Title

The biological age linked to oxidative stress modifies breast cancer aggressiveness

Authors

Sáez-Freire, María Del Mar

Blanco-Gómez, Adrián

Castillo-Lluva, Sonia

et al.

Publication Date

2018-05-01

DOI

10.1016/j.freeradbiomed.2018.03.012

Peer reviewed

Free Radical Biology and Medicine 120 (2018) 133–146

Contents lists available at ScienceDirect

Free Radical Biology and Medicine

journal homepage: www.elsevier.com/locate/freeradbiomed

Original article

The biological age linked to oxidative stress modifies breast cancer aggressiveness

María del Mar Sáez-Freire[a,b,c,1], Adrián Blanco-Gómez[a,b,1], Sonia Castillo-Lluva[a,b,d,2], Aurora Gómez-Vecino[a,b], Julie Milena Galvis-Jiménez[a,b,e], Carmen Martín-Seisdedos[b,f], María Isidoro-García[b,f], Lourdes Hontecillas-Prieto[a,b,3], María Begoña García-Cenador[b,g], Francisco Javier García-Criado[b,g], María Carmen Patino-Alonso[b,h], Purificación Galindo-Villardón[b,i], Jian-Hua Mao[i], Carlos Prieto[j], Andrés Castellanos-Martín[a,b,*,4,5], Lars Kaderali[k,**,5], Jesús Pérez-Losada[a,b,*,5]

[a] Instituto de Biología Molecular y Celular del Cáncer (IBMCC-CIC), Universidad de Salamanca/CSIC, Salamanca, Spain
[b] Instituto de Investigación Biosanitaria de Salamanca (IBSAL), Salamanca, Spain
[c] Departamento de Fisiología y Farmacología, Universidad de Salamanca, Salamanca, Spain
[d] Departamento de Bioquímica y Biología Molecular I, Facultad de Biología, Universidad Complutense de Madrid, Madrid, Spain
[e] Instituto Nacional de Cancerología, Bogotá, D.C., Colombia
[f] Servicio de Bioquímica Clínica, Hospital Universitario de Salamanca, Salamanca, Spain
[g] Departamento de Cirugía, Universidad de Salamanca, Salamanca, Spain
[h] Departamento de Estadística, Universidad de Salamanca, Spain
[i] Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[j] Bioinformatics Service, Nucleus, University of Salamanca (USAL), Salamanca, Spain
[k] Institute for Bioinformatics, University Medicine Greifswald, Greifswald, Germany

ARTICLE INFO

Keywords:
Aging
Breast cancer
Biological age
Mouse genetics
Oxidative stress
Subphenotypes

ABSTRACT

The incidence of breast cancer increases with age until menopause, and breast cancer is more aggressive in younger women. The existence of epidemiological links between breast cancer and aging indicates that both processes share some common mechanisms of development. Oxidative stress is associated with both cancer susceptibility and aging. Here we observed that ERBB2-positive breast cancer, which developed in genetically heterogeneous ERBB2-positive transgenic mice generated by a backcross, is more aggressive in chronologically younger than in older mice (differentiated by the median survival of the cohort that was 79 weeks), similar to what occurs in humans. In this cohort, we estimated the oxidative biological age using a mathematical model that integrated several subphenotypes directly or indirectly related to oxidative stress. The model selected the serum levels of HDL-cholesterol and magnesium and total AKT1 and glutathione concentrations in the liver. The grade of aging was calculated as the difference between the predicted biological age and the chronological age. This comparison permitted the identification of biologically younger and older mice compared with their chronological age. Interestingly, biologically older mice developed more aggressive breast cancer than the biologically younger mice. Genomic regions on chromosomes 2 and 15 linked to the grade of oxidative aging were identified. The levels of expression of *Zbp1* located on chromosome 2, a gene related to necroptosis and inflammation, positively correlated with the grade of aging and tumour aggressiveness. Moreover, the pattern of gene expression of genes linked to the inflammation and the response to infection pathways was enriched in the

* Corresponding authors at: Instituto de Biología Molecular y Celular del Cáncer (IBMCC-CIC), Universidad de Salamanca/CSIC, Salamanca, Spain
** Corresponding author. Institute for Bioinformatics, University Medicine Greifswald, Greifswald, Germany.
E-mail addresses: maria_del_mar@usal.es (M.d.M. Sáez-Freire), adrianblanco@usal.es (A. Blanco-Gómez), scastillolluva@usal.es (S. Castillo-Lluva), auroragvrm@gmail.com (A. Gómez-Vecino), jmgalvis@usal.es (J.M. Galvis-Jiménez), mcmartin@saludcastillayleon.es (C. Martín-Seisdedos), misidoro@usal.es (M. Isidoro-García), lhontecillas-ibis@us.es (L. Hontecillas-Prieto), mbgc@usal.es (M.B. García-Cenador), fjgc@usal.es (F.J. García-Criado), carpatino@usal.es (M.C. Patino-Alonso), pgalindo@usal.es (P. Galindo-Villardón), jhmao@lbl.gov (J.-H. Mao), cprietos@usal.es (C. Prieto), andres.castellanos@irbbarcelona.org (A. Castellanos-Martín), lars.kaderali@uni-greifswald.de (L. Kaderali), jperezlosada@usal.es (J. Pérez-Losada).
[1] Equal contribution as first authors.
[2] Current address: Departamento de Bioquímica y Biología Molecular I, Facultad de Biología, Universidad Complutense de Madrid, Madrid, Spain.
[3] Current address: Departamento de Patología Molecular, Instituto de Biomedicina de Sevilla (IBiS), Sevilla, Spain.
[4] Current address: Institute for Research in Biomedicine (IRB), Barcelona, Spain.
[5] Equal contribution as senior authors.

https://doi.org/10.1016/j.freeradbiomed.2018.03.012
Received 8 August 2017; Received in revised form 7 March 2018; Accepted 8 March 2018
Available online 14 March 2018

livers of biologically old mice. This study shows part of the complex interactions between breast cancer and aging.

1. Introduction

Breast cancer is a complex disease that results from the interaction of environmental factors with many genes with weak effects, which helps to explain the differences in cancer susceptibility and the evolution among individuals [1]. Although many factors contribute to an increased risk of breast cancer, the epidemiological factor most consistently associated with breast cancer susceptibility, and with most epithelial tumours, is aging. The incidence of breast cancer increases with age and doubles about every ten years until women reach menopause. After this occurs, the incidence rate dramatically decreases and may even be suppressed after the age of 80 [2,3]. Also, numerous epidemiologic studies have shown that younger women with breast tumours have a worse prognosis than older women, which is often related to more aggressive tumour biology. Therefore, early-onset tumours are more likely to be oestrogen receptor-negative, higher grade, and have increased lymphovascular invasion and molecular markers of aggressiveness [4]. Also, young age alone is per se an independent negative predictor of cancer-specific survival [5,6]. The reasons why the incidence of breast cancer increases with age and why younger women develop more aggressive tumours than older women are not known.

Cancer and aging are both multifactorial processes influenced by environmental and endogenous factors. Among them, some factors regulate intracellular functions such as proliferation, apoptosis and senescence. Other factors control extracellular components, such as the stroma, the immune system, and the endocrine system, all of which contribute to controlling processes such as angiogenesis, tissue growth, and tissue repair in normal tissues and also in tumours [7]. Thus, both cancer and aging susceptibility are genetically influenced by multiple genetic determinants, mostly low penetrance genes that control the behaviour of various intermediate phenotypes. Consequently, the different predisposition to both cancer and aging among individuals could be explained, at least in part, by polygenic inheritance [8,9]. Both the epidemiological and pathogenic relationships between epithelial cancer, including breast cancer, and aging suggest that they may share intermediate phenotypes and molecular and genetic determinants. One of the intermediate phenotypes associated with cancer and aging is oxidative stress [3,10,11]. Indeed, oxidative stress is associated with breast cancer susceptibility and aggressiveness [3] and also to the degree of aging [12]. Biological age estimates the functional status of an individual by comparing the functional status of other people of the same chronological age [12,13]. In 1969, Alex Comfort proposed that the biological changes related to age could be measured [14]. These changes were considered as aging biomarkers that could be integrated using mathematical models. Hence, investigators interpret the age predicted by these models as biological age [15,16]. Biological age is a complex term that must reflect the functional status of the individual and their physiological reserve [12–15]. It is very difficult to find the precise biomarkers that define the global biological age, but it is possible to identify a “partial biological age," which defines the functional status of a specific organ or specific physiological processes. Evidently, a "partial biological age" would better reflect the overall aging when studying the organs or systems that are more central in the overall aging of the organism, like the cardiovascular system. In relation to this, the physician Thomas Sydenham (1624–1689) previously proposed that "a man is as old as his arteries" [17]. In this study, we focused on oxidative stress to generate a “partial biological age” because of its importance, as shown in the literature, in the pathogenesis of the global aging of the organism, as well as in tumour susceptibility [3,10,11,18,19]. Thus, in a heterogeneous cohort of mice generated by a backcross, we observed that biologically young mice in terms of oxidative stress outnumbered the chronologically old mice that developed, or not, a less aggressive breast tumour.

2. Material and methods

2.1. Mice

In this study, a cohort of mice with different ERBB2-induced breast cancer susceptibility and evolution (N = 147) was generated. A backcross strategy was carried out between a mouse strain resistant to breast cancer (C57BL/6 genetic background) and a sensitive strain (FVB genetic background) overexpressing the *cNeu/ErbB2* proto-oncogene under the control of the Mouse Mammary Gland Tumour Virus (MMTV) promoter, *FVB/N-Tg(MMTVneu)202Mul/J*. The transgene was inserted on chromosome 3 (see Fig. 1**A** in Data in Brief [20]), and the mice developed a luminal-type breast tumour within a median of 7 months in the FVB genetic background [21]. The transgene-positive F1 mice were mated with FVB non-transgenic mice to generate the first generation of backcross mice [22]. Regarding the resultant genetic backgrounds, since the 147 mice were generated by a backcross, each had a variable mixture of the genetic backgrounds of B6 and FVB. The genotype of all the mice in the backcross population of the 244 SNPs that differed between the B6 and FVB strains is shown (see Fig. 1**B** in Data in Brief [20]). The mice were housed in the Animal Research Facility of the University of Salamanca and all practices carried out were previously approved by the Institutional Animal Care and Bioethics Committee. *FVB/N-Tg(MMTVneu)202Mul/J* mice were obtained from the Jackson laboratories and wild-type FVB/N and C57BL/6 mice were purchased from Charles River. All mice were maintained in ventilated filter cages under specific pathogen-free conditions and with free access to food and water. Body weight was determined when the mice were three, six and twelve months old, respectively; and the final weight was determined at the necropsy after subtracting the tumour weight. The number of mice available for determining each pathophenotype and intermediate phenotype is shown in (see **Table 1** in Data in Brief [20]).

2.2. Quantification of the liver signalling pathways associated with pro-oxidant activity

Some of the specific signalling molecules from pathways related to metabolism were measured in liver tissue [22] (see material and methods in Data in Brief [20]).

2.3. Quantification of serum metabolites and biochemical parameters determined at disease-free stage

Serum samples were collected when the mice were 3–4 months old, at a disease-free stage, and metabolic and biochemical parameters were determined with the purpose of identifying putative biomarkers. To do this, we quantified biochemical parameters routinely determined in clinical practice such as alpha1-antitrypsin, albumin, alkaline phosphatase, C3-complement, C4-complement, creatinine, ferritin, urea, IgG, total proteins, glucose, total cholesterol, cholinesterase, HDL, AST, ALT, LDH, Ca, Mg, P, Fe, and uric acid. All of these compounds were determined using a modular clinical chemistry analyser in the Department of Clinical Biochemistry at the University Hospital of Salamanca. Also, 139 metabolites were quantified by mass spectrometry [22].

2.4. Quantification of molecules of the antioxidant defence system and biomarkers of oxidative stress

The different components of the antioxidant defence system in the liver were quantified, including *catalase activity*, using Catalase fluorometric detection (ADI-907-027, Enzo Life Sciences) and *total glutathione levels* using the enzymatic recycling method [23]. To evaluate *lipid damage*, the level of 4-hydroxynonenal (4-HNE) in liver was determined using the *OxiSelect™ HNE Adduct ELISA (STA-338, Cell Biolabs)* kit. *Protein oxidation* was determined by measuring the levels of carbonyl groups in liver using the *Protein Carbonyl ELISA kit (#ALX-850-312-KI01, Enzo, LifeScience)*. To quantify *DNA damage* by oxidative stress, the 8-hydroxy-2-deoxy Guanosine (8-OH-dG) levels in liver were determined using a specific EIA kit (StressMarq, SKT-120).

3. Data and expression analysis

Multivariate models were generated using the oxidative stress phenotypes and the genetic information to predict breast cancer susceptibility and evolution. Regarding the *model of prediction of tumour latency*, the Cox proportional hazards model was used to identify independent prognostic factors in our cohort of mice. A prognostic index was constructed with the variables that predict tumour latency. To generate the coefficients of the prognostic index, *B coefficients*

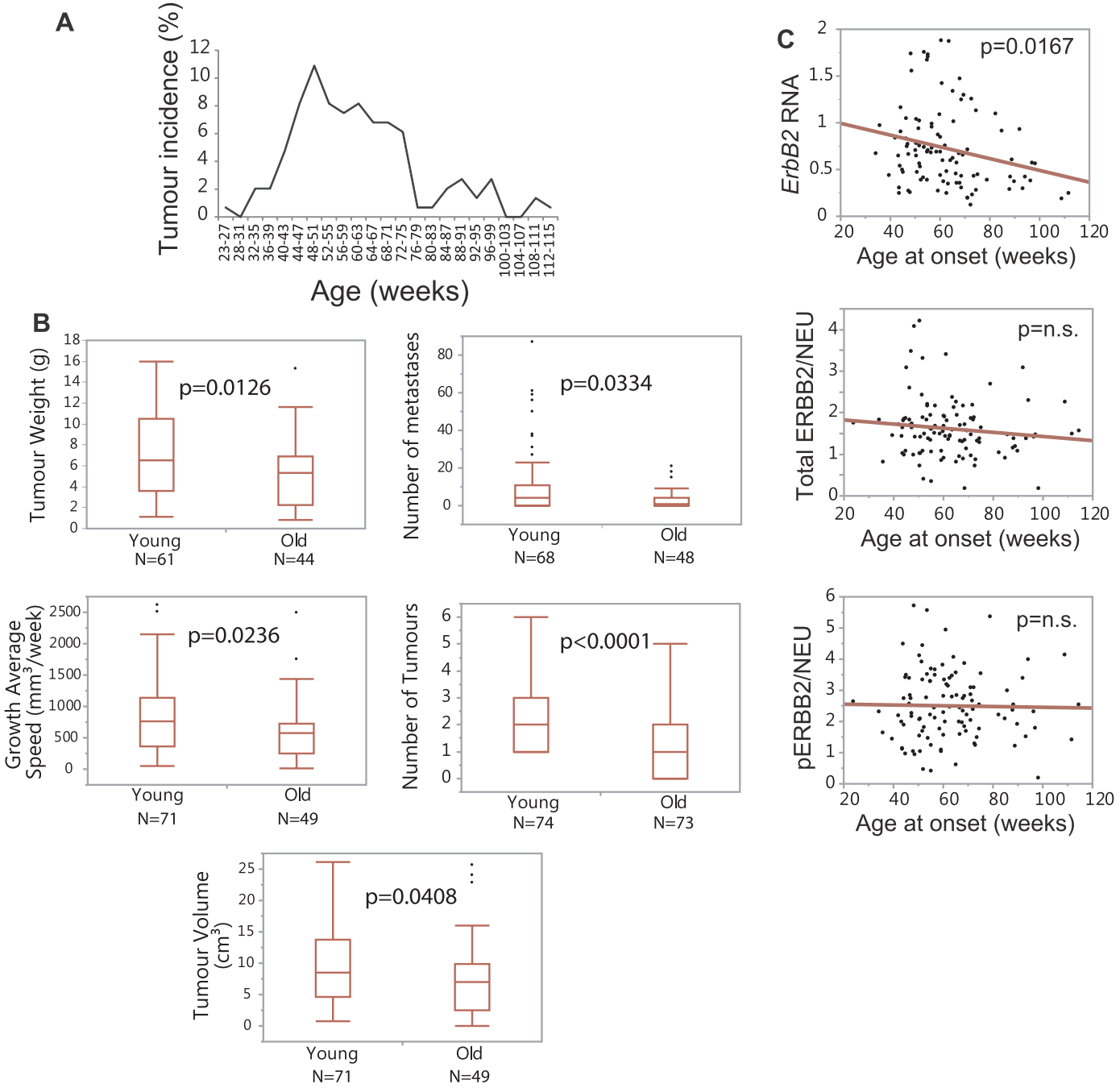


Fig. 1. The behaviour of breast cancer in the population of backcross mice. A) Tumour incidence in mice increased with age as in human populations. The values of this figure are presented in the Supplementary Table 3A in Data in Brief. **B**) The behaviour of breast cancer was more aggressive in the group of chronologically younger mice than in chronologically older mice, similar to what occurs in humans. **C**) The RNA levels of *ErbB2/Neu* determined by QPCR decreased with age, but did not correlate with total and phosphorylated ERBB2/NEU protein levels quantified by a fluorescent western blot. The values of panel C are presented in Table 3B in Data in Brief [20].

derived from the Cox model were used. Mice were ranked according to their risk score and divided into two groups using the median risk score as the cut-off [24]. Regarding the *model of prediction of the incidence of metastasis*, the *Support Vector Machine (SVM)* method and the *leave-one-out cross-validation* method were used to determine the incidence of metastasis. We used ROC and Precision-Recall curves to select the best model.

To design the *linear regression model of biological age,* we used the *lm function* implemented in *R*. Training and predictions were done using the *leave-one-out cross-validation* method. Previous studies have used the difference between the predicted/biological age and the chronological age as the grade of aging [12,25]. When the predicted age was older than chronological age, the mouse was considered biologically older. Conversely, when the predicted age was younger than the chronological age, the mouse was classified as biologically younger [12]. We generated *biplots* associated with the principal component analysis [26]. In particular, the HJ-*biplot*, an exploratory method for data analysis [27], was used, and then the linear least square regression method was applied. We used the Ward's method for performing the cluster analysis.

3.1. Expression array analysis

Normalization of microarray signal data across chips was performed with the Robust Multichip Analysis (RMA) algorithm in the Affymetrix Expression Console v. 1.4.1. Spearman correlation between probesets was calculated with the function cor implemented in the Stats package

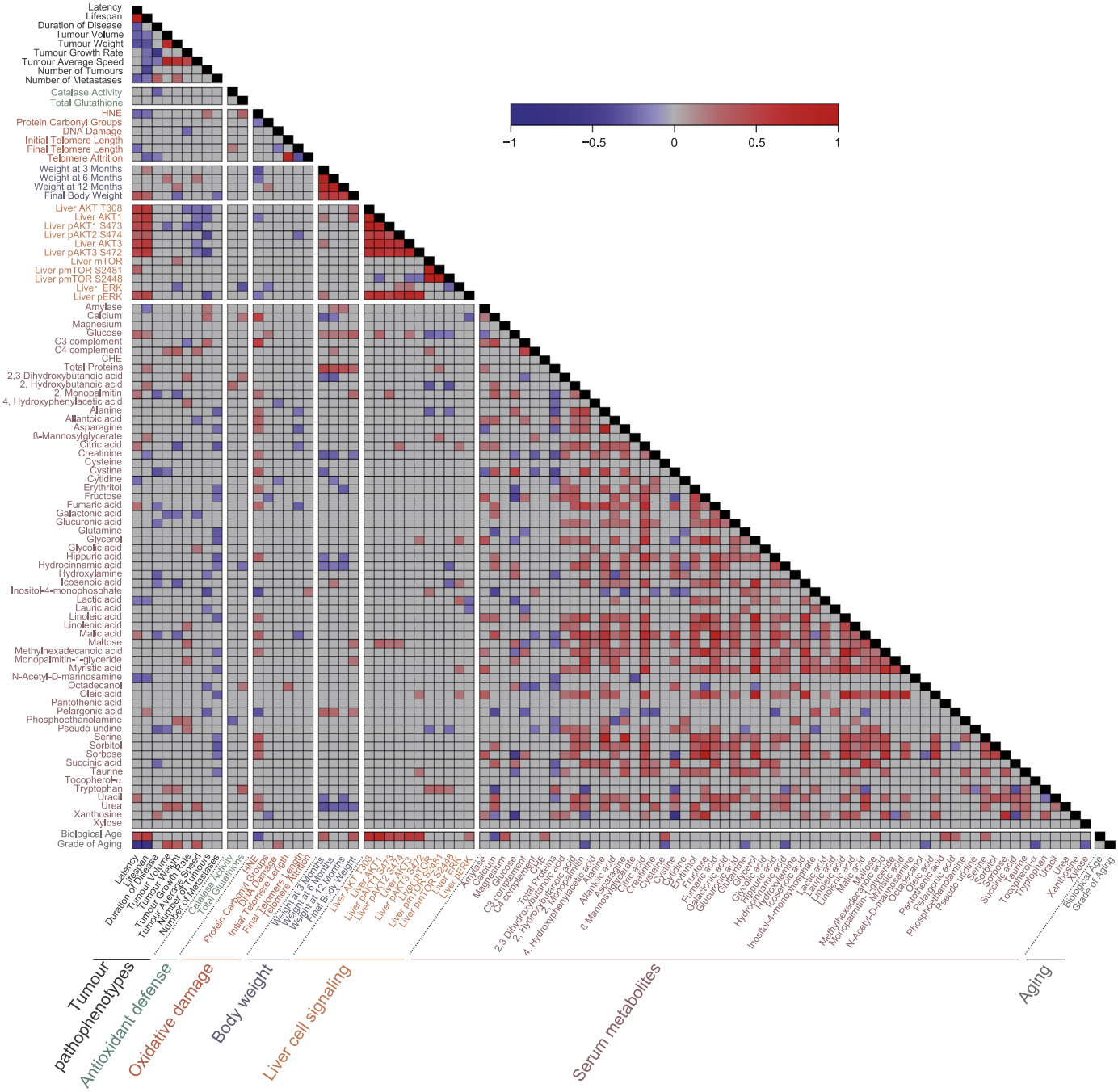
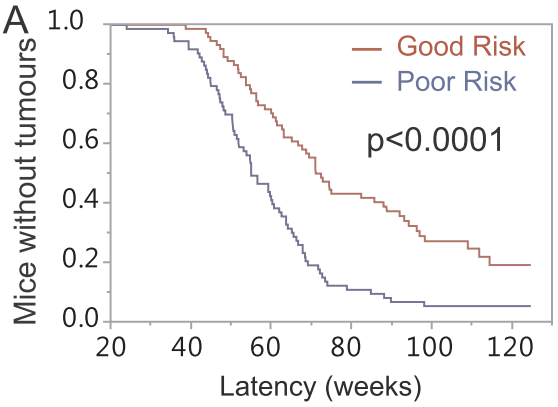


Fig. 2. Statistically significant associations between pathophenotypes and different subtypes related to oxidative stress. Only statistically significant associations were included ($P < 0.05$). The red colour indicates a positive correlation and the blue colour indicates a negative correlation. The values of this figure are presented in Table 4 in Data in Brief [20]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

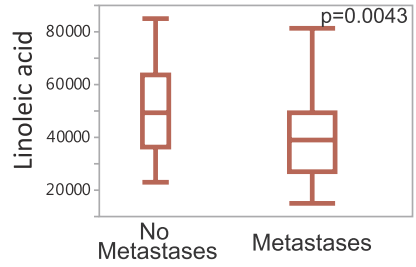
A

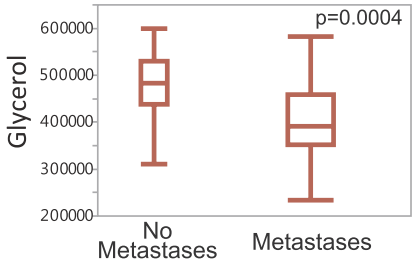


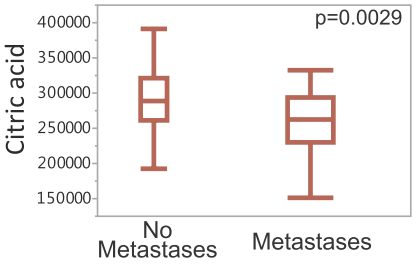
PROGNOSTIC INDEX	
Coefficient	Variable
8.58	2-monopalmitin
4.66	Uracil
8.31	Weight at 3 Months
9.33	Lactic acid
7.22	N-acetyl-D-mannosamine
-1.62	Initial Telomere Length
12.36	rs6193859 Chr.2 (89.52 cM)
9.70	rs4231934 Chr.18 (55.31 cM)

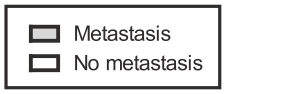
B

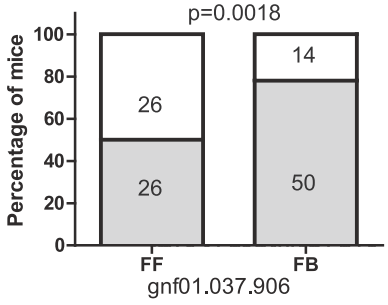
Variables	
linoleic acid	gnf01.037.906
glycerol	cel-5_79059370
citric acid	rs6159963
	gnf05.069.163
	rs3658401
	rs6256504
	rs13478222

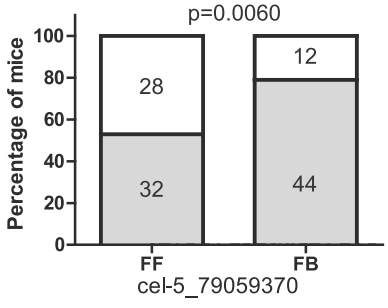


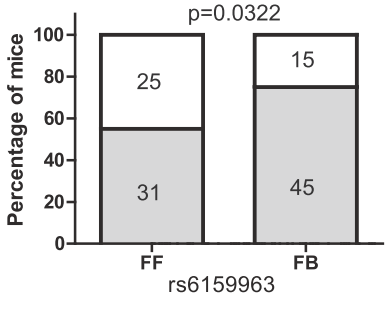


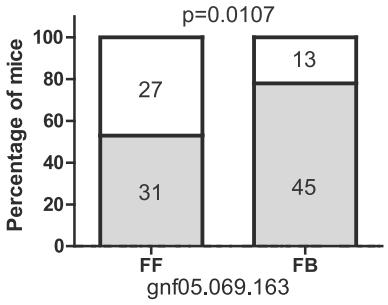


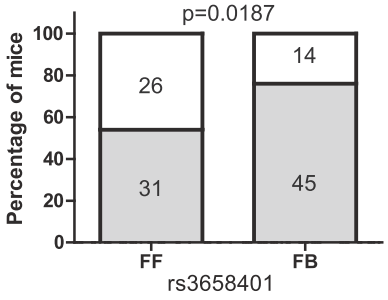


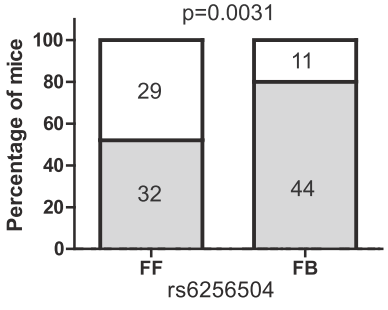


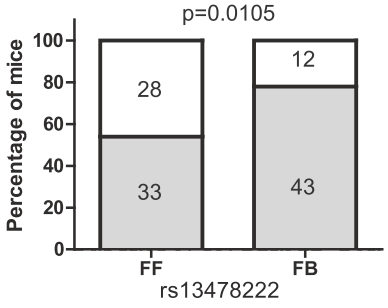












(caption on next page)

of R (version 3.2.3) [28]. Gene expression data for mouse livers are available through Gene Expression Omnibus GSE 99962.

Linkage analysis was carried out using the Interval mapping with the expectation maximization (EM) algorithm and the R/QTL software. The criteria for significant and suggestive linkages for single markers was chosen from Lander and Kruglyak [29] (see material and methods in Data in Brief [20]). To develop multiple Quantitative Trait Loci (QTL) models, the *fitQTL* function with Haley-Knott regression in R/QTL was used to fit and compare the models based on the LOD score, and the percentage of variance explained [30]. All the genetic markers used in the linkage analysis are shown (see **Table 2** in Data in Brief [20]). Genotypes were classified as FVB/FVB (FF) or FVB/C57BL/6 (FB).

Gene set enrichment analysis (GSEA) [31] was used to evaluate the enrichment of pathways identified in the differential expression analysis. The pathways used as gene sets were downloaded from the NCBI Biosystem database, which contains the records of several pathway databases (download date 07/07/2015). All probesets were ranked according to the Delta value obtained with the SAM test application between the two biological age groups. GSEA computed an Enrichment Score (ES) for each pathway and the significance was determined by permuting the data and correcting multiple testing using the family wise error rate (FEWER).

4. Results

4.1. Breast cancer is more aggressive in chronologically younger mice than in older mice

The manner in which the behaviour of breast cancer varied with respect to chronological age in a mouse backcross population exhibiting different susceptibility and evolution of ERBB2-induced breast cancer was addressed [22] (Fig. 1 and see Fig. 2 in Data in Brief [20]). To do so, a backcross was generated by crossing a breast cancer resistant mouse strain (C57BL/6) with a sensitive strain (FVB/J) overexpressing the *cNeu/ErbB2* proto-oncogene under the control of the Mouse Mammary Gland Tumour Virus (MMTV) promoter [21]. It was observed that tumour incidence increased with age, peaked at around the interval of 44–51 weeks of age, and then remained lower but relatively high until the interval of 72–75 weeks of age. After this interval, the incidence fell drastically. The resulting curve showed a bimodal behaviour which distinguished two populations of mice in terms of incidence and age (Fig. 1**A** and see **Table 3A** in Data in Brief [20]).

In addition, the cohort could be separated by the median age of survival (78.86 weeks) which was used to distinguish between young (N = 74) and old mice (N = 73) from the chronological point of view. The group of chronologically older mice included those that developed breast cancer with a long latency, plus the mice that did not develop tumours two years after the experiment was carried out. Additionally, it was observed that younger mice exhibited a more aggressive form of breast cancer and developed heavier and larger tumours, which generated more metastases than in older mice (Fig. 1**B**, additionally, see Fig. 2**A** and **Table 3B** both in Data in Brief [20]). Although tumours from younger mice expressed higher levels of *ErbB2/Neu* RNA, these levels did not correlate with total and phosphorylated ERBB2/NEU proteins (Fig. 1**C** and also see Fig. 2**B** in Data in Brief [20]). Moreover, total and phosphorylated ERBB2 protein levels did not correlate with the aggressiveness of the disease (see Fig. 2**C** in Data in Brief [20]). In sum, although younger mice showed a more aggressive disease, which is similar to what occurs in humans [4–6], the more aggressive behaviour of the disease in young mice was not due to a higher expression of the oncogenic driver in the breast tumour.

4.2. Subphenotypes related to oxidative stress were associated with heterogeneous breast cancer susceptibility and evolution

One of the pathogenic mechanisms common to cancer and aging broadly described in the literature is oxidative stress [10,11,19]. Thus we evaluated 120 intermediate phenotypes directly or indirectly related to oxidative stress (**see Table 1** in Data in Brief [20]).

The overall scenario of the associations identified is presented in Fig. 2 (also see **Table 4** in Data in Brief [20]). We observed that the levels of lipid peroxidation, determined by the quantification of 4-hydroxynonenal (4-HNE) in the liver, were associated with breast cancer with the worst prognosis. Furthermore, it was found that mice with higher levels of 4-HNE presented shorter tumour latencies, a greater number of tumours, and shorter lifespans than mice with lower levels of 4-HNE. With respect to the antioxidant defence system, mice with higher levels of catalase activity in the liver showed a shorter disease duration and tended to develop fewer metastases ($P = 0.055$). Paradoxically, higher levels of oxidative DNA damage in the liver were associated with slower tumour growth rate. Regarding telomere length, shorter tumour latencies were associated with longer final telomere length, measured in a pool of normal tissues from the tails of mice that had been sacrificed. Thus, several subphenotypes related to oxidative stress were associated with variable breast cancer pathophenotypes within this mouse cohort (Fig. 2, additionally, see **Table 4** in Data in Brief [20]).

4.3. Integration of subphenotypes related to oxidative stress and genetic markers in multivariate models to predict tumour latency and metastasis

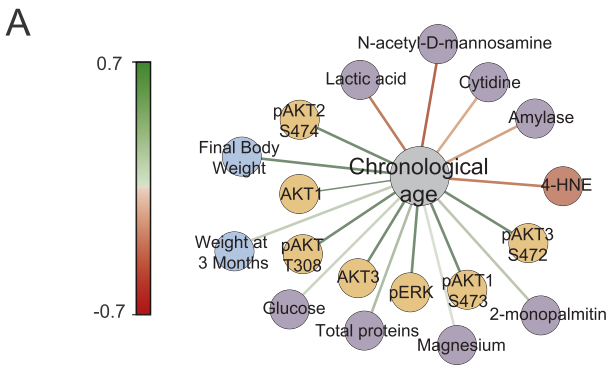
Several genetic variants have been implicated in oxidative stress [32]. Previously, we identified some Quantitative Trait Loci (QTLs) associated with ERK and AKT/mTOR signalling in the liver and the serum metabolites included within this study [22]. In addition, some QTLs associated with the variability of other subphenotypes of oxidative stress were also identified. These new QTLs were combined together with those QTLs previously identified, and presented in an overall picture (**see** Fig. 3 and **Tables 5 and 6**, all of them in Data in Brief [20].

Here it was observed that some of these QTLs identified were simultaneously associated with tumour pathophenotypes of breast cancer and other subphenotypes related to oxidative stress (see Fig. 3 and **Table 7**, both of them in Data in Brief [20]).

Later, using multivariate models, the subphenotypes related to oxidative stress and genetic markers were integrated to predict tumour latency and the incidence of metastases (Fig. 3). The Cox regression model and the corresponding prognostic index were implemented for predicting tumour latency. To do this, only the variables collected before the development of breast cancer were taken into account, which included the genetic markers and subphenotypes related to oxidative stress measured at a disease-free stage. The variables that best predicted

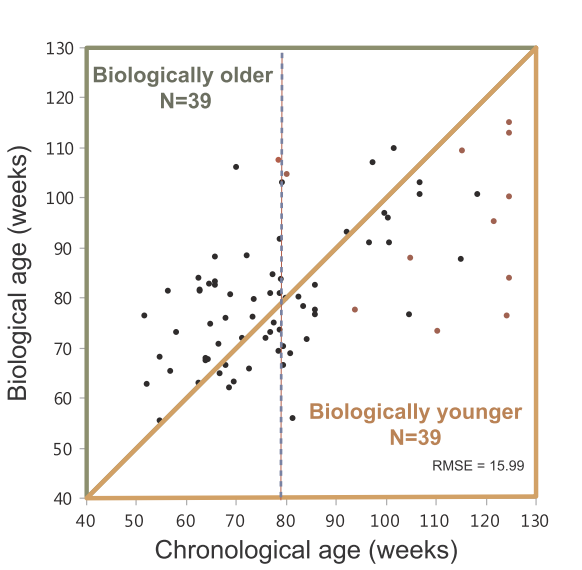
Fig. 3. Prediction of the behaviour of breast cancer by genetic markers and subphenotypes related to oxidative stress collected at a disease-free stage. A) Cox regression model to predict tumour latency. In this multivariate context, variables in green indicate a positive correlation between their values and tumour latency. Thus, high levels of 2-monopalmitin, uracil and the first weight measurement were associated with longer latencies and vice versa. Variables in red indicate a negative correlation between their values and tumour latency. Therefore, high levels of lactic acid, N-acetyl-D- mannosamine and long initial telomere length were associated with shorter tumour latencies and vice versa. The genetic marker in green indicates that heterozygosity of this marker is associated with longer tumour latency. The genetic marker in red indicates that heterozygosity of this marker is associated with a shorter latency. **B**) Multivariate model using the Support Vector Machine method to predict the incidence of metastases. The comparison of the levels of the metabolites included in the model in mice with and without metastasis is shown. The relative and absolute (numbers in bars) frequency of metastases for genetic variables included in the model is also shown. **RS-**: indicates a **r**eference **S**NP ID number, or "rs" ID, is an identification tag assigned by NCBI. **Gnf-** indicates that the marker originated at the Genomics Institute of the Novartis Research Foundation. **Cel**- refers to SNP from Celera database. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

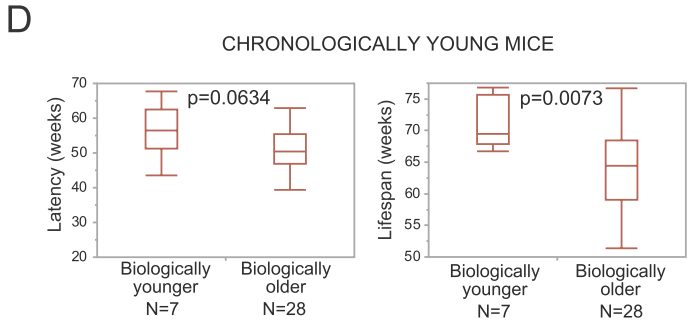
tumour latency were: the weight of the mice before the onset of the disease, the initial telomere length measured in tail tissue at three months of age, and the serum levels of 2-monopalmitin, uracil, lactic acid, and N-acetyl-D-mannosamine. There were also two SNPs, rs6193859 and rs4231934, located on chromosomes 2 and 18, respectively. In this multivariate context, the higher serum levels of 2-

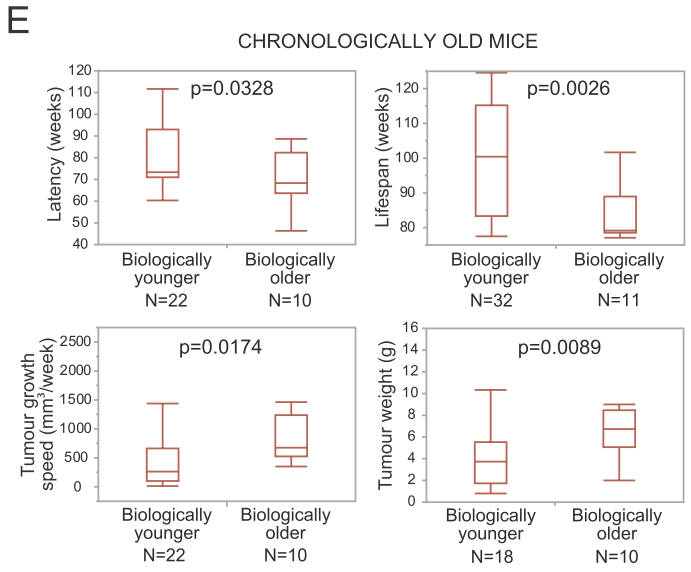


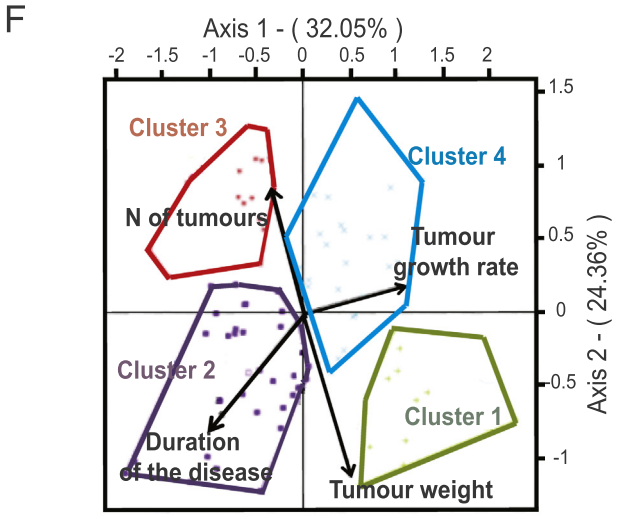
B

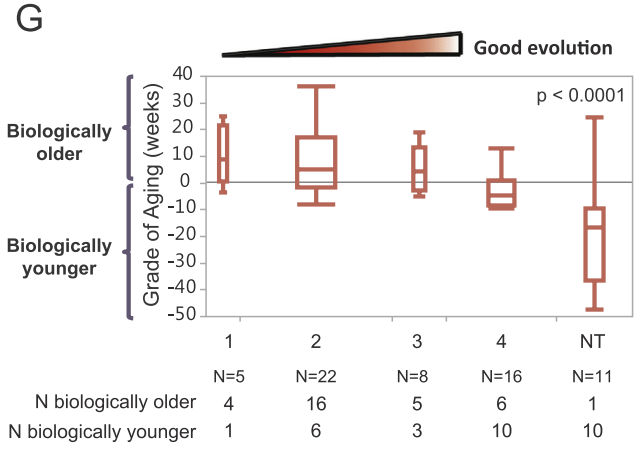
Regression model of Biological Age		
Variable	Coefficient	*P* value
AKT1	12.13	8.30 E-10
Mg	5.11	0.0036
Total glutathione	-4.34	0.0127
HDL-cholesterol	4.39	0.0137











(caption on next page)

monopalmitin, uracil and the higher body weights of mice, before the onset of the disease, were associated with longer tumour latencies. By contrast, the higher serum levels of lactic acid, N-acetyl-D-mannosamine and longer initial telomere length were associated with shorter tumour latencies. Also, mice heterozygous for the rs6193859 SNP on chromosome 2 (one allele from C57BL/6J and the other one from FVB, or F/B) showed higher tumour latencies than mice with both alleles from the FVB (F/F) background. In addition, the rs4231934 SNP on chromosome 18, heterozygous mice (F/B), showed shorter tumour latencies. The measurement of these variables allowed a prognostic index to be calculated, which divided the population of mice into two groups of good and poor risk according to their tumour latency (Fig. 3**A**).

To predict the incidence of metastases, the Support Vector Machine method was applied using the variables obtained before the onset of breast cancer (Fig. 3**B**). The oxidative stress variables that best predicted the incidence of metastases were serum linoleic acid, glycerol, and citric acid levels. Mice without metastases showed higher levels of all of these compounds in serum at a disease-free stage. Among the genetic markers included within the model, all were located near the peaks of tQTLs (tumour QTLs) associated with the metastases pathophenotypes that were identified in a previous study [22] (see **Table 5** in Data in Brief [20]). Moreover, gnf01.037.906 located on chromosome 1 overlapped with tQTL14, and the rest of the markers located on chromosome 5 overlapped with tQTL12. Mice with a C57BL/6 J (F/B) allele in all of the SNPs included in the model showed a higher incidence of metastases incidence and/or absolute number (Fig. 3**B**).

4.4. Subphenotypes related to oxidative stress are associated with chronological age and allow biological age to be defined

Oxidative stress has been linked to aging susceptibility [11,19]. Therefore, a search was carried out to look for associations between the subphenotypes related to oxidative stress and the chronological age of the mice in our cohort (Fig. 4**A**). It was observed that chronologically older mice showed higher levels of total and phospho-AKT and ERK in the liver, higher serum levels of 2-monopalmitin, glucose, magnesium and total proteins, and lower serum levels of cytidine, lactic acid, N-acetyl-D-mannosamine, and amylase at a disease-free stage. Regarding the specific markers of oxidative damage, a negative correlation was found between the chronological age and the levels of 4-HNE. As expected, somatic telomere length, again quantified using a pool of tail tissues, showed telomere shortening with age (see Fig. 4**A–E** in Data in Brief [20]). All this data show an association between several subphenotypes related to oxidative stress and the chronological age of mice.

Following on, the possibility of estimating a partial biological age based on a global status of oxidative stress was investigated. To do so, first a mathematical model was used to identify which of the subphenotypes related to oxidative stress best predicted chronological age (Fig. 4**B**). As expected, the age computed by the model did not perfectly match with the real chronological age, but instead reflected the estimated partial biological age associated with oxidative stress subphenotypes. Certainly, mathematical models have previously been used to predict biological age based on different aging biomarkers, contrasting biological to real chronological age [15,16,33]. Here, linear regression was used to establish a predictor of biological age, using chronological age as a dependent variable in training and performing recursive feature elimination on the independent variables. Training and predictions were carried out on all samples using leave-one-out cross-validation. The predictions were considered to be the biological age of the mice. Overall, the variables selected by the model to predict biological age were the serum levels of magnesium and HDL-cholesterol measured at a disease-free stage, and whole AKT1 protein levels and total glutathione levels in the liver at the time of necropsy. Higher levels of magnesium, HDL-cholesterol, and total AKT1 protein were associated with the prediction of higher values of biological age, while higher levels of total glutathione were associated with the prediction of lower values of biological age (Fig. 4**B**).

Second, we calculated the grade of aging, which aims to identify whether each mouse was biologically younger or older with respect their chronological age. To do so, the biological age of each mouse was compared to its real chronological age. Therefore, the grade of aging was determined by calculating the difference between the estimated biological age and the chronological age. Thus, when biological age was greater than the chronological age, the mouse was biologically older. On the contrary, when the biological age was lower than the chronological age, the mouse was biologically younger (Fig. 4**C**).

As previously indicated, the mice were divided into chronologically young and old based on the median of the chronological age of the cohort, which was 78.86 weeks (see Fig. 4**F** in Data in Brief [20]). Based on the grade of aging, we observed that mice that were biologically older were those that predominately developed breast cancer and died at a relatively young chronological age as a result of the disease. In contrast, mice that were biologically younger were among those that frequently developed the disease at a relatively old chronological age. Interestingly, nearly all of the mice that did not develop breast cancer, except for two, were among those that were biologically younger (red points in Fig. 4**C**).

Interestingly, when the disease among the chronologically younger mice was compared and mice that were biologically younger and biologically older were distinguished, it was observed that biologically older mice developed the disease earlier and had a shorter lifespan than their biologically younger counterparts (Fig. 4**D**). Similarly, when the disease among chronologically old mice was compared, it was again observed that biologically younger mice developed the disease later, had less aggressive tumours and a longer lifespan than their biologically older counterparts (Fig. 4**E**).

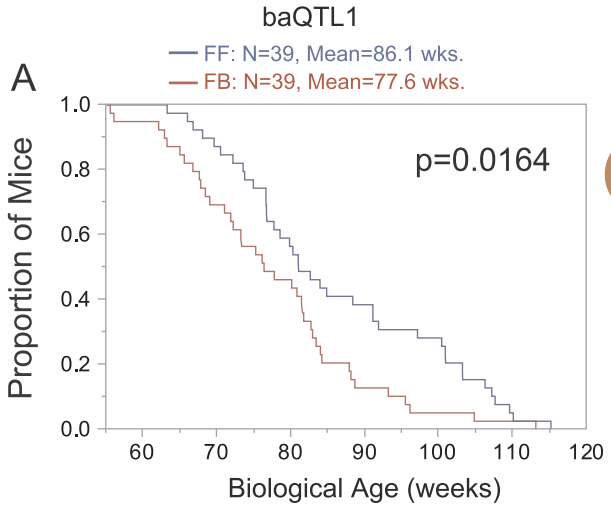
The groups of mice exhibiting similar breast cancer behaviour, regarding tumour pathophenotypes, were then classified using a principal component analysis. The result of this analysis showed how biologically younger and older mice were distributed within different clusters. Temporal phenotypes of the disease, such as latency and lifespan, were

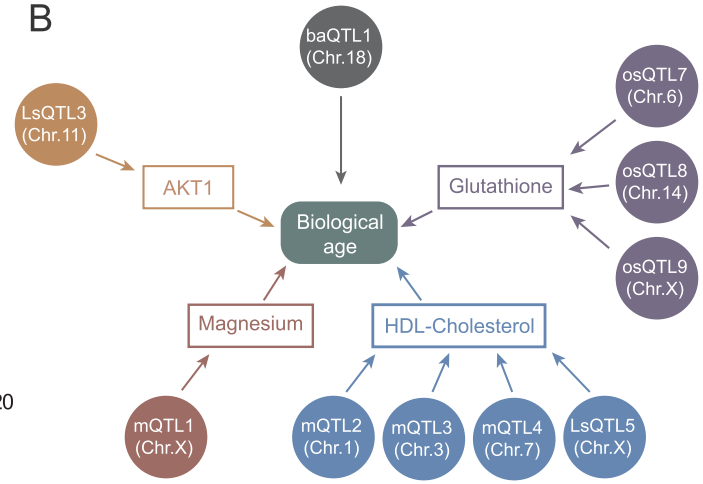
Fig. 4. Biological age and behaviour of breast cancer in backcross mice based on grade of aging. A) Subphenotypes related to oxidative stress associated with the chronological age in a statistically significant manner. The bivariate correlations are shown in Fig. 4A–E in Data in Brief [20]. Red colour indicates negative associations. Green colour indicates a positive association. The intensity of the colour is proportional to the correlation coefficient. **B**) Multivariate model to define the biological or estimated age as a function of the sub-phenotypes related to oxidative stress selected by the model. **C**) A representation of the chronological age and biological age in weeks is shown. Each point represents the value of both ages for each mouse. The distance from the diagonal indicates the grade of aging. The points on the left, above the diagonal, show animals with a positive grade of aging, indicating they are biologically older than the chronological age. Points on the right, below the diagonal, show mice with a negative grade of aging, indicating they are biologically younger than the chronological age. The red dots represent the mice that did not develop breast cancer. The red line on the x-axis indicates the median of the chronological age (78.86 weeks). RMSE, *root mean square error*. **D**) Mice chronologically young at the time of death (those mice younger than the median of 78.86 weeks), and defined as being biologically young by the model (those mice whose biological age was smaller than their chronological age), presented less aggressive breast cancer than those defined as biologically old (those mice whose biological age was bigger than their chronological age). **E**) Similarly, within chronologically old mice at the time of death (those mice older than the median of 78.86 weeks), the ones described by the model as being biologically younger (those mice whose biological age was smaller than their chronological age) also developed a less aggressive disease than those defined as biologically old (those mice whose biological age was bigger than their chronological age). **F**) Clusters of mice with similar evolution of breast cancer after excluding latency and lifespan as variables (HJ-biplot analysis). **G**) Distribution of the number of biologically young and old mice in each cluster. The percentage of biologically young mice increases from left to right in the groups with the best prognosis, where the non-tumour developing mice in cluster 5 have the highest percentage. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

excluded to remove the effect of their strong correlation with both chronological and biological age. After doing so, four clusters associated with the evolution of breast cancer (Fig. 4**F**) were identified. It is worth noting that even though the variables of latency and lifespan were eliminated from the analysis, the mice were still classified from a chronological perspective, reflecting the strong correlation between chronological age and some pathophenotypes of the disease (see Fig. 5 in Data in Brief [20]). Interestingly, biologically older mice were mainly grouped in the clusters (1 and 2) associated with the worst evolution of the breast cancer, while biologically younger mice were grouped in the clusters (3 and 4) associated with the best prognosis. Additionally, the group of mice that did not develop tumours was the one that was comprised of more biologically younger mice (Fig. 4**G**). In conclusion, these findings indicate that biologically older mice develop more aggressive breast cancer regardless of their chronological age.

4.5. *Identification of QTLs associated with biological age through genetic markers linked to the subphenotypes*

The next step was to search for QTLs related to biological age





C

Model	LOD Score	Percentage of Variability
One QTL: baQTL1	2.34	12.90%
Two QTLs: baQTL1 + LsQTL3	3.52	18.79%
Three QTLs: baQTL1 + LsQTL3 + osQTL9	4.59	23.74%

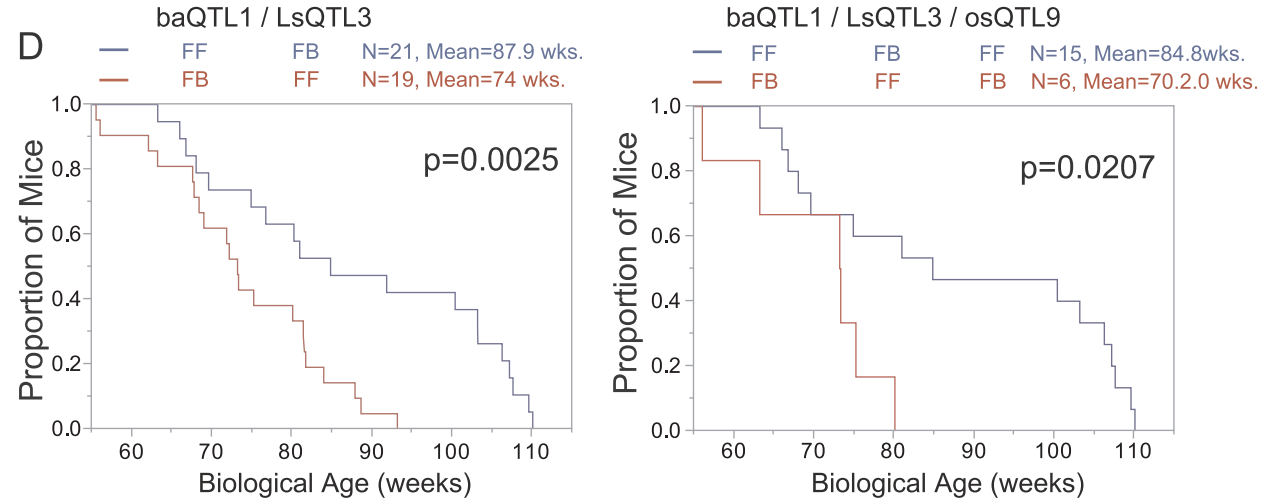


Fig. 5. Genetic markers linked to biological age. A) Kaplan-Meier analysis showing how the homozygous (F/F) and heterozygous (F/B) mice for baQTL1 on chromosome 18 have different biological age in a statistically significant manner. **B)** Scheme showing the oxidative stress subphenotypes that define the biological age and the QTLs linked to them. baQTL, biological aging QTL; LsQTL, liver signalling QTL; osQTL, oxidative stress QTL. The number after QTL is the identification number as shown in Table 5 of Data in Brief. **C)** Table showing the LOD score reached and the different proportion of phenotypic variability of the biological age explained by baQTL1 alone and in combination with two or three QTLs linked to the subphenotypes associated with biological age. **D)** Kaplan-Meier analysis showing how the addition of other QTLs to baQTL-1 improves discrimination between biologically older and younger mice. The effect of one QTL is represented on the left and the effect of two QTLs is represented on the right.

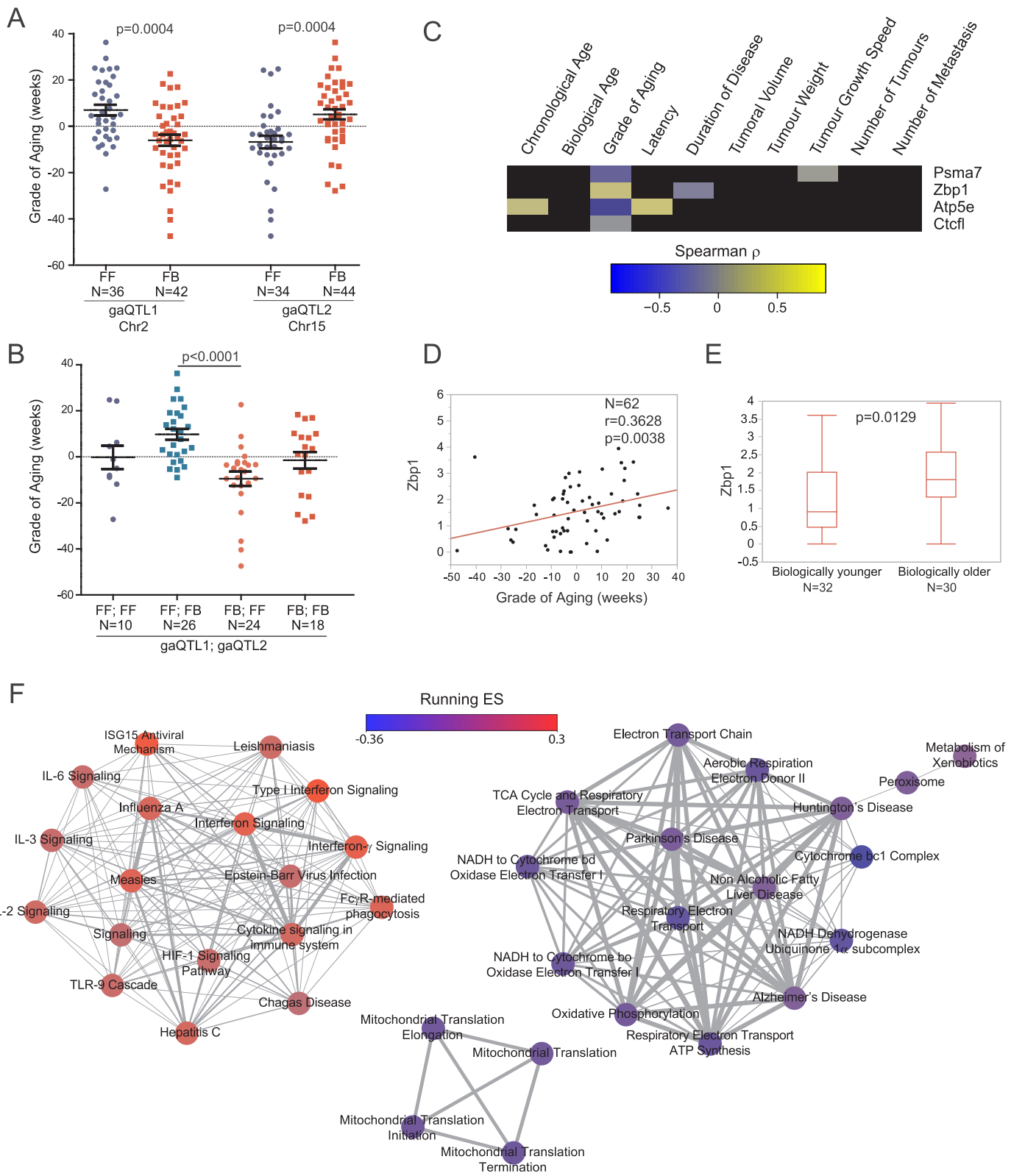


Fig. 6. The grade of aging. A) The grade of aging in the backcross mice depending on their genotype for gaQTL1 and gaQTL2. **B**) The combination of both gaQTLs allowed for better discrimination between biologically younger and older mice. **C**) The heatmap shows transcripts from gaQTL1 whose levels correlated with the grade of aging and other phenotypes. The scale is based on the correlation coefficient. Yellow colour means positive correlation and blue colour indicates negative correlations. Spearman test. **D**) Correlation between the levels of *Zbp1* expression in the liver and the grade of aging in backcross mice. **E**) Biological older mice had higher levels of *Zbp1* expression in the liver than biological younger mice. **F**) The network shows the enriched pathways in biologically old mice (red nodes) and in biologically young mice (blue nodes). The colour intensity of each node is proportional to the mean of the running enrichment score (ES) of the genes of that pathway, reflecting the degree to which the expression of these genes is overrepresented in the liver samples. The thickness of the connections is proportional to the number of enriched genes common among the pathways. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(baQTLs) (**see Table 5** in Data in Brief [20]). From this search, a QTL located on chromosome 18 (baQTL1) was identified with a peak at 31.59 cM. Mice homozygous for the FVB allele on this QTL peak had significantly higher biological age than heterozygous FVB/B6 mice (Fig. 5**A**). This QTL could be related to the tQTL4, previously identified by our group, which is associated with tumour latency [22] (see Fig. 3 in Data in Brief [20] and **Tables 5** and **7** also in Data in Brief [20]).

Additionally, QTLs associated with the subphenotypes that defined biological age were identified: serum levels of HDL-cholesterol and magnesium, and AKT1 and glutathione levels in the liver (Fig. 5**B**). Surprisingly, none of these QTLs were directly associated with the biological age (see **Tables 6** and **7** in Data in Brief [20]). The lack of association of a complex trait with the genetic determinants linked to the subphenotypes of that complex trait has been related to the missing heritability [8]. Thus, we hypothesized that a genetic model which includes QTLs related to some of the subphenotypes associated with oxidative stress together with baQTL1 should explain more phenotypic variability than baQTL1 alone (Fig. 5**B**). Indeed this was the case, as generating a model with two QTLs which included baQTL1 and LsQTL3 (associated with the levels of total AKT1 in the liver) increased the LOD score from 2.34 to 3.52. The model also explained 18.79% of the variability of the phenotype versus 12.90% with the baQTL1 alone. Similarly, a model with three QTLs, the two previously mentioned plus osQTL9 which is associated with glutathione levels in the liver, was able to increase the LOD score to 4.59 and explained 23.74% of the variability (Fig. 5**C**). These QTL combinations also resulted in a clear separation of the mouse subpopulations in the Kaplan-Meier analyses (Fig. 5**D**). In conclusion, these results show that it is possible to increase the explained phenotypic variability of a complex trait through QTLs linked to subphenotypes, which in turn are associated with the complex trait. The application of this strategy on a larger scale with regards to the specific complex trait could identify part of the missing heritability [8].

4.6. The pattern of gene expression in the liver of biologically old mice was enriched in inflammation and response to infection pathways

Later on, a search for QTLs related to the grade of aging, that is, genetic regions linked to being either biologically older or younger with respect to chronological age (see **Table 5** in Data in Brief [20]), was carried out. Consequently, two QTLs associated with the grade of aging were identified and were called gaQTL1 (grade of aging QTL1) located on chromosome 2 (LOD score = 3.07), and gaQTL2 located on chromosome 15 (LOD score = 2.69). The marker peak for gaQTL1 was rs13476913. Mice homozygous for the FVB allele for this marker tended to be biologically older, whereas heterozygous FVB/B6 mice tended to be biologically younger (Fig. 6**A** and, additionally, see **Table 5** in Data in Brief [20]). In addition, heterozygous FVB/B6 mice for gaQTL2 tended to be biologically older than homozygous FVB mice, which tended to be biologically younger (Fig. 6**A**, and also see **Table 5** in Data in Brief [20]). When the two gaQTLs were considered together, the mice were better distinguished with respect to the grade of aging (Fig. 6**B**).

The linkage association of gaQTL1 was relatively high (LOD score = 3.1), and the confidence interval was not very large (88.09–102.57 cM) (see **Table 5** in Data in Brief [20]). Hence, we then searched for the liver expression of genes located on this genomic region by expression arrays to identify possible candidate genes associated with the grade of aging and breast cancer pathophenotypes (see **Table 8** in Data in Brief [20]). Twenty-seven transcripts in this gaQTL1 associated with the grade of aging and breast cancer pathophenotypes were identified, and 24 of them were uncharacterized genes. Seven out of these 24 uncharacterized genes had a predicted structure that encoded for a Krüppel associated box (KRAB) domain. Also, there were only 4 transcripts that originated from known genes: *Atp5e, Zbp1, Psma7* and *Ctcf l* (Fig. 6**C**). An attempt was made to validate all four genes by QPCR, but only *Zbp1*, a gene related to necroptosis and inflammation [34], was confirmed to be associated with the grade of aging ($P = 0.0038$, $r = 0.363$) (Fig. 6**D** and also see **Table 8** in Data in Brief [20]). Also, biologically younger mice had lower levels of *Zbp1* expression in the liver than biologically older mice ($P = 0.013$) (Fig. 6**E**). The association of *Zbp1* with increased inflammation [34,35] supports the hypothesis that *Zbp1* might be one of the genes responsible for connecting oxidative stress and grade of aging.

Since the expression of *Zbp1* was found to be elevated in the liver of biologically old mice and *Zbp1* was identified as a key gene in the inflammatory response against infections [34,35], the extent to which the pattern of gene expression in the liver of biologically old mice was enriched in pathways related to these functions was analysed. Interestingly, 131 pathways were identified as being significantly enriched in their gene expression patterns in the livers of biologically old mice (FWER p-value lower than 0.05). Among 20 of the more significant pathways identified, there were 7 pathways related to inflammation and the response to infections (see **Table 9A** in Data in Brief [20]). Moreover, an abundance of signalling pathways related to these functions was observed. This finding suggests that there are common genes that are differentially enriched and likely to be involved in various functionally related signalling pathways. To check this, a search was carried out to identify genes common to some of the pathways associated with inflammation and response to infections, such as some interleukins, Toll-like receptor, and response to viral infections such as influenza A, Epstein-Barr virus or measles, and non-viral infections like leishmaniosis or Chagas disease (Fig. 6**F**, and also see Fig. 6**B** in Data in Brief [20]).

On the contrary, in biologically young mice, only 21 pathways significantly enriched in the gene expression pattern in the liver (FWER p-value lower than 0.05) were identified. Interestingly, among the first 20 most significant pathways, up to 15 pathways were related to mitochondrial function (see **Table 9B** in Data in Brief [20]). In addition, genes common to these mitochondrial pathways and differentially enriched (Fig. 6**F** and also see Fig. 6**A** in Data in Brief [20]) were also found. This data indicate that expression patterns in the liver show a connection between oxidative stress and the grade of aging.

5. Discussion

The incidence of breast cancer increases with age until a particular time point that coincides with menopause [1,3]. This type of tumour tends to be more aggressive in young patients; actually, early age is an independent factor of poor prognosis in the evolution of breast cancer [5,6]. Age is, in fact, the strongest demographic risk factor for most human malignancies [36,37]. The reason for the epidemiological association between age and the incidence and aggressiveness of breast cancer is unknown. In this study it was observed that the behaviour of breast cancer, developed in a genetically heterogeneous population of mice generated by a backcross, was similar to the behaviour of breast cancer in humans with respect to aggressiveness. Furthermore, breast cancer in younger mice was seen to present greater tumour aggressiveness (Fig. 1**A and B**, and also see Fig. 2**A** in Data in Brief [20]). Hence, it was thought that these mice would provide a suitable model for studying the pathogenic reasons for the epidemiological association between breast cancer behaviour and age. However, the different behaviour of the disease with age was not due to ERBB2 protein levels (Fig. 1**C** and also see Fig. 2**C** in Data in Brief [20]). Therefore, although ERBB2 is the driver of the disease, it does not appear to be responsible for most of the different levels of phenotypic expression of the disease among mice. The absence of the participation of ERBB2 in the variability of tumour aggressiveness indicates that this is due to other genetic determinants and intermediate phenotypes. In this sense, it is known that the quantitative inheritance not only contributes to the heterogeneity of complex traits but also contributes, to a greater or lesser extent, to the phenotypic heterogeneity of traits with Mendelian

inheritance [38]. This fact is in agreement with the number of QTL regions associated with the aggressiveness of different pathophenotypes found (see Fig. 3 in Data in Brief [20]).

The epidemiological link between breast cancer and aging indicate the existence of common pathogenic processes and intermediate phenotypes. One of the intermediate phenotypes common to breast cancer and aging, abundantly supported by the literature, is oxidative stress. Thus, the higher the degree of oxidative stress, the greater the susceptibility to aging in different organisms [3,5,10,11,19]. Also, oxidative stress has been associated with tumour incidence and aggressiveness [3]. In this work, we studied the association of oxidative stress with both the chronological age and the susceptibility and evolution of breast cancer. Also, it was took into account that oxidative stress is a complex trait whose global phenotypic manifestations could be influenced by multiple intermediate phenotypes [8].

Using multivariate models, the extent to which a combination of several subphenotypes related to oxidative stress would be associated with the different susceptibility and evolution to breast cancer and aging was evaluated. We think that multivariate models help to better define the variability of complex traits than individual subphenotypes. For example, if we consider telomere length alone, shorter tumour latencies were associated with longer final telomere length, measured in a pool of normal tissues from the tails of mice that had been sacrificed. This result was expected since mice with short tumour latencies were chronologically younger by definition and would have a longer final telomere length. But it could be surprising that a longer initial telomere length was selected by the Cox regression model to define poor prognosis in terms of latency (Fig. 3**A**). Although this may seem counterintuitive, first, we must consider that the variability of the initial telomere length between mice does not reflect the chronological age, since all of them were 3 months old at the time of telomere quantification. Thus, the variability of telomere length between mice at that age would be more related to other aspects, such as genetic background [39]. Second, we must also consider that the initial telomere length contributed to explain tumour latency in a multivariate context, although not by itself, since it was not associated with tumour latency in the univariate analysis (Fig. 2).

Denham Hartman previously proposed the hypothesis that oxidative stress was a cause of aging [40]. Here, it was observed that some of the subphenotypes related to oxidative stress were independently associated with the chronological age of mice (Fig. 4**A** and also see Fig. 4**A–E** in Data in Brief [20]). The possibility of establishing a global measure of oxidative stress was investigated, by generating a variable that encompassed the values of the oxidative stress subphenotypes associated with age. To do this, several variables of oxidative stress were integrated in a multivariate mathematical model (Fig. 4**B**). When applying this mathematical model on the data collected from each mouse, a predicted age was obtained and considered as the biological age regarding oxidative stress. Different mathematical models have been described in the literature to predict the biological age of an individual [12–16]. Biological age is a complex concept which, fundamentally, refers to the physiological capacity and the physiological reserve of the organs and systems in an individual. This physiological capacity decreases with age, but the degree of decline varies within individuals and may be greater or lower than the chronological age. Hence, the individual would be considered as being biologically older or younger, depending on the grade of aging, in comparison to their actual chronological age. It is difficult to identify molecular markers that can encompass, in a general sense, something as complex as the functional capacity of different organs and systems of an organism. Given the importance of oxidative stress in aging [3,10,11,19], some subphenotypes related to oxidative stress were used to define biological age. The purpose of this study was not to determine the overall functional status of the mice, but to assess to which extent the degree of oxidative stress was reflected in the different subphenotypes simultaneously, which was also associated with greater or less tumour susceptibility. However, the objective of obtaining a global biological age that would indicate the total physiological state of an individual might be difficult or even impossible. We propose the utility of defining what we could call "partial biological ages" obtained by integrating several physiological features and biomarkers in multivariate models. The integrated value would be compared with the chronological age to ascertain whether the individual was biologically older or younger than their actual chronological age concerning the physiological phenotype under analysis. Interestingly, our model permitted us to integrate breast cancer behaviour, aging, and oxidative stress. In this study, it was observed that most of the relative chronologically young mice that had more aggressive breast cancer were also biologically older in terms of oxidative stress. On the contrary, most part of those mice that developed breast cancer at more advanced chronological ages, were younger based on oxidative stress (Fig. 4**D–G**). These is in agreement with the fact that oxidative stress has been associated with aging and tumour incidence and aggressiveness [3].

Finally, we identified two QTLs associated with the grade of aging. Among the genes studied in gaQTL1, *Zbp1* correlated with the grade of aging. ZBP1 is a cytosolic sensor of pathogen-associated DNA that activates the downstream interferon regulatory factor (IRF) and NF-kappa B transcription factors, leading to type-I interferon production. ZBP1 has been related to inflammation and necroptosis of virus-infected cells [34,35]. Moreover, ZBP1 has recently been identified as a critical mediator of inflammation, a function beyond its already known role in antiviral defence [34,35]. Due to the well-known role of inflammation in cancer and aging pathogenesis [7,41,42], it is easy to speculate that elevated basal levels of ZBP1 together with other genes could contribute to increased basal inflammation and increased oxidative stress, increasing susceptibility to cancer and aging but perhaps with more efficient protection against viruses. In agreement with this fact, a large number of signalling pathways related to inflammation and response to viral and other infections were observed to be enriched in the liver of biologically old mice (Fig. 6**F**, additionally, see also Fig. 6**B** and **Table 9A,** both in Data in Brief [20]). Within these pathways and genes, ZBP1 could be one of the most important genes because it was detected at a genetic level as a putative candidate responsible for the effect of gaQTL1 (more studies are needed to demonstrate this point). Certainly, infections have been one of the factors that have had the greatest selection pressure on human populations [43]. A better response to infectious agents through a more potent inflammatory response and necroptosis may favour the higher probability of survival in the early and middle stages of life. However, this may be counterproductive later on, favouring the onset of cancer and other chronic inflammatory or autoimmune components [44]. Lastly, biologically young mice showed enriched pathways related to mitochondrial function in the liver (Fig. 6**F**, additionally, see Fig. 6**A** and **Table 9B,** both in Data in Brief [20]). Certainly, mitochondria are one of the key regulators of longevity. It has been proposed that respiratory chain-deficient cells are more predisposed to suffer apoptosis and an increased cell loss is probable central in the age-associated mitochondrial dysfunction [45].

6. Conclusion

Breast cancer is more aggressive in young patients than in old ones. This behaviour is reproduced in a mouse cohort generated by a backcross. In this work, we observed that mice with more aggressive breast cancer are older regarding oxidative stress, according to the mathematical model of biological age generated. The use of strategies similar to those employed in this study may help to further understand the relationship between mechanisms of inflammation, oxidative stress, and susceptibility to cancer and aging.

Acknowledgements

JPL was partially supported by FEDER and the MICINN (SAF2014-

56989-R and SAF2017-88854R), the Instituto de Salud Carlos III (PIE14/00066), "Proyectos Integrados IBSAL 2015" (IBY15/00003), the Sandra Ibarra Foundation "de Solidaridad Frente al Cáncer" Foundation and "We can be heroes" Foundation. JHM was supported by the National Institutes of Health, a National Cancer Institute grant (R01 CA116481), and the Low-Dose Scientific Focus Area, Office of Biological & Environmental Research, US Department of Energy (DE-AC02-05CH11231). We thank Trent Northen for his help in mass spec determination, Chris Lauber for some R code corrections and Emma Keck for her help in editing the English.

Competing interest

The authors declare no potential conflict of interest.

Authors' contributions

M.M.S.F. and A.B.G. helped to coordinate the overall work, helped with necropsy studies and participated in tissue processing, contributed to data analysis and interpretation, and carried out the linkage analysis in R-software. They performed some protein expression studies in livers. M.M.S.F. helped to develop the multivariate aging model. S.C.L. carried out general protein analyses and contributed to the processing of mouse tissues. A.G.V. and M.G.J. helped with multiple QPCR and expression arrays validation. C.M.S. and M.I.G. carried out the clinical biochemistry studies. L.H.P. helped to carry out the dissection of the tissues and other tasks at the animal facility. M.B.G.C. and J.G.C. advised and helped to implement the dissections and necropsy studies. C.P.A. and P.G.V. advised and helped perform the statistical studies, in particular the HJ-biplot. J.H.M. helped with the linkage studies. A.C.M. helped to conceive and coordinate the overall work, helped with necropsy studies and participated in tissue processing. He helped all the laboratory and animal facility technicians, and contributed to data analysis and interpretation. C.P. performed the bioinformatic analyses with the expression arrays. L.K. conceived the strategy to develop the multivariate model to define aging and performed the bioinformatics analyses. J.P.L. conceived and coordinated the overall work, helped in the necropsy and biochemical studies, and processed the tissues. The general association studies, including those with metabolites, were done in the laboratory of J.P.L. and he wrote the manuscript with the input and help of the other co-authors.

References

[1] M.R. Stratton, N. Rahman, The emerging landscape of breast cancer susceptibility, Nat. Genet. 40 (1) (2008) 17–22.
[2] K. McPherson, C.M. Steel, J.M. Dixon, ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics, BMJ 321 (7261) (2000) 624–628.
[3] C.C. Benz, C. Yau, Ageing, oxidative stress and cancer: paradigms in parallax, Nat. Rev. Cancer 8 (11) (2008) 875–879.
[4] N. Klauber-DeMore, Tumor biology of breast cancer in young women, Breast Dis. 23 (2005) 9–15.
[5] C. Yau, V. Fedele, R. Roydasgupta, J. Fridlyand, A. Hubbard, J.W. Gray, K. Chew, S.H. Dairkee, D.H. Moore, F. Schittulli, S. Tommasi, A. Paradiso, D.G. Albertson, C.C. Benz, Aging impacts transcriptomes but not genomes of hormone-dependent breast cancers, Breast Cancer Res.: BCR 9 (5) (2007) R59.
[6] S.A. Narod, Breast cancer in young women, Nat. Rev. Clin. Oncol. 9 (8) (2012) 460–470.
[7] L.M. Coussens, Z. Werb, Inflammation and cancer, Nature 420 (6917) (2002) 860–867.
[8] A. Blanco-Gomez, S. Castillo-Lluva, M. Del Mar Saez-Freire, L. Hontecillas-Prieto, J.H. Mao, A. Castellanos-Martin, J. Perez-Losada, Missing heritability of complex diseases: enlightenment by genetic variants from intermediate phenotypes, Bioessay.: News Rev. Mol. Cell. Dev. Biol. 38 (7) (2016) 664–673.
[9] D.Y. Cho, Y.A. Kim, T.M. Przytycka, Chapter 5: network biology approach to complex diseases, PLoS Comput. Biol. 8 (12) (2012) e1002820.
[10] S. De Flora, A. Izzotti, F. D'Agostini, R.M. Balansky, Mechanisms of N-acetylcysteine in the prevention of DNA damage and cancer, with special reference to smoking-related end-points, Carcinogenesis 22 (7) (2001) 999–1013.
[11] A. Mitsui, J. Hamuro, H. Nakamura, N. Kondo, Y. Hirabayashi, S. Ishizaki-Koizumi, T. Hirakawa, T. Inoue, J. Yodoi, Overexpression of human thioredoxin in transgenic mice controls oxidative stress and life span, Antioxid. Redox Signal. 4 (4) (2002) 693–696.
[12] D. Karasik, S. Demissie, L.A. Cupples, D.P. Kiel, Disentangling the genetic determinants of human aging: biological age as an alternative to the use of survival measures, J. Gerontol. Ser. A Biol. Sci. Med. Sci. 60 (5) (2005) 574–587.
[13] D.W. Belsky, A. Caspi, R. Houts, H.J. Cohen, D.L. Corcoran, A. Danese, H. Harrington, S. Israel, M.E. Levine, J.D. Schaefer, K. Sugden, B. Williams, A.I. Yashin, R. Poulton, T.E. Moffitt, Quantification of biological aging in young adults, Proc. Natl. Acad. Sci. USA 112 (30) (2015) E4104–E4110.
[14] A. Comfort, Test-battery to measure ageing-rate in man, Lancet 2 (7635) (1969) 1411–1414.
[15] M.E. Levine, Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age? J. Gerontol. Ser. A Biol. Sci. Med. Sci. 68 (6) (2013) 667–674.
[16] I.H. Cho, K.S. Park, C.J. Lim, An empirical comparative study on biological age estimation algorithms with an application of Work Ability Index (WAI), Mech. Ageing Dev. 131 (2) (2010) 69–78.
[17] M. Barton, Aging and endothelin: determinants of disease, Life Sci. 118 (2) (2014) 97–109.
[18] M. Giorgio, E. Migliaccio, F. Orsini, D. Paolucci, M. Moroni, C. Contursi, G. Pelliccia, L. Luzi, S. Minucci, M. Marcaccio, P. Pinton, R. Rizzuto, P. Bernardi, F. Paolucci, P.G. Pelicci, Electron transfer between cytochrome c and p66Shc generates reactive oxygen species that trigger mitochondrial apoptosis, Cell 122 (2) (2005) 221–233.
[19] N. Treiber, P. Maity, K. Singh, M. Kohn, A.F. Keist, F. Ferchiu, L. Sante, S. Frese, W. Bloch, F. Kreppel, S. Kochanek, A. Sindrilaru, S. Iben, J. Hogel, M. Ohnmacht, L.E. Claes, A. Ignatius, J.H. Chung, M.J. Lee, Y. Kamenisch, M. Berneburg, T. Nikolaus, K. Braunstein, A.D. Sperfeld, A.C. Ludolph, K. Briviba, M. Wlaschek, L. Florin, P. Angel, K. Scharffetter-Kochanek, Accelerated aging phenotype in mice with conditional deficiency for mitochondrial superoxide dismutase in the connective tissue, Aging Cell 10 (2) (2011) 239–254.
[20] M. Del Mar. Saez-Freire, A. Blanco-Gomez, S. Castillo-Lluva, A. Gomez-Vecino, J. Galvis-Jimenez, C. Martin-Seisdedos, M. Isidoro-Garcia, L. Hontecillas-Prieto, B. Garcia-Cenador, J. Garcia-Criado, C. Patino-Alonso, P. Galindo-Villardon, J.H. Mao, C. Prieto, A. Castellanos-Martin, L. Kaderali, J. Perez-Losada, Supplementary data for the biological age linked to oxidative stress modifies breast cancer aggressiveness, Data in Brief Submitted.
[21] C.T. Guy, M.A. Webster, M. Schaller, T.J. Parsons, R.D. Cardiff, W.J. Muller, Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease, Proc. Natl. Acad. Sci. USA 89 (22) (1992) 10578–10582.
[22] A. Castellanos-Martin, S. Castillo-Lluva, M. Saez-Freire Mdel, A. Blanco-Gomez, L. Hontecillas-Prieto, C. Patino-Alonso, P. Galindo-Villardon, L. Perez Del Villar, C. Martin-Seisdedos, M. Isidoro-Garcia, M.M. Abad-Hernandez, J.J. Cruz-Hernandez, C.A. Rodriguez-Sanchez, R. Gonzalez-Sarmiento, D. Alonso-Lopez, J. De Las Rivas, B. Garcia-Cenador, J. Garcia-Criado, D.Y. Lee, B. Bowen, W. Reindl, T. Northen, J.H. Mao, J. Perez-Losada, Unraveling heterogeneous susceptibility and the evolution of breast cancer using a systems biology approach, Genome Biol. 16 (2015) 40.
[23] I. Rahman, A. Kode, S.K. Biswas, Assay for quantitative determination of glutathione and glutathione disulfide levels using enzymatic recycling method, Nat. Protoc. 1 (6) (2006) 3159–3165.
[24] M.K.B.A.M. Parmar, D, Survival analysis: A practical approach, Chichester, England, 1995.
[25] D. Karasik, M.T. Hannan, L.A. Cupples, D.T. Felson, D.P. Kiel, Genetic contribution to biological aging: the Framingham study, J. Gerontol. Ser. A Biol. Sci. Med. Sci. 59 (3) (2004) 218–226.
[26] K.R. Gabriel, The biplot graphic display of matrices with application to principal component analysis, Biometrika 58 (3) (1971) 453–467.
[27] M. Galindo, The alternative for simultaneous representations, HJ-Biplot Qüestiió 10 (1) (1986) 13–23.
[28] R.D.C. Team, R: A Language and Environment for Statistical Computing, Vienna, Austria, 2014.
[29] E. Lander, L. Kruglyak, Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results, Nat. Genet. 11 (3) (1995) 241–247.
[30] K.W. Broman, H. Wu, S. Sen, G.A. Churchill, R/qtl: QTL mapping in experimental crosses, Bioinformatics 19 (7) (2003) 889–890.
[31] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. USA 102 (43) (2005) 15545–15550.
[32] L. Forsberg, U. de Faire, R. Morgenstern, Oxidative stress, human genetic variation, and disease, Arch. Biochem. Biophys. 389 (1) (2001) 84–93.
[33] F. Bormann, M. Rodriguez-Paredes, S. Hagemann, H. Manchanda, B. Kristof, J. Gutekunst, G. Raddatz, R. Haas, L. Terstegen, H. Wenck, L. Kaderali, M. Winnefeld, F. Lyko, Reduced DNA methylation patterning and transcriptional connectivity define human skin aging, Aging Cell 15 (3) (2016) 563–571.
[34] J. Lin, S. Kumari, C. Kim, T.M. Van, L. Wachsmuth, A. Polykratis, M. Pasparakis, RIPK1 counteracts ZBP1-mediated necroptosis to inhibit inflammation, Nature 540 (7631) (2016) 124–128.
[35] K. Newton, K.E. Wickliffe, A. Maltzman, D.L. Dugger, A. Strasser, V.C. Pham, J.R. Lill, M. Roose-Girma, S. Warming, M. Solon, H. Ngu, J.D. Webster, V.M. Dixit, RIPK1 inhibits ZBP1-driven necroptosis during development, Nature 540 (7631) (2016) 129–133.
[36] B.K. Edwards, H.L. Howe, L.A. Ries, M.J. Thun, H.M. Rosenberg, R. Yancik, P.A. Wingo, A. Jemal, E.G. Feigal, Annual report to the nation on the status of cancer, 1973–1999, featuring implications of age and aging on U.S. cancer burden,

Cancer 94 (10) (2002) 2766–2792.
[37] U.S.C.S.W. Group, United States Cancer Statistics: 1999–2013 Incidence and Mortality Web-based Report, (2016) ⟨www.cdc.gov/uscs ⟩.
[38] W.G. Hill, Quantitative genetics in the genomics era, Curr. Genom. 13 (3) (2012) 196–206.
[39] E.L. Manning, J. Crossland, M.J. Dewey, G. Van Zant, Influences of inbreeding and genetics on telomere length in mice, Mamm. Genome.: Off. J. Int. Mamm. Genome. Soc. 13 (5) (2002) 234–238.
[40] D. Harman, Aging: a theory based on free radical and radiation chemistry, J. Gerontol. 11 (3) (1956).
[41] C. Franceschi, M. Bonafe, S. Valensin, F. Olivieri, M. De Luca, E. Ottaviani, G. De Benedictis, Inflamm-aging. An evolutionary perspective on immunosenescence, Ann. N.Y. Acad. Sci. 908 (2000) 244–254.
[42] A. Salminen, K. Kaarniranta, A. Kauppinen, Inflammaging: disturbed interplay between autophagy and inflammasomes, Aging 4 (3) (2012) 166–175.
[43] M. Fumagalli, M. Sironi, Human genome variability, natural selection and infectious diseases, Curr. Opin. Immunol. 30 (2014) 9–16.
[44] L.B. Barreiro, L. Quintana-Murci, From evolutionary genetics to human immunology: how selection shapes host defence genes, Nat. Rev. Genet. 11 (1) (2010) 17–30.
[45] I. Bratic, A. Trifunovic, Mitochondrial energy metabolism and ageing, Biochim. Biophys. Acta 1797 (6–7) (2010) 961–967.