Evidence from Case-control and Longitudinal Studies Supports Associations of Genetic Variation in *APOE, CETP* and *IL6* with Human Longevity

AGE

Mette Soerensen*, Serena Dato, Qihua Tan, Mikael Thinggaard, Rabea Kleindorp, Marian Beekman, H. Eka D. Suchiman, Rune Jacobsen, Matt McGue, Tinna Stevnsner, Vilhelm A. Bohr, Anton J.M. de Craen, Rudi G.J. Westendorp, Stefan Schreiber, Eline Slagboom, Almut Nebel, James W. Vaupel, Kaare Christensen and Lene Christiansen

*Author for correspondence: Mette Soerensen, msoerensen@health.sdu.dk, The Danish Aging Research Center, Epidemiology, Institute of Public Health, University of Southern Denmark, J.B. Winsloews Vej 9B, 5000 Odense C, Denmark and Department of Clinical Genetics and Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Sdr. Boulevard 29, 5000 Odense C, Denmark

Supplementary Material

Selection of candidate genes, candidate SNPs and tagging SNPs

The 16 candidate genes under study were chosen as part of a large candidate gene study performed based on comprehensive literature/data base searches for candidate genes, candidate SNPs and tagging SNPs.

For identification of candidate genes the following data bases were applied: www.ncbi.nlm.nih.gov (/sites/entrez, /OMIM, /sites/entrez?Db=gap), www.genecards.org, www.hgvbaseg2p.org, genomics.senescence.info and geneticassociationdb.nih.gov, employing search terms such as 'human longevity', 'human aging', 'age related disease' and 'premature ageing syndrome'. The chromosomal regions were ascertained through the www.ncbi.nlm.nih.gov (/gene and /mapview) and genome.ucsc.edu/ databases, using the same gene ID, Sequence Accession IDs, NCBI Reference Sequence, assembly build 36 and SNP build 129. In case of more than one isoform, the longest one was selected. 5000 bp upstream and 1000 bp downstream was included for each gene, and in case of another gene within these regions, only the region until the neighboring gene was included.

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 latter the www.ncbi.nlm.nih.gov/SNP, koreanbio.org/Variome, snpper.chip.org, snps3d.org and manticore.niehs.nih.gov data bases were used. The identities of the SNPs were confirmed by linking back to the NCBI data base.

Moreover, to cover the majority of the common genetic variation within each gene region, we added a number of tagging SNPs in each gene region 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. This genotype data was analyzed by the HaploView software (http://www.broadinstitute.org/haploview/haploview, (Barrett *et al.* (2005)) using 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 the best possible genotyping, SNPs known to perform poorly on the Illumina GoldenGate genotyping platform were excluded. In total 118 SNPs were chosen and 102 were successfully genotyped.

Genotyping in the replication cohorts

Genotyping of the German replication samples were conducted by Sequenom MassARRAY iPLEX®Gold technology (Sequenom®, Inc., CA, USA), using iPlex primers designed via the MassARRAY Design Software. 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).