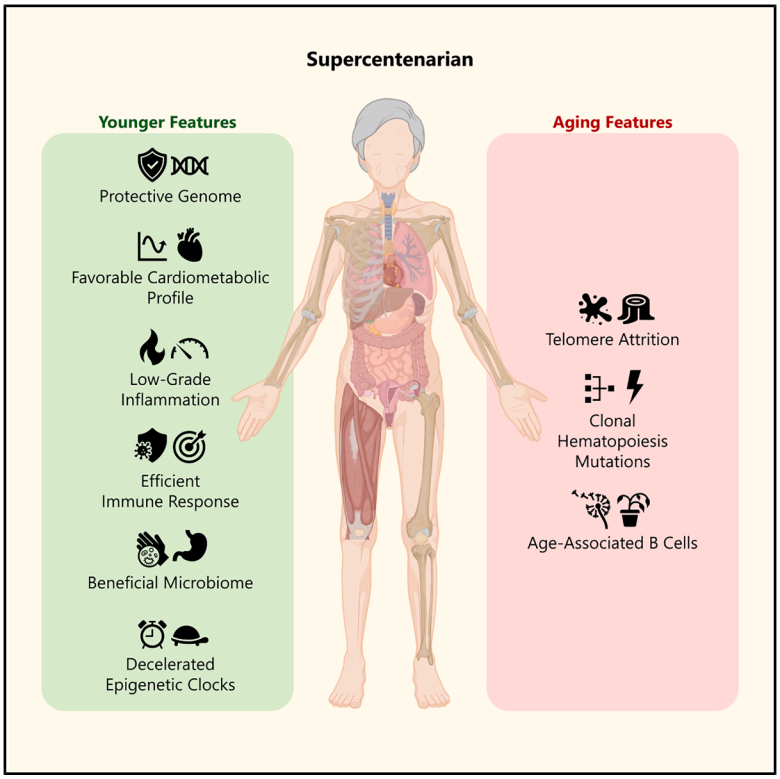
Cell Reports
Medicine

Article

The multiomics blueprint of the individual with the most extreme lifespan

Graphical abstract



Highlights

- (Epi)genome, transcriptome, metabolome, proteome, and microbiome study of the oldest human
- Despite molecular hallmarks of aging, absence of major age-associated diseases
- Resilient genetic variants and low-inflammation metabolic profile reduce aging risks
- Bacteria occurrence and epigenome profile resembling younger individuals

Authors

Eloy Santos-Pujol,
Aleix Noguera-Castells,
Marta Casado-Pelaez, ..., Hafid Laayouni,
Arcadi Navarro, Manel Esteller

Correspondence

mesteller@carrerasresearch.org

In brief

In brief, Santos-Pujol and colleagues characterize the multiomics landscape of the human with the most extreme longevity. The study enables the association of advantageous genetic variants, an engaged lipid metabolism, low inflammation levels, a proficient immune system, a rejuvenated microbiome composition, and a younger epigenetic age with the extraordinary lifespan.

Santos-Pujol et al., 2025, Cell Reports Medicine 6, 102368
October 21, 2025
https://doi.org/10.1016/j.xcrm.2025.102368

CellPress

Cell Reports Medicine

CellPress
OPEN ACCESS

Article

The multiomics blueprint of the individual with the most extreme lifespan

Eloy Santos-Pujol,[1,46] Aleix Noguera-Castells,[1,2,3,46] Marta Casado-Pelaez,[1] Carlos A. García-Prieto,[1,4] Claudia Vasallo,[5,6] Ignacio Campillo-Marcos,[1,3] Carlos Quero-Dotor,[1,3] Eva Crespo-García,[1] Alberto Bueno-Costa,[1] Fernando Setién,[1] Gerardo Ferrer,[1,3] Veronica Davalos,[1,3] Elisabetta Mereu,[7] Raquel Pluvinet,[8,9] Carles Arribas,[8] Carolina de la Torre,[10] Francisco Villavicencio,[11] Lauro Sumoy,[9] Isabel Granada,[12] Natalie S. Coles,[13] Pamela Acha,[14] Francesc Solé,[14] Mar Mallo,[14,15] Caterina Mata,[16] Sara Peregrina,[4,17] Toni Gabaldón,[4,17,18,19] Marc Lirós,[20] Meritxell Pujolassos,[20] Robert Carreras-Torres,[21] Aleix Lluansí,[21,22] Librado Jesús García-Gil,[21,23] Xavier Aldeguer,[21,24] Sara Samino,[25] Pol Torné,[25] Josep Ribalta,[26,27,28] Montse Guardiola,[26,27,28] Núria Amigó,[25,26,28]

(Author list continued on next page)

[1]Cancer Epigenetics Group, Josep Carreras Leukaemia Research Institute (IJC), Badalona, Barcelona, Catalonia, Spain
[2]Department of Biosciences, Faculty of Science, Technology and Engineering, University of Vic - Central University of Catalonia (UVic-UCC), Vic, Barcelona, Catalonia, Spain
[3]Centro de Investigacion Biomedica en Red Cancer (CIBERONC), Madrid, Spain
[4]Life Sciences Department, Barcelona Supercomputing Center (BSC), Barcelona, Catalonia, Spain
[5]BarcelonaBeta Brain Research Center, Pasqual Maragall Foundation, Barcelona, Catalonia, Spain
[6]Institute of Evolutionary Biology (IBE, UPF-CSIC), Department of Medicine and Life, Barcelona, Catalonia, Spain
[7]Cellular Systems Genomics Group, Josep Carreras Leukaemia Research Institute (IJC), Badalona, Barcelona, Catalonia, Spain
[8]Genomics Unit, Josep Carreras Leukaemia Research Institute (IJC), Badalona, Barcelona, Catalonia, Spain
[9]High Content Genomics and Bioinformatics (HCGB)-Germans Trias I Pujol Research Institute (IGTP), Badalona, Barcelona, Catalonia, Spain
[10]Proteomics Unit, Josep Carreras Leukaemia Research Institute (IJC), Badalona, Barcelona, Catalonia, Spain
[11]Department of Economic, Financial and Actuarial Mathematics, University of Barcelona, Barcelona, Catalonia, Spain
[12]Hematology Department, Hospital Germans Trias i Pujol, Institut Català d'Oncologia, Josep Carreras Leukaemia Research Institute, Barcelona, Spain
[13]Los Angeles Gerontology Research Group, Los Angeles, CA, USA
[14]Myelodysplastic Syndromes Research Group, Josep Carreras Leukaemia Research Institute (IJC), Badalona, Barcelona, Catalonia, Spain
[15]Microarrays Unit, Josep Carreras Leukaemia Research Institute (IJC), Badalona, Barcelona, Catalonia, Spain
[16]Single Cell Unit, Josep Carreras Leukaemia Research Institute (IJC), Badalona, Barcelona, Catalonia, Spain
[17]Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Barcelona, Catalonia, Spain
[18]Institució Catalana de Recerca i Estudis Avançats (ICREA) and Universitat Pompeu Fabra, Barcelona, Catalonia, Spain
[19]CIBER de Enfermedades Infecciosas (CIBERINFEC), Instituto de Salud Carlos III, Madrid, Spain
[20]Bioinformatics and Bioimaging (BI-SQUARED) Research Group, Biosciences Department, Faculty of Science, Technology and Engineering, Universitat de Vic – Universitat Central de Catalunya, Vic, Catalunya, Spain
[21]Digestive Diseases and Microbiota Group, Institut d'Investigació Biomèdica de Girona Dr. Josep Trueta (IDIBGI), Girona, Catalonia, Spain
[22]Infectious diseases and Microbiome Unit, Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Catalunya, Spain
[23]Biology Department, Universitat de Girona, Girona, Catalunya, Spain
[24]Digestive Service, Hospital Universitari de Girona Dr Josep Trueta, Girona, Catalunya, Spain

(Affiliations continued on next page)

SUMMARY

Extreme human lifespan, exemplified by supercentenarians, presents a paradox in understanding aging: despite advanced age, they maintain relatively good health. To investigate this duality, we have performed a high-throughput multiomics study of the world's oldest living person, interrogating her genome, transcriptome, metabolome, proteome, microbiome, and epigenome, comparing the results with larger matched cohorts. The emerging picture highlights different pathways attributed to each process: the record-breaking advanced age is manifested by telomere attrition, abnormal B cell population, and clonal hematopoiesis, whereas absence of typical age-associated diseases is associated with rare European-population genetic variants, low inflammation levels, a rejuvenated bacteriome, and a younger epigenome. These findings provide a fresh look at human aging biology, suggesting biomarkers for healthy aging, and potential strategies to increase life expectancy. The extrapolation of our results to the general population will require larger cohorts and longitudinal prospective studies to design potential anti-aging interventions.

Oscar Yanes,[27,28,29] Paula Martínez,[30] Raúl Sánchez-Vázquez,[30] Maria A. Blasco,[30] Jose Oviedo,[31,32] Bernardo Lemos,[31,32,33] Julia Rius-Bonet,[5,34] Marta Torrubiano,[34] Marta Massip-Salcedo,[5,34] Kamal A. Khidir,[35,36] Thong Huy Cao,[37,38,39] Paulene A. Quinn,[37,39] Donald J.L. Jones,[39,40] Salvador Macip,[5,34,35] Eva Brigos-Barril,[6,41] Mauricio Moldes,[42] Fabio Barteri,[5] Gerard Muntané,[41,43,44] Hafid Laayouni,[5,41] Arcadi Navarro,[5,6,18,41,42] and Manel Esteller[1,3,18,45,47,*]

[25]Biosfer Teslab, 43201 Reus, Tarragona, Spain
[26]Universitat Rovira i Virgili, Department of Basic Medical Sciences, Reus, Spain
[27]Institut d'Investigació Sanitària Pere Virgili, Reus, Spain
[28]Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), Instituto de Salud Carlos III (ISCIII), Madrid, Spain
[29]Universitat Rovira i Virgili, Department of Electronic Engineering, Tarragona, Spain
[30]Telomeres and Telomerase Group - Fundación Humanismo y Ciencia, Molecular Oncology Program, Spanish National Cancer Centre (CNIO), Madrid, Spain
[31]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[32]Department of Pharmacology and Toxicology, R. Ken Coit College of Pharmacy, The University of Arizona, Tucson, AZ, USA
[33]Coit Center for Longevity and Neurotherapeutics, The University of Arizona, Tucson, AZ, USA
[34]epi4health, Faculty of Health Sciences, Universitat Oberta de Catalunya, 08018 Barcelona, Catalonia, Spain
[35]Mechanisms of Cancer and Aging Laboratory, Department of Molecular and Cell Biology, University of Leicester, Leicester LE1 7RH, UK
[36]University of Raparin, Kurdistan Region, Iraq
[37]Department of Cardiovascular Sciences, College of Life Sciences, University of Leicester, Leicester LE1 7RH, UK
[38]National Institute for Health and Care Research Leicester Biomedical Research Centre, University Hospitals of Leicester NHS Trust, Glenfield Hospital, Leicester LE3 9QP, UK
[39]Leicester van Geest MultiOMICS Facility, University of Leicester, Leicester LE1 7RH, UK
[40]Leicester Cancer Research Centre, University Hospitals of Leicester NHS Trust, Leicester Royal Infirmary, University of Leicester, Leicester LE1 5WW, UK
[41]Department of Medicine and Life Sciences, Universitat Pompeu Fabra, PRBB, Barcelona, Spain
[42]Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Barcelona, Catalonia, Spain
[43]Hospital Universitari Institut Pere Mata (HUIPM), Institut d'Investigació Sanitària Pere Virgili (IISPV), Universitat Rovira i Virgili (URV), Catalonia, Spain
[44]Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Instituto de Salud Carlos III, Madrid, Spain
[45]Physiological Sciences Department, School of Medicine and Health Sciences, University of Barcelona (UB), Barcelona, Catalonia, Spain
[46]These authors contributed equally
[47]Lead contact
*Correspondence: mesteller@carrerasresearch.org
https://doi.org/10.1016/j.xcrm.2025.102368

INTRODUCTION

The indexed individual, from now on termed M116, was the world's oldest verified living person from January 17th, 2023, until her passing on August 19th, 2024, reaching the age of 117 years and 168 days (https://www.grg-supercentenarians.org/wop/). She was a Caucasian woman born on March 4th, 1907, in San Francisco, USA, from Spanish parents and settled in Spain since she was 8. A timeline of her life events and her genealogical tree are shown in Figures S1A and S1B. Although centenarians are becoming more common in the demographics of human populations, the so-called supercentenarians (over 110 years old) are still a rarity. In Catalonia, the historic nation where M116 lived, the life expectancy for women is 86 years, so she exceeded the average by more than 30 years (https://www.idescat.cat). In a similar manner to premature aging syndromes, such as Hutchinson-Gilford Progeria and Werner syndrome, which can provide relevant clues about the mechanisms of aging, the study of supercentenarians might also shed light on the pathways involved in lifespan. To unfold the biological properties exhibited by such a remarkable human being, we developed a comprehensive multiomics analysis of her genomic, transcriptomic, metabolomic, proteomic, microbiomic, and epigenomic landscapes in different tissues, as depicted in Figure 1A, comparing the results with those observed in non-supercentenarian (NSC) populations. The picture that emerges from our study, although derived only from this one exceptional individual, shows that extremely advanced age and poor health are not intrinsically linked and that both processes can be distinguished and dissected at the molecular level.

RESULTS

Genomics studies unveil genetic variants related to immune fitness, efficient mitochondrial function, cognition retention, cardioprotection, and longevity

Samples from the subject were obtained from four different sources: total peripheral blood, saliva, urine, and stool at different times. Most of the analyses were performed in the blood material at the time point of 116 years and 74 days, unless otherwise specifically indicated (Table S1). The simple karyotype of the supercentenarian did not show any gross chromosomal alteration (Figure S1C). Since many reports indicate the involvement of telomeres in aging and lifespan,[1] we interrogated the telomere length of the M116 individual using high-throughput quantitative fluorescence *in situ* hybridization (HT-Q-FISH) analysis.[2] Illustrative confocal images with DAPI staining and the telomeric probe (TTAGGG) for M116 and two control

CellPress
OPEN ACCESS

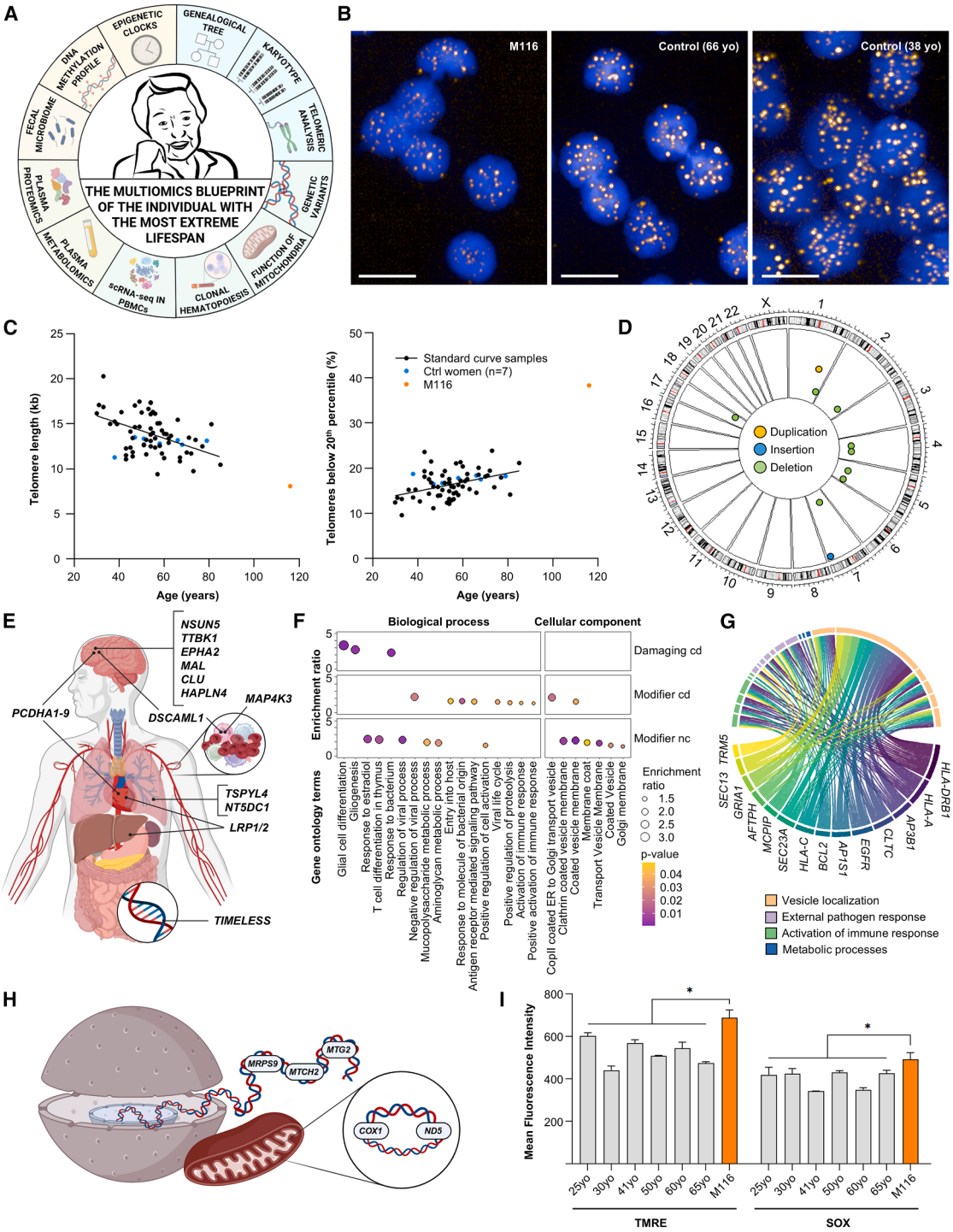


Figure 1. Chromosomes and genes: Genomics studies for telomeres, structural variations, and genetic variants of interest in the supercentenarian

(A) Schematic representation of all -omics studied in the supercentenarian.

(B) Telomeres marked with Cy3 (yellow) in nuclei stained with DAPI (blue) observed in HT-qFISH from M116 and younger women's PBMCs. Scale bars: 20 μm.

(C) Telomere length (Kb) calculation (left) and percentage of extremely short telomeres (below the 20th percentile) (right) in M116 (orange) using standard curve from samples previously analyzed (black) and control women (blue).

(legend continued on next page)

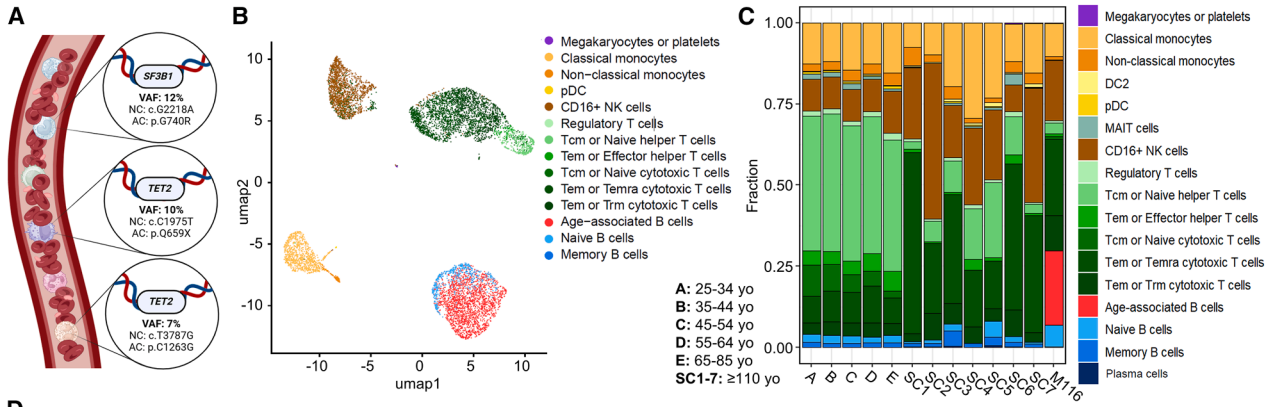
samples are shown in Figure 1B. Strikingly, we observed that the supercentenarian exhibited the shortest mean telomere length among all healthy volunteers[3] with a value of barely 8 kb (Figure 1C). Even more noticeably, the M116 individual displayed a 40% of short telomeres below the 20th percentile of all the studied samples (Figure 1C). Thus, the observed far reach longevity of our case occurred in the chromosomal context of extremely short telomeres. Interestingly, because the M116 individual presented an overall good health status, it is tempting to speculate that, in this setting, telomere attrition behaves more as a chromosomal clock for aging rather than a predictor of age-linked diseases such as neurodegeneration or diabetes. As an idea worth considering, the huge erosion of the telomere sequences in the supercentenarian could also relate to the absence of any diagnosed cancer because it could limit the replicative lifespan of any malignant cell.

Beyond telomeric regions, we assessed the presence of structural variants (SVs) at the chromosomal scale that could not be detected by regular karyotyping. Using optical genome mapping (OGM), we observed the presence of ten rare SVs in M116 as described in Table S2. The largest SVs were a 3,312.4 kb deletion in chromosome 4 and a 93.5 kb deletion in chromosome 17, as illustrated in the Circos plot of Figure 1D. Extreme longevity is believed to be influenced by a complex interplay of genetic variants, both common and rare, that affect healthy aging. This study sought to investigate rare genetic variants in M116's genome that may contribute to her exceptional lifespan. We performed whole genome sequencing (WGS) of the M116 individual where we identified approximately 3.8 million single nucleotide variants (SNVs) across her genome. Variants were filtered and annotated, and rare variants were defined as having a variant allele frequency (VAF) below 0.015 in European populations (1000 Genomes and gnomAD). Then, they were further annotated attending to their potential impact on protein function, and if applicable, classified as variants of interest (VOIs), ending up with 91,666 VOIs affecting 25,146 genes. The analysis focused on identifying functional variants that might impact genes or gene sets associated with longevity or disease resistance. The comparison was made between M116's genome and a control set of 75 Iberian women from the 1000 Genomes Project to identify "extreme" variants potentially linked to her longevity. We identified 7 homozygous variants in M116's genome, affecting 16 protein-coding and 3 non-coding genes (Table S3). Remarkably, none of these rare homozygous variants were found in the control European populations, suggesting that these variants could contribute to her exceptional longevity. Examples include homozygous variants detected on *DSCAML1*, a gene associated with immune function and cognition retention[4]; *MAP4K3*, linked to lifespan regulation of *Caenorhabditis elegans*[5] and to autoimmune disease, cancer, and aging[6]; *TSPYL4* and *NT5DC1*, linked to homeostatic pulmonary function[7] and the protocadherin alpha cluster (*PCDHA1-9*), related to aging brain health and heart disease[8,9] (Figure 1E; Table S3).

We also conducted an overrepresentation of functional categories analysis to evaluate whether specific biological processes or pathways were disproportionately affected by the rare variants in M116's genome. In addition to glial functionality—potentially linked to neuroprotection against degeneration—and the classic hormonal factors associated with aging associated features ("response to estradiol"),[10] one of the strongest enrichments was found in immune system-related categories, such as "T cell differentiation in thymus," "response to bacterium," "regulation of viral process," or "antigen receptor-mediated signaling pathway" (Figure 1F; Table S4). Furthermore, we observed VOI affected genes with highly pleiotropic functions across several overrepresented immune system-related pathways (Figure 1G). These functions are essential for controlling infections, autoimmune regulation, and possibly cancer surveillance, which could have contributed to M116's longevity. Categories related to heart function, such as "atrial septum morphogenesis" and "myocardial fibrosis" cholesterol metabolism and insulin signaling, and neuroprotection and axonal function, were also significantly enriched (Table S4). In this regard, rare variants were found in genes involved in lipid metabolism and heart function, such as *LRP1* and *LRP2*,[11] and in *NSUN5* and *TTBK1*, genes linked to neuroprotection that could potentially be contributing to the preservation of cognitive function in extreme old age (Figure 1E). Additionally, we found a rare variant in the *TIMELESS* DNA repair gene, linked to the evolutionary divergence in longevity across *Drosophila* populations[12] (Figure 1E). Interestingly, due to the long studied effect of reactive oxygen species in aging, rare variants in many genes involved in mitochondrial function, such as *ND5*, *COX1*, *MTG2*, *MTCH2*, and *MRPS9*, were detected (Figure 1H). These variants could affect mitochondrial oxidative phosphorylation, as illustrated in Figure S2; a process crucial for energy production, aging and longevity.[13] To experimentally interrogate this possibility, peripheral blood mononuclear cells (PBMCs) from M116 were isolated to assess mitochondrial membrane potential and to measure superoxide ion levels. We observed that the values of the markers were even higher than the ones observed in PBMCs from younger women (Figure 1I),

(D) Circos plot with chromosomal alterations detected through optical genome mapping in supercentenarian. Each type of structural variation (deletion, insertion, or duplication) is plotted at a fixed radius, with all variants of the same type positioned equidistantly from the center. The placement of each dot along the chromosome track approximates the actual genomic location of the structural variation.
(E) Variants of interest (VOIs)-harboring genes found in supercentenarian's genomic DNA contributing to immune function, cardiovascular health, neuroprotection, metabolism, and DNA dynamics.
(F) Significantly enriched functions of VOI-harboring genes in the supercentenarian.
(G) VOIs-harbouring genes significantly contributing to enriched functions.
(H) VOIs-harboring genes found in supercentenarian's genomic and mitochondrial DNA contributing to mitochondrial function.
(I) Mean fluorescence intensity of TMRE (a marker of mitochondrial membrane potential) and SOX (a marker of mitochondrial superoxide ion) in PBMCs from the supercentenarian (orange) and healthy controls across various ages (gray). Unpaired t test was used to statistically compare M116 to the mean of all control women. $^{*}p < 0.05$.
Figures 1A–1E and 1H were created with BioRender.com.

CellPress
OPEN ACCESS

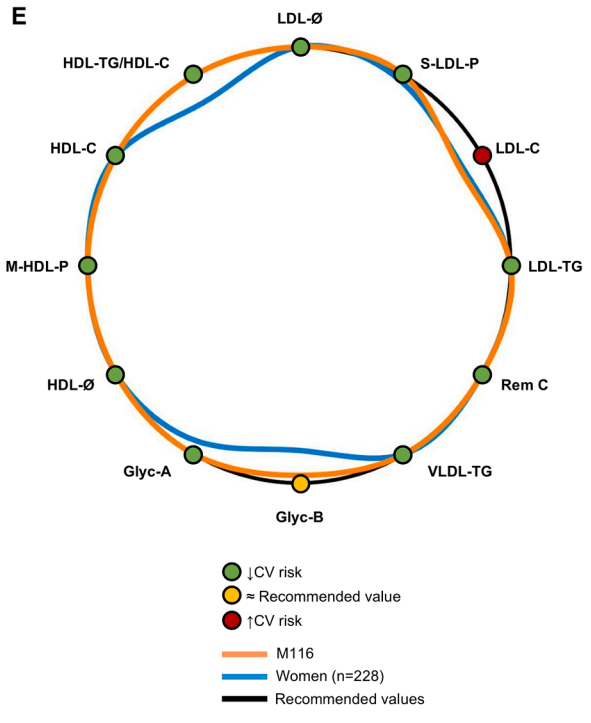


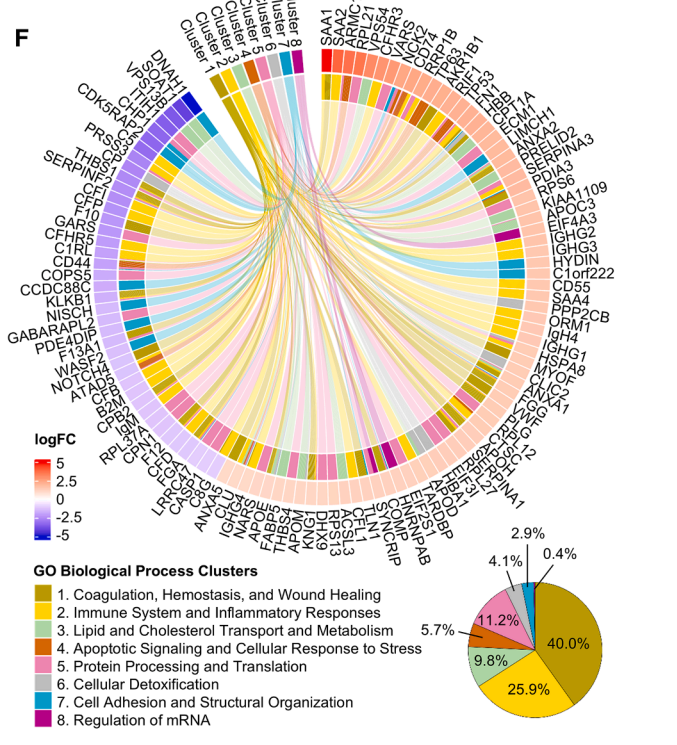
D

Parameter	Result	Recommended target value	Percentiles of the reference population 25%	50%	75%
Cholesterol content in principal lipoproteins					
VLDL-C (mg/dl)	6	< 22	6	11	17
HDL-C (mg/dl)	72	> 50♀ > 40♂	48	56	64
TG/HDL-C	1.1	< 3.0	1.2	1.8	3.0
HDL-TG/HDL-C	0.18	< 0.25	0.16	0.21	0.27
HDL-TG (mg/dl)	13	< 12	9	12	15
VLDL-TG (mg/dl)	39	< 98	39	54	78
Particle number					
Large (L-HDL-P) (µmol/l)	0.41	> 0.24	0.25	0.28	0.32
Medium (M-HDL-P) (µmol/l)	16.9	> 8.2	8.5	9.7	11
Small (S-HDL-P) (µmol/l)	10	> 15	15	18	21
Large (L-LDL-P) (nmol/l)	228	< 180	170	200	230
Protein-sugar bond concentration					
GLYC-A (µmol/l)	621	< 650	633	766	900
GLYC-B (µmol/l)	352	< 340	336	375	417

Parameter	Result	Recommended target value	Percentiles of the reference population 25%	50%	75%
Lipid concentration					
Saturated fatty acids (mmol/l)	7.99	7.8 – 9.7	8.28	8.74	9.19
Esterified cholesterol (mmol/l)	1.94	2.5 – 3.3	2.6	2.9	3.1
Free cholesterol (mmol/l)	2.2	1.52 – 2.19	1.65	1.82	2.0
Linoleic acid (mmol/l)	2.8	3.2 – 5.3	3.6	4.1	4.6
Low molecular weight metabolites concentration					
Acetone (µmol/l)	8	14 – 59	18	25	38
Glycine (µmol/l)	116	169 – 292	188	214	247
Histidine (µmol/l)	24	54 – 86	61	70	79
Valine (µmol/l)	115	130 – 206	145	164	186
Leucine (µmol/l)	52	75 – 126	84	96	109
Lactate (µmol/l)	551	232 – 572	287	359	452
Creatinine (µmol/l)	58	27 – 57	33	39	48

↓CV risk
≈ Recommended value
↑CV risk
M116
Women (n=228)





(legend on next page)

indicating not only preserved but also robust mitochondrial function in the supercentenarian.

Finally, we conducted three burden tests to identify genes or gene sets with a significantly higher burden of rare variants in M116 compared to controls, which were considered as differentiating genes, and performed an over-representation analysis. Using this approach, key rare variants in genes involved in neuroprotection and longevity, such as *EPHA2*, *MAL*, *CLU*, and *HAPLN4* were also highlighted (Figure 1E; Table S5). For genetic variants previously associated with increased longevity, we have analyzed the 91 SNPs variants in 88 genes[14,15] shown in Table S6, including those in *APOE* and *FOXO3A*.[16] We have observed that M116 exhibited a homozygous genotype of the allele associated with increased longevity for the genes *RABGAP1L*, *TMEM43*, *LSM3*, and *SOX6* (Table S6). Interestingly, the interrogated supercentenarian did not present in homozygosity any deleterious variants associated with short lifespan for all the analyzed genes. For *APOE* we have observed that the supercentenarian exhibited the GG genotype at position chr19:45410002 (rs769449), a variant that is opposed to the "A" allele, which is the one associated with aging-associated diseases and decreased lifespan. For *FOXO3A*, the supercentenarian did not harbor the gene variant associated with increased lifespan, the A allele at chr6:108867031 (rs6911407). Overall, our study indicates that there was not a single biological process targeted by only one rare genetic variant solely associated with the healthy aging and extended lifespan of M116, but rather a combination of rare variants in multiple genes and pathways (immune system, cardioprotection, brain activity, and mitochondrial metabolism) likely working together that promoted her remarkable longevity.

Blood analyses demonstrate the occurrence of clonal hematopoiesis, expanded age-associated B cells, low inflammation, and efficient lipid metabolism

To dwell further in the molecular and cellular landscapes associated with the extreme lifespan of the individual M116, we studied in deep detail the characteristics of the subject's blood. This is an important issue since this source of biological material is the one commonly used in the laboratories of the hospitals around the world; and there are exciting data from parabiotic experiments in mice suggesting that sharing bloodstream can affect aging.[17] We first assessed the existence of clonal hematopoiesis of indeterminate potential (CHIP), also denominated age-related clonal hematopoiesis that it is associated with the advanced age.[18] The occurrence of CHIP is thought to be a precursor of myeloid malignancies and other cancers and fosters the development of cardiovascular disease.[18] Defining CHIP by the presence of mutations in any leukemia driver gene (Table S7) with a VAF $\geq$2%, we found that M116 harbored a CHIP genotype defined by one mutation in the RNA splicing factor *SF3B1* and two different mutations in the DNA demethylase *TET2* (Figure 2A). Importantly, despite these predisposing mutations, the supercentenarian did not experience any tumorigenic process or cardiovascular disorder in her lifetime. The 5′-end regulatory CpG sites for both genes were fully unmethylated in M116 and NSC cohort (Figure S3), thus, epigenetic regulation for these targets was not observed in our study. The reasons why a significant fraction of individuals with CHIP mutations, such as M116, remain tumor-free are unknown.[18] However, potential explanations include the absence of necessary cooperating somatic hits, restricted clonal expansion, or enhanced immune surveillance.[18] We next characterized the distinct immune populations in M116 by analyzing her PBMCs using single-cell transcriptomics (scRNA-seq). This approach enabled the identification of major lymphoid and myeloid populations, including naive and memory B cells, natural killer cells, classical and non-classical monocytes, and distinct T cell subpopulations. The distribution of these populations is visualized in the UMAP (uniform manifold approximation and projection) plot in Figure 2B, where each cell cluster is colored according to its cell type annotation. A striking feature is the presence of an expanded cluster of age-associated B cells (ABCs). This subpopulation is known to accumulate with age contributing to a pro-inflammatory environment related to autoimmunity and infection response. The ABC cluster in the M116 individual is distinguished by the high expression of ABC associated markers, such as *SOX5* and *FCRL2*. Interestingly, we found that the top ranked overexpressed gene in ABC cells compared to naive and memory B cells was the oncogene *MYC* (Table S8). This observation could suggest that the ABC cells in M116 are transitioning to the recently identified age-associated clonal B cells that depend on MYC activity to proliferate and precede B cell malignancies in mouse models.[19] In addition, gene ontology (GO) analyses for the differential ABC genes showed an enrichment for pathways related to mitochondrial function, the regulation of apoptosis and major histocompatibility complex (MHC) activity (Figure S4).

We next sought to compare the immune cell type proportions of the M116 case with those of healthy individuals spanning

Figure 2. Blood genomics, metabolomics, and proteomics: Analyses of clonal hematopoiesis, single-cell RNA expression, and cardiolipidic profiles in the supercentenarian

(A) Mutated genes contributing to supercentenarian's CHIP. VAF: variant AF; NC: nucleotide change in cDNA; AC: amino acid change in protein. Created with https://BioRender.com.

(B) UMAP of PBMCs from supercentenarian colored by cell type annotation after single-cell RNA-seq.

(C) PBMCs' cell type proportions comparison of M116 with healthy controls of five different age ranges and other previously studied supercentenarians.

(D) Characteristic metabolic signatures in supercentenarian's plasma represented alongside recommended target values established from reference population. Percentiles in reference population are represented in bars, where arrows indicate levels in supercentenarian (orange) and control population (black). Those variables clearly associated with cardiovascular (CV) risk appear in a color scale, with green representing low CV risk and red representing high CV risk.

(E) Lipidic contour of supercentenarian (orange) and control population (blue) modeling 12 lipid and inflammatory metabolism variables associated with cardiovascular (CV) risk. Recommended values in black.

(F) Proteomics data showing differentially expressed proteins between M116 and controls and their associated functional category (p value and q value <0.05, with at least three associated proteins per functional category). Protein colors range from red (upregulated) to blue (downregulated), according to fold change in expression. 1–8 gene ontology (GO) biological process clusters (colored) were obtained from clustering GO terms based on semantic similarity.

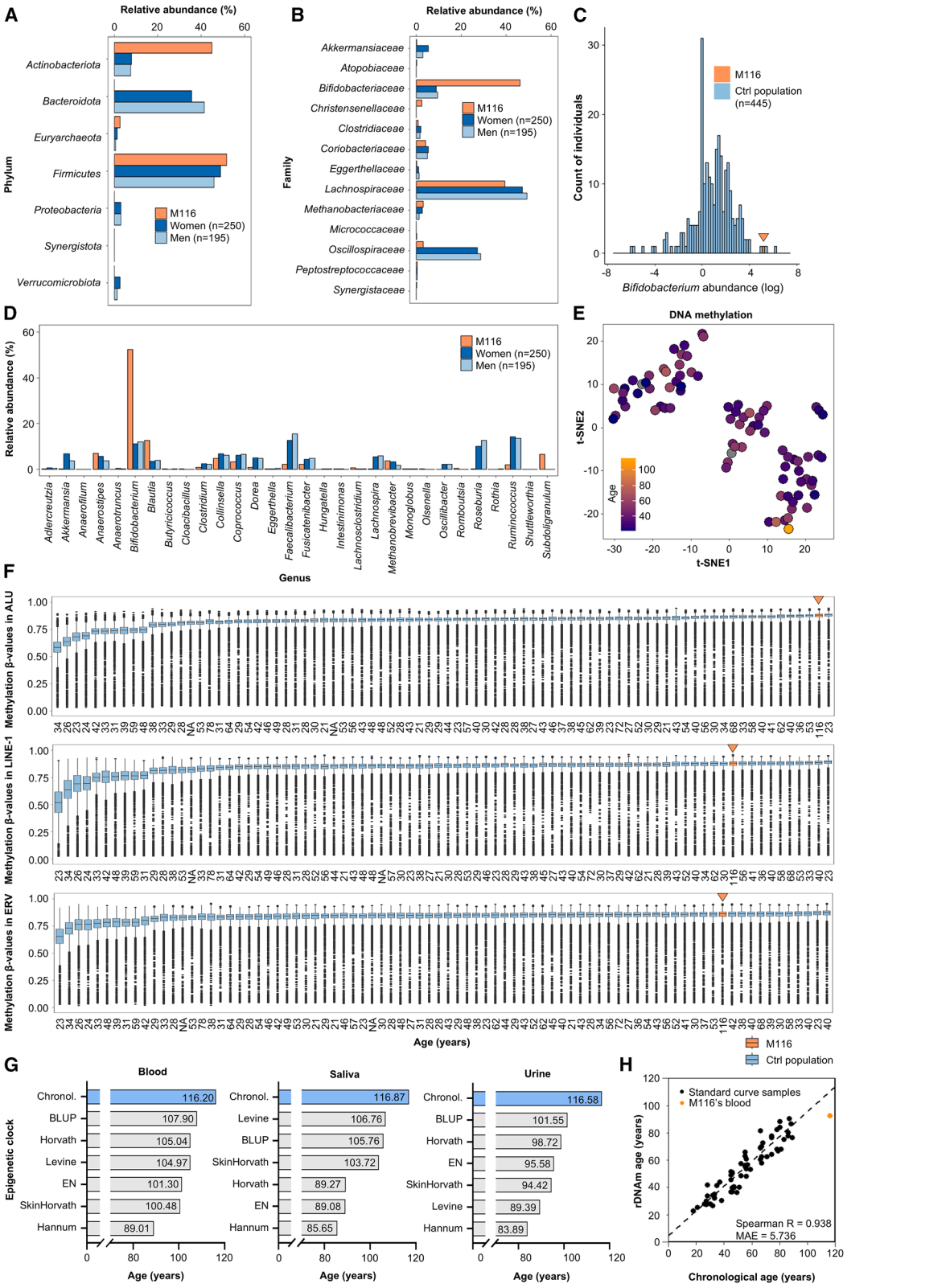
different age groups, from young adults (25 years) to other supercentenarians. Figure 2C illustrates the relative abundance of immune cell populations across five previously characterized age groups,[20] seven supercentenarians,[21] and the oldest supercentenarian M116. In younger age groups, the immune profile is characterized by higher fractions of regulatory and naive T cells. In supercentenarians, these populations are reduced and accompanied by an expansion of other lymphoid populations. Notably, in our longest-living supercentenarian M116, effector and memory T cells dominate the T cell compartment, particularly cytotoxic T cell subsets. This pattern aligns with previous observations of cytotoxic T cell enrichment in supercentenarians.[21] In this regard, we also observed a progressive increase in senescent T cell populations ($CD28^-/CD57^+$ and $CD28^-/CD57^+/KLRG1^+$) and a decrease in naive T cells ($CD45RA^+/CCR7^+$) with aging, our supercentenarian exhibiting the highest and lowest values, respectively, for these populations (Figures S5A–S5C). These last T cell populations shifts in aging have been previously described.[22,23] However, the pronounced expansion of the ABC cluster in the M116 individual is not observed in other supercentenarians, illustrating the singularity of her exceptional longevity. We also obtained the differentially expressed genes at the single-cell level for each cell type in M116 vs. other supercentenarians (SC group)[21]; and M116 vs. non-supercentenarians (NSC group)[20] (Table S9). Using the scRNA-seq data as pseudo-bulk, we also interrogated another aging hallmark, autophagy.[1] The study of an extensive list of autophagy markers (Table S10)[24] showed that the M116 sample shared a similar autophagy-related transcriptome profile with the young group[24] in comparison to other supercentenarians,[21] clustering in the NSC arm (Figure S6). These data reinforce, also for this hallmark, the notion of an enriched presence of young-like features in our case of extreme longevity.

Using the scRNA-seq data as pseudobulk-RNA we have also obtained the overall differentially expressed genes between M116 vs. SC and M116 vs. NSC (Table S11), which showed again how M116 clustered in the younger group (Figure S7A). These targets were then used to devise the main transcription factors regulation networks for M116 (Table S11; Figure S7B). Among the highly ranked transcription factors involved in the particular expression profile of M116, we found modulators of the immune and inflammatory response such as the widely studied nuclear factor κB/RELA (Rel avian reticuloendotheliosis viral oncogene homolog A) network, the MHC regulator RFX5 or the STAT1/STAT2 complex that controls the expression of interferon-stimulated genes (Table S11).

The observed occurrence of CHIP and the observed scRNA-seq landscape of the blood in M116 prompted us to study further the molecular composition of this clinically applicable biological source. To achieve this aim, we performed whole proton nuclear magnetic resonance (^{1}H-NMR) analyses to obtain the lipid and lipoprotein landscape, the glycoprotein profile and the low molecular weight metabolites (LMWMs) spectrum. All the values derived from these analyses are shown in Figures S8A and S8B. The lipoprotein and glycoprotein values derived for the supercentenarian were compared with 6,022 individuals of two Spanish population-base cohorts,[25,26] thus, with similar background to M116. For the comparison of concentrations in aqueous metabolome and lipid families, we used the 1,965 individuals of the Di@bet.es study,[25] where these data are available. The unfolded metabolomic landscape of the oldest human reflects an interesting dichotomy: she showed values of many metabolites and biomolecules associated with a healthy life and having had a long lifespan, but at the same time, a few biomarkers were suggesting that the end of her life was near. Probably the most relevant finding was that she presented one of the most efficient lipid metabolisms reported, a trait that the UK Biobank publications link to extended longevity[27] and absence of dementia,[28] as it occurred in the M116 subject. In this regard, M116 displayed extremely low levels of VLDL-cholesterol and triglycerides, whereas HDL-cholesterol (the "*good*" cholesterol) was very high (Figure 2D). Additionally, the high number of medium and large HDL-cholesterol particles and large LDL-cholesterol particles, combined with a low number of small HDL-cholesterol particles, support the idea of effective lipoprotein maturation and overall enhanced lipid metabolism[29] (Figure 2D).

The supercentenarian's highly efficient lipid metabolism is also reflected in the low levels of other lipid-related biomarkers typically associated with poor health and increased mortality, such as saturated fatty acids,[30] esterified cholesterol,[30] linoleic acid,[29] and acetone,[27] as well as in the high levels of free cholesterol,[29] associated with good health and survival (Figure 2D). Beyond the lipid metabolism, Figure 2D highlights two key inflammatory markers: GlycA and GlycB, which are composite nuclear magnetic resonance (NMR) signals derived from the N-acetyl methyl groups of circulating acute-phase glycoproteins. These markers have emerged as robust indicators of systemic inflammation and cardiometabolic risk. GlycA, in particular, reflects a combination of several glycosylated acute-phase proteins (e.g., α1-acid glycoprotein, haptoglobin, α1-antitrypsin, and transferrin), and has been independently associated with cardiovascular events and mortality in large population studies.[31] Unlike single inflammatory markers, GlycA integrates both the circulating concentration and glycosylation status of multiple proteins, offering a stable and reproducible measure of systemic inflammation. GlycB, a complementary but distinct signal, provides additional inflammatory context and may reflect different glycoprotein dynamics.[32] Low levels of GlycA and GlycB, as observed in our supercentenarian, suggest minimal acute-phase response and low-grade inflammation, consistent with a favorable cardiometabolic profile and a lower risk of subclinical atherosclerosis and cardiovascular events.[32] Furthermore, the study of the metabolites associated with cardiovascular risk support the idea of an excellent cardiovascular health in the supercentenarian (Figure 2E). Interestingly, the only metabolites that suggest her real very advanced chronological age and the risk of death are the low levels of the amino acids glycine, histidine, valine, and leucine, and the high concentration of lactate and creatinine[27] (Figure 2D). Overall, these metabolomic findings suggest that a highly engaged lipid metabolism together with very low levels of inflammation could explain the excellent health and extreme longevity observed in the studied supercentenarian, despite signs of functional decay in other pathways.

Since the composition of extracellular vesicles (ECVs) has also been linked to longevity,[33] we performed an additional



(legend on next page)

proteomic analysis of M116 ECVs and compared it to a baseline of younger post-menopausal women (yPM; n = 4, ages: 49–65 years). Using an orthogonal partial least squares discriminant analysis (OPLS-DA) model, we observed proteomic distinctions between both groups (Figure S9). Overall, 231 proteins had significant expression differences among M116 and the post-menopausal women set (Table S12). GO enrichment analysis clustered in eight groups (coagulation, immune system, lipid metabolism, apoptosis, protein processing, cellular detoxification, cellular adhesion, and mRNA regulation) (Figure 2F), showing an enrichment of pathways related to complement activation, B cell immunity, acute inflammatory response, adaptive immune response, humoral response and, potentially, a reduced inflammatory response in the supercentenarian (Table S13). Furthermore, we also found an increased lipid and cholesterol transport, higher lipoprotein remodeling, enhanced fatty acid transport, lipoprotein clearance, and high regulation of lipoprotein levels in the M116 sample (Table S13) in addition to enhanced oxidant detoxification and response to oxidative stress (Figure 2F; Table S13). These differences suggest an increase in protective mechanisms that could contribute to a healthy aging in the interrogated individual. Interestingly, these protein biomarkers of well-being in the M116 individual coexisted with a potentially dangerous pathophysiological sign: the protein most elevated in the supercentenarian compared to the yPM women was serum amyloid A-1 protein (SAA1) (Table S12). SAA1 is linked to Alzheimer's disease,[34] but despite her advanced age, M116 showed no evidence of any neurodegenerative disorder. We also performed an integrated analysis of the transcriptomic data (using scRNA-seq as pseudobulk-RNA) and the aforementioned proteomic data in M116 compared with non-supercentenarians. This approach revealed twenty-one genes with a distinct expression for both RNA and protein levels in our supercentenarian (Table S14; Figures S10A and S10B). Interestingly, M116 displayed an upregulation of immunoglobulin G genes, such as *IGHG2* and *IGHG4*, that have been linked to efficient generation of immunological memory responses in aging.[35]

Microbiome and epigenome characterization show a younger biological age according to bacterial profiles and DNA methylation clocks

The aforementioned discoveries drove us to learn about the microbiome composition of the studied record-breaking supercentenarian since microorganisms are critical in determining not only the metabolite composition of our body, but also inflammation, intestinal permeability, cognition, and bone and muscle health.[36] Thus, we determined the fecal microbiota composition by 16S rDNA analysis of the M116 individual and compared the results with 445 samples from control individuals (250 women and 195 men) aged 61–91 years that were not under antibiotic treatment from the curatedMetagenomicData dataset. First, we analyzed the microbiome diversity within the samples from M116 and compared it to that of control women individuals. Focusing on the α-diversity, a measure of diversity within-sample, we observed that the M116 sample exhibited a higher value compared to the mean alpha-diversity of the control female population (Shannon alpha-diversity values: 6.78 vs. 3.05, respectively). Next, we examined the beta-diversity, which measures dissimilarity between conditions. As shown in the principal coordinate analysis plot (Figure S11), the M116 sample is positioned outside the 90% confidence interval of the control population's distribution (STAR Methods). At the phylum level, the most startling finding was the high levels of *Actinobacteriota* in comparison with the control populations from both genders (Figure 3A). Zooming at the family level we observed that the increase was mostly due to the elevated amount of *Bifidobacteriaceae* (Figure 3B), particularly *Bifidobacterium* when we move to the genus classification (Figures 3C and 3D). This finding contrasts sharply with the typical decline of this bacterial genus in older individuals[37]; however, it has also been reported at elevated levels in centenarians[38] and, most importantly, in supercentenarians.[39]

Bifidobacterium is thought to be a beneficial bacterium contributing, among other processes, to anti-inflammatory responses, an observation that links with the low levels of inflammation markers in the metabolomics study (Figure 2D). High content of *Bifidobacterium* has also been associated with the production of short-chain fatty acids and conjugated linoleic acid,[40] observations that relate to the "healthy" lipid-related biomarker profile detected by ^{1}H-NMR (Figures 2D and 2E). Importantly, the use of *Bifidobacterium* as a probiotic that could slow down the progression of many aging-associated disorders is gaining momentum. Intriguingly, our healthy supercentenarian ingested around 3 yogurts every day containing *Streptococcus thermophilus* and *Lactobacillus delbrueckii* subsp. *bulgaricus*, known to favor the growth of the described bacteria in the gut (a representative 3-week menu of the diet that the supercentenarian followed during her last 20 years of life is shown in Table S15). Thus, this could be an example of a dietary intervention that, acting in the gut microbiota, is associated with healthy aging and long lifespan. Other bacteria might also contribute to

Figure 3. Microbiome and epigenetics: Profiles of bacterial composition, DNA methylation status of repetitive elements, and epigenetic clocks in the supercentenarian

(A–C) Bar plots representing the percentage of relative abundance of bacterial populations found in supercentenarian's fecal samples at (A) phylum level, (B) family level, and (C) genus level.

(D) Abundance of *Bifidobacterium* in supercentenarian (orange) and the control population (blue) represented as the count of individuals. Orange arrow points to supercentenarian.

(E) tSNE representation of differentially methylated CpGs between supercentenarian and healthy control population.

(F) Boxplot representing predicted methylation values in ALU, LINE-1, and ERV repetitive element sequences for supercentenarian (orange) and healthy control population (blue). Orange arrows point to supercentenarian. Boxplots are ordered increasingly according to median methylation value for each sample.

(G) Chronological (blue) and estimated biological age (gray) of supercentenarian's tissues according to different publicly available epigenetic clocks.

(H) Estimated rDNA methylation age in supercentenarian's blood (orange) and control population (black) according to rDNAm clock.

the described phenotype. For example, lower levels of the phylum *Proteobacteria* and *Verrucomicrobiota,* as we have found in our M116 sample vs. the control groups (Figure 3A), have been linked to older adults without age-related frailty.[41] In a similar manner, the diminished levels of the pro-inflammatory genus *Clostridium*[37] in the supercentenarian vs. controls (Figure 3D) would support the absence of aging-associated inflammatory diseases. Overall, these results suggest that the studied extreme supercentenarian possesses a microbiome that confers an increased likelihood for a healthy extended lifespan. This is also consistent with her adherence to a Mediterranean diet (Table S15), which might have contributed to the described microbiome composition.[42]

Finally, we added the layer of the epigenetic setting in the studied longest-living human, as a factor (like microbiome composition) that might provide a dynamic, plastic and adaptable readout of M116 biology. DNA methylation is probably the most studied epigenetic mark in cell biology and disease,[43] being also disrupted as we age,[44] as it has been shown, for example, comparing newborns vs. nonagenarians[45] and in Hutchinson-Gilford progeria.[46] Thus, we first interrogated the CpG methylation status of the supercentenarian case using a comprehensive microarray that analyzed more than 850K loci,[47] comparing the results to a large collection of control individuals previously studied in our laboratory[48] (n = 81, age range = 21–78 years old) and interrogated with the same epigenetic platform. We observed that 69 CpG sites were differentially methylated (β-value at least 50% distinct) in the supercentenarian vs. the rest of individuals (Table S16). According to CpG content, only 26.09% (18 out of 69) of these sites were placed in CpG islands, whereas 73.91% (51 out of 69) were located in regions of the genome with lower CpG density (Table S17). Related to genomic structure, 33.33% (23 out of 69) of CpG sites were located within the 5′-regulatory gene regions (transcription start site −1,500 base pairs, transcription start site −200 base pairs, 5′-untranslated region and 1st exon), while the remaining 66.67% (46 out of 69) of CpG sites were associated with the gene body, 3′-untranslated region, or intergenic regions (Table S17). Importantly, 68.12% (47 out of 69) of the differently methylated loci corresponded to a loss of CpG methylation in the supercentenarian sample. We also compared the localization of hypomethylated and hypermethylated CpG sites in relation to CpG islands and observed no remarkable differences (Table S17). However, when focusing on their localization within gene loci, we found a higher proportion of hypomethylated CpG sites (21 out of 47; 44.68%) within gene bodies compared to hypermethylated CpG sites (6 out of 22; 27.27%). This loss of CpG methylation in the supercentenarian sample is in line with the previously described waves of DNA hypomethylation shifts taking place in human aging.[45,46] Interestingly, a global DNA hypomethylation status is also commonly observed in human cancer,[49] a disease with an increased incidence in older people, highlighting some of the crosstalks and commonalities between the aging and tumorigenesis pathways.[49]

To provide further functional impact to the identified CpG sites, we crossed their methylation status with RNA expression data in human cell lines.[50] For those CpG sites included in the described databases within an associated gene with available RNA values, we found that the M116 CpG hypomethylation status was associated with low expression of *EGFL7* (EGF-like domain multiple 7) and *ADCY3* (adenylate cyclase 3), and also hypermethylation of *PLEKHA1* (pleckstrin homology domain containing A1) (for all cases, p value < 0.05 and rho > |0.3|) (Table S18). For an additional gene, *VASN* (vasorin), the hypomethylation event at the supercentenarian occurred in the 5′-UTR and, as expected for this regulatory region,[51] was associated with high levels of the RNA transcript (Table S18). The identified DNA methylation-controlled genes are *bona fide* candidates to play a role in the supercentenarian's biology, since they are involved in vascular stemness (*EGFL7*), body mass index (*ADCY3*), macular degeneration (*PLEKHA1*), and bone turnover (*VASN*). All of these activities considered hallmarks of aging or tightened to aging-associated disorders. Using the 69 differentially methylated loci in a supervised hierarchical clustering, we can observe the distinct DNA methylation profile between M116 and the controls (Figure S12). As described previously, most of these differential CpG sites represented hypomethylation events in M116 compared with the other samples (Figure S12; Table S17). Further dimensionality reduction analysis with the 69 differentially methylated CpG sites by t-distributed stochastic neighbor embedding (t-SNE) also shows the supercentenarian as a DNA methylome outlier (Figure 3E). Importantly, because the loss of DNA methylation in repetitive sequences has been reported through the aging process,[52] we carefully studied the DNA methylation content of three families of repeats: LINE-1, ALU, and ERV (endogenous retrovirus). Interestingly, we found that the supercentenarian did not undergo major hypomethylation events in these loci, instead retained hypermethylated CpGs even at higher levels than most of the younger individuals (Figure 3F). Thus, these findings suggest that a disruption of the DNA methylation balance (hypermethylation/hypomethylation) in gene 5′-regulatory CpG islands is linked to the aging trajectory, but keeping epigenetically silent DNA repetitive sequences could confer an advantage associated with healthy longevity, as it occurred in our case.

DNA methylation analyses provided another important clue that might explain the amazing lifespan of our supercentenarian. The last years have witnessed the development of the so-called "epigenetic clocks"[53] that are utilized as a proxy to calculate "biological age" of a given tissue or specimen that, although in most cases is expected to match the "chronological age" (the amount of time that has elapsed from birth), it is not always the case. In this regard, several pathologies can accelerate the process and, for example, premature biological aging determined by these epigenetic clocks occurs among carriers of viral infections.[54,55] Since these technologies use a comparable DNA methylation platform used herein, we were able to calculate the biological age of our supercentenarian using six different epigenetic clocks (Figure 3G). Remarkably, all the distinct algorithms of age based on DNA methylation yielded the same result. Our supercentenarian exhibited a much younger biological age than her real chronological age and this occurred in the three different tissues analyzed (Figure 3G). These results were reinforced when we used a completely different epigenetic clock: CpG methylation status of ribosomal DNA[56] (Figure 3H). Importantly, this approach does not use the mentioned DNA

methylation microarray data, but whole genome bisulfite sequencing (WGBS). Thus, we carried out WGBS of our supercentenarian sample and compared the obtained data with a set of 70 cases studied using the same experimental and bioinformatic pipeline. Our results showed that while the control group showed concordance between both types of age, the M116 case exhibited a much lower biological age than her chronological age, 23.17 years difference (Figure 3H). Interestingly, the age pace (the vertical distance between M116's predicted age and the reference cohort's baseline regression line, Figure 3H; STAR Methods) was −17.34 years, indicating a biological age deceleration determined by the rDNAm clock in comparison to her chronological age. This further validates the findings of the six epigenetic clocks derived from the DNA methylation microarray test (Figure 3G). Overall, these data suggest that one of the reasons that our supercentenarian reached such a world record age was that her cells "felt" or "behaved" as younger cells.

DISCUSSION

The occurrence of extreme longevity and the onset of aging associated diseases reflects the balance between nature and nurture. A resilient genome, such as the one carried by M116 with potential beneficial variants and exclusion of damaging ones provides an advantage for a healthy life and increased lifespan, such as the harboring of homozygous status for pro-longevity alleles of genes, such as *APOE*, and many other gene variants related to cardiovascular protection, brain health, immune system-related pathways, and mitochondrial oxidative phosphorylation (Tables S3 and S4). Interestingly, we have identified seven homozygous variants in M116's genome that have not been described in any European control population (Table S3), suggesting that these variants could contribute to her lifespan. However, the habits and environment could have shaped these features as well. For example, related to this last issue, our supercentenarian ingested a high amount of yogurts, a trait associated with reduced body weight and type-2 diabetes incidence[57] and diminished body fat and insulin resistance.[58] In our case, whether the dominance of the *Bifidobacterium* related genus is fully attributable or not to the yogurt diet cannot be completely confirmed since that would have required a longitudinal study with sample collection over several years. However, we believe that it is likely that a beneficial effect of yogurt ingestion via modulation of the gut ecosystem could have contributed to her well-being and advanced age. In addition to the dietary influence, there is an increasing interest in fecal microbiota transplantation studies that through promoting gut microbial rejuvenation could also foster healthy aging.[59] The type of diet may also increase longevity by metabolomic tuning such as causing calorie restriction or limiting particular nutrients such as methionine.[60] Metabolic intervention in aging could also be achieved by agents that mimic calorie restriction, mTOR inhibitors (e.g., rapalogs) or metformin (many functions including adenosine monophosphate-activated protein kinase [AMPK] activation).[60] Other behavioral components beyond diet can affect human lifespan, such as exercise. In this regard, even low-intensity physical activities, such as walking, exerts anti-aging effects[61] and structured exercise can also improve survival in age-associated diseases such as cancer.[62] Beyond lifestyle choices, some of the observed molecular features could also be potential intervention targets for delaying aging. For example, the observation in M116 of very short telomeres could suggest that drugs protecting these chromosome-end regions may have anti-aging effects. It is more complicated for the epigenetic setting. The supercentenarian maintains repeat DNA sequences in a hypermethylated status, as in much younger individuals, preventing the activation of these endoparasitic sequences and protecting against genomic breaks. Thus, an epigenetic treatment using the available DNA-demethylating agents could be counterproductive. However, other drugs targeting distinct layers of epigenetic regulation, such as histone modifications, may be worth exploring.

Overall, our findings suggest that extreme human longevity may be characterized by the coexistence of two distinct and potentially unrelated sets of features within the same individual. On one hand, there are characteristic biomarkers of very advanced age, such as shortened telomeres, clonal hematopoiesis-associated mutations, or an aged B cell population. On the other hand, there are simultaneously preserved healthy (epi)genetic and functional tissue environment traits. This last scenario is evidenced by the presence of genetic variants protective against common diseases (e.g., cardiovascular disorders, diabetes, and neurodegeneration), an efficient lipid metabolism, an anti-inflammatory gut microbiome, and an epigenome associated with chromosomal stability and decelerated epigenetic aging. All these findings illustrate how aging and disease can, under certain conditions, become decoupled, challenging the common perception that they are inextricably linked.

Limitations of the study

Aging and the achievement of extreme longevity are probably highly individualized processes, shaped by a confluence of genetic, environmental, and stochastic factors. Thus, despite our study has interrogated several multiomics layers, and compared with many population datasets, drawing broadly applicable conclusions from a single subject should be taken with caution. We have also assessed telomere length using HT-Q-FISH on PBMCs, a heterogeneous cell population composed of various lymphocyte and monocyte subsets, each with distinct telomere dynamics. Thus, this is a limitation to interpret telomere length data as reflective of systemic aging. Further detailed studies for aging hallmarks such as inflammation, senescence, and autophagy would also be necessary. A final limitation of our work is that we have not studied the effect of exercise, metabolic tuning, or assessed the effects of drugs targeting some of the observed features to explore their potential anti-aging effects. We expect that these endeavors will be the focus of prospective research in the field.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Manel Esteller (mesteller@carrerasresearch.org).

Materials availability

This study did not generate new reagents.

Data and code availability

- All data reported in this paper has been deposited at the European Genome-Phenome Archive, under accession number EGAS50000000884 (https://ega-archive.org/studies/EGAS50000000884) and under restricted access to protect individual information. Access will be granted for appropriate use for researchers by the lead contact.
- The codes developed can be accessed via https://github.com/mesteller-bioinfolab/Supercentenarian/ and https://github.com/FusterTormo/AUP/.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank Yoana Veselinova for the drawing of the graphical abstract. Research in the M.E. group is funded by the CERCA Programme/Generalitat de Catalunya (MCIN/AEI/10.13039/501100011033/); the European Regional Development Fund, "A way to make Europe" ERDF (project PID2021-125282OB-I00 and PID2024-159192OB-I00); Departament de Recerca i Universitats/Generalitat de Catalunya (2021 SGR 01494); European Union under THRIVE grant agreement no. 101136622; "La Caixa" Research Foundation; and the Cellex Foundation (CEL007). E.S.-P. is a fellow of the Spanish Ministry of Science, Innovation and Universities, under FPI contract no. PRE2022-105015. G.F. is a recipient of Ayuda Investigador AECC 2023 (INVES234765FERR), Fundación Científica AECC. C.Q.-D. is a fellow of the Spanish Ministry of Science, Innovation and Universities, under FPU contract no. FPU22/01655. E.M.'s work in this publication was supported by the Ramón y Cajal fellowship RYC2021-032359-I, funded by the Spanish Ministry of Science and the Catalan Agency for Management of University and Research Grants (AGAUR, 2021 SGR 01586). Research in the A.N. group is supported by the I+D+i project PID2021-127792NB-I00, funded by MCIN/AEI/10.13039/501100011033 (FEDER Una manera de hacer Europa), and "Unidad de Excelencia María de Maeztu," funded by the AEI (CEX2018-000792-M) and Departament de Recerca i Universitats de la Generalitat de Catalunya (GRC 2021 SGR 0467). Research in the R.C.-T.. group is funded by the Departament de Recerca i Universitats/Generalitat de Catalunya (2021 SGR 01366). Research in the S.M. group was supported by the Instituto de Salud Carlos III (PI20/00328) and the M. C. Andreu Memorial Fund. J. R.-B. was supported by a doctoral grant from the Universitat Oberta de Catalunya. Research in the M.A.B. laboratory is funded by the European Union, project ERC-AvG Shelterins (GA 882385), Horizon 2020 Programme. The mass spectrometry work was supported by the John and Lucille Van Geest Foundation and the National Institute for Health and Care Research Leicester Biomedical Research Centre. M.E. is an ICREA Research Professor. Special thanks are given to personnel taking care of M116 and to all of the biobank technicians involved in and contributing to keeping samples and procedures for the current research, with special mention to Gerard Pardo Albiñana.

AUTHOR CONTRIBUTIONS

E.S.-P., A.N.-C., M.C.-P., C.A.G.-P., I.C.-M., C.Q.-D., E.C.-G., A.B.-C., F. Setién, G.F., and V.D., conceptualization, performing experiments, formal analysis, and writing. E.M., scRNA-seq analysis supervision. R.P. and C.A., DNA methylation experiments. L.S., genetic analysis. C.d.l.T., proteomics analysis. F.V., statistics. I.G., cytogenetic analysis. N.S.C., sample collection. P.A., F. Solé, and M. Mallo, genetic analysis. C.M., scRNA-seq analysis supervision. S.P., T.G., M.L., M.P., R.C.-T., A.L., L.J.G.-G., and X.A., microbiome analysis. S.S., P.T., J.R., M.G., N.A., and O.Y., metabolome analysis. P.M., R.S.-V., and M.A.B., telomeric analysis. J.O. and B.L., rDNAm analysis. J.R.-B., M.T., M. M.-S., K.A.K., T.H.C., P.A.Q., D.J.L.J., and S.M., proteomics analysis. C.V., E.B.-B., M. Moldes, F.B., G.M., H.L., and A.N., genetic analysis. M.E., conceptualization, writing, and scientific supervision.

DECLARATION OF INTERESTS

M.E. declares past grants from Ferrer International and Incyte and personal fees from Quimatryx and Eucerin, outside of the submitted work.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Sample and data acquisition
- METHOD DETAILS
  - PBMCs extraction
  - Karyotype
  - High-Throughput Quantitative Fluorescence In Situ Hybridization (HT-Q-FISH) analysis
  - Optical Genome Mapping
  - DNA extraction from peripheral blood
  - DNA extraction from urine
  - DNA extraction from saliva
  - Whole Genome Sequencing
  - Mitochondrial analyses
  - Clonal hematopoiesis analysis
  - Single-cell transcriptomics
  - Metabolome
  - Proteomics
  - Transcriptomics and proteomics integration
  - Transcription factor regulatory network analysis
  - 16S meta-genomics sequencing
  - Infinium MethylationEPIC BeadChip
  - Whole Genome Bisulfite Sequencing
  - rDNA clock assessment
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrm.2025.102368.

Received: April 2, 2025
Revised: July 11, 2025
Accepted: August 26, 2025
Published: September 24, 2025

REFERENCES

1. López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M., and Kroemer, G. (2023). Hallmarks of aging: An expanding universe. Cell *186*, 243–278. https://doi.org/10.1016/j.cell.2022.11.001.
2. Canela, A., Vera, E., Klatt, P., and Blasco, M.A. (2007). High-throughput telomere length quantification by FISH and its application to human population studies. USA *104*, 5300–5305. https://doi.org/10.1073/pnas.0609367104.
3. Sanchez-Vazquez, R., Guío-Carrión, A., Zapatero-Gaviria, A., Martínez, P., and Blasco, M.A. (2021). Shorter telomere lengths in patients with severe COVID-19 disease. Aging *13*, 1–15. https://doi.org/10.18632/aging.202463.
4. Li, P., Marshall, L., Oh, G., Jakubowski, J.L., Groot, D., He, Y., Wang, T., Petronis, A., and Labrie, V. (2019). Epigenetic dysregulation of enhancers in neurons is associated with Alzheimer's disease pathology and cognitive symptoms. Nat. Commun. *10*, 2246–14. https://doi.org/10.1038/s41467-019-10101-7.

5. Khan, M.H., Hart, M.J., and Rea, S.L. (2012). The role of MAP4K3 in lifespan regulation of Caenorhabditis elegans. Biochem. Biophys. Res. Commun. *425*, 413–418. https://doi.org/10.1016/j.bbrc.2012.07.113.

6. Chuang, H.C., and Tan, T.H. (2019). MAP4K3/GLK in autoimmune disease, cancer and aging. J. Biomed. Sci. *26*, 82–88. https://doi.org/10.1186/s12929-019-0570-5.

7. Guo, Y., Gong, Y., Shi, G., Yang, K., Pan, C., Li, M., Li, Q., Cheng, Q., Dai, R., Fan, L., and Wan, H. (2012). Single-nucleotide polymorphisms in the TSPYL-4 and NT5DC1 genes are associated with susceptibility to chronic obstructive pulmonary disease. Mol. Med. Rep. *6*, 631–638. https://doi.org/10.3892/mmr.2012.964.

8. Hirayama, T., and Yagi, T. (2017). Regulation of clustered protocadherin genes in individual neurons. Semin. Cell Dev. Biol. *69*, 122–130. https://doi.org/10.1016/j.semcdb.2017.05.026.

9. Teekakirikul, P., Zhu, W., Gabriel, G.C., Young, C.B., Williams, K., Martin, L.J., Hill, J.C., Richards, T., Billaud, M., Phillippi, J.A., et al. (2021). Common deletion variants causing protocadherin-α deficiency contribute to the complex genetics of BAV and left-sided congenital heart disease. HGG Adv. *2*, 100037. https://doi.org/10.1016/j.xhgg.2021.100037.

10. Kumar, R.S., and Goyal, N. (2021). Estrogens as regulator of hematopoietic stem cell, immune cells and bone biology. Life Sci. *269*, 119091. https://doi.org/10.1016/j.lfs.2021.119091.

11. May, P., Woldt, E., Matz, R.L., and Boucher, P. (2007). The LDL receptor-related protein (LRP) family: An old family of proteins with new physiological functions. Ann. Med. *39*, 219–228. https://doi.org/10.1080/07853890701214881.

12. Hunt, L.C., Jiao, J., Wang, Y.D., Finkelstein, D., Rao, D., Curley, M., Robles-Murguia, M., Shirinifard, A., Pagala, V.R., Peng, J., et al. (2019). Circadian gene variants and the skeletal muscle circadian clock contribute to the evolutionary divergence in longevity across Drosophila populations. Genome Res. *29*, 1262–1276. https://doi.org/10.1101/gr.246884.118.

13. Li, Y., Zhang, H., Yu, C., Dong, X., Yang, F., Wang, M., Wen, Z., Su, M., Li, B., and Yang, L. (2024). New Insights into Mitochondria in Health and Diseases. Int. J. Mol. Sci. *25*, 9975. https://doi.org/10.3390/ijms25189975.

14. Wright, K.M., Rand, K.A., Kermany, A., Noto, K., Curtis, D., Garrigan, D., Slinkov, D., Dorfman, I., Granka, J.M., Byrnes, J., et al. (2019). A Prospective Analysis of Genetic Variants Associated with Human Lifespan. G3 (Bethesda) *9*, 2863–2878. https://doi.org/10.1534/g3.119.400448.

15. Liu, X., Song, Z., Li, Y., Yao, Y., Fang, M., Bai, C., An, P., Chen, H., Chen, Z., Tang, B., et al. (2021). Integrated genetic analyses revealed novel human longevity loci and reduced risks of multiple diseases in a cohort study of 15,651 Chinese individuals. Aging Cell *20*, e13323. https://doi.org/10.1111/acel.13323.

16. Caruso, C., Ligotti, M.E., Accardi, G., Aiello, A., Duro, G., Galimberti, D., and Candore, G. (2022). How Important Are Genes to Achieve Longevity? Int. J. Mol. Sci. *23*, 5635.

17. Pálovics, R., Keller, A., Schaum, N., Tan, W., Fehlmann, T., Borja, M., Kern, F., Bonanno, L., Calcuttawala, K., Webber, J., et al. (2022). Molecular hallmarks of heterochronic parabiosis at single-cell resolution. Nature *603*, 309–314. https://doi.org/10.1038/s41586-022-04461-2.

18. Dunn, W.G., McLoughlin, M.A., and Vassiliou, G.S. (2024). Clonal hematopoiesis and hematological malignancy. J. Clin. Investig. *134*, e180065. https://doi.org/10.1172/JCI180065.

19. Castro, J.P., Shindyapina, A.V., Barbieri, A., Ying, K., Strelkova, O.S., Paulo, J.A., Tyshkovskiy, A., Meinl, R., Kerepesi, C., Petrashen, A.P., et al. (2024). Age-associated clonal B cells drive B cell lymphoma in mice. Nat. Aging *4*, 1403–1417. https://doi.org/10.1038/s43587-024-00671-7.

20. Terekhova, M., Swain, A., Bohacova, P., Aladyeva, E., Arthur, L., Laha, A., Mogilenko, D.A., Burdess, S., Sukhov, V., Kleverov, D., et al. (2023). Single-cell atlas of healthy human blood unveils age-related loss of NKG2C+GZMB−CD8+ memory T cells and accumulation of type 2 memory T cells. Immunity *56*, 2836–2854.e9. https://doi.org/10.1016/j.immuni.2023.10.013.

21. Hashimoto, K., Kouno, T., Ikawa, T., Hayatsu, N., Miyajima, Y., Yabukami, H., Terooatea, T., Sasaki, T., Suzuki, T., Valentine, M., et al. (2019). Single-cell transcriptomics reveals expansion of cytotoxic CD4 T cells in supercentenarians. USA *116*, 24242–24251. https://doi.org/10.1073/pnas.1907883116.

22. Márquez, E.J., Chung, C.H., Marches, R., Rossi, R.J., Nehar-Belaid, D., Eroglu, A., Mellert, D.J., Kuchel, G.A., Banchereau, J., and Ucar, D. (2020). Sexual-dimorphism in human immune system aging. Nat. Commun. *11*, 751. https://doi.org/10.1038/s41467-020-14396-9.

23. Mittelbrunn, M., and Kroemer, G. (2021). Hallmarks of T cell aging. Nat. Immunol. *22*, 687–698. https://doi.org/10.1038/s41590-021-00927-z.

24. Bordi, M., De Cegli, R., Testa, B., Nixon, R.A., Ballabio, A., and Cecconi, F. (2021). A gene toolbox for monitoring autophagy transcription. Cell Death Dis. *12*, 1044. https://doi.org/10.1038/s41419-021-04121-9.

25. Valdés, S., Doulatram-Gamgaram, V., Maldonado-Araque, C., García-Escobar, E., García-Serrano, S., Oualla-Bachiri, W., García-Vivanco, M., Garrido, J.L., Gil, V., Martín-Llorente, F., et al. (2024). Association between exposure to air pollution and blood lipids in the general population of Spain. Eur. J. Clin. Invest. *54*, 1–11. https://doi.org/10.1111/eci.14101.

26. Puig-Jové, C., Castelblanco, E., Falguera, M., Hernández, M., Soldevila, B., Julián, M.T., Teis, A., Julve, J., Barranco-Altirriba, M., Franch-Nadal, J., et al. (2022). Advanced lipoprotein profile in individuals with normal and impaired glucose metabolism. Rev. Esp. Cardiol. *75*, 22–30. https://doi.org/10.1016/j.recesp.2021.02.013.

27. Zhang, S., Wang, Z., Wang, Y., Zhu, Y., Zhou, Q., Jian, X., Zhao, G., Qiu, J., Xia, K., Tang, B., et al. (2024). A metabolomic profile of biological aging in 250,341 individuals from the UK Biobank. Nat. Commun. *15*, 8081. https://doi.org/10.1038/s41467-024-52310-9.

28. Wang, Z., Yin, Z., Sun, G., Zhang, D., and Zhang, J. (2025). Genetic evidence for the liver-brain axis: lipid metabolism and neurodegenerative disease risk. Lipids Health Dis. *24*, 41. https://doi.org/10.1186/s12944-025-02455-3.

29. Lian, J., and Vardhanabhuti, V. (2024). Metabolic biomarkers using nuclear magnetic resonance metabolomics assay for the prediction of aging-related disease risk and mortality: a prospective, longitudinal, observational, cohort study based on the UK Biobank. Geroscience *46*, 1515–1526. https://doi.org/10.1007/s11357-023-00918-y.

30. Wang, F., Tessier, A.J., Liang, L., Wittenbecher, C., Haslam, D.E., Fernández-Duval, G., Heather Eliassen, A., Rexrode, K.M., Tobias, D.K., Li, J., et al. (2023). Plasma metabolomic profiles associated with mortality and longevity in a prospective analysis of 13,512 individuals. Nat. Commun. *14*, 5744. https://doi.org/10.1038/s41467-023-41515-z.

31. Chiesa, S.T., Charakida, M., Georgiopoulos, G., Roberts, J.D., Stafford, S.J., Park, C., Mykkänen, J., Kähönen, M., Lehtimäki, T., Ala-Korpela, M., et al. (2022). Glycoprotein Acetyls: A Novel Inflammatory Biomarker of Early Cardiovascular Risk in the Young. J. Am. Heart Assoc. *11*, e024380. https://doi.org/10.1161/JAHA.121.024380.

32. Cediel, G., Teis, A., Codina, P., Julve, J., Domingo, M., Santiago-Vacas, E., Castelblanco, E., Amigó, N., Lupón, J., Mauricio, D., et al. (2022). GlycA and GlycB as Inflammatory Markers in Chronic Heart Failure. Am. J. Cardiol. *181*, 79–86. https://doi.org/10.1016/j.amjcard.2022.07.019.

33. Zhang, X., Ma, S., Naz, S.I., Soderblom, E.J., Aliferis, C., and Kraus, V.B. (2025). Plasma extracellular vesicles carry immune system-related peptides that predict human longevity. Geroscience *47*, 1455–1469. https://doi.org/10.1007/s11357-024-01454-z.

34. Yang, J., Zhou, Y., Wang, T., Li, N., Chao, Y., Gao, S., Zhang, Q., Wu, S., Zhao, L., and Dong, X. (2024). A multi-omics study to monitor senescence-associated secretory phenotypes of Alzheimer's disease. Ann. Clin. Transl. Neurol. *11*, 1310–1324. https://doi.org/10.1002/acn3.52047.

35. de Jong, B.G., Ijspeert, H., Marques, L., van der Burg, M., van Dongen, J.J., Loos, B.G., and van Zelm, M.C. (2017). Human IgG2- and IgG4-expressing memory B cells display enhanced molecular and phenotypic signs of maturity and accumulate with age. Immunol. Cell Biol. *95*, 744–752.

36. Ticinesi, A., Tana, C., and Nouvenne, A. (2019). The intestinal microbiome and its relevance for functionality in older persons. Curr. Opin. Clin. Nutr. Metab. Care *22*, 4–12. https://doi.org/10.1097/MCO.0000000000000521.

37. Kossowska, M., Olejniczak, S., Karbowiak, M., Mosiej, W., Zielińska, D., and Brzezicka, A. (2024). The Interplay between Gut Microbiota and Cognitive Functioning in the Healthy Aging Population: A Systematic Review. Nutrients *16*, 852. https://doi.org/10.3390/nu16060852.

38. Ku, S., Haque, M.A., Jang, M.J., Ahn, J., Choe, D., Jeon, J.I., and Park, M.S. (2024). The role of Bifidobacterium in longevity and the future of probiotics. Food Sci. Biotechnol. *33*, 2097–2110. https://doi.org/10.1007/s10068-024-01631-y.

39. Biagi, E., Franceschi, C., Rampelli, S., Severgnini, M., Ostan, R., Turroni, S., Consolandi, C., Quercia, S., Scurti, M., Monti, D., et al. (2016). Gut Microbiota and Extreme Longevity. Curr. Biol. *26*, 1480–1485. https://doi.org/10.1016/j.cub.2016.04.016.

40. Arboleya, S., Watkins, C., Stanton, C., and Ross, R.P. (2016). Gut bifidobacteria populations in human health and aging. Front. Microbiol. *7*, 1204–1209. https://doi.org/10.3389/fmicb.2016.01204.

41. Wen, N.-N., Sun, L.-W., Geng, Q., and Zheng, G.-H. (2024). Gut microbiota changes associated with frailty in older adults: A systematic review of observational studies. World J. Clin. Cases *12*, 6815–6825. https://doi.org/10.12998/wjcc.v12.i35.6815.

42. Barber, T.M., Kabisch, S., Pfeiffer, A.F.H., and Weickert, M.O. (2023). The Effects of the Mediterranean Diet on Health and Gut Microbiota. Nutrients *15*, 2150–17. https://doi.org/10.3390/nu15092150.

43. Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. Nat. Biotechnol. *28*, 1057–1068. https://doi.org/10.1038/nbt.1685.

44. Seale, K., Horvath, S., Teschendorff, A., Eynon, N., and Voisin, S. (2022). Making sense of the ageing methylome. Nat. Rev. Genet. *23*, 585–605. https://doi.org/10.1038/s41576-022-00477-6.

45. Heyn, H., Li, N., Ferreira, H.J., Moran, S., Pisano, D.G., Gomez, A., Diez, J., Sanchez-Mut, J.V., Setien, F., Carmona, F.J., et al. (2012). Distinct DNA methylomes of newborns and centenarians. USA *109*, 10522–10527. https://doi.org/10.1073/pnas.1120658109.

46. Bejaoui, Y., Razzaq, A., Yousri, N.A., Oshima, J., Megarbane, A., Qannan, A., Potabattula, R., Alam, T., Martin, G.M., Horn, H.F., et al. (2022). DNA methylation signatures in Blood DNA of Hutchinson–Gilford Progeria syndrome. Aging Cell *21*, e13555–11. https://doi.org/10.1111/acel.13555.

47. Moran, S., Arribas, C., and Esteller, M. (2016). Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics *8*, 389–399. https://doi.org/10.2217/epi.15.114.

48. Joshi, R.S., Rigau, M., García-Prieto, C.A., Castro de Moura, M., Piñeyro, D., Moran, S., Davalos, V., Carrión, P., Ferrando-Bernal, M., Olalde, I., et al. (2022). Look-alike humans identified by facial recognition algorithms show genetic similarities. Cell Rep. *40*, 111257. https://doi.org/10.1016/j.celrep.2022.111257.

49. Esteller, M., Dawson, M.A., Kadoch, C., Rassool, F.V., Jones, P.A., and Baylin, S.B. (2024). The Epigenetic Hallmarks of Cancer. Cancer Discov. *14*, 1783–1809. https://doi.org/10.1158/2159-8290.CD-24-0296.

50. Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Gonçalves, E., Barthorpe, S., Lightfoot, H., et al. (2016). A Landscape of Pharmacogenomic Interactions in Cancer. Cell *166*, 740–754. https://doi.org/10.1016/j.cell.2016.06.017.

51. Murtha, M., and Esteller, M. (2016). Extraordinary Cancer Epigenomics: Thinking Outside the Classical Coding and Promoter Box. Trends Cancer *2*, 572–584. https://doi.org/10.1016/j.trecan.2016.08.004.

52. Cardelli, M. (2018). The epigenetic alterations of endogenous retroelements in aging. Mech. Ageing Dev. *174*, 30–46. https://doi.org/10.1016/j.mad.2018.02.002.

53. Bell, C.G., Lowe, R., Adams, P.D., Baccarelli, A.A., Beck, S., Bell, J.T., Christensen, B.C., Gladyshev, V.N., Heijmans, B.T., Horvath, S., et al. (2019). DNA methylation aging clocks: Challenges and recommendations. Genome Biol. *20*, 249–24. https://doi.org/10.1186/s13059-019-1824-y.

54. Esteban-Cantos, A., Rodríguez-Centeno, J., Barruz, P., Alejos, B., Saiz-Medrano, G., Nevado, J., Martin, A., Gayá, F., De Miguel, R., Bernardino, J.I., et al. (2021). Epigenetic age acceleration changes 2 years after antiretroviral therapy initiation in adults with HIV: a substudy of the NEAT001/ANRS143 randomised trial. Lancet. HIV *8*, e197–e205. https://doi.org/10.1016/S2352-3018(21)00006-0.

55. Cao, X., Li, W., Wang, T., Ran, D., Davalos, V., Planas-Serra, L., Pujol, A., Esteller, M., Wang, X., and Yu, H. (2022). Accelerated biological aging in COVID-19 patients. Nat. Commun. *13*, 2135–2137. https://doi.org/10.1038/s41467-022-29801-8.

56. Wang, M., and Lemos, B. (2019). Ribosomal DNA harbors an evolutionarily conserved clock of biological aging. Genome Res. *29*, 325–333. https://doi.org/10.1101/gr.241745.118.

57. Mozaffarian, D., Hao, T., Rimm, E.B., Willett, W.C., and Hu, F.B. (2011). Changes in diet and lifestyle and long-term weight gain in women and men. N. Engl. J. Med. *364*, 2392–2404.

58. Chen, Y., Feng, R., Yang, X., Dai, J., Huang, M., Ji, X., Li, Y., Okekunle, A.P., Gao, G., Onwuka, J.U., et al. (2019). Yogurt improves insulin resistance and liver fat in obese women with nonalcoholic fatty liver disease and metabolic syndrome: a randomized controlled trial. Am. J. Clin. Nutr. *109*, 1611–1619.

59. Novelle, M.G., Naranjo-Martínez, B., López-Cánovas, J.L., and Díaz-Ruiz, A. (2025). Fecal microbiota transplantation, a tool to transfer healthy longevity. Ageing Res. Rev. *103*, 102585. https://doi.org/10.1016/j.arr.2024.102585.

60. López-Otín, C., Galluzzi, L., Freije, J.M.P., Madeo, F., and Kroemer, G. (2016). Metabolic Control of Longevity. Cell *166*, 802–821. https://doi.org/10.1016/j.cell.2016.07.031.

61. Ungvari, Z., Fazekas-Pongor, V., Csiszar, A., and Kunutsor, S.K. (2023). The multifaceted benefits of walking for healthy aging: from Blue Zones to molecular mechanisms. Geroscience *45*, 3211–3239. https://doi.org/10.1007/s11357-023-00873-8.

62. Courneya, K.S., Vardy, J.L., O'Callaghan, C.J., Gill, S., Friedenreich, C.M., Wong, R.K.S., Dhillon, H.M., Coyle, V., Chua, N.S., Jonker, D.J., et al. (2025). Structured Exercise after Adjuvant Chemotherapy for Colon Cancer. N. Engl. J. Med. *393*, 13–25. https://doi.org/10.1056/NEJMoa2502760.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Lymphoprep density gradient centrifugation	Serumwerk Bernburg	Cat. 00122
RPMI medium	Gibco	Cat. 11875093
Poly-lysine	Greiner Bio-One, Inc.	Cat. 655087
DAPI (4′,6-diamidino-2-phenylindole, dihydrochloride)	Life Technologies	Cat. D-1306
PrepIT-L2P	DNA Genotek	Cat. PT-L2P
MitoTracker Green	Invitrogen	Cat. M7514
Red Tetramethylrhodamine, ethyl ester (TMRE)	Invitrogen	Cat. T669
MitoSOX Red	Invitrogen	Cat. M36008
BioTracker ATP-Red	Sigma-Aldrich	Cat. SCT045
Ammonium bicarbonate	Sigma-Aldrich	Cat. 1066-33-7
1% Ammonium deoxycholate	Sigma-Aldrich	Cat. K2755-1MG
Dithiothreitol	Sigma-Aldrich	Cat. D9779
Iodoacetamide	Sigma-Aldrich	Cat. I6125
Trypsin	Roche	Cat. RTRYP-RO
Formic acid	ThermoFisher scientific	Cat. 28905
Critical commercial assays		
Direct Label & Stain kit (DLS)	Bionano Genomics	Cat. 80005
Qubit BR kit	ThermoFisher Scientific	Cat. Q32850
KAPA HyperCap	Roche	Custom
Chromium Next GEM Single Cell 3′ Kit v3.1	10X Genomics	Cat. PN-1000268
DNeasy Powersoil Pro kit	Qiagen	Cat. 47014
EZ DNA Methylation-Gold™ Kit	Zymo Research	Cat. D5005
Deposited data		
Multiomics data derived from this study	This paper	Upon request
Software and algorithms		
Harmony High-Content Analysis Software	PerkinElmer	https://content.perkinelmer.com/
Prism v9.1.1	GraphPad	https://www.graphpad.com/updates/prism-900-release-notes
Bionano Solve 3.7	Bionano Genomics	https://bionano.com/
Bionano Access 1.7.2	Bionano Genomics	https://bionano.com/
R	The R Project for Statistical Computing	https://www.r-project.org/
Ensembl Variant Effect Predictor	Ensembl	https://www.ensembl.org/info/docs/tools/vep/index.html
BWA 0.7.15.	Github	https://github.com/lh3/bwa
GATK 4.1.8.0	Broad Institute	https://gatk.broadinstitute.org/hc/en-us/articles/360045673592–GATK-4-1-8-0-release
ANNOVAR 20200607	ANNOVAR	http://annovar.openbioinformatics.org/en/latest/
Seurat 5.1.0	CRAN	https://cran.r-project.org/web/packages/Seurat/index.html
DoubletFinder 2.0.4	Github	https://github.com/chris-mcginnis-ucsf/DoubletFinder

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
scDblFinder 1.20.0	Bioconductor	https://bioconductor.org/packages/devel/bioc/vignettes/scDblFinder/inst/doc/scDblFinder.html
CellTypist 1.6.3	Github	https://github.com/Teichlab/celltypist/releases/tag/1.6.3
Lipspin	Github	https://github.com/rbarri/LipSpin
MassLynx 4.1.	Waters	https://www.waters.com/nextgen/us/en.html
Progenesis software 4.2	Nonlinear Dynamic	https://www.nonlinear.com/
NOISeq	Bioconductor	https://www.bioconductor.org/packages/release/bioc/html/NOISeq.html
ARSyNseq	RDRR.IO	https://rdrr.io/bioc/NOISeq/man/ARSyNSeq.html
Ropls	Bioconductor	https://bioconductor.org/packages/release/bioc/html/ropls.html
limma	Bioconductor	https://bioconductor.org/packages/release/bioc/html/limma.html
clusterProfiler	Bioconductor	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
GOSemSim	Bioconductor	https://bioconductor.org/packages/release/bioc/html/GOSemSim.html
dada2	Bioconductor	https://www.bioconductor.org/packages/release/bioc/html/dada2.html
phyloseq	Bioconductor	https://www.bioconductor.org/packages/release/bioc/html/phyloseq.html
Vegan	CRAN	https://cran.r-project.org/web/packages/vegan/index.html
Ape	CRAN	https://cran.r-project.org/web/packages/ape/index.html
MicrobiotaProcess	Bioconductor	https://bioconductor.org/packages/release/bioc/html/MicrobiotaProcess.html
UpsetR	CRAN	https://cran.r-project.org/web/packages/UpSetR/index.html
Mia	Bioconductor	https://www.bioconductor.org/packages/release/bioc/html/mia.html
curatedMetagenomicData	Bioconductor	https://www.bioconductor.org/packages/release/data/experiment/html/curatedMetagenomicData.html
Minfi 1.48.0	Bioconductor	http://bioconductor.jp/packages/3.18/bioc/html/minfi.html
REMP	Bioconductor	https://www.bioconductor.org/packages/release/bioc/html/REMP.html
Decouple R	Bioconductor	https://saezlab.github.io/decoupleR/
Igraph	CRAN	https://cran.r-project.org/web/packages/igraph/index.html
ggraph	CRAN	https://cran.r-project.org/web/packages/ggraph/index.html

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Sample and data acquisition

Peripheral blood, urine, saliva and stool were collected from the supercentenarian and age at time of sampling was annotated and considered for subsequent analyses. The protocol of the study was approved by the institutional ethics review board (PI-23-19). An extensive interview was conducted addressing her clinical history and main lifestyle habits such as sleep time, diet, exercise, and social interactions. M116 was the world's oldest living person at the time of starting the present study, and was the 8th world's

longest-living person in history on verified records (https://www.grg-supercentenarians.org/wop/). She was a Caucasian woman born on March 4th 1907 in San Francisco, USA, to Spanish parents. She moved to New Orleans, USA, in August 1910, and in 1915, when she was 8 years old, her father passed away and she moved to Barcelona, Spain. From then on, she spent her life settled in Spain. From 2001 until death, she lived in a residence for the elderly in Olot, Catalonia, Spain. In spite of several emotionally painful events during her last years of life, like her son's death, she kept a strong physical and mental health throughout life with good sleep habits, balanced Mediterranean diet, and active social life. She largely enjoyed from quality time with family and friends, playing with dogs, reading books, growing a garden, walking, and playing the piano. She suffered from COVID-19 and chronic age-related diseases like bronchiectasis, esophagus diverticulum, and osteoarthritis, with limited movement and high dependency during her last months of life. She never suffered from other prevalent age-related diseases like cancer or neurodegenerative diseases, unlike siblings. She passed away on August 19th 2024, while sleeping, at the age of 117 years, 5 months, and 18 days.

METHOD DETAILS

PBMCs extraction

Freshly collected peripheral blood was centrifuged at 2,000*g* for 6 min at room temperature (RT) to remove plasma, and the cellular fraction was diluted with sterile PBS-EDTA 2 mM. Later, peripheral blood mononuclear cells (PBMCs) were isolated by Lymphoprep density gradient centrifugation (Serumwerk Bernburg; Cat. 00122) at 800*g* for 30 min at RT (acceleration and deceleration 3/9) and transferred to new tubes. PBMCs were centrifuged at 3,200*g* for 10 min at RT, and pellets were resuspended in 90% fetal bovine serum (FBS) + 10% DMSO for cryopreservation in liquid nitrogen.

Karyotype

Conventional G-banding karyotype was carried out by standard procedures. Briefly, 15 million B-cells from M116 were plated in RPMI medium (Gibco, Cat. 11875093) supplemented with 10% FBS. Colcemid was added to the culture followed by 20 minutes-incubation to arrest cells in metaphase. After centrifugation, a hypotonic solution (KCl, 0.075M) was added followed by 25 minutes-incubation at 37°C. Cells were centrifuged thrice and washed with Carnoy's solution (methanol and glacial acetic acid, 3:1) to break cytoplasmic membranes. Chromosomes were fixed in the spindle and, along with the different fixer passes, membranes, cytoplasm, and different organelles were removed to obtain a pellet with the nuclei of the cells that were in interphase and metaphase. Finally, the G band pattern was achieved with Wright dye.

High-Throughput Quantitative Fluorescence In Situ Hybridization (HT-Q-FISH) analysis

PBMCs (2 x 10^5 cells per well) from M116 and control women (IJC-Campus ICO-GTP Biological Sample Collection) were plated in 96-well plate coated with poly-lysine (Greiner Bio-One, Inc.; Cat. 655087) and fixed in Methanol:Acetic acid (3:1) for 15 min. Plates were stored at −20°C with the fixative solution. For Q-FISH hybridization, plates were dried at 37°C overnight and cells were then rehydrated with PBS and fixed with 4% formaldehyde, followed by digestion with pepsin/HCl and a second fixation with 4% formaldehyde. Samples were dehydrated with increasing concentrations of EtOH (70%, 90%, 100%) and incubated with the telomeric probe (TTAGGG) labeled with Cy3 at 85°C for 3 min followed by 1h RT in a wet chamber. Samples were extensively washed with 50% formamide and 0.08% TBS-Tween 20 followed by TBST containing 1 μg/mL DAPI (4′,6-diamidino-2-phenylindole, dihydrochloride) (Life Technologies; Cat. D-1306) to stain the nuclei. Next, the plate was washed with 0.08% TBS-Tween20 in a plate shaker for 5 min. Confocal images were captured using the Opera Phenix High-Content Screening System (PerkinElmer). Images were analyzed with Harmony High-Content Analysis Software (PerkinElmer).

Optical Genome Mapping

High molecular weight DNA (>250 Kb) isolated from the supercentenarian B-cells was marked with Direct Label & Stain kit (DLS) (Bionano Genomics, Cat. 80005), which labels specific sequences of 6 nucleotides (CTTAAG) repeated throughout the genome using an enzyme (DLE-1) attached to a fluorophore. Images were obtained using the Saphyr instrument (Bionano Genomics), with a minimum coverage of 300X. For result analysis, the Rare Variant Analysis algorithm included in the Bionano Solve 3.7 software, with the hg19 genome version, was used. Visualization was performed using Bionano Access 1.7.2. Alterations obtained were considered after applying the Bionano Genomics confidence filters and discarding benign/polymorphic alterations.

DNA extraction from peripheral blood

Freshly collected peripheral blood was centrifuged at 2,000*g* for 6 min at RT to remove plasma, and cellular fraction was diluted with sterile PBS-EDTA 2 mM and centrifuged at 3,200*g* for 10 min at RT. Supernatant was discarded, and cells were washed twice in Ammonium-Chloride-Potassium (ACK) lysing buffer (centrifugations at 3,200*g* for 10 min at RT). Pellets with washed whole-blood leukocytes were stored at −80°C.

Lysis buffer was freshly prepared (TrisHCl 2.5 mM pH 8 + EDTA 1.25 mM + NaCl 50 mM) and used to prepare lysis mix (1 mL lysis buffer +100 μL 10% SDS +10 μL proteinase K 18.6 mg/mL). Pellet with whole-blood leukocytes was thawed on ice and resuspended in 1 mL lysis mix. After 3h-incubation in the thermoblock at 55°C with agitation, NaCl was added to a final concentration of 1.5M. The tube was centrifuged at maximum speed for 10 min at RT, and supernatant was transferred to a new tube. Cold 100% isopropanol

was added and pipetted up and down for proper homogenization. The tube was incubated overnight at −20°C and centrifuged at maximum speed for 15 min at 4°C. The supernatant was discarded, and the pellet was resuspended in 70% ethanol. The tube was centrifuged again at maximum speed for 15 min at 4°C, the supernatant was discarded, and the pellet was dried at RT until complete evaporation of ethanol. DNA was resuspended in Tris-EDTA (TE) buffer and incubated in the thermoblock for 10 min at 55°C with agitation for proper rehydration. Quality and quantity of DNA were evaluated using 2200 TapeStation (Agilent) and Nanodrop One spectrophotometer (ThermoFisher Scientific).

DNA extraction from urine

Freshly collected urine was transferred to 50 mL tubes and centrifuged at 3,000*g* for 10 min at RT. Supernatant was discarded, and pellet was washed in PBS prior to storage at −80°C. DNA extraction was performed as for whole-blood leukocytes.

DNA extraction from saliva

Saliva was freshly collected into Oragene 500 DNA tubes, and DNA was extracted according to manufacturer's instructions (DNA Genotek). In brief, the tube was incubated in an air incubator overnight at 50°C, and the content was transferred to a 15 mL tube. 1/25 volumes of PrepIT-L2P (DNA Genotek, Cat. PT-L2P) were added, and the tube was vortexed and incubated on ice for 10 min. The tube was centrifuged at maximum speed for 10 min at RT, and the supernatant was transferred to a new 15 mL tube. 1.2x volumes of 100% ethanol were added, and the sample was gently mixed and incubated for 10 min at RT to allow complete DNA precipitation. The tube was centrifuged at maximum speed for 10 min at RT, the supernatant was discarded, and the pellet was washed with 1 mL of 70% ethanol for 1 min. DNA was rehydrated with 500 μL TE buffer and quantified using Qubit BR kit (ThermoFisher Scientific, Cat. Q32850) on a Qubit2.0 fluorimeter (ThermoFisher Scientific), and DNA integrity was measured through 2200 Tapestation (Agilent).

Whole Genome Sequencing

WGS was performed using DNA extracted from M116's blood, saliva, and urine samples. DNA was fragmented to 350 bp, followed by library preparation, paired-end 150 bp WGS, and variant calling, all conducted by Novogene. Sequencing was carried out with a 10X coverage (approximately 60 Gb and 200 million reads per sample) using the NovaSeq X Plus platform (Illumina). After Quality control (QC) performed by Novogene (QD < 2.0, FS > 60.0, MQ < 40.0, HaplotypeScore>13.0, MappingQualityRankSum<−12.5, and ReadPosRankSum<−8.0), approximately 3.8 million SNVs were identified. Variants were further filtered to include only those present in at least 2 out of the 3 samples, thereby minimizing the impact of technical artifacts. A control cohort consisting of 75 women from the Iberian population in Spain, was used for comparison. The same filters applied to the M116 samples were used for this cohort. All samples were processed and analyzed uniformly.

SNVs were annotated using the Ensembl Variant Effect Predictor to determine transcript location, variant class, and additional attributes such as allele frequency (AF) from various populations, including 1000 Genomes Project (1000G) and gnomAD EUR exomes. Functional variant effect annotations were obtained from SIFT, PolyPhen-2, and Combined Annotation-Dependent Depletion scores. Rare variants were defined as those with an AF < 0.015 in both the 1000G 30X Illumina NovaSeq sequencing dataset (2,504 unrelated individuals) and gnomAD v4.1 genome dataset. These rare variants were further classified based on their potential impact on protein function or their predicted effects according to sequence annotations, and were designated as VOIs. To compare the rarity of M116's rare variants with the control cohort, three burden tests were performed: Cohort Allelic Sum Test, Rare Variant Test, and Rank Test of Proportions. Rare variants were considered differentially enriched if p-value <0.05 and significant in at least 2 out of the 3 tests. Functional analyses, including over-representation analysis, were conducted using genes carrying VOIs or differentiating genes. These analyses employed ontology gene sets from the Human MSigDB collection. Pathways were deemed over-represented with a p-value <0.05.

Mitochondrial analyses

PBMCs from supercentenarian and healthy women controls (obtained from Banc de Sang i Teixits, Barcelona, Spain) were washed and plated in 96 well plates (25,000 cells/well) in RPMI (Gibco, Cat. 11875093). PBMCs were simultaneously stained with 100 nM MitoTracker Green (Invitrogen, Cat. M7514) for 30 min at RT together with either 50 nM Red Tetramethylrhodamine, ethyl ester (TMRE) (Invitrogen, Cat. T669) for 30 min at RT, 5 μM MitoSOX Red (Invitrogen, Cat. M36008) for 10 min at 37°C, or 7.5 μM BioTracker ATP-Red (Sigma-Aldrich, Cat. SCT045) for 15 min at 37°C following manufacturer's instructions. Cells were washed with PBS and analyzed by flow cytometry (FACSCanto II, BD Biosciences) using FITC (MitoTracker Green) and PE (Red TMRE, MitoSOX Red, and BioTracker ATP-Red) channels. Only MitoTracker Green positive cells were considered in the subsequent analytical assessments. Unpaired t-test was performed to statistically compare M116 to the mean of controls.

Clonal hematopoiesis analysis

Genomic DNA was obtained from peripheral blood samples and was used for targeted deep sequencing studies. Barcoded libraries were prepared according to the manufacturer's instructions, using a probe-based panel (KAPA HyperCap, Roche) targeting frequently mutated regions of 50 myeloid-related genes (Table S7). Samples were run on a MiSeq (Illumina) sequencer for paired-end 2x75 bp reads with a mean coverage of 1000X. Sequencing data were analyzed using an in-house pipeline. Reads were aligned

to human genome build 19 (hg19/GRCh37) using BWA 0.7.15. Post-alignment and base recalibration were performed using the tools in GATK 4.1.8.0 software package. Variant calling was performed with Mutect2 4.1.8.0, which is included in GATK software package. ANNOVAR 20200607 was used for variant annotation. The variant filtering process was based on the criteria proposed by the Spanish Group of myelodysplastic syndromes. Briefly, variants were filtered according to location (exonic and splicing), variant type (nonsynonymous single-nucleotide variants and small insertions/deletions), read depth (>100x), minor AF (MAF <0.01 according to dbSNP, ExAC, Exome Variant Server and 1000 Genomes project population databases) and variant AF (VAF $\geq$2%).

Single-cell transcriptomics

ScRNA-seq was performed using Chromium Next GEM Single Cell 3′ Kit v3.1 (10X Genomics, Cat. PN-1000268) according to the manufacturer's instructions. Briefly, PBMCs were thawed and quantified in order to calculate the number of cells to be loaded. Then, barcoded Single Cell 3' Gel Beads, a master mix containing PBMCs, and Partitioning Oil were combined onto the Chromium Next GEM Chip G to generate Gel Beads-In-Emulsion (GEMs), where polyadenylated mRNAs were reverse transcribed. This way, all generated cDNAs from the same cell shared a common 10X Barcode. Following the reverse transcription, GEMs were broken and cDNAs were amplified and cleaned up with SPRIselect beads (Beckman Coulter). Next, a portion of these cDNAs were enzymatically fragmented and subjected to adaptor ligation before using them as a PCR template for the incorporation of i5/i7 indexes. Finally, libraries were purified with SPRIselect beads, quantified and quality checked by using the 2200 TapeStation (Agilent), and subjected to paired-end 150-bp sequencing (Novaseq systems, Illumina) following standard practices at external facilities (Novogene).

The gene count matrices were analyzed using the Seurat package (version 5.1.0) in R (version 4.4.1). QC was performed to ensure that high-quality cells were retained for downstream analysis. Genes expressed in less than 3 cells were filtered out. Mitochondrial gene expression was quantified using PercentageFeatureSet function, and cells with mitochondrial content >20% were excluded. Cells with less than 300 detected genes were also filtered out as low quality. To exclude potential artificial doublets generated during library construction, we identified them applying two independent methods: DoubletFinder (version 2.0.4) and scDblFinder (version 1.20.0). Cells classified as doublets by both methods were filtered out for downstream analysis. Data were normalized (NormalizeData function: "LogNormalize" method), followed by dimensionality reduction performed via principal component (PC) analysis (RunPCA function) on the scaled expression (ScaleData function) of the top 2,000 highly variable genes (FindVariableFeatures function: variance-stabilizing transformation method). 15 PCs were retained for downstream analysis, by inspecting the elbow plots. These top PCs were used to create a UMAP embedding (RunUMAP function) and cluster the cells (FindClusters function) in a 20-nearest neighbor graph (FindNeighbors function). Cell type annotations were predicted with automatic label transfer using CellTypist (version 1.6.3) (with the human Immune_All_High or Immune_All_Low models). This study analyzed two publicly available scRNA-seq datasets from human PBMCs: Terekhova et al.,[20] Synapse database under accession code syn49637038; and Hashimoto et al.[21] at http://gerg.gsc.riken.jp/SC2018/. Normalized data were subsampled to retain 10,000 cells from each age group. In both datasets, CellTypist was used to annotate the cells, as above described. Using scRNA-seq as pseudo-bulk, differential gene expression analysis was performed to compare transcriptomic profiles between M116 and the two described age groups.[20,21] Genes were considered differentially expressed (DEGs) using a zero-inflated MAST (test with the following thresholds: absolute $\log_2$ fold change (logFC) $\geq$ 0.50, minimum detection fraction (min.pct) $\geq$ 10% in either group, and a minimum of 3 cells per group required for testing. Statistical significance was determined using adjusted *p*-values (p_adj) with Bonferroni's correction for multiple testing. The analysis was implemented through Seurat's FindMarkers() function, comparing each group pair while controlling for cellular detection rates as covariates in the MAST model.

Metabolome

300 μL of serum sample was used for the whole proton NMR (1H-NMR) analysis, which includes the lipoprotein profile based on the Liposcale test, the glycoprotein profile and the LMWM profile from intact serum 1H-NMR spectra; and the specific lipid species characterization, from the lipid serum extract 1H-NMR spectra. High-resolution 1H-NMR spectroscopy data were acquired on a Bruker 600 MHz Spectrometer: 1D Nuclear Overhauser Effect Spectroscopy (NOESY), Carr-Purcell-Meiboom-Gill (CPMG) was used to characterize small molecules such as amino acids and sugars; and LED Diffusion (Diff) experiments, to detect larger molecules such as lipoproteins, glycoproteins and choline compounds. All the sequences were run at 37°C in quantitative conditions. We obtained the lipid extract using a biphasic extraction with methanol/methyl-*tert*-butyl ether. For NMR measurements, the lipid extract was dried and reconstituted in 0.01% tetramethylsilane (TMS) solution (0.067 mM) and deuterated solvents. From intact serum, the Liposcale test (IVD-CE marked) was used to obtain the composition, the mean size and the number of lipoproteins particles of nine subtypes (large, medium and small) of the main lipoprotein types (VLDL, LDL and HDL). From the same NMR spectra, we obtained the general measurement of circulating glycoproteins by deconvoluting with analytical functions the specific region where glycoproteins resonate, to quantify the area, proportional to the concentration of the acetyl groups of N-acetylglucosamine and N-acetylgalactosamine (GlycA) and acetyl groups of N-acetylneuraminic acid (GlycB). Complementary, we applied a CPMG filter to profile the 1H-NMR spectra to obtain the concentration of a set of LMWM such as amino acids and sugars from the same serum sample. Finally, we obtained the concentration of the major lipid classes (fatty acids, glycerolipids, phospholipids and sterols) and some individual species from the 1H-NMR spectra of the previously extracted serum sample using the BUME protocol. Lipid quantification, based on the Lipspin software, relied on lineshape fitting analysis of spectral regions. Since the spectral area is equivalent to the molecular abundance, individual signal areas were converted into molar concentrations by normalizing by external

measurements. The lipid species obtained by this NMR approach included: cholesterol (free and esterified), unsaturated fatty acids (omega-6, omega-7, omega-9, omega-3), saturated fatty acids, monounsaturated fatty acids, linoleic acid, docosahexaenoic acid, arachidonic acid and eicosapentaenoic acid; glycerides and phospholipids (total cholines, triglycerides, phosphoglycerides, phosphatidylcholine, sphingomyeline and lysophosphatidylcholine).

The concentration of metabolites was compared with the population values (median, 25–75 percentiles) using the same methodological approach and NMR equipment. Specifically, the lipoprotein and glycoprotein profiles were compared with the general population values from a total of 6,022 individuals across two Spanish cohorts: the Di@bet.es Study and the Mollerussa Study. These are population-based Spanish cohorts comprising individuals aged 18 years and older, with a composition of 55% women. Similarly, the concentration of the aqueous metabolome and lipid families were compared with the general population values from a total of a subset of 1,965 individuals from the Di@bet.es study (57% women).

Proteomics

100 μL plasma of each sample was pre-cleared by centrifugation at 3000*g* for 10 min. For ECVs extraction, 1M ammonium acetate was added to precipitate ECVs on ice for 45 min. Then, 100 mM ammonium acetate was added to the mixture, and ECVs were precipitated by centrifugation at 20,000*g* for 30 min. ECVs were washed with 50 mM ammonium bicarbonate (Sigma-Aldrich, Cat. 1066-33-7). Then, 600 μL of 1% ammonium deoxycholate (Sigma Aldrich, Cat. K2755-1MG) were added. The concentration of protein in each sample was measured using a bicinchoninic acid assay (BCA assay). Ammonium bicarbonate was used to dilute 500 μg of protein into a final volume of 500 μL. Next, dithiothreitol (Sigma-Aldrich, Cat. D9779) was added to obtain a final concentration of 20 mM, followed by iodoacetamide (Sigma-Aldrich, Cat. I6125) to a final concentration of 40 mM. Next, trypsin (Roche, Cat. RTRYP-RO) was added to the sample in a 1:25 protein ratio and incubated at 37°C overnight. Next day, formic acid (FA) (ThermoFisher Scientific, Cat. 28905) in a final concentration of 0.1% was added and extraction of proteins was done using Empore Solid Phase Extraction Cartridges (3M), following manufacturer's instructions. The eluted samples were then centrifuged for 90 min using a speed vacuum centrifuge (Thermo, RC1010), followed by snap freezing in liquid nitrogen. Then, the samples were kept in a freeze dryer (LyoDry Compact Benchtop, MechaTech) overnight. Next, the samples were reconstituted in 30 μL of 0.1% FA and an *o*-Phthaladehyde (Oparil) assay was performed to determine the concentration of each sample. After that, the sample was prepared in a concentration of 0.5 μg/μL using 0.1% FA and alcohol dehydrogenase. The samples were prepared in glass mass spectrometry vials for proteomic analysis using a Waters Synapt G2Si High-Definition Mass Spectrometry (Waters Corporation) operated by the MassLynx 4.1., 110 min running time with 2 μL of an injection containing 1 μg of peptide. Quality controls were also run along with samples to guarantee consistency. Pooled quality controls were made from all samples, in which the samples were run at the beginning, middle and end of the mass spectrometry run. Samples were randomized before running the experiment. The proteomic data were then imported into Progenesis software 4.2 (Nonlinear Dynamic, UK) to identify and quantify peptides and proteins.

Data were analyzed in R Studio. The intensity data protein abundances were normalized using Trimmed Mean of M-values (TMM) normalization using the NOISeq package. Technical replicates effects were corrected by using ARSyNseq, and log2 transformation was applied to stabilize variance. OPLS-DA was performed using ropls package to identify group separations. For differential expression analysis, the limma package was used. Statistical significance was assessed using empirical Bayes moderation (eBayes), with differentially expressed proteins identified based on an adjusted p-value <0.05 and $|\log FC| > 1$.

To identify BP associated with differentially expressed proteins, GO enrichment analysis was performed using clusterProfiler. The enrichGO function was applied selecting the BPs GO terms with a p-value and q-value cutoff of 0.05. To classify GO terms into general BP clusters, Wang's semantic similarity method was used with the GOSemSim package to calculate term similarities. A heatmap was created to visualize clustering, which was further redefined into eight functional clusters. A Circos plot was used to visualize the relationships between differentially expressed genes and the GO clusters, using the circlize package.

Transcriptomics and proteomics integration

To integrate both transcriptomic and proteomic data, scRNA-seq pseudobulk data and ECV proteomics data were used. First, overlapping genes between the transcriptomic and proteomic datasets were identified. Genes and proteins were considered significant if they exhibited an absolute logFC $\geq$0.5 and showed consistent regulation (i.e., both upregulated or both downregulated) in both analyses. The selected genes and proteins were then used to generate a heatmap.

Transcription factor regulatory network analysis

TF activity was inferred from scRNA-seq pseudobulk differential expression analyses comparing M116 to non-supercentenarians and M116 to supercentenarians, using the decoupleR R package. Briefly, TF-target gene interactions were obtained from the DoRothEA database, and TF activity was estimated using a univariate linear model applied to TFs and their target genes present among the differentially expressed genes. TFs were considered differentially activated if they had a p-value $\leq$0.01. To construct the TF regulatory networks, differentially activated TFs and differentially expressed target genes were mapped and visualized using the igraph and ggraph R packages.

16S meta-genomics sequencing

Stool was collected from 3 different days (3 biological replicates) and kept at −20°C until DNA extraction (Biobanc IDIBGI and Goodgut). Total genomic DNA was extracted from 150 to 200 mg of each stool sample after homogenization using the DNeasy Powersoil Pro kit (Qiagen, Cat. 47014) according to manufacturer's instructions. Quality and quantity of DNA were evaluated using Qubit BR kit on a Qubit2.0 fluorimeter (ThermoFisher Scientific) and on a Nanodrop ND-2000 UV-Vis spectrophotometer (ThermoFisher Scientific). The v3-v4 region of the bacterial 16S rRNA gene was amplified and sequenced (paired-end 250-bp) following standard practices at external facilities (Novogene) using primers 515F (5′-GTG CCA GCM GCC GCG GTA A-3′) and 806R (5′-GGA CTA CHV GGG TWT CTA AT-3′).

Obtained reads were processed with the dada2 pipeline and phyloseq R package. Default settings were used for filtering and trimming reads. Built-in training models were used to learn error rates for the amplicon dataset. Identical sequencing reads were combined through dada2's dereplication functionality, and the dada2 sequence-variant inference algorithm was applied to each dataset. Subsequently, paired-end reads were merged, a sequence table was constructed, taxonomy was assigned, and abundance was calculated at all possible taxonomic levels using the Silva's 138 dada2-formatted database (www.arb-silva.de). Diversity indices (i.e., Chao1 and Shannon) together with beta-diversity matrices (e.g., weighted unifrac and principal coordinate analyses (PCoA)) were computed using phyloseq, vegan, ape, MicrobiotaProcess, UpsetR and mia R packages.

As independent validation, we considered the publicly available datasets collected and curated in the curatedMetagenomicData R package. Of the available datasets, we selected those comprising samples matching the following criteria: (i) gut samples collected from healthy individuals (''disease = = healthy''); (ii) samples with available age value and classified as senior (''age_category = = senior''), which comprised between 65 and 91 years old; and finally (iii) those samples without antibiotic use on the day of sampling (available information; antibiotics_current_use ! = "yes"). A total of 445 samples were analyzed from the curatedMetagenomicData dataset. To better characterize microbiome variation between the M116 sample and control cohort, we generated a PCoA plot with the taxonomic relative abundances with 90% confidence ellipse. This approach was selected to optimize trend visualization given the exploratory nature of the analysis, the large reference cohort providing stable covariance estimation, and the need to clearly visualize the position of the supercentenarian sample relative to population variation.

Infinium MethylationEPIC BeadChip

Genome-wide DNA methylation profiling was performed at the Genomics Unit of the Josep Carreras Leukemia Research Institute. Briefly, DNA samples were quantified with Qubit BR kit on a Qubit2.0 fluorimeter (ThermoFisher Scientific), and integrity was checked using agarose gel electrophoresis. DNA from samples passing the QC was bisulfite-converted using the EZ DNA Methylation-Gold Kit (Zymo Research, Cat. D5005) following the manufacturer's instructions. Then, bisulfite-converted DNA was hybridized into the Methylation EPICv1.0 BeadChip (Illumina) array interrogating >850,000 CpG sites according to manufacturer's instructions. Fluorescent signal was detected by the Illumina iScan confocal laser scanner with the Autoloader system. Illumina methylation raw data (idat files) were preprocessed using the R environment (version 4.3.2) with the minfi package (version 1.48.0). Raw signal intensities were normalized applying a background correction method with dye-bias normalization using ssNoob method. Then, probes with a detection p-value >0.01, cross-reacting probes and probes that overlapped SNVs and sex-chromosomal probes were removed as QC steps. The DNA methylation levels were represented as β-values ranging from 0 to 1 (corresponding to 0–100% methylation). Supervised methylation analysis was performed comparing the difference between M116 saliva samples to the mean methylation of the control cohort hybridized and pre-processed as mentioned above. Differentially methylated CpGs were considered if $\Delta\beta > |0.50|$ as stated in the manuscript. Differentially methylated CpG probes were used to perform hierarchical clustering analysis with canberra distance and ward.D2 clustering method. The same CpG probes were used for the t-SNE analysis. To correlate methylation and expression values, the common CpGs differentially methylated located within a gene were selected in a panel of human cell lines[50] and correlated with the expression of the same gene. We considered a significant correlation between methylation and expression if p-value <0.05 and Spearman Rho > |0.3|. To determine the methylation status in repetitive elements, the REMP R package was used. The methylation values in repetitive elements were predicted for ALU, LINE-1 and ERV applying the default parameters, using a random forest model for the prediction with 1,000 bp windows within the genome. The median methylation level of each repetitive element for individual was represented in a boxplot. Finally, epigenetic clocks, with the exception of the ribosomal DNA (rDNA) methylation clock, were calculated using the methylclock package in R following the recommended pipeline.

Whole Genome Bisulfite Sequencing

DNA fragmentation (350 bp), library preparation, bisulfite conversion, and paired-end-150 WGBS was conducted by Novogene. Sequencing was performed as for WGS.

rDNA clock assessment

The reference population consisted of DNA from human blood obtained from Reprocell (reprocell.com) with no known history of disease. We used 62 samples (28 males and 34 females) aged between 18 and 88 years old. Data were collected with both RRBS and WGBS. After evaluating sequencing quality using FastQC, we used Trim_Galore! to trim the 3′ adaptors as well as low quality bases (BAQ <20). The '–rrbs' option was used for RRBS reads to remove the filled-in bases. We then used Bismark, which invokes bowtie2 v2.3.1 to map the bisulfite sequencing reads onto the genome. The rDNA clock has previously been fully described in Wang et al.,[56]

including CpGs and coefficients. The rRNAm clock consists of 72 CpGs selected from RRBS mouse datasets. Initially, 816 CpGs within the rDNA region (with read depth $\geq$50 in all samples) were retained. The dataset was randomly split 10,000 times into training and testing subsets, and an elastic-net regression model was applied to each to select predictive CpGs and estimate site-specific weights. In addition, the model exerts extra constraint on the coefficients of predict variables by adding penalty to the coefficients using the combination of lasso and ridge regulation methods, with re-adjustment of coefficients to fit the reference population. Additionally, model performance was evaluated using median absolute error (MAE) and Spearman's correlation (ρ) between predicted and chronological age. From 736 CpGs selected across all iterations, the most stable CpGs (top 200 appearing in 95.3% of models) were prioritized. The best-performing model (lowest MAE = 1.63 months, highest ρ = 0.98) comprised 72 CpGs distributed across the rDNA region. These sites were later found to be evolutionarily conserved in humans. Confidence interval (95%) for age estimates of M116 was calculated by generating 30 subsamples with random sets of reads. All sites had a minimum read depth >25. M116 age pace was calculated as the vertical distance between M116 predicted age by the rDNA methylation clock and the expected age derived from the regression model fitted with the control cohort (intercept). Positive and negatives value indicates accelerated or decelerated aging, respectively.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were carried out using R software (v4.4.1) or GraphPad Prism (v10.0). Normality was assessed for continuous variables. For comparisons between two groups, a t-test a t-test was used if the data were normally distributed, and a Mann-Whitney U test was used if they were not. All statistical tests were two-sided unless otherwise stated. Bar plots display the mean value of the different measures within a condition and whiskers represent the standard error of the mean (SEM). Sample sizes (n) are indicated in the figures, results or methods details. Asterisks indicate the level of statistical significance: $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***).