

LongevityMap: a database of human genetic variants associated with longevity

Arie Budovsky^{1,2*}, Thomas Craig^{3*}, Jingwei Wang^{3*}, Robi Tacutu³, Attila Csordas⁴, Joana Lourenço³, Vadim E. Fraifeld¹, and João Pedro de Magalhães^{3*}

¹The Shraga Segal Department of Microbiology, Immunology and Genetics, Center for Multidisciplinary Research on Aging, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

²Judea Regional Research and Development Center, Carmel 90404, Israel

³Integrative Genomics of Ageing Group, Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK

⁴European Bioinformatics Institute (EMBL-EBI), European Molecular Biology Laboratory, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK

Understanding the genetic basis of human longevity remains a challenge but could lead to life-extending interventions and better treatments for age-related diseases. Toward this end we developed the LongevityMap (<http://genomics.senescence.info/longevity/>), the first database of genes, loci, and variants studied in the context of human longevity and healthy ageing. We describe here its content and interface, and discuss how it can help to unravel the genetics of human longevity.

Given the worldwide ageing of the population, studying the genetics of human longevity is of widespread importance [1,2]. Longevity is moderately heritable in humans (~25%), with increasing heritability with age [1], and exceptional longevity and healthy ageing in humans is an inherited phenotype [3]. Hundreds of longevity association studies have been performed in recent years and some genes associated with human longevity may be suitable targets for drug development [4]. Nonetheless, the heritability of human longevity remains largely unexplained in part due to the complexity of this phenotypic trait [1]. Thanks to advances in next-generation sequencing and genome-wide approaches, the capacity of longevity association studies is increasing. The growing amounts of data being generated also increase the complexity of the data analysis and the difficulty of placing findings in context of previous studies. We created the LongevityMap (<http://genomics.senescence.info/longevity/>), the first catalogue of human genetic variants associated with longevity, to serve as a reference to help researchers navigate the rising tide of data related to human longevity.

The LongevityMap is a new addition to our already highly successful collection of online databases and tools on the biology and genetics of ageing, the Human Ageing Genomic Resources (<http://genomics.senescence.info/>) [5]. GenAge, our existing database of ageing-related genes,

focuses mostly on genes modulating longevity in model organisms plus the few genes associated with human progeroid syndromes [5], and thus there is an unmet need for a database of human genetic variants associated with longevity. As such, we followed the high standards and rigorous procedures of GenAge to develop the LongevityMap. Briefly, all entries in the LongevityMap were manually curated from the literature. Studies were selected following an in-depth literature survey. The LongevityMap is an inclusive database in which both large and small studies are included; different types of study are featured, from cross-sectional studies to studies of extreme longevity (e.g., centenarians). However, studies focused on cohorts of unhealthy individuals at baseline, such as cancer patients, were excluded. Details on study design are provided for each entry, including a brief description of the type of study, population ethnicity, sample size, age of probands and controls, and any gender bias. Negative results are also integrated in the LongevityMap to provide visitors with as much information as possible regarding each gene, variant, and locus previously studied in the context of longevity. Each entry refers to a specific observation from a study. This means that studies, and large-scale studies in particular, can have multiple entries in the LongevityMap, reflecting different results and observations. Each entry also includes a brief description of the major conclusions. Entries are flagged regarding whether results were statistically significant or not, though many studies have marginal or indicative results that require a brief explanation of the findings. Our policy concerning controversial and subjective results is to detail the facts concerning the controversy and let users form their own opinions. A link to the primary publication in PubMed is always included in each entry.

We developed an intuitive, user-friendly interface for the LongevityMap that allows users to query genes, variants (including by reference SNP ID number), studies, and cytogenetic locations (Figure 1A). Users can browse/filter the data by association (i.e., significant or non-significant), population, and chromosome. For each single nucleotide polymorphism (SNP) and gene, additional annotation was retrieved from the US National Center for Biotechnology Information (NCBI) databases dbSNP and RefSeq [6] to provide further information on

Corresponding author: de Magalhães, J.P. (jp@senescence.info).

Keywords: ageing; genetics; GWAS; humans; lifespan; polymorphisms.

* These authors contributed equally to this work.



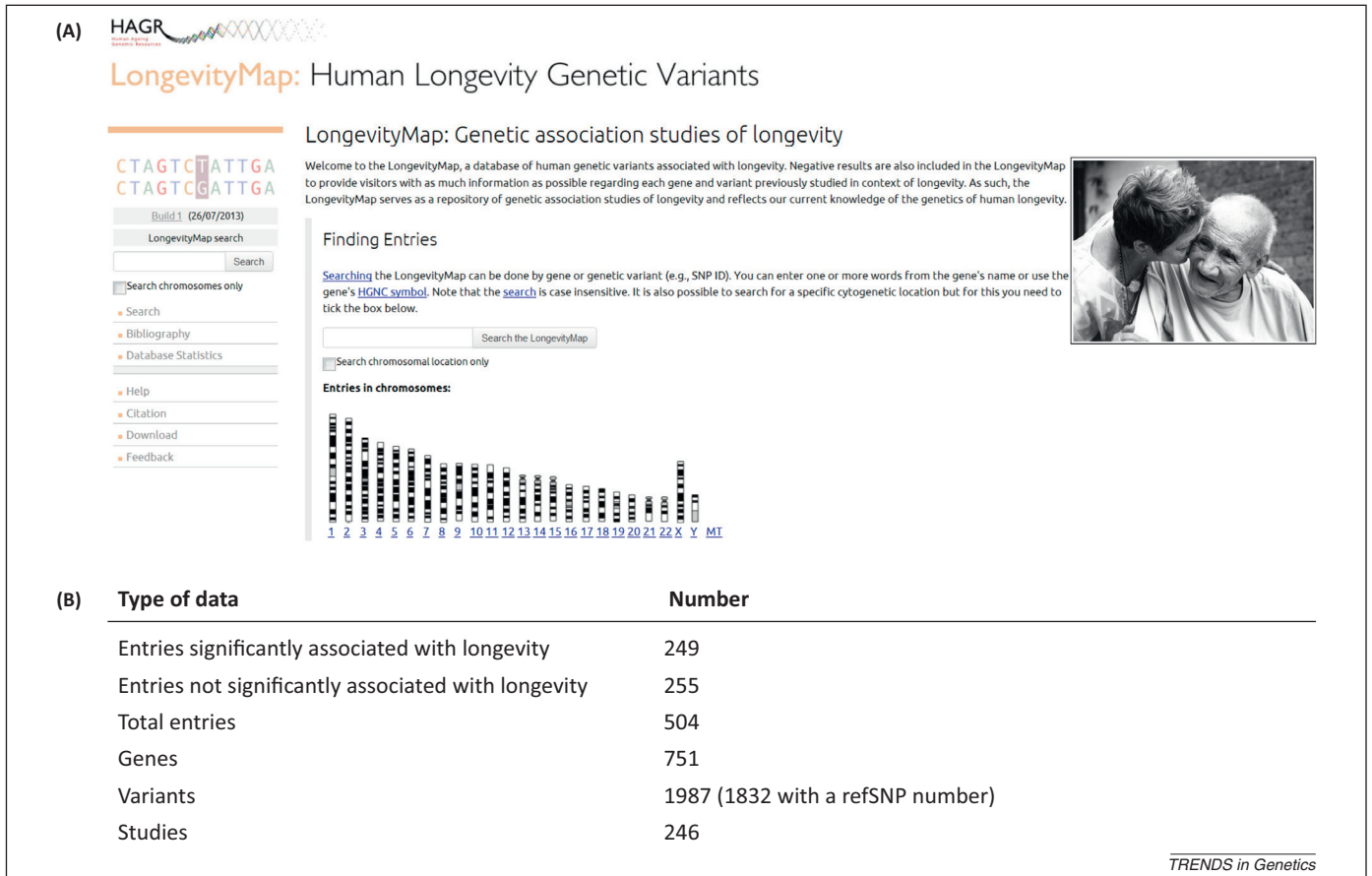


Figure 1. LongevityMap home page which showcases the design and layout of the website as well as its multiple search options and links (A); old couple picture by Jonel Hanopol. Types and amount of data in the LongevityMap (B).

genes associated with SNPs and gene function, respectively. Homologues in model organisms were obtained from the InParanoid database [7]. Links are widely implemented to allow users to identify quickly other entries related to a given study, gene, or variant. In fact, each gene in the LongevityMap has a gene-centric page that aggregates and condenses the information on the database taken from different studies. In addition, the LongevityMap is fully integrated with our other ageing-related databases to provide users with selected, relevant information. In particular, crosslinks to GenAge are included to indicate genes associated with progeroid syndromes and those with homologues in model organisms known to modulate ageing/longevity. If appropriate, links to other major databases, such as Ensembl, Swiss-Prot, dbSNP, HapMap, and NCBI Entrez, are included for each entry. At time of writing, the LongevityMap includes data from 246 studies, featuring 751 different genes and 1987 variants (Figure 1B). Similarly to our other ageing-related databases, the LongevityMap is freely available online under a Creative Commons Attribution license. The full dataset is available for download and third-party use. It is our hope that the LongevityMap will serve as a novel database to help researchers decipher the genetics of human longevity.

Acknowledgements

The authors wish to thank Joana Costa, Daniel Wuttke, and Alex Freitas for helping to collate data and for comments and suggestions. This work was funded by a Wellcome Trust grant (ME050495MES) to J.P.M. This work was also funded in part by the European Union Framework Program (FP) 7 Health Research Grant number HEALTH-F4-2008-202047 (to V.E.F.) and the Israel Ministry of Science and Technology (to A.B.). J.P.M. is also grateful for support from the Ellison Medical Foundation and R.T. is supported by a Marie Curie Intra-European Fellowship within FP7.

References

- Christensen, K. *et al.* (2006) The quest for genetic determinants of human longevity: challenges and insights. *Nat. Rev. Genet.* 7, 436–448
- Chung, W.H. *et al.* (2010) The role of genetic variants in human longevity. *Ageing Res. Rev.* 9 (Suppl. 1), S67–S78
- Atzmon, G. *et al.* (2005) Biological evidence for inheritance of exceptional longevity. *Mech. Ageing Dev.* 126, 341–345
- de Magalhaes, J.P. *et al.* (2012) Genome–environment interactions that modulate aging: powerful targets for drug discovery. *Pharmacol. Rev.* 64, 88–101
- Tacutu, R. *et al.* (2013) Human ageing genomic resources: integrated databases and tools for the biology and genetics of ageing. *Nucleic Acids Res.* 41, D1027–D1033
- NCBI Resource Coordinators (2013) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 41, D8–D20
- Ostlund, G. *et al.* (2010) InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Res.* 38, D196–D203