



A simple blood biomarker based on gene expression describes cardiovascular health-related biological age

Natalia Hernández-Bellido · Adrián Hernández-Vicente · Laura García-Mendivil ·
Estel Ramos-Marquès · David Hernando · Alberto Cebollada · Ralf Köhler · Nuria Garatachea ·
Esther Pueyo · Laura Ordovás

Received: 11 February 2025 / Accepted: 26 June 2025
© The Author(s) 2025

Abstract The need to monitor the aging process as a risk factor for disease and mortality beyond chronological age (CA) has led to numerous investigations into the estimation of the biological age (BA) of individuals. However, the accuracy of BA estimation tools is often judged by their ability to approximate CA, questioning their value in capturing the variance in health status and thus correctly estimating BA. Their biological relevance is often assessed in relation to health outcomes or mortality, underexploiting their potential for real-time monitoring of BA.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11357-025-01784-6>.

N. Hernández-Bellido (✉) · L. García-Mendivil ·
E. Ramos-Marquès · D. Hernando · E. Pueyo ·
L. Ordovás (✉)
Biomedical Signal Interpretation and Computational
Simulation (BSICoS) Group, Instituto de Investigación
en Ingeniería de Aragón (I3A), Universidad de Zaragoza,
Zaragoza, Spain
e-mail: nhbellido@unizar.es

L. Ordovás
e-mail: lordovas@unizar.es

N. Hernández-Bellido · L. García-Mendivil ·
E. Ramos-Marquès · D. Hernando · E. Pueyo · L. Ordovás
BSICoS Group, Instituto de Investigación Sanitaria (IIS)
Aragón, Zaragoza, Spain

A. Hernández-Vicente · N. Garatachea
Growth, Exercise, Nutrition and Development (GENUD)
Group, University of Zaragoza, Zaragoza, Spain

Furthermore, their complexity may limit their clinical translation to large populations. Here, we describe the gene expression-based age monitoring Clock (GamC), a simple biomarker of aging (BOA), and characterize its biological relevance with synchronous cardiovascular (CV) health-related functional data. GamC is calculated from the expression levels of three genes consistently dysregulated with age in blood (*ABLIM1*, *CCR7*, and *LEF1*). GamC shows moderate but reliable association with CA in three independent cohorts, supported by transcriptome-wide changes. It demonstrates specialized biological meaning, as it specifically describes current physical activity levels, but poorly correlates with autonomic

A. Hernández-Vicente · N. Garatachea
Department of Physiatry and Nursing, Faculty of Health
and Sport Sciences (FCSD), University of Zaragoza,
Huesca, Spain

D. Hernando · E. Pueyo · L. Ordovás
Centro de Investigación Biomédica en Red de
Bioingeniería, Biomateriales y Nanomedicina (CIBER-
BBN), Zaragoza, Spain

A. Cebollada · R. Köhler
Instituto Aragonés de Ciencias de La Salud (IACS),
Zaragoza, Spain

R. Köhler · L. Ordovás
Fundación Agencia Aragonesa Para La Investigación y El
Desarrollo (ARAID), Zaragoza, Spain

nervous system function, both age-related factors associated with CV health. Finally, it expresses BA monitoring capacity by modestly responding to an effective exercise-based intervention in centenarians. In conclusion, GamC is proposed as a simple and affordable candidate BOA for reporting the individuals' current CV health-related BA, thus promoting its broad translation and application into measures aimed at promoting healthy aging in relation to CV health, the leading cause of death worldwide.

Keywords Biological age · Biomarker of aging · Physical activity · Autonomic nervous system modulation · Exercise intervention

Introduction

Aging, a major risk factor for disease, is an organismic process resulting from the progressive and cumulative decline in bodily function over time, which is built up from the molecular level. The molecular, cellular, and physiological changes that occur with aging are not uniform within and between individuals or populations. Therefore, aging is a much more complex biological process than just chronological age (CA). Biological age (BA), the age of the individual defined by the extent of age-dependent biological changes [1], captures the effect of both genetic and environmental factors acting over time. This effect is expressed in the individuals' aging rate (AR), the difference between BA and CA. Fast and slow agers with early and late functional decline, respectively, translate into different disease susceptibilities and mortality risks [2]. Promoting healthy aging aims to achieve slow aging rates. Therefore, determining BA is crucial to gain insight into the role of specific mechanisms in the aging process, to predict the age-related disease/mortality risk of individuals and, given its environmentally modifiable nature [3, 4], to guide personalized preventive measures and potential

therapeutic interventions that promote healthy aging, but, more importantly, to objectively evaluate their efficacy. This has led to a growing interest in the field of geroscience in describing biomarkers of aging (BOA) to estimate BA.

Many age prediction methods, clocks and predictors, have been reported to explore the association between age and age-related biological markers (reviewed extensively in [5]). However, the terms clock and predictor are used ambiguously and/or include mixed concepts that are difficult to separate. For example, their accuracy is often judged by their ability to approximate CA, being actually CA predictors, or their biological relevance is usually determined by their ability to predict mortality or disease, making them more appropriate as death/disease timers rather than current BA indicators [6]. In other cases, some methods even include CA as an age-related marker [2, 7, 8], in which case they are not strictly speaking aging clocks. Ideally, an aging clock should be an age-related index based on biological data, a BOA, related to an individual's current age-related health status, their BA. The aging clock can also be a predictive biomarker of disease/mortality risk and, ideally, a response biomarker able to monitor changes in BA induced by healthy aging interventions. Instead, an age predictor attempts to accurately estimate CA using biological data, somehow losing its ability to capture the variance in health status of individuals and being more useful for other purposes such as forensics.

Different types of biological data have been used to estimate BA (reviewed in [1, 6]). Among these, BA indices based on gene and/or protein expression markers may be more sensitive to detect changes in the aging rates than, for example, clinical/physiological features, due to the rapidity of such changes and the sensitivity of the techniques used to quantify them [9]. Both protein and RNA-based indices have been successfully used to estimate BA [9–11]. Machine learning techniques and molecular omics data have also emerged as popular options for constructing age predictors and aging clocks, as they offer a wide range of applications even across different tissues, organs, and systems [6, 12]. However, they are highly complex and costly tools, which may limit their clinical translation. Furthermore, comparisons of BA estimators within the same [13–15] or independent study populations [1] have shown variability

N. Garatachea
Centro de Investigación Biomédica en Red de
Fisiopatología de La Obesidad y Nutrición (CIBER-OBN),
Madrid, Spain

N. Garatachea
Instituto Agroalimentario de Aragón (IA2), CITA-
Universidad de Zaragoza, Zaragoza, Spain

in the predictive power and biological information explained by each BA estimator, highlighting their complementary nature and the need for validation in independent populations. This is of concern because differences in age ranges, gender, ethnicity, sample size, and the data analysis strategy chosen can significantly affect the predictions and estimates of model performance [16, 17]. In any case, blood biomarkers offer significant advantages over current methods in terms of simplicity, cost-effectiveness, and scalability for clinical translation [18, 19].

Aging is one of the major risk factors for cardiovascular diseases (CVD), which are the leading cause of death worldwide. Thus, predicting cardiovascular (CV) health in relation to biological age could help to reduce the socioeconomic burden of CVD. Autonomic nervous system (ANS) regulation and physical activity (PA) levels are significant independent predictors of cardiovascular (CV) health and mortality [20, 21]. ANS regulation plays an important role during aging, leading to a gradual deterioration of body organs and functions, disrupting the homeostasis of the CV system and subsequently affecting the level of CV health of individuals [22]. Heart rate variability (HRV) is recognized as a robust method to assess the status of the ANS [23, 24] and is also an indicator of arrhythmic complications and a strong predictor of mortality and sudden death [25]. Indeed, there is a correlation of HRV markers with age, highlighting the intricate relationship between aging and the ANS [26]. In relation to PA, it is well known that PA declines with age, with older adults being the most physically inactive population group. PA levels are associated with the prevalence of risk factors that predispose individuals to the development of CV diseases [27] and with all-cause mortality [28]. Importantly, age modifies the relationship between PA and mortality, with its inverse association being stronger in the elderly [28]. In addition, Grässler et al. demonstrated the improvements in HRV markers, CV health, and CV risk factors in response to different levels of PA by using data from 26 studies [24], linking exercise and ANS modulation to CV health. Overall, the evidence suggests that indices reporting the ANS function and PA levels of individuals are relevant age-related phenotypes, so they are suitable for assessing the biological significance of novel BOAs in relation to CV health. Consequently, ANS- and/or PA-related BOAs can adequately monitor the current BA of individuals.

Here, we develop and characterize a novel and simple blood gene expression-based BOA, the Gene expression-based Age Monitoring Clock (GamC). GamC demonstrates an association with CA in three independent cohorts supported by whole transcriptome changes, thus confirming its age-related nature. Consistently, GamC associates with synchronous indicators of CV health of individuals (with differential sensitivity for variables that quantify the ANS activity and PA levels, which is indicative of the specific biological meaning of GamC) and shows a moderate sensitivity to an effective healthy aging intervention in centenarians, collectively demonstrating its biological relevance as an aging clock. Overall, this study proposes a simple and affordable BOA that gathers information about the current CV health status of individuals.

Materials and methods

Identification of consistently dysregulated age-related genes in blood

A literature search was conducted in PubMed to identify transcriptomic studies that have published lists of age-related differentially expressed genes in human blood. A cross-comparison of the lists resulted in the proposal of candidate genes for the construction of GamC (Fig. 1).

Donors and samples of the study cohort

The study cohort is formed by a total of 87 Caucasian individuals from the region of Aragón (Spain) (Tables 1 and S1). The study was approved by the ethical committee for clinical research of Aragón (ID of the approval: PI17/0409 and PI18/381), and donors signed a written informed consent prior to enrollment. The study was conducted by adhering to the Declaration of Helsinki and in compliance with the European Union General Data Protection Regulation (EU 2016/679).

Subjects self-reported their date of birth and current medical conditions and medications. They were divided into four age groups (Table 1). Young, adult, and elderly individuals were excluded from the study if they suffered from acute illness, heart disease (such as heart failure or atrial fibrillation), were taking

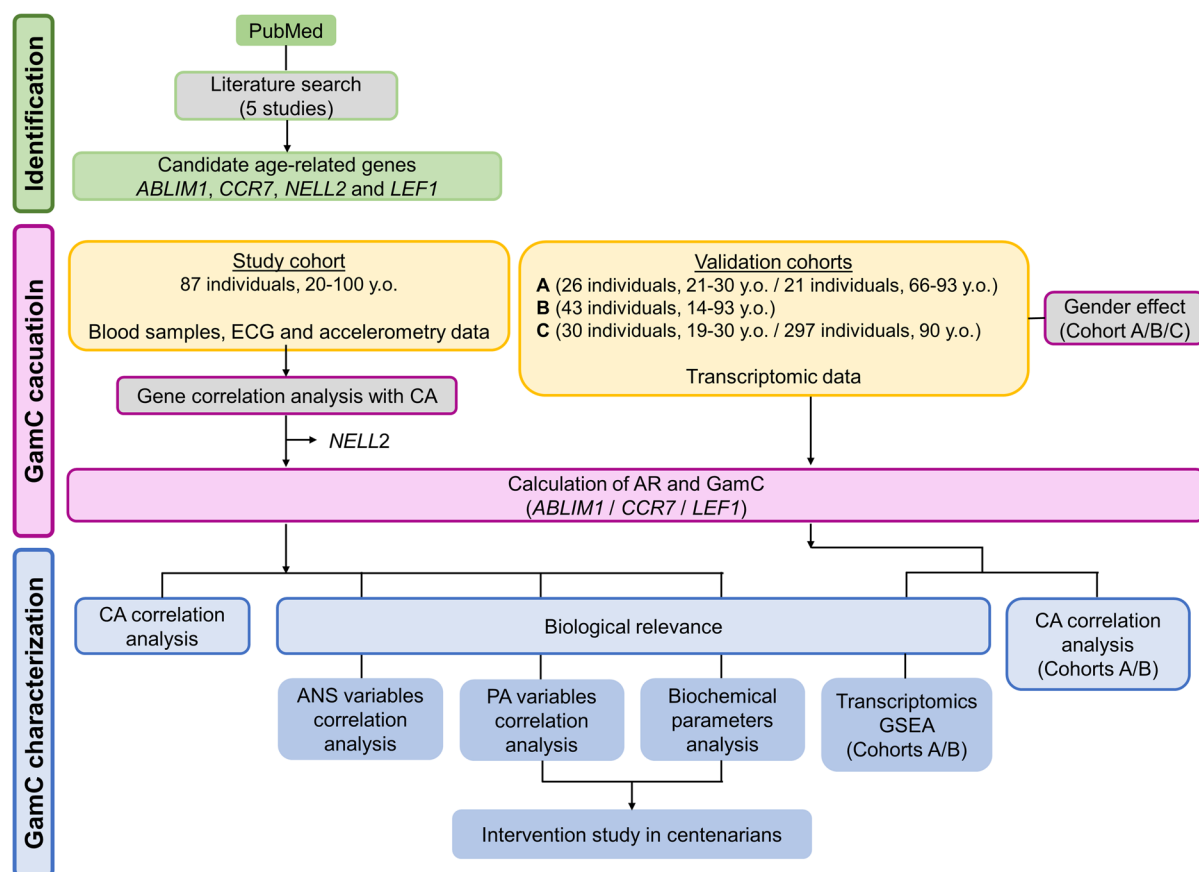


Fig. 1 Workflow for the description and characterization of GamC. The study cohorts used for GamC calculation, validation and characterization are described in the yellow boxes, indicating the number of individuals, age range, and collected

material or data. CA: chronological age, ECG: electrocardiogram, GEO: Gene Expression Omnibus; GSEA: Gene Set Enrichment Analysis, y.o.: years old

Table 1 Individuals in the study cohort: age range in years old (y.o.), total number of individuals, and percentage of males

Group	Age range (y.o.)	Number (% male)
Young	20–30	22 (100)
Adult	40–50	22 (100)
Elder	60–70	21 (100)
Centenarian	> 100	19 (26)

cardiac medication, or had any other clinical condition that contraindicated exercise. However, centenarians and donors who were overweight, sedentary, or had chronic diseases such as hypertension, diabetes, or hypercholesterolemia were included in the study, since they are representative of extreme longevity or have a

high prevalence in society, respectively, and thus contribute to the power of GamC. Further demographic, clinical, and biometric data are shown in Table S1.

Whole blood samples, electrocardiogram (ECG) recordings, and PA data were synchronously collected from each individual along with their CA.

Gene expression analysis by qPCR and generation of GamC

Whole blood samples were immediately processed to obtain peripheral blood mononuclear cells (PBMC) by using standard density gradient centrifugation procedures and stored at -80°C until use. RNA was extracted using the AllPrep DNA/RNA/miRNA Universal Kit (Quiagen). Next, 200 ng of RNA were

reverse transcribed using the PrimeScript™ RT Reagent Kit (Takara). Quantitative PCR (qPCR) was performed using the NZYSupreme qPCR Green Master Mix Kit (Nzytech) with gene-specific oligonucleotide pairs (Table S2) in a Viia7 instrument (Thermo Fisher Scientific). The expression level of each age-related candidate gene was calculated as $2^{-\Delta Ct}$, where $\Delta Ct = Ct_{\text{Gene}} - Ct_{\text{Reference}}$, being Ct_{Gene} the expression value of the gene and $Ct_{\text{Reference}}$ the mean expression value of two reference genes, namely, *UBE2D2* and *GUSB2*.

The GamC of each individual was calculated as the sum of the AR and CA [29]. The AR was calculated according to previously described methods [30]. Briefly, the candidate genes with significant Pearson correlation coefficient (R) between their expression level and CA were selected to construct GamC (Fig. 1). The AR of each individual (Δ_i) was obtained by linear regression of the expression levels ($2^{-\Delta Ct}$) of each gene G against CA. The regression analysis returned the residual expression level of gene G in each individual i ($\sigma_{G,i}$) and the regression factor associated with the gene G expression levels (a_G). The Δ_i is calculated as the integration of all (N) gene-specific $\Delta_{G,i}$ (which is in turn obtained as the ratio between $\sigma_{G,i}$ and a_G) (Fig. 1) (Eq. (1)):

$$AR(\Delta_i) = \frac{1}{N} \sum_{G=1}^N \Delta_{G,i} = \frac{1}{N} \sum_{G=1}^N \frac{\sigma_{G,i}}{a_G} \quad (1)$$

ECG data recording and HRV analysis

ECG recordings were obtained for all individuals in the study cohort (Fig. 1). A 12-lead high-resolution Holter ECG was acquired at rest with the 10 electrodes placed according to the manufacturer's instructions (H12+, Mortara Instrument, Milwaukee, WI, USA) and digitized at a sampling rate of 1000 Hz.

From the ECG recording, the heartbeats were detected and the RR interval time series were extracted using a multi-lead wavelet-based approach [31]. An operator manually verified each beat detection using a dedicated interface.

The 5-minute (min) RR interval series of each young, adult, and elderly subject was trimmed by removing 0.5 min from the beginning and the end, resulting in a final time series of 4 min. For the centenarian series, 4-min windows were defined within the

10-min recording in order to select the window with the lowest standard deviation (SD).

Linear and non-linear HRV indices were derived using algorithms developed and previously published in our research. These algorithms were implemented in MATLAB version R2021a (MATLAB, MathWorks Inc., Natick, MA, USA) [31–34]. The HRV markers extracted from the ECG are explained below:

Linear domain indices

Temporal domain indices The time series of normal RR intervals after correction for ectopic beats [32] were denoted as NN and used for subsequent analyses. Several temporal HRV indices were derived from the NN interval series including mean heart rate (MHR) calculated as the inverse of the mean NN intervals, standard deviation of NN intervals (SDNN), standard deviation of the differences between adjacent NN intervals (SDSD), root mean square of consecutive differences of adjacent NN intervals (RMSSD), and percentage of consecutive NN intervals differing by more than 50 ms (ms) divided by the total number of all NN intervals (pNN50). SDNN is considered a long-term variability measure of total HRV power, representing the variability of the ANS activity. RMSSD, SDSD, and pNN50 are short-term variability measures of parasympathetic nervous system (PNS) activity.

Frequency domain indices From the NN interval series, the instantaneous HR signal ($dHR(n)$) was derived using the integral pulse frequency modulation (IPFM) model, taking into account the presence of ectopic beats [32], and sampled at 4 Hz. This signal was high pass filtered (0.03 Hz) to remove the very low frequency components ($dMHR(n)$) and corrected as displayed in Eq. (2) [33]. This modulating signal ($m(n)$) carries information about the ANS activity.

$$m(n) = \frac{dHR(n) - dMHR(n)}{dMHR(n)} \quad (2)$$

To estimate the spectral properties of the HRV signal, the Welch period programme was applied to $m(n)$ using a Hamming window of 60 seconds (s) length with 30 s overlap. The power in the low frequency (LF) band, P_{LF} , and the power in the high frequency (HF) band, P_{HF} , were extracted by integrating the

power in their respective bands: 0.04–0.15 Hz and 0.15–0.4 Hz, respectively. Normalized P_{LF} with the total power, named as P_{LFn} , and the ratio P_{LF}/P_{HF} , as an indicator of the activity of both ANS branches, were also calculated.

Non-linear indices

Non-linear HRV indices, as non-stationary measures, were also derived from the RR interval series between the individual heartbeats. Detrended fluctuation analysis (DFA) was performed by extracting the correlations between successive RR intervals at different time scales in order to measure fluctuations at different scales to detect short- and long-term correlations (α_1 and α_2 , respectively) [35]. These markers are used to assess the complexity and adaptability of the ANS.

Finally, SD1 and SD2 were extracted from the Poincaré plot by plotting each RR interval against the previous interval to create a scatter plot. SD1 and SD2 measure the short-term and long-term beat-to-beat variability of the RR interval series, respectively. The ratio SD1/SD2 was also calculated, as the ratio between the two variabilities in the RR time series [35]. These measures are used to investigate the structure and complexity of the interbeat intervals.

Accelerometer data recording and analysis

Accelerometer data recordings were obtained for all individuals in the study cohort (Fig. 1). All individuals were asked to wear GENEActiv triaxial accelerometers (ActivInsights Ltd., Cambridgeshire, UK) for 24 h over seven consecutive days. These accelerometers were placed on the non-dominant wrist and programmed to record accelerations at 10 Hz, a frequency previously validated to classify daily activities (Zhang et al. 2012). Initialization of the GENEActiv accelerometers and retrieval of data in binary format were performed using GENEActiv PC (version 3.2) (ActivInsights Ltd., Cambridgeshire, UK).

The GGIR 3.0–2 package of the statistical programming language R v.4.3.2 [36] was used to perform accelerometer data analysis. Non-wear time detection and minimum valid time requirements for each accelerometer register were evaluated using GGIR default settings to facilitate comparability with previous studies. The minimum valid hours per day

were set to sixteen, while the minimum valid days per record were set to two. Table S3 shows the relationship between the nomenclature used in this article and the variable names from the GGIR output.

Accelerometer variables were calculated [37], including the MX metrics, the average acceleration (AA), and the intensity gradient (IG). The MX metrics [38] assess the acceleration of the most active X min of a participant's daily activity (e.g., M30 refers to the acceleration at which the most active 30 min were spent). These metrics can be used to describe the distribution of intensities over different time periods and allow for direct comparison with health-related PA guidelines. Intensity levels corresponding to the most active 1, 2, 5, 10, 15, 20, 30, 45, 60, 120, 240, 360, 480, 600, and 720 min were recorded. AA reflects the average acceleration throughout the entire measurement period and can be used as a proxy for total daily PA-related energy expenditure or PA volume [37, 39]. IG, calculated as the slope (negative) of the linear regression between the natural logarithm of time and acceleration intensity, captures the distribution of PA intensity.

Despite knowing the limitations of traditional cut-points [40, 41] due to their widespread use, the following indices were also calculated: Inactive time (IT: < 30 mg), time spent in light intensity PA (LPA: 30–100 mg), moderate intensity PA (MPA: > 100–400 mg), vigorous intensity PA (VPA: > 400 mg), and moderate-vigorous intensity PA (MVPA: > 100 mg) [42].

Exercise intervention

For the exercise intervention (Fig. 1), centenarians were randomly assigned to the control or intervention group using a computer-generated shuffle list. Participants in the control group maintained their usual PA. The intervention group performed twice per week, non-consecutive, resistance training sessions over 12 weeks (24 sessions in total). All the sessions were performed one-on-one in the gym of the geriatric nursing home and were supervised by an experienced (5 years) strength and conditioning trainer (MSc in Sports Science). After the initial assessment, participants were enrolled in a resistance exercise routine consisting of 8 different exercises according to their Functional Ambulation Classification (FAC) level [43].

Each session lasted 40–60 min, including a 10-min warm-up and 30–50 min of resistance training. The warm-up consisted of one set of eight single-joint seated exercises without resistance, 10 repetitions, and 30-s rest between exercises. Resistance training consisted of 1–3 sets of the 8 exercises routine adapted to FAC (8–10 repetitions as fast as possible, at 50–70% of the one-repetition maximum) with resting periods of 1 min between exercises and 3–5 min between sets [44]. Load, number of sets, and type of exercise routine were adjusted to the new level of physical capacity every 2 weeks.

Whole blood samples and accelerometer variables were obtained at the end of the 3-month intervention for both control and intervention groups as described above. Biochemical analysis for total cholesterol, low-density lipoproteins (LDL), triglycerides, creatine kinase (CK) and glucose levels were conducted by Centro Inmunológico de la Comunidad Valenciana (CIALAB). The accelerometer variables were obtained as described before. For each PA variable, biochemical parameter and GamC, the change in the value of each parameter after the intervention period in each individual (P_i) was calculated as the difference between after (II) and before (I) the intervention (ΔP_i). Then, the median and interquartile range (IQR) of ΔP_i for each PA variable, biochemical parameter and GamC were calculated for the control and intervention groups.

Validation cohorts and transcriptomic analysis

The literature search of point 2.1 and a search conducted in the GEO DataSets database, were filtered to extract studies containing readily accessible bulk tissue transcriptomic datasets from PBMC (as the study cohort) from young to elder individuals. Three independent transcriptomic datasets were obtained for analysis: cohort A (study from the USA, two age groups being 21–30 and > 70 y.o.) [45], cohort B (study from Spain, continuous ages from 19 to 93 y.o.) [46] and cohort C (study from Finland, two age groups being 19–30 and 90 y.o.) [47] (Fig. 1 and Table S4). Data were retrieved using the GEOquery package in R software, and gene annotation was applied to assign biological identifiers and filter out unannotated genes. Expression values were normalized and standardized by a quantile method for further analysis.

GamC was calculated as described above in cohorts A and B, but using the normalized and standardized expression values of the GamC-building genes. Cohort C was excluded from this calculation due to the lack of age variability in the older group.

Differential expression analysis (DEA) was performed in cohorts A and B using the Limma package in R software. Individuals were ranked by chronological age and GamC values, then divided into quartiles: young/first quartile (Q_1) and elderly/fourth quartile (Q_4) (Table S4). A linear model was fitted to the normalized and standardized expression data, including contrast matrices to define age-based comparisons. Log fold-change (logFC) values were estimated, and gene set enrichment analysis (GSEA) was performed using logFC values to identify enriched or depleted biological processes (GO terms) with a false discovery rate (FDR) ≤ 0.05 . The top 50 GO terms with the highest absolute normalized enrichment score (NES) were manually classified into broader functional categories.

Statistical analysis

Association analysis between two variables was performed using Pearson correlation analysis for the calculation of GamC [30], its association with CA, and the association of AR with CA. Spearman correlation analysis was used for the assessment of the association between independent HRV markers and PA variables with CA or GamC and also between HRV markers with PA variables.

Comparisons between unpaired observations in two independent groups were assessed using the Mann–Whitney U test.

As normality could not be assumed in the intervention study due to the small group sizes, the Wilcoxon matched-pairs signed rank test was performed in values of control and intervention groups independently to examine differences on ΔP_i of PA variables, biochemical parameters and GamC. The magnitude of the effect size ($|r|$) was calculated as described in Eq. (3), where Z represents the Z -score for the Wilcoxon matched-pairs signed rank test and n is the total number of observations [48].

$$|r| = \frac{Z}{\sqrt{n}} \quad (3)$$

The lrl was considered small if $lrl=0.1$, medium if $lrl=0.3$ and large if $lrl=0.5$ [49].

All statistical analyses were performed using GraphPad Prism 8.0. The significance level was set at $P \leq 0.05$ or, when adjusted for multiple-testing correction, at $FDR \leq 0.05$, as conveniently indicated in the text and figure legends. The numerical value of P or FDR of every statistical test can be found in the supplementary tables.

Results

GamC is associated with CA

The calculation of the BOA started by identifying genes whose expression was consistently and significantly dysregulated with CA across independent study cohorts in PBMC or whole blood (Fig. 1). We performed a literature search seeking bulk tissue transcriptomic studies from different geographic origins (in some cases of different ancestries), hypothesizing that such genes, when combined in the BOA, would be able to account for genetic and environmental differences, including cultural and lifestyle, in relation to CA. A total of five independent studies were identified, which included variable geographic origins and age ranges from young to centenarians [11, 47, 50–52] (Table S5). A manual cross-comparison of the published lists of age-related differentially expressed genes (DEG) revealed four consistently dysregulated

genes across all studies: *ABLIM1*, *CCR7*, *LEF1* and *NELL2*. These were selected as candidates for the construction of GamC.

Their potential to define a relevant BOA was evaluated in blood samples from a study cohort composed of 87 individuals with varying overall health and ages ranging from 20 to over 100 y.o. (Fig. 1 and Tables 1–S1). In agreement with the studies of the identification section of the workflow (Fig. 1 and Table S5) [11, 47, 50–52], *ABLIM1*, *CCR7*, and *LEF1* showed a significant negative correlation with CA in the study cohort ($R = -0.41$, -0.33 , and -0.55 , respectively and all $P \leq 0.05$), confirming their association with age (Fig. 2a and Table S1). *NELL2* instead was not significantly correlated with CA ($R = -0.18$, $P = 0.1$) (Fig. S1) and therefore, it was excluded from the following steps of the study.

Next, *ABLIM1*, *CCR7*, and *LEF1* were used to construct the AR and GamC for each individual (Fig. 1). GamC was positively and significantly correlated with CA ($R = 0.59$). The degree of association exceeded that observed for the individual genes (Fig. 2b), demonstrating an improved explanatory power of this gene combination to describe CA. The AR was not significantly correlated with CA (Figure S2a), consistent with its expected individual-dependent nature.

These results indicate that GamC is associated with CA and, consequently, that the molecular data used to construct it contributes to describing this association.

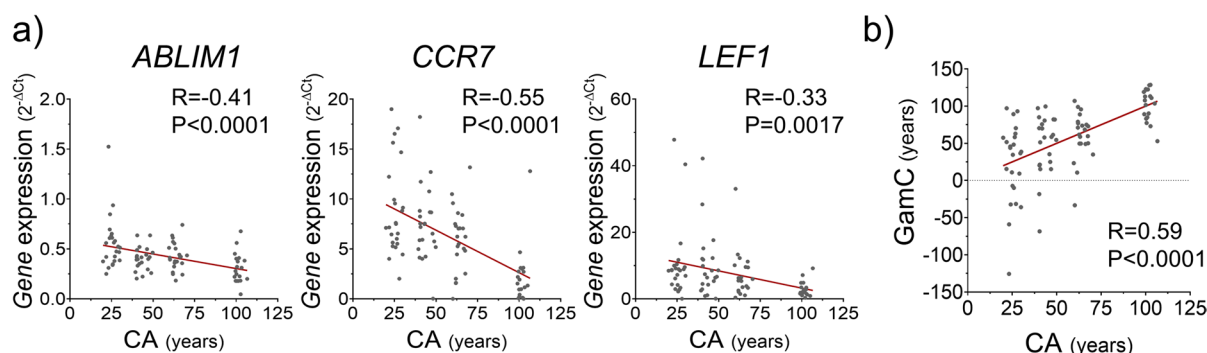


Fig. 2 Evaluation of the relation with age of the candidate genes and GamC. **a** Pearson's correlation analysis of the gene expression levels in PBMC (obtained by qPCR) of the age-related candidate genes and CA in the study cohort. **b** Pear-

son's correlation analysis of GamC with CA. Pearson's correlation coefficient (R), P -value (P), and linear regression (red line) are shown

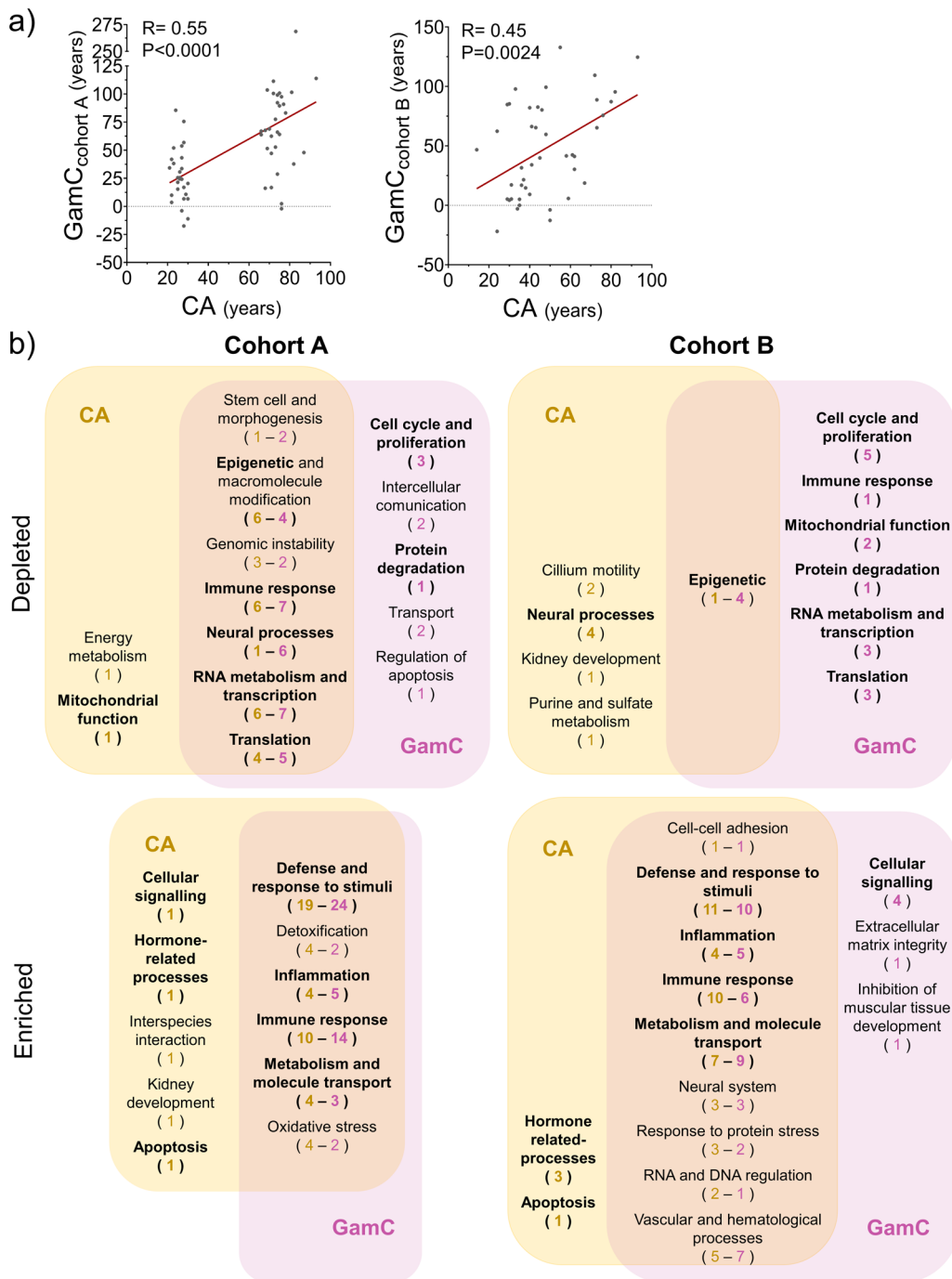


Fig. 3 Validation of GamC in independent cohorts and evaluation of its biological relevance at the molecular level. **a** Evaluation of the association between GamC (constructed from transcriptomic data) and CA in the validation cohorts A (left) and B (right). Pearson's correlation coefficient (R), P -value (P), and linear regression (red line) are shown. **b** Gene set enrichment analysis in the transcriptomic datasets of the validation

cohorts A and B. Enriched and depleted functional categories in the elderly classified by CA or GamC are shown (bold letters are used for categories common between CA and GamC and/or cohorts A and B). The number of gene ontologies belonging to each functional category is indicated in brackets and color coded (CA in yellow or GamC in pink)

GamC association with age is maintained across independent study cohorts and at the transcriptome-wide level

The capacity to describe the age of GamC was analyzed in independent validation cohorts (Fig. 1). As the study cohort was mainly composed of male donors (Table S1), first, the influence of gender on this relationship was studied (Fig. 1). Using bulk PBMC transcriptomic data from three different cohorts of variable geographic origin (cohorts A – 34 % of males-, B - 51 % of males and C - 29 % of males) (Table S4) [45–47], a general downregulation of *ABLM1*, *CCR7*, and *LEF1* with CA was observed in both genders in each dataset, but also globally. The association was significant in all cases except for *ABLM1* and *CCR7* in the segregated groups of cohort B, for which only a moderate non-significant correlation was observed (Figure S3 and Table S4). Thus, *ABLM1*, *CCR7*, and *LEF1* were generally associated with CA in a cohort- and gender-independent manner. Consequently, GamC was calculated without gender segregation for cohorts A and B (Fig. 1). Calculation of GamC for cohort C was not mathematically possible and was therefore excluded from further analysis.

Consistent with the results observed in the study cohort, GamC maintained a significant positive correlation with CA in the validation cohorts A and B with $R=0.55$ and 0.45 , respectively (Fig. 3a) that was independent of the AR (Figure S2b-c and Tables S4–S6), demonstrating its age-related relevance across populations.

Then, the biological significance of GamC was examined by analyzing its ability to capture age-related transcriptomic changes (Fig. 1). GSEA revealed that, in general, the degree of overlap of functional processes dysregulated with CA and GamC between cohorts was high. Only the depleted processes of cohort B seemed differential, but those of GamC mostly matched the depleted ones of cohort A, suggesting that GamC in cohort B may better describe the age-related changes in gene expression than CA itself (Fig. 3b and Table S7). Overall, both cohorts retrieved processes previously reported to be altered with aging (Table S4) [53] and/or found dysregulated in the studies of the identification section of the workflow (Table S5). Enriched processes included, for example, defense and response to

stimuli, immune response, inflammation or metabolism and molecule transport, whereas depleted processes related to mitochondrial function, RNA metabolism and transcription, translation and proteostasis or epigenetics. Therefore, GamC is able to capture the transcriptomic changes that occur with CA, confirming its age-related nature in different cohorts also at the molecular level.

Thus, the relationship between GamC and age is robust across independent cohorts of different geographic origin and is supported by a molecular basis. Of note, GamC can also be calculated from different sources of gene expression data, namely qPCR and transcriptomic.

GamC marginally describes age-related changes in the ANS function

Once the relationship between GamC and CA was established, we studied the biological relevance of GamC. For this, we observed the ability to predict the current CV health of individuals in terms of ANS function and PA levels (Fig. 1). Because of the established relationship between ANS and PA with age and, in turn, in defining CV health, we analyzed the general relationship between HRV and PA markers hypothesizing that those correlated would be stronger age-related functional indicators. Accelerometer and HRV markers were only partially significantly associated (116 out of 308 tested correlations had $FDR \leq 0.05$). In particular, some frequency domain (P_{LFn} , P_{LF}/P_{HF}) and non-linear domain indices (α_1 , α_2 or SD1 and SD2) had no relationship with virtually any of the PA markers (Figure S4 and Table S9).

Consistent with this observation and the established hypothesis, CA was only significantly associated with PA-related HRV markers, especially linear temporal and absolute frequency domain markers, including SDNN, SDSD, RMSSD, pNN50, P_{LF} , P_{HF} , and SD1/SD2 (R around -0.6 and $FDR \leq 0.05$) (Fig. 4 and Table S10–11). This suggests that CA can moderately explain HRV data in the study cohort, which are influenced by the individuals' PA levels. GamC largely replicated these findings, but with even weaker power than CA (R between -0.2 and -0.4) (Fig. 4 and Table S10–11), indicating that GamC marginally described the ANS function.

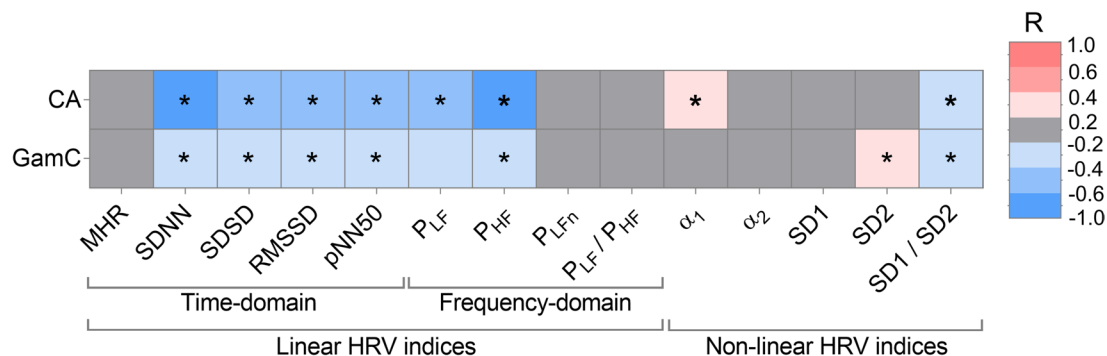


Fig. 4 Evaluation of the association of age with HRV parameters. Spearman's correlation coefficients (R) of HRV parameters with CA or GamC are colour coded according to the scale. Significant correlations ($FDR \leq 0.05$) are marked by an asterisk. The groups of HRV parameters are delimited below as linear (time domain or frequency domain) and non-linear (MHR: mean heart rate; SDNN: standard deviation of NN intervals; SDSD: standard deviation of the differences between adjacent NN intervals; RMSSD: root mean square of consecu-

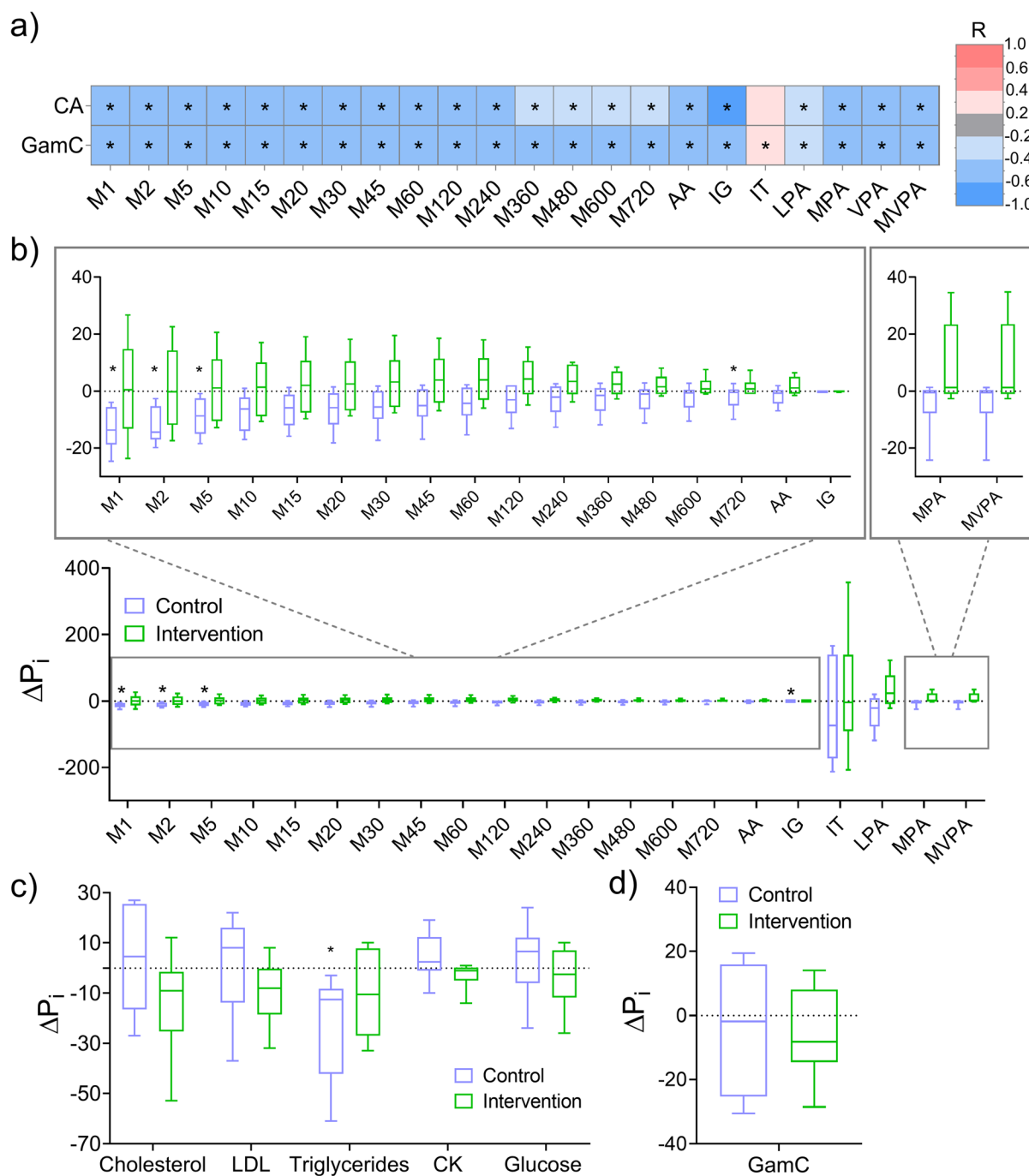
tive differences of adjacent NN intervals; pNN50: percentage of consecutive NN intervals differing by more than 50 ms divided by the total number of all NN intervals; P_{LF} : power in the low frequency (LF) band; P_{HF} : power in the high frequency (HF) band; P_{LFn} : normalized P_{LF} with the total power; P_{LF}/P_{HF} : ratio P_{LF}/P_{HF} ; α and α : short- and long-term correlations between successive RR intervals, respectively; SD1 and SD2: short-term and long-term beat-to-beat variability of the RR interval series, respectively; SD1/SD2: ratio SD1/SD2)

GamC is associated with PA levels and modestly responds to a healthy-aging exercise intervention in centenarians

The relationship of GamC with accelerometer variables, as a direct measure of PA levels and of CV health was then analyzed (Fig. 1). CA was significantly correlated with all PA markers ($FDR \leq 0.05$ and $R < -0.4$ for most markers) except for IT ($R=0.20$). GamC largely replicated this relationship, with an even stronger significant association with IT ($R=0.39$) (Fig. 5a and Table S12–13). CA and GamC showed comparable sensitivity with respect to short-term metrics (from M1 to M240). However, GamC was more sensitive than CA to long-duration MX metrics (from M360 to M720), demonstrating a better description of the acceleration at longer active periods. AA and IG, as standardized variables to describe the volume and intensity of PA, also showed their moderately negative and significant association with GamC ($R = -0.44$ and -0.52 , respectively). In terms of PA intensities, GamC showed a weaker correlation with LPA ($R = -0.33$) than with MPA, VPA, and MVPA ($R = -0.47$, -0.46 , and -0.46 , respectively), suggesting that GamC may reflect PA of higher intensities with greater precision. This fact is also observed in the associations with CA, which shows stronger associations with MPA, VPA, and

MVPA ($R = -0.49$, -0.51 , and -0.48 , respectively) than with LPA ($R = -0.29$). Therefore, GamC is able to describe individual PA levels in terms of MX metrics, AA, IG, and IT to the same extent as CA, but with improved sensitivity in some aspects over CA. In addition, the biological nature of GamC is better at describing activities of longer duration and higher intensity than those of shorter duration and lower intensity.

We then examined the potential of GamC to monitor the effect of a healthy aging intervention in centenarians (Fig. 1), since the significance of the findings in this age group is potentially higher than in younger one [54]. Individuals in the centenarian group were randomized to follow (intervention group) or not follow (control group) a 3 months of strength exercise training. The intervention was based on the fact that exercise, including strength training [55], promotes CV health and is associated with a lower risk of all-cause mortality [56], but also that the effects of aging are modifiable [4, 57]. The intervention showed a global beneficial effect in the trained group (Fig. 5b-c and Table S14-15). As expected for the age-related physical decline, the PA variables mostly worsened after the intervention in the control group with a general decrease of the median ΔP_i of PA variables (some of them being significant) with large average effect sizes ($|f| = 0.6$) (Fig. 5b and Table S13). However, the



intervention group seemed able to buffer and moderately counteract this decline after the exercise-training period. Namely, an opposite trend of the median ΔP_i of the PA variables was observed (general increased values) that was accompanied by medium average effect sizes ($|f| = 0.3$) (Fig. 5b and Table S15). In addition,

well established biochemical parameters associated with increased risk of CVD, including cholesterol, LDL, CK and glucose, also worsened after the intervention in the control group showing moderate average effect sizes ($|f| = 0.4$) (Fig. 5c and Table S15), but generally improved in the intervention group.

Fig. 5 Evaluation of the association of age with parameters of physical activity and healthy aging intervention study in centenarians. **a** Evaluation of the association between accelerometer-derived variables indicative of PA levels with CA and GamC by Spearman correlation analysis. Correlation coefficients (R) are colour coded according to the scale and significant independent associations ($FDR \leq 0.05$) are marked with an asterisk. **b–d** Effect of the healthy-aging intervention (strength exercise training for 3 months) in centenarians on the **b** PA levels, **c** biochemical parameters related to cardiovascular health and **d** GamC in each control and intervention group. Box plots and whiskers represent the median and 2.5–97.5 percentiles of the differences after (II)-before (I) intervention (ΔP_i). Significance results from Wilcoxon paired test ($P \leq 0.05$) for each control and intervention groups are marked with an asterisk. (M1 to M270: MX metrics from 1 to 720 min; AA: average acceleration; IG: intensity gradient; IT: inactive time; LPA: light intensity PA; MPA: moderate intensity PA; VPA: vigorous intensity PA; MVPA: moderate-vigorous intensity PA)

Decreased median ΔP_i values of cholesterol, LDL, CK and glucose with large average effect sizes ($|r|=0.5$) were observed in the intervened group (Fig. 5c and Table S15). Interestingly, the levels of triglycerides decreased in both control and intervention groups, which would need further investigation.

The improvement in the PA metrics and biochemical parameters in the intervention group was translated into an apparent decrease of the GamC ΔP_i median value (median (IQR) = -8.2 (12.8) years) with a medium effect size ($|r|=0.4$), that was greater and less dispersed than the differences observed in the control group (median (IQR) = -1.8 (31.3) years) with a small effect size ($|r|=0.2$) (Fig. 5c and Table S15).

Overall, GamC describes the current PA levels of individuals with some improvements over CA and shows a capacity to respond to an effective healthy aging intervention in centenarians, highlighting its biological relevance.

Discussion

In our aging societies, the need to promote healthy aging claims for objective biomarkers that monitor the aging process beyond the simple recognition of age as indicated by the date of birth. This need is underpinned by the individual AR, which this work confirms is consistently not associated with CA. Here we report GamC, a simple blood-based BOA that describes the current BA of individuals in terms

of CV health-related markers, specifically those that report the PA levels of individuals.

The age-related mechanisms observed in different populations may vary considerably [16, 17]. However, the association between GamC and CA remains consistent across study populations of different geographic origins. This is important because reliable BOAs with potential broad translational application are expected to demonstrate such consistency [58].

The observed association between GamC and CA is not as high as other established age predictors (e.g., PhenoAge and GrimAge [2, 8]). However, a partial correlation with CA is consistent with an important consideration in the development of BOAs. They should be strongly, but not perfectly, related to CA in order to represent the age-related physiological variability of each individual [1]. However, the cut-off point for this distinction is unknown. Our group previously reported AppAge, a cardiac transcriptomic BOA calculated as GamC, but using almost 3000 genes [29]. AppAge had a correlation coefficient value with CA similar to that of GamC in the cohorts used in this study, namely, $R=0.47$, but described age-related cardiac transcriptomic changes and an aging phenotype. Similar but less sensitive results were observed for a single gene (*CDKN2A*) based BOA ($R=0.24$ with CA), suggesting that even a single functionally relevant age-related gene can describe molecular changes of cardiac aging. In other words, a high correlation of BOAs with CA does not necessarily imply a better ability to explain the biology of aging. Conversely, a less-than-perfect association may actually leave room for a biologically based personalized description of an individual's aging status. In this regard, GamC, which has correlation coefficients with CA ranging from 0.45 to 0.59 in the studied cohorts, but which describes both changes in biological processes at the transcriptomic level and phenotypes related to the CV health of individuals (specifically PA levels), is calculated based on the expression levels of only three age-related genes (*ABLIM*, *CCR7*, and *LEF1*). These genes are consistently dysregulated in CA in all the independent cohorts considered in this work [11, 45–47, 50–52]. In addition, our data show that they contribute synergistically to the meaning of GamC, as their independent association with CA is improved by their combination in the BOA. GamC, in turn, demonstrates a differential nature, since it does not equally describe all age-related functional data of

the study cohort. It describes the PA levels even better than CA for some PA-related variables, but is only marginally sensitive for ANS functional markers. This can most likely be explained by the age-related gene-specific contribution of *ABLIM*, *CCR7*, and *LEF1* to the biological meaning of GamC. Whether GamC can describe other age-related functional data requires further investigation. But indeed, our results support the idea that based on the biological data used to construct BOAs, each BOA is expected to describe different and/or complementary aspects of aging [1]. Taken together, our data suggest that *ABLIM*, *CCR7*, and *LEF1* are functionally relevant age-related genes able to capture different aspects of an individual's aging biology in GamC, specifically PA-related, and, consequently, GamC is able to describe synchronous functional aspects of individuals regardless of (or perhaps exploiting) its partial correlation with CA. GamC is therefore able to estimate the CV-related BA of individuals in real time in terms of their PA levels.

Although the significance of the outcome of our intervention study in centenarians is limited by the small size of the study groups, it encourages further research into the ability of GamC to monitor BA in response to healthy aging interventions. This is supported by the non-significant decrease of the median difference of GamC in the intervention group after the intervention, which accompanies the overall beneficial effect of exercise observed in their PA-related variables and biochemical parameters. Therefore, GamC's sensitivity to exercise training seems remarkable considering that the size of the effect of the intervention on the PA levels is mostly medium too. Whether an intervention with a greater effect size (perhaps a longer training) would translate in greater changes in GamC in centenarians and whether the findings in centenarians hold true for other age groups remains to be explored.

For such purpose and beyond, the simplicity and consequent affordability of GamC (a blood sample and PCR test accessible to any basic molecular laboratory), together with the fact that its biological relevance is determined in relation to synchronous functional data rather than with mortality risk or disease outcome, makes GamC a valuable translational tool with potential application in real-time clinical decision-making and intervention monitoring. GamC is therefore a relevant candidate BOA for promoting healthy aging in relation to CV health, the leading cause of death worldwide.

Limitations of the study

We show that *ABLIM*, *CCR7*, and *LEF1* expression is related to age in a gender-independent manner; however, the low presence of females in the study cohort may limit our conclusions.

The small number of centenarians in the study limits the conclusions of the healthy-aging intervention, which was performed only in this age group.

GamC shows BOA potential in the study in relation to PA levels, but this needs to be tested in independent and larger populations.

Conclusion

We develop and characterize a simple and cost-effective BOA based on gene expression in blood, GamC, which meets capabilities of an aging clock. Specifically, it is able to describe CA on a biological (molecular) basis in independent cohorts; it associates with age-related functional aspects relevant to the current CV health status of individuals (PA levels) and shows moderate responsiveness to a healthy aging intervention. In conclusion, we propose GamC as a suitable candidate BOA for future research to determine the CV health risk in relation to BA in real time, as well as to monitor the effect of healthy aging interventions, which is beyond the capabilities of CA. Its simplicity enhances its clinical translational value to large populations.

Acknowledgements The authors would like to acknowledge the collaboration of Servicio General de Apoyo a la Investigación-SAI, Universidad de Zaragoza, and Servicios Científico-Técnicos (Secuenciación y Genómica Funcional y Biocomputación) of CIBA (IISA-IACS-Universidad de Zaragoza).

Author contribution N H-B and LO conceived the study with contribution from E R-M and EP. LO supervised the study with contribution from EP, NG and RK. N H-B and A H-V performed the experiments and analyzed the data. DH and L G-M contributed to the ECG signal processing. AC contributed to the bioinformatics analysis. N H-B wrote the manuscript with contributions from A H-V and LO, and all the authors carefully reviewed the data and the manuscript.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work was supported by grants PID2022-140556OB-I00 and PID2022-139859OB-I00 funded by MCIU/AEI/10.13039/501100011033 and ERDF, EU;

grant TED2021-130459B-I00 funded by MCIU/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR; grant G.A. 638284 (European Research Council); and grant PROY_B08_24 and BSICoS group T39_23R funded by Gobierno de Aragón. NH-B was supported by the predoctoral fellowships CPB_09_20 and MVB_03_23 funded by Gobierno de Aragón.

Data Availability Data availability statement is not included, as all data is included in the supplementary files.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Moqri M, Herzog C, Poganik JR, et al. Biomarkers of aging for the identification and evaluation of longevity interventions. *Cell*. 2023;186:3758. <https://doi.org/10.1016/J.CELL.2023.08.003>.
2. Liu Z, Kuo PL, Horvath S, et al. A new aging measure captures morbidity and mortality risk across diverse subpopulations from NHANES IV: a cohort study. *PLOS Med*. 2018;15:e1002718. <https://doi.org/10.1371/JOURNAL.PMED.1002718>.
3. Bland JS. Age as a modifiable risk factor for chronic disease. *Integr Med*. 2018;17:16–9.
4. Vitetta L, Anton B. Lifestyle and nutrition, caloric restriction, mitochondrial health and hormones: scientific interventions for anti-aging. *Clin Interv Aging*. 2007;2:537. <https://doi.org/10.2147/CIA.S866>.
5. Jia L, Zhang W, Chen X. Common methods of biological age estimation. *Clin Interv Aging*. 2017;12:759. <https://doi.org/10.2147/CIA.S134921>.
6. Galkin F, Mamoshina P, Aliper A, et al. Biohorology and biomarkers of aging: Current state-of-the-art, challenges and opportunities. *Aging Res Rev*. 2020;60:101050. <https://doi.org/10.1016/J.ARR.2020.101050>.
7. Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging* (Albany NY). 2018;10:573–91. <https://doi.org/10.18632/AGING.101414>.
8. Lu AT, Quach A, Wilson JG, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* (Albany NY). 2019;11:303–27. <https://doi.org/10.18632/AGING.101684>.
9. Oh HS, Rutledge J, Nachun D, Pálovics R, Abiose O, Moran-Losada P, Channappa D, Urey DY, Kim K, Sung YJ, Wang L. Organ aging signatures in the plasma proteome track health and disease. *Nature*. 2023;624(7990):164–72. <https://doi.org/10.1038/s41586-023-06802-1>.
10. Bortz J, Guariglia A, Klaric L, et al. (2023) Biological age estimation using circulating blood biomarkers. *Commun Biol*. 2023;6:1–10. <https://doi.org/10.1038/s42003-023-05456-z>.
11. Peters MJ, Joeannes R, Pilling LC, et al. The transcriptional landscape of age in human peripheral blood. *Nat Commun* 2015;6:. <https://doi.org/10.1038/ncomms9570>
12. Nie C, Li Y, Li R, et al. Distinct biological ages of organs and systems identified from a multi-omics study. *Cell Rep*. 2022;38:110459. <https://doi.org/10.1016/J.CELREP.2022.110459>.
13. Jansen R, Han LKM, Verhoeven JE, et al. An integrative study of five biological clocks in somatic and mental health. *Elife*. 2021;10:1–20. <https://doi.org/10.7554/ELIFE.59479>.
14. Macdonald-Dunlop E, Taba N, Klarić L, et al. A catalogue of omics biological aging clocks reveals substantial commonality and associations with disease risk. *Aging* (Albany NY) 2022;14:623–659. <https://doi.org/10.18632/AGING.203847>
15. Belsky DW, Moffitt TE, Cohen AA, et al. Eleven telomere, epigenetic clock, and biomarker-composite quantifications of biological aging: do they measure the same thing? *Am J Epidemiol*. 2018;187:1220–30. <https://doi.org/10.1093/AJE/KWX346>.
16. Cohen AA, Morissette-Thomas V, Ferrucci L, Fried LP. Deep biomarkers of aging are population-dependent. *Aging* (Albany NY) 2016;8:2253–2255. <https://doi.org/10.18632/AGING.101034>
17. Cohen AA, Legault V, Fuellen G, et al. The risks of biomarker-based epidemiology: associations of circulating calcium levels with age, mortality, and frailty vary substantially across populations. *Exp Gerontol*. 2018;107:11. <https://doi.org/10.1016/J.EXGER.2017.07.011>.
18. Vidal-Bralo L, Lopez-Golan Y, Gonzalez A. Simplified assay for epigenetic age estimation in whole blood of adults. *Front Genet* 2016;7:. <https://doi.org/10.3389/FGENE.2016.00126>
19. Wolinsky H. Testing time for telomeres. Telomere length can tell us something about disease susceptibility and aging, but are commercial tests ready for prime time? *EMBO Rep*. 2011;12:897–900. <https://doi.org/10.1038/EMBOR.2011.166>.
20. Thayer JF, Yamamoto SS, Brosschot JF. The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors. *Int J Cardiol*. 2010;141:122–31. <https://doi.org/10.1016/J.IJCARD.2009.09.543>.
21. Nocon M, Hiemann T, Müller-Riemenschneider F, et al. Association of physical activity with all-cause and cardiovascular mortality: a systematic review and meta-analysis.

- Eur J Cardiovasc Prev Rehabil. 2008;15:239–46. <https://doi.org/10.1097/HJR.0B013E3282F55E09>.
22. Oluwagbamigbe Fajemiroye J, Carlos da Cunha L, Saavedra-Rodríguez R, et al. Aging-induced biological changes and cardiovascular diseases. 2018 <https://doi.org/10.1155/2018/7156435>
 23. Russoniello CV, Zhirnov YN, Pougatchev VI, Gribkov EN. Heart rate variability and biological age: implications for health and gaming. 2013;16:302–308. <https://home.liebertpub.com/cyber> <https://doi.org/10.1089/CYBER.2013.1505>
 24. Grässler B, Thielmann B, Böckelmann I, Hökelmann A. Effects of different training interventions on heart rate variability and cardiovascular health and risk factors in young and middle-aged adults: a systematic review. *Front Physiol* 2021;12:. <https://doi.org/10.3389/fphys.2021.657274>
 25. Bonnemeier H, Wiegand UKH, Brandes A, et al. Circadian profile of cardiac autonomic nervous modulation in healthy subjects. *J Cardiovasc Electrophysiol*. 2003;14:791–9. <https://doi.org/10.1046/J.1540-8167.2003.03078.X>.
 26. Reardon M, Malik M. Changes in heart rate variability with age. *Pacing Clin Electrophysiol*. 1996;19:1863–6. <https://doi.org/10.1111/J.1540-8159.1996.TB03241.X>.
 27. Isath A, Koziol KJ, Martinez MW, et al. Exercise and cardiovascular health: A state-of-the-art review. *Prog Cardiovasc Dis*. 2023;79:44–52. <https://doi.org/10.1016/J.PCAD.2023.04.008>.
 28. Martinez-Gomez D, Luo M, Huang Y, et al. Physical activity and all-cause mortality by age in 4 multinational megacohorts. *JAMA Netw open*. 2024;7:e2446802. <https://doi.org/10.1001/JAMANETWORKOPEN.2024.46802>.
 29. Ramos-Marquès E, García-Mendivil L, Pérez-Zabalza M, et al. Chronological and biological aging of the human left ventricular myocardium: Analysis of microRNAs contribution. *Aging Cell*. 2021;20:e13383. <https://doi.org/10.1111/ACEL.13383>.
 30. Rhinn H, Abeliovich A. Differential aging analysis in human cerebral cortex identifies variants in TMEM106B and GRN that regulate aging phenotypes. *Cell Syst*. 2017;4:404–415.e5. <https://doi.org/10.1016/j.cels.2017.02.009>.
 31. Martínez JP, Almeida R, Olmos S, et al. A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Trans Biomed Eng*. 2004;51:570–81. <https://doi.org/10.1109/TBME.2003.821031>.
 32. Mateo J, Laguna P. Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. *IEEE Trans Biomed Eng*. 2003;50:334–43. <https://doi.org/10.1109/TBME.2003.808831>.
 33. Bailón R, Laouini G, Grao C, et al. The integral pulse frequency modulation model with time-varying threshold: application to heart rate variability analysis during exercise stress testing. *IEEE Trans Biomed Eng*. 2011;58:642–52. <https://doi.org/10.1109/TBME.2010.2095011>.
 34. Hernandez D, Bailón R, Almeida R, Hernández A. QRS detection optimization in stress test recordings using evolutionary algorithms. *Comput Cardiol* 2014;41:737–740. <https://doi.org/10.26754/JJI-I3A.201401652>
 35. Dimitriev DA, Saperova E V, Dimitriev AD. State anxiety and nonlinear dynamics of heart rate variability in students. 2016 <https://doi.org/10.1371/journal.pone.0146131>
 36. Migueles JH, Rowlands AV, Huber F, et al. GGIR: a research community-driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. *J Meas Phys Behav*. 2019;2:188–96. <https://doi.org/10.1123/JMPB.2018-0063>.
 37. Migueles JH, Aadland E, Andersen LB, et al. GRANADA consensus on analytical approaches to assess associations with accelerometer-determined physical behaviours (physical activity, sedentary behaviour and sleep) in epidemiological studies. *Br J Sports Med*. 2022;56:376–84. <https://doi.org/10.1136/BJSPORTS-2020-103604>.
 38. Rowlands AV, Dawkins NP, Maylor B, et al. Enhancing the value of accelerometer-assessed physical activity: meaningful visual comparisons of data-driven translational accelerometer metrics. *Sport Med - Open*. 2019;5:1–11. <https://doi.org/10.1186/S40798-019-0225-9/FIGURES/6>.
 39. Rowlands AV, Edwardson CL, Davies MJ, et al. Beyond cut points: accelerometer metrics that capture the physical activity profile. *Med Sci Sports Exerc*. 2018;50:1323–32. <https://doi.org/10.1249/MSS.0000000000001561>.
 40. Hernández-Vicente A, Marín-Puyalto J, Pueyo E, et al. (2022) Physical activity in centenarians beyond cut-point-based accelerometer metrics. *Int J Environ Res Public Heal*. 2022;19:11384. <https://doi.org/10.3390/IJERP191811384>.
 41. Rowlands AV, Sherar LB, Fairclough SJ, et al. A data-driven, meaningful, easy to interpret, standardised accelerometer outcome variable for global surveillance. *J Sci Med Sport*. 2019;22:1132–8. <https://doi.org/10.1016/J.JSAMS.2019.06.016>.
 42. Schwendinger F, Wagner J, Knaier R, et al. Accelerometer metrics: healthy adult reference values, associations with cardiorespiratory fitness, and clinical implications. *Med Sci Sports Exerc*. 2024;56:170–80. <https://doi.org/10.1249/MSS.0000000000003299>.
 43. Serra-Rexach JA, Bustamante-Ara N, Hierro Villarán M, et al. Short-term, light- to moderate-intensity exercise training improves leg muscle strength in the oldest old: a randomized controlled trial. *J Am Geriatr Soc*. 2011;59:594–602. <https://doi.org/10.1111/J.1532-5415.2011.03356.X>.
 44. Fragala MS, Cadore EL, Dorgo S, et al. Resistance training for older adults: position statement from the national strength and conditioning association. *J Strength Cond Res*. 2019;33:2019–52. <https://doi.org/10.1519/JSC.0000000000003230>.
 45. Thakar J, Mohanty S, Phillip West A, et al. Aging-dependent alterations in gene expression and a mitochondrial signature of responsiveness to human influenza vaccination. *Aging (Albany NY)*. 2015;7:. <https://doi.org/10.18632/aging.100720>
 46. Irizar H, Goñi J, Alzualde A, et al. Age gene expression and coexpression progressive signatures in peripheral blood leukocytes. *Exp Gerontol*. 2015;72:. <https://doi.org/10.1016/j.exger.2015.09.003>
 47. Marttila S, Jylhävä J, Nevalainen T, et al. Transcriptional analysis reveals gender-specific changes in the aging of the human immune system. *PLoS One* 2013;8:. <https://doi.org/10.1371/JOURNAL.PONE.0066229>

48. Tomczak M, Tomczak E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *TRENDS Sport Sci.* 2014;1:19–25. <https://doi.org/10.1186/s13054-016-1208-6>.
49. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. New York: Routledge; 2013.
50. Harries LW, Hernandez D, Henley W, et al. Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. *Aging Cell.* 2011;10:868–78. <https://doi.org/10.1111/J.1474-9726.2011.00726.X>.
51. Passtoors WM, Boer JM, Goeman JJ, et al. Transcriptional profiling of human familial longevity indicates a role for ASF1A and IL7R. *PLoS ONE.* 2012;7:e27759. <https://doi.org/10.1371/JOURNAL.PONE.0027759>.
52. Zubakov D, Liu F, Kokmeijer I, et al. Human age estimation from blood using mRNA, DNA methylation, DNA rearrangement, and telomere length. *Forensic Sci Int Genet.* 2016;24:33–43. <https://doi.org/10.1016/J.FSIGEN.2016.05.014>.
53. López-Otín C, Blasco MA, Partridge L, et al. Hallmarks of aging: an expanding universe. *Cell.* 2023;186:243–78. <https://doi.org/10.1016/J.CELL.2022.11.001>.
54. Chaudhari PS, Ermolaeva MA. Too old for healthy aging? Exploring age limits of longevity treatments. *Npj Metabolic Health and Disease* 2024 2:1. 2024;2(1):1–9. <https://doi.org/10.1038/s44324-024-00040-3>.
55. Shailendra P, Baldock KL, Li LSK, et al. Weight training and risk of all-cause, cardiovascular disease and cancer mortality among older adults. *Int J Epidemiol* 2024;53:. <https://doi.org/10.1093/IJE/DYAE074>
56. Pálovics R, Keller A, Schaum N, Tan W, Fehlmann T, Borja M, Kern F, Bonanno L, Calcuttawala K, Weber J, McGeever A. Molecular hallmarks of heterochronic parabiosis at single-cell resolution. *Nature.* 2022Mar 10;603(7900):309–14. <https://doi.org/10.1038/s41586-022-04461-2>.
57. Herzog CMS, Goeminne LJE, Poganik JR, et al. Challenges and recommendations for the translation of biomarkers of aging. *Nat Aging.* 2024;410(4):1372–83. <https://doi.org/10.1038/s43587-024-00683-3>.
58. Paluch AE, Boyer WR, Franklin BA, et al. Resistance exercise training in individuals with and without cardiovascular disease: 2023 update: A Scientific Statement From the American Heart Association. *Circulation.* 2024;149:E217–31. <https://doi.org/10.1161/CIR.0000000000001189>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.