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,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. Nost s, Elisabete Weiderpass Vainio s,l,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 Épidémiologie 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 Melbourne, 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).
DNA methylation; Epigenomics; Age acceleration; Breast cancer; Biomarkers; Prospective studies

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.

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.

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.

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

ª 2017 Elsevier Ltd. All rights reserved.
1. Introduction

Aging 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 hypothesized 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 aging 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 multissue biomarker of aging [4]. The DNAm-based marker of aging (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 aging, 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 sub-sites (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 \( \beta \) 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 (BM IQ) 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 leukocytes). 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>
<thead>
<tr>
<th><strong>Table 1</strong></th>
<th>Characteristics of incident breast cancer and control participants at baseline (i.e. time of blood collection).</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All samples</td>
</tr>
<tr>
<td></td>
<td>Controls (%)</td>
</tr>
<tr>
<td>Sample size</td>
<td>451</td>
</tr>
<tr>
<td>Mean methylation (in %)</td>
<td>51.86</td>
</tr>
<tr>
<td>Age (years)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td></td>
<td>Median</td>
</tr>
<tr>
<td>Alcohol consumption (g/d)</td>
<td>Mean(SD)</td>
</tr>
<tr>
<td>Age at menarche</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>BMI</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Physical activity (Cambridge index)</td>
<td></td>
</tr>
<tr>
<td>Sedentary</td>
<td>99 (22.0)</td>
</tr>
<tr>
<td>Moderately sedentary</td>
<td>187 (41.5)</td>
</tr>
<tr>
<td>Moderately active</td>
<td>76 (16.9)</td>
</tr>
<tr>
<td>Active</td>
<td>78 (17.3)</td>
</tr>
<tr>
<td>Missing</td>
<td>11 (2.4)</td>
</tr>
<tr>
<td>Hormone receptor status</td>
<td></td>
</tr>
<tr>
<td>ER(^+)/PR(^+)/Her2(^+)</td>
<td>–</td>
</tr>
<tr>
<td>ER(^+)/PR(^+)/Her2(^-)</td>
<td>–</td>
</tr>
<tr>
<td>ER(^+)/Her2(^-)</td>
<td>–</td>
</tr>
<tr>
<td>Country</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>160 (35.5)</td>
</tr>
<tr>
<td>Spain</td>
<td>27 (6.0)</td>
</tr>
<tr>
<td>UK</td>
<td>38 (8.4)</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>66 (14.6)</td>
</tr>
<tr>
<td>Greece</td>
<td>25 (5.5)</td>
</tr>
<tr>
<td>Germany</td>
<td>135 (29.9)</td>
</tr>
</tbody>
</table>

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 for variation to 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/SD, 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>
<thead>
<tr>
<th>Context</th>
<th># CpGs</th>
<th>Std. dev.</th>
<th>OR (95% CI)$^a$</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All CpG sites</td>
<td>421,583</td>
<td>3.45E-04</td>
<td>1.09 (0.94–1.25)</td>
<td>0.21</td>
</tr>
<tr>
<td>Islands</td>
<td>130,982</td>
<td>5.87E-04</td>
<td>1.20 (1.03–1.40)</td>
<td>0.02</td>
</tr>
<tr>
<td>Open Sea</td>
<td>150,852</td>
<td>4.50E-03</td>
<td>1.49 (0.36–6.24)</td>
<td>0.58</td>
</tr>
<tr>
<td>CpG</td>
<td>40,948</td>
<td>4.88E-04</td>
<td>0.89 (0.78–1.02)</td>
<td>0.10</td>
</tr>
<tr>
<td>context</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shore</td>
<td>98,801</td>
<td>5.40E-04</td>
<td>1.06 (0.87–1.16)</td>
<td>0.97</td>
</tr>
<tr>
<td>Distal promoter</td>
<td>19,990</td>
<td>5.24E-04</td>
<td>1.06 (0.92–1.21)</td>
<td>0.44</td>
</tr>
<tr>
<td>Distal sequence</td>
<td>7,828</td>
<td>6.68E-04</td>
<td>0.96 (0.84–1.09)</td>
<td>0.52</td>
</tr>
<tr>
<td>Genic</td>
<td>168,460</td>
<td>3.80E-04</td>
<td>1.02 (0.89–1.18)</td>
<td>0.76</td>
</tr>
<tr>
<td>Gene Body</td>
<td>54,467</td>
<td>3.53E-04</td>
<td>1.02 (0.89–1.17)</td>
<td>0.76</td>
</tr>
<tr>
<td>context</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intergenic</td>
<td>56,403</td>
<td>5.35E-04</td>
<td>1.04 (0.89–1.21)</td>
<td>0.44</td>
</tr>
<tr>
<td>Proximal promoter</td>
<td>168,337</td>
<td>5.26E-04</td>
<td>1.15 (0.99–1.34)</td>
<td>0.07</td>
</tr>
</tbody>
</table>

$^a$ 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.

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 1.073). 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 1.073).

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).

---

### Table 3

**Logistic regression analysis of IEAA for incident breast cancer status.**

<table>
<thead>
<tr>
<th></th>
<th>Univariate analysis</th>
<th>Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>All samples</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IEAA</td>
<td>1.04 (1.007–1.075)</td>
<td>1.04 (1.007–1.076)</td>
</tr>
<tr>
<td>Premenopausal samples</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IEAA</td>
<td>1.00 (0.957–1.06)</td>
<td>1.00 (0.9510–1.056)</td>
</tr>
<tr>
<td>Postmenopausal samples</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IEAA</td>
<td>1.06 (1.019–1.11)</td>
<td>1.07 (1.020–1.11)</td>
</tr>
</tbody>
</table>

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

---

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.)
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, 0.951–0.94) (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 DNA methylation 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
Role of funding resource

The funders of the study had no role in study design, data collection, data analysis, or writing of the manuscript.

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

This work was supported by grants from the Institut National du Cancer (INCa, France) (2012-070) to IR and ZH and the European Commission (EC) Seventh Framework Programme (FP7 Exposomics) (308610-2) Translational Cancer Research (TRANS201301184) Framework and the Fondation Association pour la Recherche contre le Cancer (ARC, France) to ZH. ZH was also supported by the EC FP7 EurocanPlatform: A European Platform for Translational Cancer Research (grant number: 260791). Funding for the work also comes from a grant from: “Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy” and the Regional Government of Asturias. EPIC Greece is supported by the Hellenic Health Foundation. SH was supported by NIH/NIA 1U34AG051425-01. The funders of the study had no role in study design, data collection, data analysis, data interpretation or writing of the manuscript.

Acknowledgements

The work reported in this article was undertaken during the tenure of a Postdoctoral Fellowship (to SA) from the International Agency for Research on Cancer, partially supported by the EC FP7 Marie Curie Actions — People — Co-funding of regional, national and international programmes (COFUND). A.S. is supported by the PhD fellowship from the Fonds National de la Recherche, Luxembourg (AFR Code: 10100060). This study depended on the participation of the women in the EPIC cohort, to whom we are grateful.

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


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.


