Original Research

DNA methylome analysis identifies accelerated epigenetic ageing associated with postmenopausal breast cancer susceptibility

Srikant Ambatipudi [a], Steve Horvath [b], Flavie Perrier [a], Cyrille Cuenin [a], Hector Hernandez-Vargas [a], Florence Le Calvez-Kelm [a], Geoffroy Durand [a], Graham Byrnes [a], Pietro Ferrari [a], Liacine Bouaoun [a], Athena Sklias [a], Véronique Chajes [a], Kim Overvad [c], Gianluca Severi [d,e,f], Laura Baglietto [d,f], Françoise Clavel-Chapelon [d], Rudolf Kaaks [g], Myrto Barrdahl [g], Heiner Boeing [h], Antonia Trichopoulou [i,j], Pagona Lagiou [i,j,k], Androniki Naska [i,j], Giovanna Masala [l], Claudia Agnoli [m], Silvia Polidoro [e], Rosario Tumino [n], Salvatore Panico [o], Martijn Dollé [p], Petra H.M. Peeters [q,r], N. Charlotte Onland-Moret [q], Torkjel M. Sandanger [s], Therese H. Nøst [s], Elisabete Weiderpass Vainio [s,t,u,v], J. Ramón Quirós [w], Antonio Agudo [x], Miguel Rodriguez-Barranco [y,z], José María Huerta Castaño [aa,z], Aurelio Barricarte [ab,ac,z], Ander Matheu Fernández [ad,ae], Ruth C. Travis [af], Paolo Vineis [ag], David C. Muller [ag], Elio Riboli [ag], Marc Gunter [a], Isabelle Romieu [a], Zdenko Herceg [a,*]

[a] *International Agency for Research on Cancer (IARC), Lyon, France*
[b] *Human Genetics and Biostatistics, University of California Los Angeles, Los Angeles, CA 90095-7088, USA*
[c] *Section for Epidemiology, Department of Public Health, Aarhus University, Aarhus, Denmark*
[d] *Inserm, Centre de Recherche en Epidémiologie et Santé des Populations (CESP, U1018), Université Paris-Saclay, Université Paris-Sud, UVSQ, Institut Gustave Roussy, Villejuif, France*
[e] *Human Genetics Foundation (HuGeF), Torino, Italy*
[f] *Cancer Epidemiology Centre, Cancer Council Victoria and Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourn, Australia*
[g] *Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany*
[h] *Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbrücke, Nuthetal, Germany*
[i] *Hellenic Health Foundation, Athens, Greece*
[j] *WHO Collaborating Center for Nutrition and Health, Unit of Nutritional Epidemiology and Nutrition in Public Health, Department of Hygiene, Epidemiology and Medical Statistics, University of Athens Medical School, Athens, Greece*
[k] *Department of Epidemiology, Harvard School of Public Health, Boston, USA*

---

* *Corresponding author*: Epigenetics Group, International Agency for Research on Cancer (IARC), 150 Cours Albert Thomas, F-69008, Lyon, France. Fax: +33 4 72 73 83 22.
E-mail address: herceg@iarc.fr (Z. Herceg).

[l] Molecular and Nutritional Epidemiology Unit, Cancer Research and Prevention Institute – ISPO, Florence, Italy
[m] Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale Tumori, Milano, Italy
[n] Cancer Registry and Histopathology Unit, "Civic M.P. Arezzo" Hospital, ASP Ragusa, Italy
[o] Dipartimento di Medicina Clinica e Chirurgia, Federico II University, Naples, Italy
[p] Centre for Health Protection, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
[q] Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands
[r] MRC-PHE Centre for Environment and Health, Dept of Epidemiology and Biostatistics, School of Public Health, Imperial College, London, UK
[s] Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, The Arctic University of Norway, Tromsø, Norway
[t] Department of Research, Cancer Registry of Norway, Institute of Population-Based Cancer Research, Oslo, Norway
[u] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[v] Genetic Epidemiology Group, Folkhälsan Research Center, Helsinki, Finland
[w] Public Health Directorate, Asturias, Spain
[x] Unit of Nutrition and Cancer, Cancer Epidemiology Research Program, Catalan Institute of Oncology-IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain
[y] Escuela Andaluza de Salud Pública, Instituto de Investigación Biosanitaria ibsn Granada, Hospitales Universitarios de Granada/Universidad de Granada, Granada, Spain
[z] CIBER de Epidemiología y Salud Pública (CIBERESP), Spain
[aa] Department of Epidemiology, Murcia Regional Health Council, IMIB-Arrixaca, Murcia, Spain
[ab] Navarra Public Health Institute, Pamplona, Spain
[ac] Navarra Institute for Health Research (IdiSNA) Pamplona, Spain
[ad] Cellular Oncology Group, Biodonostia Health Research Institute, Paseo Dr. Beguiristain s/n, San Sebastian, Spain
[ae] IKERBASQUE, Basque Foundation, Spain
[af] Cancer Epidemiology Unit, Nuffield Department of Population Health University of Oxford, Oxford UK
[ag] School of Public Health, Imperial College London, London, UK

Abstract   *Aim of the study:* A vast majority of human malignancies are associated with ageing, and age is a strong predictor of cancer risk. Recently, DNA methylation-based marker of ageing, known as 'epigenetic clock', has been linked with cancer risk factors. This study aimed to evaluate whether the epigenetic clock is associated with breast cancer risk susceptibility and to identify potential epigenetics-based biomarkers for risk stratification.

Methods: Here, we profiled DNA methylation changes in a nested case−control study embedded in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort (n = 960) using the Illumina HumanMethylation 450K BeadChip arrays and used the Horvath age estimation method to calculate epigenetic age for these samples. Intrinsic epigenetic age acceleration (IEAA) was estimated as the residuals by regressing epigenetic age on chronological age.

Results: We observed an association between IEAA and breast cancer risk (OR, 1.04; 95% CI, 1.007−1.076, P = 0.016). One unit increase in IEAA was associated with a 4% increased odds of developing breast cancer (OR, 1.04; 95% CI, 1.007−1.076). Stratified analysis based on menopausal status revealed that IEAA was associated with development of postmenopausal breast cancers (OR, 1.07; 95% CI, 1.020−1.11, P = 0.003). In addition, methylome-wide analyses revealed that a higher mean DNA methylation at cytosine-phosphate-guanine (CpG) islands was associated with increased risk of breast cancer development (OR per 1 SD = 1.20; 95 %CI: 1.03−1.40, P = 0.02) whereas mean methylation levels at non-island CpGs were indistinguishable between cancer cases and controls.

Conclusion: Epigenetic age acceleration and CpG island methylation have a weak, but statistically significant, association with breast cancer susceptibility.

1. Introduction

Ageing is a major risk factor for most neoplasms [1]. In particular, breast cancer is an age-associated disease whose incidence rises sharply after menopause [1]. This increased risk was hypothesised to be the consequence of accumulation of genetic changes (mutations) associated with deregulation of cellular processes and genomic instability. However, accumulation of genetic changes exhibits striking interindividual differences [2], and differences in biological ageing processes may only be partly explained by genetic determinants [3].

A recent study demonstrates that DNA methylation (DNAm) data lend themselves for developing a highly accurate multitissue biomarker of ageing [4]. The DNAm-based marker of ageing (known as 'epigenetic clock') derived from several tissues can be used to accurately estimate the chronological age of all tissues and cell types [4]. This composite biomarker of ageing, which is defined as a weighted average across 353 specific CpG sites, produces an estimate of age (in units of years), referred to as 'epigenetic age' or 'DNA methylation age (DNAm age)'. Recent studies demonstrate that DNAm age is at least a passive biomarker of biological age: the epigenetic age of blood has been found to be predictive of all-cause mortality [5–9], frailty [10], cognitive and physical functioning [5]. Further, the utility of the epigenetic clock method using various tissues and organs has been demonstrated in applications surrounding Alzheimer disease [11], centenarian status [8], pre-natal and early life influences [12], Down syndrome [13], HIV infection [14], Huntington disease [15], obesity [16], lifetime stress [17], menopause [18], and Parkinson disease [19]. Departures of methylation-estimated age from chronological age can be used to define intrinsic epigenetic age acceleration (IEAA) that measures cell-intrinsic ageing effects that are independent of chronological age and blood cell composition.

A recent study suggests that IEAA can be used to predict lung cancer risk [20]. However, it is not yet known whether IEAA lends itself for predicting breast cancer susceptibility in a prospective case−control study. To test this hypothesis, we analysed blood methylation data from incident breast cancer cases and matching controls of a large prospective study within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort.

2. Materials and methods

2.1. Selection of incident cancer and control participants

The present study was conducted on nested case−control samples from the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort, a large prospective study conducted in 23 centres across ten European countries (Denmark, France, Germany, Greece, Italy, Norway, Spain, Sweden, The Netherlands, and the United Kingdom), aiming to investigate the relationship between diet, lifestyle, metabolism and cancer risk [21]. In brief, the EPIC cohort includes a total of about 315,000 women and 200,000 men. At baseline recruitment, all study participants provided extensive questionnaire information about nutrition and other lifestyle factors. All study participants also provided a blood sample, which was processed, divided into aliquots of plasma, serum and buffy coat and frozen at −196 °C (under liquid nitrogen) for later use in specific research projects. In all EPIC centres, an identical protocol for subject recruitment, sample collection and storage was followed. Detailed information on the subject recruitment, baseline data, and blood collection protocols have been reported previously [22]. All participants gave written, informed consent for data and biospecimen collection and storage, as well as follow-up. The study was approved by the local ethics committees and the Institutional Review Board of the International Agency for Research on Cancer (IARC, Lyon, France). During prospective follow-up of the EPIC cohort, a very large number (>11,000) of newly diagnosed, invasive breast cancer cases were confirmed histologically or cytologically as primary breast cancers according to the International Classification of Diseases for Oncology, Second Edition (ICD-O-2) and included all breast cancer subsites (ICD C50.0-C50.9). A representative subset of these cases was used for studies comparing a variety of biomarker measurements with a set of control subjects, matching the cases by recruitment centre. Incident patients with cancer were identified at regular intervals through population-based cancer registries (in Denmark, Italy except Naples, the Netherlands, Norway, Spain, Sweden, and the United Kingdom) or by active follow-up (France, Germany, Greece, and Naples), which involved a combination of methods, including a review of health insurance records, cancer and pathology registries, and direct contact with participants and their next-of-kin.

For the purpose of this study, we included 960 females from the EPIC cohort including 480 incident breast cancer cases. Our main criteria for selection of case/control pairs included: (1) a balanced representation of the main subtypes of breast cancer, and (2) representation of recruiting centres. One control participant was randomly assigned for each patient with breast cancer from appropriate risk sets consisting of all cohort participants alive and free of cancer (except for non-melanoma skin cancer) at the time of diagnosis (and hence, age) of the index case. Matching criteria were: centre, length of follow-up, age at blood collection (3 months relaxed up to 2 years for sets without available controls), time of blood collection, fasting status, menopausal status, menstrual cycle day and current use of contraceptive pill/hormone replacement therapy.

Twenty technical replicates were included to compare inter- and intra-array batch variation. Technical replicates and 38 samples or their matched counterparts which failed the quality control criteria were excluded from the analysis leaving 902 participants (451 controls and 451 cases) (Table 1).

2.2. Bisulfite conversion and genome-wide DNA methylation analysis

The DNA was isolated as per the standard DNA extraction procedure from the from the buffy coat samples (Autopure LS, Qiagen). DNA methylome profiling was carried out using Illumina Infinium HumanMethylation450 (HM450) as previously described [23].

2.3. Bioinformatics analysis

Data preprocessing and analyses were performed using R 3.2.3 (https://www.r-project.org/) and Bioconductor 3.2 [24] as described before [23]. DNAm level was described as a β value, which is a continuous variable ranging between 0 (no methylation) and 1 (full methylation). To avoid spurious associations, we excluded the cross-reactive probes and probes overlapping with a known single nucleotide polymorphism (SNPs) with a minor allele frequency of at least 5% in the overall population (European ancestry, [25]), leaving 423,066 probes. In any given sample, probes with a detection P-value (a measure of an individual probe's performance) of more than 0.05 were assigned missing status. If a probe was missing in more than 5% of samples, it was excluded from all samples. According to this criterion, we excluded 1483 probes, leaving 421,583 probes available for the analyses. We applied colour bias correction followed by quantile and beta-mixture quantile normalisation (BMIQ) to align Type I and Type II probe distributions [26].

2.4. White blood cell count estimates

Quantile normalised data were used to infer blood cell proportions. We estimate blood cell counts using two different software tools. First, Houseman's estimation method [27] was used to estimate the proportions of CD8+ T cells, CD4+ T, natural killer, B cells, and granulocytes (also known as polymorphonuclear leucocytes). Second, the advanced analysis option of the epigenetic clock software [4,14] was used to estimate the percentage of exhausted CD8+ T cells (defined as CD28-CD45RA-) and the number (count) of naïve CD8+ T cells (defined as CD45RA + CCR7+). We and others have shown that the estimated blood cell counts have moderately high correlations with corresponding flow cytometric measures [27,28]. For example, flow cytometric measurements correlate strongly with DNAm-based estimates: $r = 0.63$ for CD8+ T cells, $r = 0.77$ for CD4+ T cells, $r = 0.67$ for B cell, $r = 0.68$ for naïve CD8+ T cell, $r = 0.86$ for naïve CD4+ T, and $r = 0.49$ for exhausted CD8+ T cells [28].

2.5. Global and mean methylation analysis

For the global DNAm analyses, mean methylation of the DNAm probes (421,583) was calculated for cases and control samples. Human cancers are characterised by global hypomethylation and a loci-specific DNA hypermethylation [29]. We hypothesised that DNA methylation of probes would vary based on their physical location. To this end, the probes were classified into different categories either reflecting their physical location in relation to CpG islands (island, shore, shelf and open sea) or based on a functional criterion (DP: distal promoter, DS: distal sequence, GB: gene body, IG: intergenic, and PP: proximal promoter) as previously described [30]. A CpG shore is defined as the area 2 kb on either side of the CpG island, and a CpG shelf is defined as the area 2 kb outside of the CpG shore [31,32]. While the regions in the genome containing isolated CpG sites outside CpG islands, shores and

Table 1
Characteristics of incident breast cancer and control participants at baseline (i.e.time of blood collection).

	All samples	
	Controls (%)	Cases (%)
Sample size	451	451
Mean methylation (in %)	51.86	51.82
Age (years)		
Mean (SD)	52.3 (8.94)	52.3 (8.97)
Median	53.4	53.5
Alcohol consumption (g/d)		
Mean(SD)	8.2 (11.82)	10.0 (12.98)
Age at menarche		
Mean (SD)	12.9 (1.34)	12.7 (1.59)
BMI		
Mean (SD)	25.5 (4.22)	26.0 (4.72)
Physical activity (Cambridge index)		
Sedentary	99 (22.0)	121 (26.8)
Moderately sedentary	187 (41.5)	178 (39.5)
Moderately active	76 (16.9)	87 (19.3)
Active	78 (17.3)	62 (13.7)
Missing	11 (2.4)	3 (0.7)
Hormone receptor status		
ER+/PR+/Her2+	–	85 (18.8)
ER+/PR+/Her2−	–	290 (64.3)
ER−/PR−/Her2−	–	76 (16.9)
Country		
Italy	160 (35.5)	160 (35.5)
Spain	27 (6.0)	27 (6.0)
UK	38 (8.4)	38 (8.4)
The Netherlands	66 (14.6)	66 (14.6)
Greece	25 (5.5)	25 (5.5)
Germany	135 (29.9)	135 (29.9)

SD: Standard deviation; ER: oestrogen receptor; PR: progesterone receptor; Her2: human epidermal growth factor receptor 2; BMI: body mass index.

shelves, that do not have a specific designation are referred to as open seas [33].

2.6. Epigenetic clock of ageing

The epigenetic clock is a prediction method of chronological age based on the DNAm levels of 353 CpGs [4]. The predicted (estimated) age resulting from the epigenetic clock is referred to as 'DNA methylation age'. In IEAA, epigenetic age acceleration is defined as the DNAm age left unexplained by chronological age where intrinsic denotes a modification to this concept. In addition to adjusting for chronological age, IEAA also adjusts the DNAm age estimate for blood cell count estimates, arriving at a measure that is unaffected by both variation in chronological age and blood cell composition.

We focussed on IEAA in our blood-based methylation study as this measure of age acceleration is significantly correlated with epigenetic age acceleration in (non-malignant) female breast tissue [9].

Formally, IEAA is defined by regressing DNAm age on chronological age and seven measures of blood cell count abundances: naive CD8 T cells, exhausted CD8 T cells (defined as CD28-CD45RA-), plasma blasts, CD4 T cells, NK cells, monocytes, granulocytes. IEAA is automatically calculated using the advanced analysis option of the epigenetic clock software (where IEAA is denoted as 'AAHOAdjCellCounts'). A positive or negative value of IEAA indicates that the woman is either older or younger than expected based on chronological age at the time of the blood draw.

2.7. Statistical analysis

For the mean methylation analysis, average methylation over all probes within each category was calculated and the odds ratios (per one standard deviation of global methylation) were estimated by conditional logistic regression model with case−control status as the outcome and the epigenome-wide methylation measurement as continuous predictor adjusting for surrogate variables (technical batch effects such as sample plate, array chips), alcohol consumption (g/day) and body mass index (BMI) as continuous variable.

Odds ratios (ORs) for breast cancer and 95% CIs were calculated by using logistic regression for IEAA. Initial analysis was done using unconditional logistic regression to allow calculation of OR. Multivariate logistic regression was performed by including known breast cancer risk factors including alcohol consumption (g/day), full term pregnancy (ever/never), BMI (as continuous variable and as categorical variable: underweight, normal, overweight and obese), level of education (none, primary, technical/profession, secondary, higher education), age at menarche, Cambridge physical activity index (inactive, moderately inactive, moderately active and active) stratified by clustering variable. A stratified multivariate

conditional logistic regression analysis based on the menopausal status was performed using the aforementioned models.

3. Results

3.1. Baseline characteristics

The baseline characteristics of samples at the time of recruitment are listed in Table 1. Women were between 26 and 73 years of age with a mean age of 52.3 years for cases and controls. The majority of breast cancer cases were hormone receptor (ER and PR) positive (83%) while 17% of the breast cancers were triple negative (Table 1). There was a very high correlation between the intra- and interplate technical replicates (average correlation coefficient $r^2 = 0.98$ and 0.97, respectively, data not shown).

3.2. Hypermethylation of CpG islands is associated with breast cancer risk

We compared the global mean methylation across 421,583 probes and observed no difference between prospectively collected cases and matched controls (51.82% versus 51.86%, $P = 0.68$). Our analysis showed that each unit (95% CI/1SD, 1.03−1.40, $P = 0.02$) increase in methylation at CpG island sites increased the risk of being a case by 20% (Table 2). While $P < 0.05$, it should be noted that the results would be marginally significant allowing for four subsets (CpG islands, CGI shores, CGI shelves, and open sea). No change in breast cancer risk was observed for other regions (shore, shelf and open sea) (Table 2), nor did we find an association of individual CpG site or region with breast cancer status.

Table 2

Association between global methylation and breast cancer risk by CpG genomic features.

	Context	# CpGs	Std. dev.	OR (95% CI)[a]	P value
	All CpG sites	421 583	3.45E-04	1.09 (0.94−1.25)	0.21
	Islands	130 982	5.87E-04	1.20 (1.03−1.40)	0.02
	Open Sea	150 852	4.50E-03	1.49 (0.36−6.24)	0.58
CpG context	Shelf	40 948	4.88E-04	0.89 (0.78−1.02)	0.10
	Shore	98 801	5.40E-04	1.00 (0.87−1.16)	0.97
	Distal promoter	19 990	5.42E-04	1.06 (0.92−1.21)	0.44
	Distal sequence	7828	6.68E-04	0.96 (0.84−1.09)	0.52
Genic context	Gene Body	168 460	3.80E-04	1.02 (0.89−1.18)	0.76
	Intergenic	56 903	5.35E-04	1.02 (0.89−1.17)	0.76
	Proximal promoter	168 337	5.26E-04	1.15 (0.99−1.34)	0.07

[a] Odds ratio and confidence interval were calculated per 1 standard deviation. Odds ratios were adjusted for body mass index (BMI) (continuous variable) and daily alcohol intake. OR- Odds ratio, CI: confidence interval.

3.3. Postmenopausal breast cancer cases exhibit DNA methylation age acceleration

Epigenetic age had a strong positive correlation with chronological age in both case and control samples (Fig. 1a). We observed a marginally significant difference in age acceleration between prospective cases compared to matched controls (Fig. 1b, $P = 0.05$, Supplementary Fig. 1). Stratified analysis based on time from blood collection to disease diagnosis revealed that prospective breast cancers exhibited age acceleration 10 years prior to diagnosis compared to matched control samples (Fig. 1c, $P = 0.01$).

A conditional logistic regression model that relates breast cancer status to IEAA showed that IEAA was associated (Table 3) with breast cancer status. The results were not attenuated after adjusting for known breast cancer factors (Supplementary Table 1). Each unit increase in IEAA led to 4% increased odds of being a breast cancer case (OR, 1.04; 95% CI, 1.007–1.076, $P = 0.016$) (Table 3). IEAA follows an approximately normal distribution with mean zero, variance = 28.2, standard deviation of 5.31. The following quantiles describe the empirical distribution of IEAA: minimum = −24.2, maximum 24.4, median = −0.12, first quartile = −3.0, third quartile = 3.0. Thus, 25% of women had an IEAA value > 3.
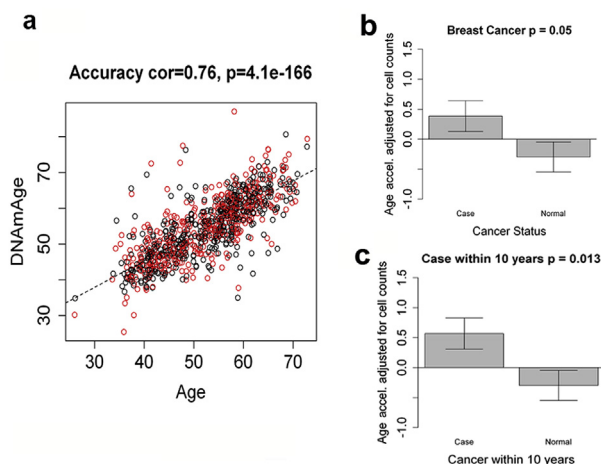

Fig. 1. Epigenetic clock analysis. a) DNA methylation age (y-axis) versus chronological age (x-axis). Points correspond to female subjects. Red indicates breast cancer case, black control. The dashed line indicates a regression line, b) epigenetic age acceleration versus breast cancer status. Each bar plot depicts the mean value, standard deviation and reports a non-parametric group comparison test p-value (Wilcoxon test), c) epigenetic age acceleration versus breast cancer status (developed within 10 years post blood draw). Each bar plot depicts the mean value, standard deviation and reports a non-parametric group comparison test p-value (Wilcoxon test). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Logistic regression analysis of IEAA for incident breast cancer status.

	Univariate analysis OR (95% CI)	Multivariate analysis[a] OR (95% CI)
All samples		
IEAA	1.04 (1.007–1.075)	1.04 (1.007–1.076)
Premenopausal samples		
IEAA	1.00 (0.9572–1.06)	1.00 (0.9510–1.056)
Postmenopausal samples		
IEAA	1.06 (1.019–1.11)	1.07 (1.020–1.11)

OR: Odds Ratio; CI: Confidence Interval; IEAA: Intrinsic Epigenetic Age Acceleration.
[a] Odds ratios were adjusted for physical activity (inactive, moderately inactive, moderately active and active).

None of the blood cell count measures were associated with disease status in prediagnostic blood samples (Supplementary Fig. 2). Interestingly, high physical activity was associated with decreased odds of being a breast cancer case (Supplementary Table 1).

A recent study demonstrated that menopause has a weak but statistically significant effect on epigenetic age acceleration. Further, menopause has been known to accelerate age-related diseases including breast cancer [34,35]. To adjust for menopausal status, we evaluated the association between IEAA and breast cancer in separate strata defined by menopausal status (premenopausal and postmenopausal). The baseline characteristics of premenopausal and postmenopausal breast samples are shown in Supplementary Table 2. We observed a positive correlation between epigenetic and chronological age in postmenopausal samples (Fig. 2a). Stratified analysis of postmenopausal breast cancers based on the lead-time between blood collection and cancer diagnosis revealed that breast cancers had a higher IEAA compared to non-cancer samples (Fig. 2b, Supplementary Fig. 3).

A very high value of IEAA = 10 is associated with a doubling of odds of developing postmenopausal breast cancer (OR = 1.97 (1.22–2.83) calculated as 1.07^{10} from our multivariate logistic regression model Table 3). Twenty-five percent of all women exhibit an age acceleration larger than 3 which is associated with 22% increase in the odds of developing postmenopausal breast cancer (OR = 1.22 (1.06–1.37) calculated as 1.07^3).

We found that breast cancer that developed within 10 years from date of recruitment had a stronger association with IEAA (Fig. 2c). However, the results of this secondary analysis should be interpreted with caution due to an inflated false positive rate resulting from multiple comparisons. We did not observe such associations in premenopausal breast samples (Supplementary Figs. 4, 5). Similar to our findings in all breast samples, high physical activity was associated with decreased odds of being a breast cancer case in postmenopausal women (Supplementary Table 3).
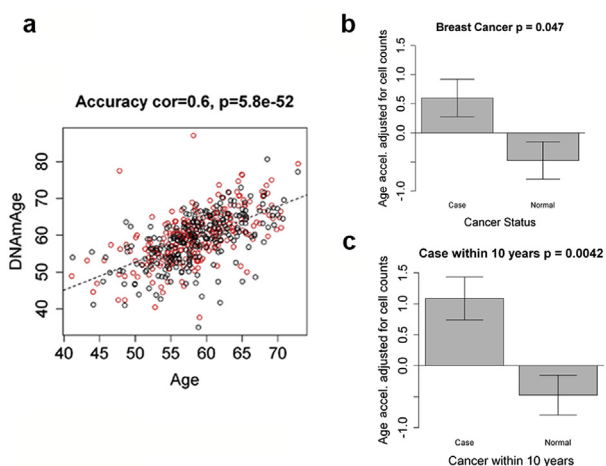
Fig. 2. Epigenetic clock analysis for postmenopausal samples. a) DNA methylation age (y-axis) versus chronological age (x-axis). Points correspond to female subjects. Red indicates breast cancer case, black control. The dashed line indicates a regression line; b) epigenetic age acceleration versus breast cancer status. Each bar plot depicts the mean value, standard deviation and reports a nonparametric group comparison test p-value (Wilcoxon test); c) epigenetic age acceleration versus breast cancer status (developed within 10 years post blood draw). Each bar plot depicts the mean value, standard deviation, and reports a non-parametric group comparison test p-value (Wilcoxon test). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Interestingly, we observed a highly significant association between IEAA and incident postmenopausal breast cancers (OR, 1.07; 95% CI, 1.020−1.11, $P = 0.003$). By contrast, no significant association could be observed for incident premenopausal breast cancers (OR, 1.00; 95% CI, 0.9510−1.056, $P = 0.94$) (Table 3).

4. Discussion

Using a rigorous and large-scale nested prospective case−control study, we demonstrate that: (1) IEAA in blood increases the odds of developing postmenopausal breast cancers and (2) genome-wide hypermethylation in CpG islands is associated with incident breast cancer cases. While several articles have studied blood methylation data versus breast cancer risk [36−39], it appears that ours is the first study to detect a weak but significant association of IEAA with breast cancer susceptibility. Our study stands out in terms of its large sample size, its use of a robust epigenome wide technology (Illumina 450K array), the careful matching of breast cancer cases with controls in a prospective case−control study, and its use of a powerful epigenetic biomarker of ageing, which is independent of blood cell counts (IEAA).

Our finding regarding the association between global CpG island methylation levels and breast cancer risk is congruent with the findings from our earlier retrospective study on breast cancer [39] and supports the notion that regulatory regions of the genome are often hypermethylated in cancer cells [29]. It is noteworthy that we observed CpG island hypermethylation in blood tissue samples of incident breast cancer patients. Several epidemiological case−control studies have reported global genomic hypomethylation in peripheral blood of cancer patients, suggesting a systemic effect of hypomethylation on disease predisposition [40,41]. In addition, two recent studies reported a lower global methylation levels in prospectively collected blood samples from breast cancer cases compared to controls [38,42]. However, we did not find any change in global DNAm levels between cases and controls. These discrepancies may be due to technical and biological variations attributable to the low power of the studies.

Epigenetic changes are ubiquitous in primary breast cancers although the role of deregulation of the epigenome is largely unknown. It has been suggested that a gradual accumulation of methylation changes ('epigenetic drift') may occur through stochastic events, resulting in clonal expansion of the stem/progenitor cells, and that this process may contribute to the age-associated increase in risk of developing breast cancer [43−45]. DNAm age is highly correlated to chronological age across sorted cell types (CD4 T cells, monocytes, B cells, glial cells, neurons), complex tissues (e.g. blood) and organs (brain, breast, kidney, liver, lung) [4]. Our findings were consistent with the previous studies in different tissues [4,16]. The epigenetic clock derived from the DNAm age is robust with respect to the batch effects and can be applied to all Illumina array platforms: the EPIC chip (850K), the Illumina 450K array and the 27K array [4] and possibly measures a cell intrinsic and tissue independent epigenetic drift [46]. For blood derived DNA measured on the Illumina 450K array, the epigenetic clock algorithm provides not only several measures of age acceleration but also estimates of blood cell counts. One of the major concerns regarding age-associated DNAm signatures is the influence of tissue's cellular composition which may alter with age. We found no differences in leucocyte subpopulations between cases and controls. By definition, our intrinsic measure of epigenetic age acceleration (IEAA) is not confounded by changes in the proportion of blood cell counts (Methods). We focussed on IEAA as it has been shown to be correlated with epigenetic age acceleration in breast tissue [9]. Future research could investigate whether epigenetic age acceleration of breast tissue is predictive of breast cancer.

We can only speculate when it comes to explaining why IEAA was only predictive of postmenopausal breast cancer but not of premenopausal breast cancer. Breast cancers developing in postmenopausal women are influenced by specific polymorphisms in endogenous steroid hormone metabolic pathways and exogenous administration of hormones at menopause (hormone

replacement therapy). Our observed age acceleration in postmenopausal breast cancers might reflect differences in hormone exposure. In this context, it is noteworthy that both natural and surgical menopause are associated with an increase in intrinsic age acceleration [18]. In addition, age-associated compromised detoxification, DNA repair mechanisms and immune surveillance may add to the endogenous factors which could lead to postmenopausal breast cancer development [1]. It is unlikely that smoking and BMI confound the relationship between epigenetic age and breast cancer risk because : (1) BMI and smoking have only a very weak effect on the epigenetic age acceleration of blood tissue (correlation r < 0.10) [16,20], and (2) we could detect accelerated ageing effects in multivariate regression models that adjusted for these potential confounders. Our results based on a prospective study cohort points to a higher rate of ageing in the blood samples from individuals who develop breast cancer compared to the controls. While the results from our epigenetic age analysis are biologically meaningful, the association between DNAm age and disease risk is probably too weak for prognostic purposes.

In the present study, we demonstrated that a surrogate tissue (blood) captures accelerated ageing effects and relates to an effector (breast cancer) of ageing. We have demonstrated that IEAA was associated with postmenopausal breast cancer susceptibility and identified potential epigenetics-based biomarkers for risk stratification. Because menopause has been known to accelerate age-related diseases including cancer, our finding also suggest potential underlying mechanism and provides biological plausibility to the association between menopause and cancer risk. Further research aimed at understanding epigenome deregulation in cancer causation, risk stratification and the mechanism underlying accelerated epigenetic clock is warranted.

Role of funding resource

Conflict of interest statement

The Regents of the University of California is the sole owner of a patent application directed at the invention of measures of epigenetic age acceleration for which Steve Horvath is a named inventor. The other authors declare no conflict of interest.

Funding

Acknowledgements

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ejca.2017.01.014.

References

[1] Benz CC. Impact of aging on the biology of breast cancer. Crit Rev Oncol Hematol 2008;66(1):65−74.

[2] Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, McLaren S, et al. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. Science 2015;348(6237):880−6.

[3] Deelen J, Beekman M, Capri M, Franceschi C, Slagboom PE. Identifying the genomic determinants of aging and longevity in human population studies: progress and challenges. Bioessays 2013;35(4):386−96.

[4] Horvath S. DNA methylation age of human tissues and cell types. Genome Biol 2013;14(10):R115.

[5] Marioni RE, Shah S, McRae AF, Ritchie SJ, Muniz-Terrera G, Harris SE, et al. The epigenetic clock is correlated with physical and cognitive fitness in the Lothian Birth Cohort 1936. Int J Epidemiol 2015;44(4):1388−96.

[6] Christiansen L, Lenart A, Tan Q, Vaupel JW, Aviv A, McGue M, et al. DNA methylation age is associated with mortality in a longitudinal Danish twin study. Aging Cell 2016;15(1):149−54.

[7] Perna L, Zhang Y, Mons U, Holleczek B, Saum KU, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. Clin Epigenetics 2016;8:64.

[8] Horvath S, Pirazzini C, Bacalini MG, Gentilini D, Di Blasio AM, Delledonne M, et al. Decreased epigenetic age of PBMCs from Italian semi-supercentenarians and their offspring. Aging (Albany NY) 2015;7(12):1159−70.

[9] Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. Aging (Albany NY) 2016;8(9):1844−65.

[10] Breitling LP, Saum KU, Perna L, Schottker B, Holleczek B, Brenner H. Frailty is associated with the epigenetic clock but not with telomere length in a German cohort. Clin Epigenetics 2016;8:21.

[11] Levine ME, Lu AT, Bennett DA, Horvath S. Epigenetic age of the pre-frontal cortex is associated with neuritic plaques, amyloid load, and Alzheimer's disease related cognitive functioning. Aging (Albany NY) 2015;7(12):1198−211.

[12] Simpkin AJ, Hemani G, Suderman M, Gaunt TR, Lyttleton O, McArdle WL, et al. Prenatal and early life influences on epigenetic age in children: a study of mother−offspring pairs from two cohort studies. Hum Mol Genet 2016;25(1):191−201.

[13] Horvath S, Garagnani P, Bacalini MG, Pirazzini C, Salvioli S, Gentilini D, et al. Accelerated epigenetic aging in Down syndrome. Aging Cell 2015;14(3):491−5.

[14] Horvath S, Levine AJ. HIV-1 infection accelerates age according to the epigenetic clock. J Infect Dis 2015;212(10):1563−73.

[15] Horvath S, Langfelder P, Kwak S, Aaronson J, Rosinski J, Vogt TF, et al. Huntington's disease accelerates epigenetic aging of human brain and disrupts DNA methylation levels. Aging (Albany NY) 2016;8(7):1485−512.

[16] Horvath S, Erhart W, Brosch M, Ammerpohl O, von Schonfels W, Ahrens M, et al. Obesity accelerates epigenetic aging of human liver. Proc Natl Acad Sci U S A 2014;111(43):15538−43.

[17] Zannas AS, Arloth J, Carrillo-Roa T, Iurato S, Roh S, Ressler KJ, et al. Lifetime stress accelerates epigenetic aging in an urban, African American cohort: relevance of glucocorticoid signaling. Genome Biol 2015;16:266.

[18] Levine ME, Lu AT, Chen BH, Hernandez DG, Singleton AB, Ferrucci L, et al. Menopause accelerates biological aging. Proc Natl Acad Sci U S A 2016;113(33):9327−32.

[19] Horvath S, Ritz BR. Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients. Aging (Albany NY) 2015;7(12):1130−42.

[20] Levine ME, Hosgood HD, Chen B, Absher D, Assimes T, Horvath S. DNA methylation age of blood predicts future onset of lung cancer in the women's health initiative. Aging (Albany NY) 2015;7(9):690−700.

[21] Bingham S, Riboli E. Diet and cancer−the European prospective investigation into cancer and nutrition. Nat Rev Cancer 2004; 4(3):206−15.

[22] Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, et al. European prospective investigation into cancer and nutrition (EPIC): study populations and data collection. Public Health Nutr 2002;5(6B):1113−24.

[23] Ambatipudi S, Cuenin C, Hernandez-Vargas H, Ghantous A, Calvez-Kelm FL, Kaaks R, et al. Tobacco smoking-associated genome-wide DNA methylation changes in the EPIC study. Epigenomics 2016;8(5):599−618.

[24] Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004; 5(10):R80.

[25] Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics 2013;8(2):203−9.

[26] Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics 2013;29(2):189−96.

[27] Houseman AE, Accomando PW, Koestler CD, Christensen CB, Marsit JC, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinforma 2012:13.

[28] Horvath S, Gurven M, Levine ME, Trumble BC, Kaplan H, Allayee H, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. Genome Biol 2016;17(1):171.

[29] Esteller M. Epigenetics in cancer. N Engl J Med 2008;358(11): 1148−59.

[30] Martin M, Ancey PB, Cros MP, Durand G, Le Calvez-Kelm F, Hernandez-Vargas H, et al. Dynamic imbalance between cancer cell subpopulations induced by transforming growth factor beta (TGF-beta) is associated with a DNA methylome switch. BMC Genomics 2014;15:435.

[31] Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics 2011;98(4):288−95.

[32] Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 2009;41(2):178−86.

[33] Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics 2011; 6(6):692−702.

[34] Blagosklonny MV. Why men age faster but reproduce longer than women: mTOR and evolutionary perspectives. Aging (Albany NY) 2010;2(5):265−73.

[35] Horiuchi S. Postmenopausal acceleration of age-related mortality increase. J Gerontol A Biol Sci Med Sci 1997;52(1):B78−92.

[36] Demetriou CA, Chen J, Polidoro S, van Veldhoven K, Cuenin C, Campanella G, et al. Methylome analysis and epigenetic changes associated with menarcheal age. PLoS One 2013;8(11):e79391.

[37] Heyn H, Carmona FJ, Gomez A, Ferreira HJ, Bell JT, Sayols S, et al. DNA methylation profiling in breast cancer discordant identical twins identifies DOK7 as novel epigenetic biomarker. Carcinogenesis 2013;34(1):102−8.

[38] Severi G, Southey MC, English DR, Jung CH, Lonie A, McLean C, et al. Epigenome-wide methylation in DNA from peripheral blood as a marker of risk for breast cancer. Breast Cancer Res Treat 2014;148(3):665−73.

[39] Xu X, Gammon DM, Hernandez-Vargas H, Herceg Z, Wetmur GJ, Teitelbaum LS, et al. DNA methylation in peripheral blood measured by LUMA is associated with breast cancer in a population-based study. FASEB J 2012;26:2657−66.

[40] Wallace K, Grau MV, Levine AJ, Shen L, Hamdan R, Chen X, et al. Association between folate levels and CpG Island hypermethylation in normal colorectal mucosa. Cancer Prev Res (Phila) 2010;3(12):1552−64.

[41] Kuchiba A, Iwasaki M, Ono H, Kasuga Y, Yokoyama S, Onuma H, et al. Global methylation levels in peripheral blood leukocyte DNA by LUMA and breast cancer: a case-control study in Japanese women. Br J Cancer 2014;110(11):2765−71.

[42] van Veldhoven K, Polidoro S, Baglietto L, Severi G, Sacerdote C, Panico S, et al. Epigenome-wide association study reveals decreased average methylation levels years before breast cancer diagnosis. Clin Epigenetics 2015;7(1):67.

[43] Issa JP. Aging, DNA methylation and cancer. Crit Rev Oncol Hematol 1999;32(1):31−43.

[44] Langevin SM, Pinney SM, Leung YK, Ho SM. Does epigenetic drift contribute to age-related increases in breast cancer risk? Epigenomics 2014;6(4):367−9.

[45] Issa JP. Aging and epigenetic drift: a vicious cycle. J Clin Invest 2014;124(1):24−9.

[46] Zheng Y, Joyce BT, Colicino E, Liu L, Zhang W, Dai Q, et al. Blood epigenetic age may predict cancer incidence and mortality. EBioMedicine 2016;5:68−73.