee S-J, Kenyon C (2007) Lifespan extension by conditions that inhibit translation in Caenorhabditis elegans. Aging Cell 6:95–110

Hayden EC (2015) Ageing pushed as treatable condition—regulators asked to consider innovative trial design. Nature 522:265–266

Herndon LA, Schmeissner PJ, Dudaronek JM, Brown PA, Listner KM, Sakano Y, Paupard MC, Hall DH, Driscoll M (2002) Stochastic and genetic factors influence tissue-specific decline in ageing C. elegans. Nature 419:808–814

Hubbard RE, Woodhouse KW (2010) Frailty, inflammation and the elderly. Biogerontology 11:635–641

Huber M, Knottnerus JA, Green L, van der Horst H, Jadad AR, Kromhout D, Leonard B, Lorig K, Loureiro MI, van der Meer JW, Schnabel P, Smith R, van Weel C, Smid H (2011) How should we define health? BMJ 343:d4163

Juhasz Gabriella, Hullam Gabor, Eszlaria Nora, Gonda Xenia, Antal Peter, Anderson Ian Muir, Hökfeltf Tomas G M, William Deakin JF, Bagdy Gyorgy (2014) Brain galanin system genes interact with life stresses in depression-related phenotypes. Proc Natl Acad Sci 111(16):E1666–E1673

Keith SA, Amrit FR, Ratnappan R, Ghazi A (2014) The C. elegans healthspan and stress-resistance assay toolkit. Methods 68:476–486

Kenyon CJ (2010) The genetics of ageing. Nature 464:504–512

Kenyon C, Chang J, Gensch E, Rudner A, Tabtlang R (1993) A C. elegans mutant that lives twice as long as wild type. Nature 366:461–464

Lapierre LR, Hansen M (2012) Lessons from C. elegans: signaling pathways for longevity. Trends Endocrinol Metab 23:637–644

Leitsalu L, Haller T, Esko T, Tammesoo ML, Alavere H, Snieder H, Perola M, Ng PC, Mägi R, Milani L, Fischer K, Metspalu A (2015) Cohort profile: estonian biobank of the

estonian genome center, University of Tartu. Int J Epidemiol 44(4):1137–1147

Meléndez A, Tallóczy Z, Seaman M, Eskelinen E-L, Hall DH, Levine B (2003) Autophagy genes are essential for dauer development and life-span extension in *C. elegans*. Science 301:1387–1391

Mitnitski A, Song X, Rockwood K (2013) Assessing biological aging: the origin of deficit accumulation. Biogerontology 14:709–717

Montesanto A, Lagani V, Martino C, Dato S, De Rango F, Berardelli M, Corsonello A, Mazzei B, Mari V, Lattanzio F, Conforti D, Passarino G (2010) A novel, population-specific approach to define frailty. Age 32:385–395

Rattan SIS (2013) Healthy ageing, but what is health? Biogerontology 14:673–677

Seals DR, Melov S (2014) Translational geroscience: emphasizing function to achieve optimal longevity. Aging (Albany NY) 6:718–730

Seals DR, Justice JN, LaRocca TJ (2015) Physiological geroscience: targeting function to increase healthspan and achieve optimal longevity. J Physiol (Lond.) 11:1–24

Shadyab AH, LaCroix AZ (2015) Genetic factors associated with longevity: a review of recent findings. Ageing Res Rev 19:1–7

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC (1998) The mini-international neuropsychiatric interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatr 59(Suppl 20):22–33; quiz 34–57

Stevenson M, Bae H, Schupf N, Andersen S, Zhang Q, Perls T, Sebastiani P (2015) Burden of disease variants in participants of the long life family study. Aging (Albany NY) 7:123–132

Stroustrup N, Ulmschneider BE, Nash ZM, López-Moyado IF, Apfeld J (2013) The Caenorhabditis elegans lifespan machine. Nat methods 10(7):665–670

Suda H (2014) Biophysical and biological meanings of healthspan from *C. elegans* cohort. Biochem Biophys Res Commun 452:36–41

Warsow G, Greber B, Falk SSI, Harder C, Siatkowski M, Schordan S, Fuellen G (2010) ExprEssence–revealing the essence of differential experimental data in the context of an interaction/regulation net-work. BMC Syst Biol 4(1):164

Willcox BJ, Donlon TA, He Q, Chen R, Grove JS, Yano K, Masaki KH, Willcox DC, Rodriguez B, Curb JD (2008) FOXO3A genotype is strongly associated with human longevity. Proc Natl Acad Sci USA 105:1398