

Published in final edited form as:

*Exp Gerontol.* 2012 May ; 47(5): 379–387. doi:10.1016/j.exger.2012.02.010.

## Human Longevity and Variation in GH/IGF-1/Insulin Signaling, DNA Damage Signaling and Repair and Pro/antioxidant Pathway Genes: Cross Sectional and Longitudinal Studies

Mette Soerensen<sup>1,2</sup>, Serena Dato<sup>1,3</sup>, Qihua Tan<sup>1,2</sup>, Mikael Thinggaard<sup>1</sup>, Rabea Kleindorp<sup>4</sup>, Marian Beekman<sup>5,6</sup>, Rune Jacobsen<sup>1</sup>, H. Eka D. Suchiman<sup>5,6</sup>, Anton J.M. de Craen<sup>7</sup>, Rudi G.J. Westendorp<sup>7</sup>, Stefan Schreiber<sup>4</sup>, Tinna Stevnsner<sup>8</sup>, Vilhelm A. Bohr<sup>8,9</sup>, P. Eline Slagboom<sup>5,6</sup>, Almut Nebel<sup>4</sup>, James W. Vaupel<sup>1,10</sup>, Kaare Christensen<sup>1,2</sup>, Matt McGue<sup>1,11</sup>, and Lene Christiansen<sup>1,2,\*</sup>

<sup>1</sup>The Danish Aging Research Center, Epidemiology, Institute of Public Health, University of Southern Denmark, J.B. Winsloews Vej 9B, 5000 Odense C, Denmark <sup>2</sup>Department of Clinical Genetics and Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Sdr. Boulevard 29, 5000 Odense C, Denmark <sup>3</sup>Department of Cell Biology, University of Calabria, Ponte Pietro Bucci cubo 4C, 87036 Rende (Cs), Italy <sup>4</sup>Institute of Clinical Molecular Biology, Christian-Albrechts-University and University Hospital Schleswig-Holstein, Schittenhelmstrasse 12, 24105 Kiel, Germany <sup>5</sup>Department of Molecular Epidemiology, Leiden University Medical Center, S5-P, PO Box 9600, 2300 RC Leiden, the Netherlands <sup>6</sup>Netherlands Consortium for Healthy Aging, Leiden University Medical Center, 2300 RC, Leiden, the Netherlands <sup>7</sup>Department of Gerontology and Geriatrics, Leiden University Medical Center, C2-R, PO Box 9600, 2300 RC Leiden, the Netherlands <sup>8</sup>The Danish Aging Research Center, Department of Molecular Biology and Genetics, Aarhus University, C. F. Moellers Allé 3, 8000 Aarhus C, Denmark <sup>9</sup>Laboratory of Molecular Gerontology, National Institute on Aging, National Institute of Health, 5600 Nathan Shock Drive, 21224-6825 Baltimore, Maryland, USA <sup>10</sup>Max Planck Institute for Demographic Research, Konrad-Zuse-Str. 1, 18057 Rostock, Germany <sup>11</sup>Department of Psychology, University of Minnesota, N218 Elliott Hall 75 East River Road, MN 55455-0344 Minneapolis, MN, USA

### Abstract

Here we explore association with human longevity of common genetic variation in three major candidate pathways: GH/IGF-1/insulin signaling, DNA damage signaling and repair and pro/antioxidants by investigating 1273 tagging SNPs in 148 genes composing these pathways. In a case-control study of 1089 oldest-old (age 92–93) and 736 middle-aged Danes we found 1 pro/antioxidant SNP (rs1002149 (*GSR*)), 5 GH/IGF-1/INS SNPs (rs1207362 (*KL*), rs2267723 (*GHRHR*), rs3842755 (*INS*), rs572169 (*GHSR*), rs9456497 (*IGF2R*)) and 5 DNA repair SNPs (rs11571461 (*RAD52*), rs13251813 (*WRN*), rs1805329 (*RAD23B*), rs2953983 (*POLB*), rs3211994 (*NTLHI*)) to be associated with longevity after correction for multiple testing.

© 2012 Elsevier Inc. All rights reserved

\*Author for correspondence: Lene Christiansen: phone: 0045-65503378, fax: 0045-65503682, lchristiansen@health.sdu.dk, Epidemiology, Institute of Public Health, University of Southern Denmark, J.B. Winsloews Vej 9B, 5000 Odense C, Denmark.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**8. CONFLICTS OF INTEREST** The authors have no conflicts to declare.

In a longitudinal study with 11 years of follow-up on survival in the oldest-old Danes we found 2 pro/antioxidant SNPs (rs10047589 (*TNXRDI*), rs207444 (*XDH*)), 1 GH/IGF-1/INS SNP (rs26802 (*GHRL*)) and 3 DNA repair SNPs (rs13320360 (*MLH1*), rs2509049 (*H2AFX*) and rs705649 (*XRCC5*)) to be associated with mortality in late life after correction for multiple testing.

When examining the 11 SNPs from the case-control study in the longitudinal data, rs3842755 (*INS*), rs13251813 (*WRN*) and rs3211994 (*NTHL1*) demonstrated the same directions of effect ( $p < 0.05$ ), while rs9456497 (*IGF2R*) and rs1157146 (*RAD52*) showed non-significant tendencies, indicative of effects also in late life survival. In addition, rs207444 (*XDH*) presented the same direction of effect when inspecting the 6 SNPs from the longitudinal study in the case-control data, hence, suggesting an effect also in survival from middle age to old age.

No formal replications were observed when investigating the 11 SNPs from the case-control study in 1613 oldest-old (age 95–110) and 1104 middle-aged Germans, although rs1157146 (*RAD52*) did show a supportive non-significant tendency (OR = 1.162, 95% CI = 0.927–1.457). The same was true for rs10047589 (*TNXRDI*) (HR = 0.758, 95% CI = 0.543–1.058) when examining the 6 SNPs from the longitudinal study in a Dutch longitudinal cohort of oldest-old (age 85+, N = 563).

In conclusion, the present candidate gene based association study, the largest to date applying a pathway approach, points to potential new longevity loci, but does also underline the difficulties of replicating association findings in independent study populations and thus the difficulties in identifying universal longevity polymorphisms.

## Keywords

human longevity; association study; case-control data; longitudinal data

## 1. INTRODUCTION

Genetic factors contribute approximately 25% to variation in human life span (Herskind et al. (1996)), an effect thought to be minimal before age 60 years and most profound from age 85 years and onwards (Hjelmborg et al. (2006)). Candidate longevity genes encode proteins involved in several biological processes including lipid-protein metabolism (*APOE* being by far the most validated longevity gene to date), growth hormone 1/insulin-like growth factor 1/insulin (GH/IGF-1/INS) signaling, DNA damage signaling, and repair and pro/antioxidant pathways (Christensen et al. (2006)). The genes composing the latter three pathways encode proteins involved in processes such as energy metabolism, growth and differentiation, and the maintenance of genetic stability (including telomere stability), and protection against free radicals produced by the general metabolism. These are all processes believed to be of key importance for aging and longevity from nematodes to humans (reviewed in e.g. (Burtner and Kennedy (2010); Cluett and Melzer (2009); Harman (1973); Kenyon (2010); Kuningas et al. (2008))). Studies using various animal models have found that either depletion or over-expression of some of these candidate genes affect the life span of the organism (e.g. *GH*, *GPX1*, *FOXO*, *KL*, *GHRHR*, *POLB*, *RAD52*, *INS* and *TP53* (Bartke (2005); Cabelof et al. (2006); Flurkey et al. (2001); Giannakou et al. (2004); Holzenberger et al. (2003); Kawano et al. (2000); Kuro-o M et al. (1997); Kurosu et al. (2005); Park et al. (1999); Tyner et al. (2002); Mockett et al. (1999))), whereas manipulating other genes leads to aging-like phenotypes (e.g. *PON1* and *CSB* (Shih et al. (1998); van der Horst et al. (1997))).

In humans rare mutations that severely affect the function of some of these genes (*WRN*, *CSB* and *RECQLA*) can cause segmental progeroid premature aging syndromes such as Werner's, Cockayne's and Rothmund-Thomson syndromes (reviewed in (Burtner and

Kennedy (2010))), whereas more common genetic variation has been reported to be associated with human lifespan. Examples from the GH/IGF-1/INS pathway include *IGF1R*, *GHR* and *FOXO3A*, from the DNA damage signaling and repair pathway *WRN*, *MLH1* and *TP53* and from the pro/antioxidant pathway *SOD1*, *SOD2* and *PON1* (Bonafe et al. (2003); Castro et al. (2000); Christiansen et al. (2004); Flachsbarth et al. (2009a); Kim et al. (2006); van Heemst et al. (2005a); van Heemst et al. (2005b); Willcox et al. (2008a); Yashin et al. (2000)).

The vast majority of the candidate genetic association studies published so far focus on a single or a few single nucleotide polymorphisms (SNPs) or candidate genes, although recently a few studies have initiated the investigation of SNPs in groups of candidate genes and pathways based on the functional roles of the gene products (Deelen et al. (2011); Flachsbarth et al. (2010); Nebel et al. (2009); Pawlikowska et al. (2009a)). In addition, previous studies have generally applied a case-control study design, thus raising concerns of bias introduced by differences in characteristics of cases and controls, for example cohort effects.

Here we investigate the association of common genetic variation in 148 genes composing the GH/IGF-1/INS signaling, DNA damage signaling and DNA repair and pro/antioxidant pathways with human longevity. The study populations are a cohort of Danes born in 1905 and followed from age 93 years until death, as well as a group of middle-aged controls. This provides a unique opportunity to investigate these variations in key candidate genes of human longevity using both case-control and longitudinal designs. Further, the SNPs identified in the case-control study are explored in the longitudinal data in order to investigate whether these SNPs also pose effects on survival during old age, and the SNPs identified in the longitudinal study are examined in the case-control data to investigate whether these SNPs also pose effects on survival from middle age to old age.

Finally, considering the importance of replicating initial results in genetic association studies, the SNPs found to be significant in the case-control study are tested for association in a German case-control sample (1613 long-lived individuals and 1140 controls), while the findings from the Danish longitudinal study are tested in the prospective Dutch Leiden 85-plus Study (563 individuals).

## 2. MATERIAL AND METHODS

### 2.1. Subjects

**2.1.1. Discovery cohorts**—The study population of oldest-old were 1200 participants from The Danish 1905 Birth Cohort Study (1905 cohort) (Nybo et al. (2001)). The study was initiated in 1998, when participants were 92.2–93.8 years of age (mean age 93.2). Survivors were re-assessed in 2000, 2003 and 2005. Vital status was followed until death or January 1<sup>st</sup> 2010, whichever came first, resulting in a mean follow-up time for survivors of 11.4 years (range: 11.2–11.6). Survival status information was retrieved from the Danish Central Person Register, which is continuously updated (Pedersen et al. (2006)). The younger control group (46.0–55.0 years of age, mean age 50.5) was 800 individuals from The Study of Middle Aged Danish Twins (MADT) (Skytthe et al. (2002)). MADT was initiated in 1998 by random selection of 2640 intact twin pairs from 22 consecutive birth years (1931–1952) via the Danish Central Person Register. Only one twin from each twin pair was included in the present study. Permission to collect blood samples and usage of register based information was granted by The Danish National Committee on Biomedical Research Ethics for both surveys. All participants gave their informed consent.

After data cleaning, the final sample size for data analysis was 1089 oldest-old and 736 middle-aged individuals. Of these 1089 oldest-old, 14 were alive by January 1<sup>st</sup> 2010, hence, for the longitudinal analysis their remaining life expectancies were imputed using [www.mortality.org](http://www.mortality.org), which, by appliance of data from the Statistics Denmark, contains cohort mortality data for the Danish population. Of the 736 MADT individuals, 35 had died since 1998. Based on period life table data for the Danish population ([www.mortality.org](http://www.mortality.org)) it is anticipated that only approximately 9% of the remaining 701 individuals will become 93 years of age.

**2.1.2. Replication samples**—For replication of findings from the Danish case-control study, two German samples of 1613 unrelated long-lived individuals (age 95–110) and 1140 middle-aged controls (mean age 67.2) were used. The participants were identified via the local registry offices and the Federal Administrative Office. They were recruited from different geographic regions of Germany and were all of German ancestry (Nebel et al. (2005)). Approval was received from the Ethics Committee of the University Hospital Schleswig-Holstein and from local data protection authorities. All participants gave their informed consent.

For replication of findings from the Danish longitudinal study, the Dutch Leiden 85-plus Study was used. This prospective cohort was initiated in 1997, when all 85-year-old inhabitants of the city of Leiden in the Netherlands were invited to participate. The sample used here consists of 563 participants, all Caucasians and members of the 1912–1914 birth cohort recruited from 1997–1999 (Bootsma-van der Wiel et al. (2002)). Mean follow up time for survivors is 10.35 years (range 9–11 years). The Medical Ethical Committee of the Leiden University Medical Centre approved the study. All participants gave their informed consent.

## 2.2 Selection of candidate genes

A comprehensive literature and data base search was conducted for the identification of longevity candidate genes in the GH/IGF-1/INS signaling, DNA damage signaling and repair and pro/antioxidant pathways. The following databases were used: [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov) (/sites/entrez, /OMIM, /sites/entrez?Db=gap), [www.genecards.org](http://www.genecards.org), [www.hgvbase2p.org](http://www.hgvbase2p.org), [genomics.senescence.info](http://genomics.senescence.info) and [geneticassociationdb.nih.gov](http://geneticassociationdb.nih.gov) using the search terms `human longevity', `human aging', `age related disease' (including specific diseases e.g. myocardial infarction) and `premature aging syndromes'. Additionally, genes previously identified through animal models were considered. In order to cover the core biological functions of the three candidate pathways, the needed genes were chosen using the [www.ncbi.nlm.nih.gov/sites/entrez](http://www.ncbi.nlm.nih.gov/sites/entrez), [www.biocarta.com](http://www.biocarta.com), [www.reactome.org](http://www.reactome.org), [www.genome.utah.edu/genesnps](http://www.genome.utah.edu/genesnps) and the [www.genome.jp/kegg/pathway.html](http://www.genome.jp/kegg/pathway.html) databases. In total 152 genes composing the three candidate pathways were identified, 148 of which remained after cleaning of the GoldenGate genotyping data (see below). The 148 genes, divided by pathway, are listed in Supplementary Table 1.

## 2.3 Selection of SNPs

The chromosomal regions composing the 152 candidate genes were identified through the [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov) (/gene and /mapview) and [genome.ucsc.edu/](http://genome.ucsc.edu/) databases, using the same gene ID, Sequence Accession IDs, NCBI Reference Sequence, assembly build 36 and SNP build 129. If more than one isoform of a gene transcript existed, the longest one was chosen. For each gene, 5000 base pairs upstream and 1000 base pairs downstream were included. If another gene was located within these regions, only the region until the neighboring gene was included. Due to the size of the genotyping platform, the large intron 2 of *IGF1R* and the large intron 2 of *INSR* were excluded. Candidate SNPs were identified

based on the following criteria: SNPs previously identified in genetic association studies, coding SNPs and SNPs having potential functional impact (coding non-synonymous SNPs, SNPs located in potential splice sites or transcription factor binding sites and SNPs potentially inducing frame shifts or nonsense-mediated mRNA decay). For the identification of potential functional SNPs, the [www.ncbi.nlm.nih.gov/SNP](http://www.ncbi.nlm.nih.gov/SNP), [koreanbio.org/Variome](http://koreanbio.org/Variome), [snpper.chip.org](http://snpper.chip.org), [snps3d.org](http://snps3d.org) and [manticore.niehs.nih.gov](http://manticore.niehs.nih.gov) data bases were used. The identity of each SNP was checked by linking back to the NCBI data base.

In order to cover as much of the common genetic variation within each gene region as possible, a number of tagging SNPs in each gene region was added, obtained through the HapMap consortium database (<http://hapmap.ncbi.nlm.nih.gov/index.html.en>) for the CEU cohort, using the HapMap Data Rel 23a/phase II Mar08, on NCBI B36 assembly, dbSNP build 126 criteria. Genotype data were analyzed using the HaploView software (<http://www.broadinstitute.org/haploview/haploview> and (Barrett et al. (2005))), by applying the 'pair wise tagging only',  $r^2 > 0.8$ , LOD=3, minor allele frequency (MAF) > 5%, and a minimum distance between SNPs = 60 bp criteria. Finally, for optimizing the quality of the genotype data, SNPs known to perform poorly on the Illumina GoldenGate genotyping platform were excluded. In total, 1415 SNPs were chosen and 1273 were successfully genotyped.

## 2.4 Genotyping and data cleaning

DNA was isolated from blood spot cards using the QIAamp DNA Mini and Micro Kits (Qiagen, Düsseldorf, Germany). The genotyping was conducted using the Illumina GoldenGate platform (Illumina Inc, San Diego, CA, USA) and data cleaning was carried out according to the manufacturers recommendations using the GenomeStudio software ([http://www.illumina.com/software/genomestudio\\_software.ilmn](http://www.illumina.com/software/genomestudio_software.ilmn)); primary exclusion criteria were sample call rate < 90% and SNP call frequency < 90%. Subsequently, SNPs clustering in the following categories were manually evaluated: a call frequency 90–95%, close SNP clusters (score < 2.3), clusters with low intensity (score < 0.2 or > 0.8), cluster having a heterozygote cluster shifting towards a homozygote cluster (score < 0.13), SNPs with excess heterozygosity (score < -0.3 or > 0.2) and SNPs located on the x chromosome (considering male hemizyosity). In total, 142 SNPs and 175 DNA samples were excluded leaving 1825 individuals, 148 candidate genes and 1273 SNPs for data analysis. Furthermore, 24 samples were included twice in order to investigate the intra-plate reproducibility and 48 samples were included twice to test the inter-plate reproducibility. Data showed an intra-plate reproducibility of 99.4% and an inter-plate reproducibility of 96.8%.

Genotyping of the German replication samples were performed using either Sequenom MassARRAY iPLEX®Gold technology (Sequenom®, Inc., CA, USA) with iPlex primers designed via the MassARRAY Design Software, or via an automated platform using predesigned TaqMan® SNP genotyping assays (Applied Biosystems) (Hampe et al. (2001)); PCR was performed using the GeneAmp PCR System 9700 (Applied Biosystems) under standard conditions and the ABI Prism 7900HT Sequence Detection System and the Sequence Detection System Software 2.0 (Applied Biosystems) were used for allelic discrimination.

The Dutch replication samples were genotyped via Sequenom MassARRAY iPLEX®Gold technology (Sequenom®, Inc., CA, USA), using primers designed via the SpectroDESIGNER (Sequenom®, Inc., CA, USA) and the high plex reaction protocol ([www.sequenom.com/iplex](http://www.sequenom.com/iplex)). Since rs207444 and rs3211994 were not compatible with the iPLEX format, genotyping was conducted using assay-on-demand TaqMan® SNP genotyping assays (Applied Biosystems); PCR was performed on the GeneAmp PCR system

9700 (Applied Biosystems, CA, USA) under universal conditions and the Applied Biosystems 7900HT Fast Real-time PCR System and SDS2.3 Software (Applied Biosystems, CA, USA) were used for allelic discrimination.

## 2.5 Statistical analysis

In the case-control study of the Danish discovery cohorts single-SNP case-control comparisons of allele frequencies (CCA) was performed using the Plink statistical program (<http://pngu.mgh.harvard.edu/purcell/plink>, (Purcell et al. (2007))) by two-tailed Chi-square test. Replication of the findings from the Danish case-control study was conducted by CCA in the German case-control samples in the same way.

Approximately 1 in 20 members of the 1905 birth cohort in Denmark had survived until 1998 (start of the Danish 1905 Birth Cohort Study) and approximately 1 in 20 of these members survived until 100 years of age, i.e. reaching age 92–93 (the inclusion age of the Danish 1905 Birth Cohort Study participants) is only “halfway to becoming a centenarian” in terms of selection. In the longitudinal survival analysis of the oldest-old from the Danish 1905 Cohort Study, two survival variables were used. First, we wanted to investigate the genetic contribution to surviving for only a short period of time as opposed to surviving for a longer period of time during the ninth decade of life. Since the distribution of survival times was clearly right skewed (data not shown), a cut-off with respect to time of death was made at December 31<sup>st</sup> 2000, when approximately half of the males and one third of the females in the cohort had died. Consequently, the early/late death variable was defined as: early = death occurred between 1998 (date at intake) and December 31<sup>st</sup> 2000 (i.e. individuals living to maximum age 95) and late = death occurred after January 1<sup>st</sup> 2001 (i.e. individuals surviving to age 95+). As the second variable, we used the number-of-days-lived defined as the exact number of days lived from date of intake in 1998 to either death or the imputed date (the latter is described in section 3.1.1). For this variable, the assumptions of normal distribution and equal variance were tested in the STATA 11.1 statistical program (Stata Corporation, College Station, TX, USA) by Shapiro-Wilk and Breusch-Pagan/Cook-Weisberg tests. Since the variable did not show normal distribution, it was transformed to the square root for a better fit. Association analysis was performed by two-tailed linear (for the number-of-days-lived variable) or logistic (for the early/late death variable) regression in Plink, adjusting for sex and age or stratifying by sex, while adjusting for age.

Replication of the findings from the Danish longitudinal study was performed in Stata 11.1 by analyzing genotype data from the Leiden 85-plus Study using a sex-adjusted Cox proportional hazard model. An additive model was used, and only in cases when inspection of Kaplan-Meier plots indicated a recessive or dominant model, such a model was applied.

In order to obtain a quantitative estimate for the differences in mean survival time between individuals homozygous for the common alleles of a SNP found to be associated with the Number-of-days-lived variable and individuals being either heterozygous or homozygous for the rare allele at the same SNP, regression analysis was performed for same sex and same age (93 years) individuals and the regression coefficients were used for estimation. Moreover, the relative risks (RRs) of being alive after age 95 (corresponding to the 1<sup>st</sup> of January 2001) were calculated for the heterozygotes and homozygotes for the rare allele of each SNP found to be associated to the early/late death variable. The RRs between genotype groups were estimated based on the ORs obtained in the logistic regression analysis. All these analyses were performed in STATA 11.1.

For all the analyses, the nominal significance level was set to 0.05. Correction for multiple testing within each pathway was performed by permutation approach applying max(T) permutation mode set at 10000 permutations for analyses carried out in the Plink software.

Correction for multiple testing in analyses performed in the STATA software was done using the Bonferroni approach.

### 3. RESULTS

By conducting a thorough literature and data base search candidate genes, candidate SNPs and tagging SNPs were identified (see section 2.2 and 2.3 for details). A total of 148 candidate genes and 1273 SNPs were successfully genotyped and are listed in Supplementary Table 1: these constitute 592 SNPs in 77 genes of the DNA damage signaling and DNA repair pathway, 370 SNPs in 33 genes of the GH/IGF-1/INS signaling pathway and 311 SNPs in 38 pro/antioxidant genes. The characteristics of the Danish discovery cohorts of middle-aged and oldest-old individuals are described in Table 1A, while the German and Dutch replication samples are described in Table 1B.

#### 3.1. Case-control comparison of middle-aged and oldest-old Danes

Exploiting the case-control allele comparisons revealed considerably more nominally significant differences than would be expected by chance (see Supplementary Table 2). After adjusting allelic comparisons for multiple testing by permutation, 10 of the 1273 SNPs remained significant; of these SNPs the minor allele of the GH/IGF-1/INS-SNPs rs1207362 (*KL*), rs2267723 (*GHRHR*) and rs572169 (*GHSR*), as well as the pro/antioxidant-SNP rs1002149 (*GSR*) and the DNA-repair-SNP rs1805329 (*RAD23B*) decreased in frequency from middle-aged to old-aged ( $OR < 1$ ), whereas the GH/IGF-1/INS-SNPs rs3842755 (*INS*) and rs9456497 (*IGF2R*) and the DNA-repair-SNPs rs11571461 (*RAD52*), rs13251813 (*WRN*) and rs2953983 (*POLB*) increased in frequency ( $OR > 1$ ). As sex-specific differences previously have been suggested for association of SNPs with longevity (Candore et al. (2006); Franceschi et al. (2000); Li et al. (2009); Soerensen et al. (2010)) a sex-stratified analysis was also performed; the analysis supported the associations of the 10 SNPs mentioned above for both genders (data not shown), and in addition identified 1 DNA-repair-SNP in males; the minor allele of rs3211994 (*NTLHI*) increased in frequency from middle-aged to old-aged. The data are summarized in Table 2.

The 11 SNPs were subsequently analyzed in replication samples of 1613 long-lived and 1104 middle-aged Germans. The MAF of the SNPs in the German and Danish case-control samples are listed in Supplementary Table 3. None of the 11 SNPs formally replicated in the German samples (see Table 3), although rs11571461 (*RAD52*) did show a non-significant tendency of association in the same direction (95% CI (OR) = [0.927–1.457]).

#### 3.2. Longitudinal study in the cohort of oldest-old Danes

As for the case-control study, analysis of survival data in the oldest-old Danes revealed more nominally significant associations than would be expected by chance, whether using the categorical early/late death or the numeric number-of-days-lived variables (see Supplementary Table 4). After correcting for multiple testing 2 SNPs were found to be associated with late life survival for both genders combined: carriers of the rare allele of the pro/antioxidant-SNP rs10047589 (*TXNRDI*) displayed a significant increase in survival when categorizing into early/late death, whereas the rare allele of the pro/antioxidant-SNP rs207444 (*XDH*) was found to be significantly associated with increased number of days lived (see Table 4A). When stratifying by sex, these results were supported in both genders (data not shown). Moreover, when applying the early/late death variable, positive effects of the rare alleles of the GH/IGF-1/INS-SNP rs26802 (*GHRL*) and the DNA-repair-SNP rs13320360 (*MLHI*) on survival were observed in males, while the DNA-repair-SNP rs2509049 (*H2AFX*) had a negative effect on survival in females. Using number-of-days-

lived as the measure of survival, the rare allele of the DNA-repair-SNP rs705649 (*XRCC5*) positively affected survival in males. The results are summarized in Table 4B.

Since we performed the longitudinal analyses as a regression analyses, we were able to estimate the quantitative effects of the 6 SNPs, represented by the relative risks (RR) for the SNPs associated with the early/late death variable, and by the additional survival time for the SNPs associated with the Number\_of\_days\_lived variable. These estimates indicated rather substantial effects: a maximum RR was seen for rs26802 (*GHRL*) CC males having a 53% higher chance for being alive after age 95 (corresponding to the 1<sup>st</sup> of January 2001) as compared to AA homozygotes, while a maximum difference of 1.66 additional years of survival was seen for rs705649 (*XRCC5*) GG males as compared to rs705649-AA males, who obtained a mean survival time of 2.15 years. The data are summarized in Supplementary Table 5A and B.

Next the 6 SNPs were evaluated in a replication cohort of 563 85+ year old Dutch individuals. The frequency data of the 6 SNPs in the Dutch and Danish longitudinal cohorts are listed in Supplementary Table 6. Applying an additive model, none of the SNP associations were replicated in the Dutch cohort (see Table 5). However, inspection of Kaplan-Meier survival plots for the 6 SNPs suggested a recessive mode of effect for rs10047589 (*TNXRDI*) in the Dutch data (data not shown), with a similar direction of effect as in the Danish longitudinal study, although the association was not significant; HR = 0.758,  $p = 0.103$ , 95%, CI = [0.543–1.058].

### 3.3. Similar directions of effect in the Danish case-control and longitudinal studies

We wanted to explore whether the 11 SNPs identified via the Danish case-control study maintained an effect *during* old age. Consequently we investigated the 11 SNPs in the longitudinal data on the oldest-old Danes. Similar directions of effects, i.e. positive effects on survival, were noted for 5 SNPs; when applying the number-of-days-lived variable, rs13251813 (*WRN*) and rs3842755 (*INS*) displayed nominally significant associations ( $p < 0.05$ ), the former in the entire group of oldest-old and the latter when restricting to males. Additionally, rs9456497 (*IGF2R*) showed a borderline significant association when restricting to males ( $p = 0.05$ ). For the early/late death variable, rs3211994 (*NTHL1*) displayed nominal significance in both genders combined, while rs1157146 (*RAD52*) showed a non-significant tendency in females only ( $p = 0.10$ ). None of the remaining six SNPs showed similar directions of effect. The data is summarized in Table 6A.

Finally, in order to explore whether the 6 SNPs, identified in the longitudinal study of oldest-old Danes, presented a similar direction of effects on survival from middle age to old age, we investigated the 6 SNPs in the Danish case-control data. Only rs207444 (*XDH*) displayed a nominally significant effect in the same direction as found in the oldest old ( $p = 0.005$ ). The data is summarized in Table 6B.

## 4. DISCUSSION

The vast majority of the candidate genetic association studies of longevity published so far has explored only a small number of variations and/or genes. The investigation of large collections of candidate variations grouped together based on the functional roles of the gene products is a much more recent strategy (Flachsbart et al. (2010); Nebel et al. (2009); Pawlikowska et al. (2009b)). In this study we investigated 592 SNPs in 77 genes of the DNA damage signaling and DNA repair pathway, 370 SNPs in 33 genes of the GH/IGF-1/INS signaling pathway and 311 SNPs in 38 pro/antioxidant genes, making it the most comprehensive study to date following this strategy. Furthermore, in addition to the widely

used case-control approach, we also present data from a longitudinal analysis of survival in the oldest-old.

In the analysis of the case-control study of 1089 oldest-old and 736 middle-aged Danes, we found 11 novel SNPs in 11 different genes to be associated with human longevity. The rare allele of variants in *INS*, *IGF2R*, *POLB*, *RAD52*, *WRN* and *NTHL1* was demonstrated to be advantageous, the latter most evident in males, whereas the rare allele of variants in *KL*, *GHRHR*, *GHSR*, *GSR*, and *RAD23B* was disadvantageous for survival.

None of the findings could be formally confirmed in a German replication cohort of comparable size, although rs11571461 (*RAD52*) did show a tendency in the same direction as found in the Danish data. In support of our findings, other variants in *KL*, *WRN* and *NTHL1* have previously been associated with longevity (Arking et al. (2005); Castro et al. (2000); Yashin et al. (2010)) and rs687045 (*KL*) which was put forward by Lunetta et al. as a mortality SNP during old age (Lunetta et al. (2007)), is in our cohorts in moderate LD ( $R^2 = 0.25$ ) with rs1207362 (*KL*) described in our study as disadvantageous for longevity. Moreover, supportive of an association of *INS*, a very recent publication of gene-based case-control analysis of genes in the IGF-1/Insulin pathway reported gene-based association of *INS* with longevity (Deelen et al. (2011)). Additionally, some of the 11 SNPs have been reported as associated with phenotypes of possible relevance to variation in lifespan; rs572169 (*GHSR*) with obesity (Gueorguiev et al. (2009)), and rs2953983 (*POLB*), rs1805329 (*RAD23*), rs11571461 (*RAD52*) and rs13251813 (*WRN*) with different types of cancers (Pan et al. (2009); Schildkraut et al. (2010); Van Ness B. et al. (2008); Yang et al. (2009)).

An interesting question was whether the genetic variants with a proposed effect on survival from middle age to old age would also demonstrate an effect for survival *during* old age. We therefore tested the 11 SNPs for association to survival in the Danish 1905 cohort using the longitudinal design and found 3 SNPs to display the same nominal significant ( $p < 0.05$ ) direction of effect in the longitudinal data, while 2 SNPs showed similar tendencies ( $0.05 < p < 0.10$ ). Interestingly, all 5 SNPs were apparent longevity SNPs, i.e. SNPs with a positive effect on survival. This is in line with a recent study suggesting that long life is not due to a decreased burden of negatively affecting disease variants but rather due to positively influencing longevity variants (Beekman et al. (2010)). A lack of similar effects in the longitudinal data for the remaining SNPs could indicate that these SNPs may be relevant at different ages, i.e. the influence of genes and gene variants may be of consequence for survival in young elderly, but lose this importance at the highest ages.

When investigating the association of the 1273 SNPs with survival at advanced age using a longitudinal study approach, we identified 6 novel SNPs in 6 genes to be associated with longevity. The rare allele of a variant in *H2AFX* appeared disadvantageous, while the rare variants in *GHRL*, *MLH1*, *TXNRD1*, *XDH* and *XRCC5* were advantageous for longevity. In line with these results an *MLH1* haplotype was previously reported to be associated with longevity in a case-control study (Kim et al. (2006)), whereas genetic variants in *GHRL* have been described as associated with phenotypes of possible relevance to variation in lifespan; high HDL-cholesterol levels (Hubacek et al. (2007)) and hypertension (Berthold et al. (2010)). Finally, the activity of the *XRCC5* protein product Ku80 has been reported to decrease from young to old age in humans (Scarpaci et al. (2003)), indicating a biological change with age.

The Danish case-control study supported an advantageous effect of rs207444 (*XDH*) also on survival from middle age to old age, and additionally some support was observed for

rs10047589 (*TNXRDI*) in the Dutch replication cohort, indicating a recessive mode of effect for this variant.

Interestingly, the longitudinal study revealed more longevity-associated SNPs in the sex-stratified analysis (4 of 6 SNPs) than the Danish case-control study (1 of 11 SNPs). This observation could indicate that the impact of genetic effect becomes more gender-specific during advanced age than at younger ages, and that the genetic effect in old age might be most pronounced for males. Such a difference between genders was recently suggested by Montesanto et al. for Italians (Montesanto et al. (2011)).

From a pathway perspective a few comments can be made on the genes holding the SNPs suggested here. For the GH/IGF1/INS pathway the present data suggests that variations in genes whose products take part in growth hormone synthesis and secretion (*GHRHR*, *GHRL* and *GHSR*) or in signaling to and effecting target organs/cells (*INS*, *KL*, *IGF2R*) might pose effects on longevity. With respect to DNA repair, the gene products of the genes holding the significant SNPs take part in different DNA repair activities, though the majority take part in base excision repair (*NTHL1*, *POLB*, and *WRN*) or in DNA double strand break repair (*RAD52*, *H2AFX*, *XRCC5* and *WRN*), thus indicating that variation in the repair of DNA damage induced by for instance reactive oxygen species and ionizing radiation could pose effects on longevity. Lastly, with respect to the pro/antioxidant pathway both antioxidants (*GSR* and *TNXRDI*) and a prooxidant (*XDH*) are observed, indicating that the role of this pathway in longevity appears complex.

The SNPs suggested here to associate to longevity might induce different functional effects; rs1002149 (*GSR*) and rs2509049 (*H2AFX*) located 5' of genes might affect the level of gene expression, while rs3842755 (*INS*), rs3211994 (*NTHL1*) and rs10047589 (*TNXRDI*) are all located 3' and could thus in principle influence decay of mRNA. The non-synonymous coding SNP rs1805329 (*RAD23B*) could potentially affect protein activity, while the intronic SNPs (rs1207362 (*KL*), rs2267723 (*GHRHR*), rs9456497 (*IGF2R*), rs11571461 (*RAD52*), rs13251813 (*WRN*), rs2953983 (*POLB*), rs207444 (*XDH*), rs26802 (*GHRL*), rs13320360 (*MLHI*) and rs705649 (*XRCC5*)) could possibly induce splicing alterations. Alternatively, the associated SNPs might simply be proxies for the other truly causative variants.

How such suggested functional effects might impact aging and/or longevity is an extremely difficult question, especially considering that the role of the three pathways in aging and/or longevity appears very complex. In any case, functional studies of the molecular effects of the associated variants are needed

Due to the emergence of high throughput techniques facilitating studies of sizeable numbers of variations, the replication of initial findings has become a major issue in genetic epidemiology. In this study we tested our findings from the case-control and the longitudinal study using German samples of 1613 long-lived individuals (95+) and 1104 middle-aged controls, and a cohort of 85+ Dutch elderly, respectively. No formal replication was observed in either case although a few supportive observations were found (see above).

A number of reasons can be given for the lack of true replication. It could be a matter of under-powered replication populations. However, the number of German samples is larger than the Danish discovery cohorts, thus arguing against this explanation. Alternatively it might be due to heterogeneity between samples, for instance differences in enrollment of the participants. As seen in Table 1, the German cases are older than the Danish oldest-old, whereas the members of the Dutch 85+ cohort are somewhat younger at intake than the Danish 1905 cohort members. Hence, if the effects of the variants are not constant throughout old age, they may not replicate well in these study samples. Change in

association effects with age has been reported previously (Jacobsen et al. (2010)). Finally, it might simply be that the associations are population specific, perhaps mediated through population specific characteristics in environment. Such populations specific effects have previously been suggested (Di Bona et al. (2009)).

In our study we only performed correction of multiple testing within each pathway, since the selection of genes and SNPs were based on a sound a priori hypothesis. Moreover, due to the design of the statistical software, we corrected for multiple testing within each type of test, and hence, we did not correct for performing the sex-stratified analysis as well as a test for both genders combined, and with respect to the longitudinal study for investigating the 2 survival variables. Such corrections should in principle be included, in order to avoid false positive findings. If correcting the p-values from each pathway using the overly conservative Bonferroni approach 8 of the 11 SNPs from the case-control study remained significant, while 1 of the 6 SNPs identified in the longitudinal study remained significant. By contrast, correction by the less conservative Benjamini and Hochberg method (Benjamini and Hochberg (1995)), left more associations significant (10 SNPs from the case-control study and 1 SNP from the longitudinal study), thus underscoring the difference in conclusions based on choice of method of correction. In any case, both of these correction methods assume independency between the tests, which is clearly not the case in our study.

In conclusion we here present new candidate variations for human longevity located in the DNA damage signaling and DNA repair, the GH/IGF-1/INS signaling and pro/antioxidant pathways. We find that 4 of the 17 identified SNP associations show significant and similar direction of effect when using the two different study approaches applied to the Danish discovery cohorts, while 2 SNP associations showed similar tendencies. However, none of the 17 SNP associations displayed replication in the independent samples from Germany and the Netherlands, although for 2 SNP associations some support was noticed.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This study was supported by the Max-Planck Institute for Demographic Research (Rostock, Germany), the INTERREG 4 A programme Syddanmark-Schleswig-K.E.R.N (by EU funds from the European Regional Development Fund), the National Institute on Aging (P01 AG08761), a Dutch grant from the Netherlands Consortium for Healthy Aging (Grant 050-060-810) in the framework of the Netherlands Genomics Initiative / Netherlands Organization for Scientific Research (NWO), the European Union's Seventh Framework Programme (FP7/2007–2011) under grant agreement n° 259679, the Novo Nordisk Foundation, the Aase and Ejnar Danielsen Foundation, the Augustinus Foundation, the Brødrene Hartmann Foundation, the King Christian the 10<sup>th</sup> Foundation and the Einer Willumsens Mindelegat Foundation. The Danish Aging Research Center is supported by a grant from the VELUX Foundation. Susanne Knudsen, Steen Gregersen, Ulla Munk, Shuxia Li, Anne Mette Hedegaard Nielsen, Marlene Graff Sørensen and Lene Elnegaard are thanked for excellent technical work.

## 9. REFERENCES

- Anselmi CV, Malovini A, Roncarati R, Novelli V, Villa F, Condorelli G, Bellazzi R, Puca AA. Association of the FOXO3A locus with extreme longevity in a southern Italian centenarian study. *Rejuvenation.Res.* 2009; 12:95–104. [PubMed: 19415983]
- Arking DE, Atzmon G, Arking A, Barzilai N, Dietz HC. Association between a functional variant of the KLOTHO gene and high-density lipoprotein cholesterol, blood pressure, stroke, and longevity. *Circ.Res.* 2005; 96:412–418. [PubMed: 15677572]
- Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005; 21:263–265. [PubMed: 15297300]

- Bartke A. Minireview: role of the growth hormone/insulin-like growth factor system in mammalian aging. *Endocrinology*. 2005; 146:3718–3723. [PubMed: 15919742]
- Beekman M, Nederstigt C, Suchiman HE, Kremer D, van der BR, Lakenberg N, Alemayehu WG, de Craen AJ, Westendorp RG, Boomsma DI, de Geus EJ, Houwing-Duistermaat JJ, Heijmans BT, Slagboom PE. Genome-wide association study (GWAS)-identified disease risk alleles do not compromise human longevity. *Proc.Natl.Acad.Sci.U.S.A.* 2010; 107:18046–18049. [PubMed: 20921414]
- Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society B.* 1995; 57:289–300.
- Berthold HK, Giannakidou E, Krone W, Tregouet DA, Gouni-Berthold I. Influence of ghrelin gene polymorphisms on hypertension and atherosclerotic disease. *Hypertens.Res.* 2010; 33:155–160. [PubMed: 20010782]
- Bonafe M, Barbieri M, Marchegiani F, Olivieri F, Ragno E, Giampieri C, Mugianesi E, Centurelli M, Franceschi C, Paolisso G. Polymorphic variants of insulin-like growth factor I (IGF-I) receptor and phosphoinositide 3-kinase genes affect IGF-I plasma levels and human longevity: cues for an evolutionarily conserved mechanism of life span control. *J.Clin.Endocrinol.Metab.* 2003; 88:3299–3304. [PubMed: 12843179]
- Bootsma-van der Wiel A, Van Exel E, de Craen AJ, Gussekloo J, Lagaay AM, Knook DL, Westendorp RG. A high response is not essential to prevent selection bias: results from the Leiden 85-plus study. *J.Clin.Epidemiol.* 2002; 55:1119–1125. [PubMed: 12507676]
- Burtner CR, Kennedy BK. Progeria syndromes and ageing: what is the connection? *Nat.Rev.Mol.Cell Biol.* 2010; 11:567–578. [PubMed: 20651707]
- Cabelof DC, Ikeno Y, Nyska A, Busuttill RA, Anyangwe N, Vijg J, Matherly LH, Tucker JD, Wilson SH, Richardson A, Heydari AR. Haploinsufficiency in DNA polymerase beta increases cancer risk with age and alters mortality rate. *Cancer Res.* 2006; 66:7460–7465. [PubMed: 16885342]
- Candore G, Balistreri CR, Listi F, Grimaldi MP, Vasto S, Colonna-Romano G, Franceschi C, Lio D, Caselli G, Caruso C. Immunogenetics, gender, and longevity. *Ann.N.Y.Acad.Sci.* 2006; 1089:516–537. [PubMed: 17261795]
- Castro E, Edland SD, Lee L, Ogburn CE, Deeb SS, Brown G, Panduro A, Riestra R, Tilvis R, Louhija J, Penttinen R, Erkkola R, Wang L, Martin GM, Oshima J. Polymorphisms at the Werner locus: II. 1074Leu/Phe, 1367Cys/Arg, longevity, and atherosclerosis. *Am.J.Med.Genet.* 2000; 95:374–380. [PubMed: 11186893]
- Christensen K, Johnson TE, Vaupel JW. The quest for genetic determinants of human longevity: challenges and insights. *Nat.Rev.Genet.* 2006; 7:436–448. [PubMed: 16708071]
- Christiansen L, Bathum L, Frederiksen H, Christensen K. Paraoxonase 1 polymorphisms and survival. *Eur.J.Hum.Genet.* 2004; 12:843–847. [PubMed: 15241482]
- Cluett C, Melzer D. Human genetic variations: Beacons on the pathways to successful ageing. *Mech.Ageing Dev.* 2009; 130:553–563. [PubMed: 19580824]
- Deelen J, Uh HW, Monajemi R, van HD, Thijssen PE, Bohringer S, van den Akker EB, de Craen AJ, Rivadeneira F, Uitterlinden AG, Westendorp RG, Goeman JJ, Slagboom PE, Houwing-Duistermaat JJ, Beekman M. Gene set analysis of GWAS data for human longevity highlights the relevance of the insulin/IGF-1 signaling and telomere maintenance pathways. *Age (Dordr.)*. 2011
- Di Bona D, Vasto S, Capurso C, Christiansen L, Deiana L, Franceschi C, Hurme M, Mocchegiani E, Rea M, Lio D, Candore G, Caruso C. Effect of interleukin-6 polymorphisms on human longevity: a systematic review and meta-analysis. *Ageing Res.Rev.* 2009; 8:36–42. [PubMed: 18930842]
- Flachsbart F, Caliebe A, Kleindorp R, Blanche H, von Eller-Eberstein H, Nikolaus S, Schreiber S, Nebel A. Association of FOXO3A variation with human longevity confirmed in German centenarians. *Proc.Natl.Acad.Sci.U.S.A.* 2009a; 106:2700–2705. [PubMed: 19196970]
- Flachsbart F, Caliebe A, Kleindorp R, Blanche H, von Eller-Eberstein H, Nikolaus S, Schreiber S, Nebel A. Association of FOXO3A variation with human longevity confirmed in German centenarians. *Proc.Natl.Acad.Sci.U.S.A.* 2009b; 106:2700–2705. [PubMed: 19196970]
- Flachsbart F, Franke A, Kleindorp R, Caliebe A, Blanche H, Schreiber S, Nebel A. Investigation of genetic susceptibility factors for human longevity - A targeted nonsynonymous SNP study. *Mutat.Res.* 2010; 694:9–13.

- Flurkey K, Papaconstantinou J, Miller RA, Harrison DE. Lifespan extension and delayed immune and collagen aging in mutant mice with defects in growth hormone production. *Proc.Natl.Acad.Sci.U.S.A.* 2001; 98:6736–6741. [PubMed: 11371619]
- Franceschi C, Motta L, Valensin S, Rapisarda R, Franzone A, Berardelli M, Motta M, Monti D, Bonafe M, Ferrucci L, Deiana L, Pes GM, Carru C, Desole MS, Barbi C, Sartoni G, Gemelli C, Lescai F, Olivieri F, Marchegiani F, Cardelli M, Cavallone L, Guerresi P, Cossarizza A, Troiano L, Pini G, Sansoni P, Passeri G, Lisa R, Spazzafumo L, Amadio L, Giunta S, Stecconi R, Morresi R, Viticchi C, Mattace R, De BG, Baggio G. Do men and women follow different trajectories to reach extreme longevity? Italian Multicenter Study on Centenarians (IMUSCE). *Aging (Milano).* 2000; 12:77–84. [PubMed: 10902049]
- Giannakou ME, Goss M, Junger MA, Hafen E, Leivers SJ, Partridge L. Long-lived *Drosophila* with overexpressed dFOXO in adult fat body. *Science.* 2004; 305:361. [PubMed: 15192154]
- Georgiev M, Lecoecur C, Meyre D, Benzinou M, Mein CA, Hinney A, Vatin V, Weill J, Heude B, Hebebrand J, Grossman AB, Korbonits M, Froguel P. Association studies on ghrelin and ghrelin receptor gene polymorphisms with obesity. *Obesity.(Silver.Spring).* 2009; 17:745–754. [PubMed: 19165163]
- Hampe J, Wollstein A, Lu T, Frevel HJ, Will M, Manaster C, Schreiber S. An integrated system for high throughput TaqMan based SNP genotyping. *Bioinformatics.* 2001; 17:654–655. [PubMed: 11448884]
- Harman D. Free radical theory of aging. *Triangle.* 1973; 12:153–158. [PubMed: 4769083]
- Herskind AM, McGue M, Holm NV, Sorensen TI, Harvald B, Vaupel JW. The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870-1900. *Hum.Genet.* 1996; 97:319–323. [PubMed: 8786073]
- Hjelmborg, J. Vb.; Iachine, I.; Skytthe, A.; Vaupel, JW.; McGue, M.; Koskenvuo, M.; Kaprio, J.; Pedersen, NL.; Christensen, K. Genetic influence on human lifespan and longevity. *Human Genetics.* 2006; 119:312–321. [PubMed: 16463022]
- Holzenberger M, Dupont J, Ducos B, Leneuve P, Geloan A, Even PC, Cervera P, Le BY. IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature.* 2003; 421:182–187. [PubMed: 12483226]
- Hubacek JA, Bohuslavova R, Skodova Z, Adamkova V. Variants within the ghrelin gene--association with HDL-cholesterol, but not with body mass index. *Folia Biol.(Praha).* 2007; 53:202–206. [PubMed: 18070417]
- Jacobsen R, Martinussen T, Christiansen L, Jeune B, ndersen-Ranberg K, Vaupel JW, Christensen K. Increased effect of the ApoE gene on survival at advanced age in healthy and long-lived Danes: two nationwide cohort studies. *Aging Cell.* 2010; 9:1004–1009. [PubMed: 20849521]
- Kawano T, Ito Y, Ishiguro M, Takuwa K, Nakajima T, Kimura Y. Molecular cloning and characterization of a new insulin/IGF-like peptide of the nematode *Caenorhabditis elegans*. *Biochem.Biophys.Res.Comm.* 2000; 273:431–436. [PubMed: 10873623]
- Kenyon CJ. The genetics of ageing. *Nature.* 2010; 464:504–512. [PubMed: 20336132]
- Kim DJ, Yi SM, Lee SY, Kang HS, Choi YH, Song YW, Park SC. Association between the MLH1 gene and longevity. *Hum.Genet.* 2006; 119:353–354. [PubMed: 16474933]
- Kuningas M, Mooijaart SP, van HD, Zwaan BJ, Slagboom PE, Westendorp RG. Genes encoding longevity: from model organisms to humans. *Aging Cell.* 2008; 7:270–280. [PubMed: 18208581]
- Kuro-o M, Matsumura Y, Aizawa H, Kawaguchi H, Suga T, Utsugi T, Ohyama Y, Kurabayashi M, Kaname T, Kume E, Iwasaki H, Iida A, Shiraki-Iida T, Nishikawa S, Nagai R, Nabeshima YI. Mutation of the mouse *klotho* gene leads to a syndrome resembling ageing. *Nature.* 1997; 390:45–51. [PubMed: 9363890]
- Kurosu H, Yamamoto M, Clark JD, Pastor JV, Nandi A, Gurnani P, McGuinness OP, Chikuda H, Yamaguchi M, Kawaguchi H, Shimomura I, Takayama Y, Herz J, Kahn CR, Rosenblatt KP, Kuro-o M. Suppression of aging in mice by the hormone *Klotho*. *Science.* 2005; 309:1829–1833. [PubMed: 16123266]
- Li Y, Wang WJ, Cao H, Lu J, Wu C, Hu FY, Guo J, Zhao L, Yang F, Zhang YX, Li W, Zheng GY, Cui H, Chen X, Zhu Z, He H, Dong B, Mo X, Zeng Y, Tian XL. Genetic association of FOXO1A

- and FOXO3A with longevity trait in Han Chinese populations. *Hum.Mol.Genet.* 2009; 18:4897–4904. [PubMed: 19793722]
- Lunetta KL, D'Agostino RB Sr, Karasik D, Benjamin EJ, Guo CY, Govindaraju R, Kiel DP, Kelly-Hayes M, Massaro JM, Pencina MJ, Seshadri S, Murabito JM. Genetic correlates of longevity and selected age-related phenotypes: a genome-wide association study in the Framingham Study. *BMC.Med.Genet.* 2007; 8(Suppl 1):S13. [PubMed: 17903295]
- Mockett RJ, Sohal RS, Orr WC. Overexpression of glutathione reductase extends survival in transgenic *Drosophila melanogaster* under hyperoxia but not normoxia. *FASEB J.* 1999; 13:1733–1742. [PubMed: 10506576]
- Montesanto A, Latorre V, Giordano M, Martino C, Domma F, Passarino G. The genetic component of human longevity: analysis of the survival advantage of parents and siblings of Italian nonagenarians. *Eur.J.Hum.Genet.* 2011 Ahead of print.
- Nebel A, Croucher PJ, Stiegeler R, Nikolaus S, Krawczak M, Schreiber S. No association between microsomal triglyceride transfer protein (MTP) haplotype and longevity in humans. *Proc.Natl.Acad.Sci.U.S.A.* 2005; 102:7906–7909. [PubMed: 15911777]
- Nebel A, Flachsbart F, Till A, Caliebe A, Blanche H, Arlt A, Hasler R, Jacobs G, Kleindorp R, Franke A, Shen B, Nikolaus S, Krawczak M, Rosenstiel P, Schreiber S. A functional EXO1 promoter variant is associated with prolonged life expectancy in centenarians. *Mech.Ageing Dev.* 2009; 130:691–699. [PubMed: 19698732]
- Nybo H, Gaist D, Jeune B, Bathum L, McGue M, Vaupel JW, Christensen K. The Danish 1905 cohort: a genetic-epidemiological nationwide survey. *J.Aging Health.* 2001; 13:32–46. [PubMed: 11503846]
- Pan J, Lin J, Izzo JG, Liu Y, Xing J, Huang M, Ajani JA, Wu X. Genetic susceptibility to esophageal cancer: the role of the nucleotide excision repair pathway. *Carcinogenesis.* 2009; 30:785–792. [PubMed: 19270000]
- Park PU, Defossez PA, Guarente L. Effects of mutations in DNA repair genes on formation of ribosomal DNA circles and life span in *Saccharomyces cerevisiae*. *Mol.Cell Biol.* 1999; 19:3848–3856. [PubMed: 10207108]
- Pawlikowska L, Hu D, Huntsman S, Sung A, Chu C, Chen J, Joyner AH, Schork NJ, Hsueh WC, Reiner AP, Psaty BM, Atzmon G, Barzilai N, Cummings SR, Browner WS, Kwok PY, Ziv E. Association of common genetic variation in the insulin/IGF1 signaling pathway with human longevity. *Aging Cell.* 2009a; 8:460–472. [PubMed: 19489743]
- Pawlikowska L, Hu D, Huntsman S, Sung A, Chu C, Chen J, Joyner AH, Schork NJ, Hsueh WC, Reiner AP, Psaty BM, Atzmon G, Barzilai N, Cummings SR, Browner WS, Kwok PY, Ziv E. Association of common genetic variation in the insulin/IGF1 signaling pathway with human longevity. *Aging Cell.* 2009b; 8:460–472. [PubMed: 19489743]
- Pedersen CB, Gotzsche H, Moller JO, Mortensen PB. The Danish Civil Registration System. A cohort of eight million persons. *Dan.Med.Bull.* 2006; 53:441–449. [PubMed: 17150149]
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC. PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics.* 2007; 81
- Scarpaci S, Frasca D, Barattini P, Guidi L, Doria G. DNA damage recognition and repair capacities in human naive and memory T cells from peripheral blood of young and elderly subjects. *Mech.Ageing Dev.* 2003; 124:517–524. [PubMed: 12714261]
- Schildkraut JM, Iversen ES, Wilson MA, Clyde MA, Moorman PG, Palmieri RT, Whitaker R, Bentley RC, Marks JR, Berchuck A. Association between DNA damage response and repair genes and risk of invasive serous ovarian cancer. *PLoS.One.* 2010; 5:e10061. [PubMed: 20386703]
- Shih DM, Gu L, Xia YR, Navab M, Li WF, Hama S, Castellani LW, Furlong CE, Costa LG, Fogelman AM, Lusis AJ. Mice lacking serum paraoxonase are susceptible to organophosphate toxicity and atherosclerosis. *Nature.* 1998; 394:284–287. [PubMed: 9685159]
- Skytte A, Kyvik K, Holm NV, Vaupel JW, Christensen K. The Danish Twin Registry: 127 birth cohorts of twins. *Twin.Res.* 2002; 5:352–357. [PubMed: 12537858]

- Soerensen M, Dato S, Christensen K, McGue M, Stevnsner T, Bohr VA, Christiansen L. Replication of an association of variation in the FOXO3A gene with human longevity using both case-control and longitudinal data. *Aging Cell*. 2010; 9:1010–1017. [PubMed: 20849522]
- Tyner SD, Venkatachalam S, Choi J, Jones S, Ghebranious N, Igelmann H, Lu X, Soron G, Cooper B, Brayton C, Hee PS, Thompson T, Karsenty G, Bradley A, Donehower LA. p53 mutant mice that display early ageing-associated phenotypes. *Nature*. 2002; 415:45–53. [PubMed: 11780111]
- van der Horst GT, van SH, Berg RJ, van Gool AJ, de WJ, Weeda G, Morreau H, Beems RB, van Kreijl CF, de Gruijl FR, Bootsma D, Hoeijmakers JH. Defective transcription-coupled repair in Cockayne syndrome B mice is associated with skin cancer predisposition. *Cell*. 1997; 89:425–435. [PubMed: 9150142]
- van Heemst D, Beekman M, Mooijaart SP, Heijmans BT, Brandt BW, Zwaan BJ, Slagboom PE, Westendorp RG. Reduced insulin/IGF-1 signalling and human longevity. *Aging Cell*. 2005a; 4:79–85. [PubMed: 15771611]
- van Heemst D, Mooijaart SP, Beekman M, Schreuder J, de Craen AJ, Brandt BW, Slagboom PE, Westendorp RG. Variation in the human TP53 gene affects old age survival and cancer mortality. *Exp.Gerontol*. 2005b; 40:11–15. [PubMed: 15732191]
- Van Ness B, Ramos C, Haznadar M, Hoering A, Haessler J, Crowley J, Jacobus S, Oken M, Rajkumar V, Greipp P, Barlogie B, Durie B, Katz M, Atluri G, Fang G, Gupta R, Steinbach M, Kumar V, Mushlin R, Johnson D, Morgan G. Genomic variation in myeloma: design, content, and initial application of the Bank On A Cure SNP Panel to detect associations with progression-free survival. *BMC.Med*. 2008; 6:26. [PubMed: 18778477]
- Willcox BJ, Donlon TA, He Q, Chen R, Grove JS, Yano K, Masaki KH, Wilcox DC, Rodriguez B, Curb JD. FOXO3A genotype is strongly associated with human longevity. *Proceedings of the National Academy of Sciences*. 2008a; 105:13987–13992.
- Willcox BJ, Donlon TA, He Q, Chen R, Grove JS, Yano K, Masaki KH, Wilcox DC, Rodriguez B, Curb JD. FOXO3A genotype is strongly associated with human longevity. *Proceedings of the National Academy of Sciences*. 2008b; 105:13987–13992.
- Yang XR, Pfeiffer RM, Wheeler W, Yeager M, Chanock S, Tucker MA, Goldstein AM. Identification of modifier genes for cutaneous malignant melanoma in melanoma-prone families with and without CDKN2A mutations. *Int.J.Cancer*. 2009; 125:2912–2917. [PubMed: 19626699]
- Yashin AI, De BG, Vaupel JW, Tan Q, Andreev KF, Iachine IA, BonaFe M, Valensin S, De LM, Carotenuto L, Franceschi C. Genes and longevity: lessons from studies of centenarians. *J.Gerontol.A Biol.Sci.Med.Sci*. 2000; 55:B319–B328. [PubMed: 10898245]
- Yashin AI, Wu D, Arbeev KG, Ukraintseva SV. Joint influence of small-effect genetic variants on human longevity. *Aging (Albany.NY)*. 2010; 2:612–620. [PubMed: 20834067]

**Highlights**

- 1) Study of 1273 tagging SNPs in 148 genes in three major longevity candidate pathways
- 2) We apply both a classical case-control and a longitudinal study approach
- 3) We test important findings in independent study populations
- 4) Effects of rs3842755, rs13251813, rs3211994 and rs207444 in both Danish studies
- 5) Non-significant tendencies of rs11571461 and rs10047589 in replication samples

**Table 1**

Characteristics of (A) the Danish discovery cohorts and (B) the German and Dutch replication samples.

<b>(A) The Danish discovery cohorts</b>	<b>MADT</b>	<b>1905</b>
Number of individuals	736	1089
Average age in years at intake (range)	50.6 (46.0–55.0)	93.2 (92.2–93.8)
Males : females ratio (%) at intake	50.4 : 49.6	28.7 : 71.3
Mean follow up time (SD)	-	3.53 (2.5)
Follow up time (years)	-	11.4

	<b>German case-control study</b>		<b>Dutch longitudinal study</b>
<b>(B) The replication samples</b>	<b>Controls</b>	<b>LLI</b>	<b>Leiden 85-plus Cohort</b>
Number of individuals	1104	1613	563
Average age at intake (SD)	67.2 (4.1)	98.4 (2.7)	85.0 (0)
Males : females ratio (%) at intake	25.6 : 74.3	26.8 : 73.2	33.4 : 66.6
Mean follow up time (SD)	-	-	5.49 (3.0)
Follow up time (years)	-	-	10.4

Notes: MADT: the Study of Middle Aged Danish Twins, 1905: the Danish 1905 birth cohort, LLI: long lived individuals, SD: standard deviation.

Table 2

SNPs significantly associated with longevity after correction for multiple testing in the case-control comparison of middle-aged and oldest-old Danish cohorts. (A) Both genders combined and (B) additional SNP in sex-stratified analyses.

(A) Both genders	Gene	SNP	Minor/Major allele	Chr./position	Position in gene	MAF MADT/1905	PCCA uncorrected	PCCA corrected	OR (95% CI)
Pro/antioxidants	GSR	rs1002149	A/C	8/30705280	5'	0.183/0.137	1.47.10 <sup>-4</sup>	0.0335	0.706 (0.589–0.845)
GH/IGF/INS	INS	rs3842755	A/C	11/2137372	3'UTR	0.292/0.424	8.38.10 <sup>-16</sup>	1.10 <sup>-4</sup>	1.789 (1.552–2.063)
	KL	rs1207362	A/C	13/32510839	Intronic	0.339/0.245	2.77.10 <sup>-9</sup>	1.10 <sup>-4</sup>	0.634 (0.545–0.737)
	GHRHR	rs2267723	G/A	7/30973467	Intronic	0.414/0.330	2.28.10 <sup>-7</sup>	1.10 <sup>-4</sup>	0.696 (0.607–0.799)
	GHSR	rs572169	A/G	3/173648421	Coding (syn)	0.342/0.274	1.65.10 <sup>-5</sup>	0.0045	0.725 (0.626–0.839)
	IGF2R	rs9456497	G/A	6/160363418	Intronic	0.191/0.252	1.86.10 <sup>-5</sup>	0.0050	1.424 (1.211–1.674)
DNA repair	RAD52	rs11571461	G/A	12/896580	Intronic	0.098/0.195	5.05.10 <sup>-15</sup>	1.10 <sup>-4</sup>	2.226 (1.815–2.730)
	WRN	rs13251813	A/G	8/31083753	Intronic	0.077/0.134	1.72.10 <sup>-7</sup>	2.10 <sup>-4</sup>	1.840 (1.460–2.318)
	POLB	rs2953983	G/A	8/42332413	Intronic	0.080/0.123	4.54.10 <sup>-5</sup>	0.0225	1.599 (1.274–2.006)
	RAD23B	rs1805329	A/G	9/109124149	Coding (non-syn)	0.245/0.190	6.74.10 <sup>-5</sup>	0.0309	0.722 (0.615–0.847)

(B) Males	Gene	SNP	Minor/Major allele	Chr./position	Position in gene	MAF MADT/1905	PCCA uncorrected	PCCA corrected	OR (95% CI)
DNA repair	NTLH1	rs3211994	A/G	16/2029029	3'	0.112/0.201	1.38.10 <sup>-5</sup>	0.0056	1.994 (1.455–2.732)

Notes: Chr: chromosome, MAF: minor allele frequency, MADT: the Study of Middle Aged Danish Twins, 1905: the Danish 1905 Birth cohort, PCCA: allelic case-control comparison P-value, OR: odds ratio, 95% CI: 95% confidence interval, UTR: untranslated region, Syn: synonymous, Non-syn: Non-synonymous.

Table 3

Replication of Danish case-control findings in the German case-control samples.

	Gene	SNP	OR (Danish study)	OR (German study)	PCCA Germans	95% CI
Pro/antioxidants	GSR	rs1002149	0.706	1.054	0.470	0.913–1.217
	INS	rs3842755	1.789	0.949	0.434	0.832–1.082
GH/IGF/INS	KL	rs1207362	0.634	0.985	0.804	0.873–1.110
	GHRHR	rs2267723	0.696	1.005	0.934	0.900–1.121
	GHSR	rs572169	0.725	1.106	0.097	0.981–1.247
	IGF2R	rs9456497	1.424	0.962	0.586	0.837–1.105
	RAD52	rs11571461	2.226	1.162	0.191	0.927–1.457
DNA repair	WRN	rs13251813	1.84	0.999	0.998	0.754–1.325
	POLB	rs2953983	1.599	1.046	0.683	0.842–1.299
	RAD23B	rs1805329	0.722	0.955	0.507	0.833–1.094
	NTHL1	rs3211994	1.994	0.931	0.627	0.699–1.241

Notes: OR: odds ratio, PCCA = Allelic case-control comparison P-value, 95% CI: 95% confidence interval.

**Table 4**

SNPs significantly associated with longevity after correction for multiple testing in the longitudinal analysis of follow-up survival

(A) Both genders combined							
Early/late death							
Pathway	Gene	SNP	Minor/major allele	Chr./position	Position in gene	Freq 1905	95% CI
Pro/antioxidants	TNFRD1	rs10047589	A/G	12:103268518	3'	0.098/0.445/0.455	1.143-2.318
						0.00018	1.458
						0.043	0.043

Number-of-days-lived							
Pathway	Gene	SNP	Minor/major allele	Chr./position	Position in gene	Freq 1905	95% CI
Pro/antioxidants	XDH	rs207444	A/G	2:31417301	Intronic	0.001/0.177/0.820	1.946-6.141
						0.00016	4.043
						0.046	4.043

(B) Stratified by sex											
Early/late death											
	Pathway	Gene	SNP	Minor/major allele	Chr./position	Position in gene	Freq 1905	P-value	P-value corrected	OR	95% CI
Males	GH/IGF- I/INS	GHRH	rs26802	C/A	3/10307365	Intronic	0.076/0.392/0.531	1.89*10 <sup>-5</sup>	0.032	2.338	1.585-3.451
Females	DNA repair	H2AFX	rs2509049	A/G	11/118471731	5'	0.156/0.527/0.315	4.77*10 <sup>-5</sup>	0.017	0.625	0.498-0.782

Number-of-days-lived											
	Pathway	Gene	SNP	Minor/major allele	Chr./position	Position in gene	Freq 1905	P-value	P-value corrected	β-coefficient	95% CI
Males	DNA repair	XRCC5	rs705649	G/A	2/216689431	Intronic	0.175/0.512/0.311	6.19*10 <sup>-5</sup>	0.029	4.669	2.416-6.921

data from the cohort of oldest-old Danes. (A) Both genders combined and (B) additional SNPs in sex-stratified analyses.  
 Notes: Chr: chromosome, Freq: frequency, 1905: the 1905 cohort, 95%CI: 95% confidence interval. The frequencies in the Freq. 1905 column refer to the frequencies of the rare homozygotes/ the heterozygotes/ the common homozygotes for each SNP.

**Table 5**  
 Replication of findings from the Danish longitudinal study in the longitudinal Dutch Leiden 85-plus Study

Both genders combined			Danish longitudinal study			Dutch longitudinal study		
Pathway	Gene	SNP	Estimate	HR	P-value	Estimate	HR	95%CI
Pro/antioxidants	TNXRD1	rs10047589	1.458 (OR)		0.969	1.458 (OR)	0.969	0.844-1.112
Pro/antioxidants	XDH	rs207444	4.043 (β)		0.886	4.043 (β)	0.404	0.667-1.177

  

Males								
Pathway	Gene	SNP	Estimate	HR	P-value	Estimate	HR	95%CI
DNA repair	GH/IGF-1/INS	rs26802	2.338 (OR)	0.730	0.962	2.338 (OR)	0.774-1.197	
	MLH1	rs13320360	3.132 (OR)	1.169	0.827	3.132 (OR)	0.288-4.727	
	XRCC5	rs705649	4.669 (β)	1.037	0.749	4.669 (β)	0.828-1.300	

  

Females								
Pathway	Gene	SNP	Estimate	HR	P-value	Estimate	HR	95%CI
DNA repair	H2AFX	rs2509049	0.625 (OR)	0.909	0.260	0.625 (OR)	0.769-1.073	

Notes: β: Beta coefficient (linear regression using the Number\_of\_days\_lived variable), OR: Odds ratio (logistic regression using the Early/late death variable), HR hazard ratio, 95%CI: 95% confidence interval.

**Table 6**

Similar directions of effect in the Danish case-control and longitudinal studies.

(A) Similar directions of effect in the Danish longitudinal study of SNPs initially identified in the Danish case-control study and (B) Similar direction of effect in the Danish case-control study of a SNP initially identified in the Danish longitudinal study.

(A) Case-control study (initial findings)				Longitudinal study (similar directions of effects)			
Pathway	Gene	SNP	OR	Number_of_days_lived variable			P-value
				$\beta$ -coef.	95% CI		
GH/IGF/Insulin	INS	rs3842755	1.789	2.73 $\sigma$	0.66–4.80 $\sigma$	0.01 $\sigma$	
	IGF2R	rs9456497	1.424	2.77 $\sigma$	–0.01–5.56 $\sigma$	0.05 $\sigma$	
DNA repair	WRN	rs13251813	1.840	2.23	0.31–4.15	0.02	

  

Pathway	Gene	SNP	OR	Early_late_death_variable	
				OR	P-value
DNA repair	RAD52	rs11571461	2.226	1.30 $\phi$	0.96–1.78 $\phi$
	NTHL1	rs3211994	1.994 $\sigma$	1.42	1.09–1.86

  

(B) Longitudinal study (initial Findings)				Case-control study (similar direction of effect)			
Pathway	Gene	SNP	$\beta$ -coef.	Number_of_days_lived variable			PCCA
				OR	95% CI		
Pro/antioxidants	XDH	rs207444	4.043	1.42	1.10–1.83	0.005	

Notes: OR: Odds ratio,  $\beta$ -coef: beta coefficient, 96% CI: 95% confidence interval

PCCA: allelic case-control comparison P-value.

$\sigma$  males separately

$\phi$  females separately