Each Pregnancy Linearly Changes Immune Gene Expression in the Blood of Healthy Women Compared with Breast Cancer Patients
Eiliv Lund, Aurelie Nakamura, Igor Snapkov, Jean-Christophe Thalabard, Karina Standahl Olsen, Lars Holden, Marit Holden

To cite this version:
Eiliv Lund, Aurelie Nakamura, Igor Snapkov, Jean-Christophe Thalabard, Karina Standahl Olsen, et al.. Each Pregnancy Linearly Changes Immune Gene Expression in the Blood of Healthy Women Compared with Breast Cancer Patients. Clinical Epidemiology, 2018, 10, pp.931–940. 10.2147/CLEP.S163208. hal-03833313

HAL Id: hal-03833313
https://hal.sorbonne-universite.fr/hal-03833313
Submitted on 2 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Each pregnancy linearly changes immune gene expression in the blood of healthy women compared with breast cancer patients

Eiliv Lund¹,²
Aurelie Nakamura¹,³,⁴
Igor Snapkov¹
Jean-Christophe Thalabard⁵
Karina Standahl Olsen¹
Lars Holden⁶
Marit Holden⁶

¹Department of Community Medicine, UiT The Arctic University of Norway, Tromso, Norway; ²The Cancer Registry of Norway, Oslo, Norway; ³Department of Social Epidemiology, Pierre Louis Institute of Epidemiology and Public Health, Sorbonne University, INSERM, Paris, France; ⁴French School of Public Health (EHESP), Doctoral Network, Rennes, France; ⁵MAPS, UMR CNRS 8145, Université Paris Descartes, Sorbonne Paris Cite, Paris, France; ⁶Norwegian Computing Center, Oslo, Norway

Background: There is a large body of evidence demonstrating long-lasting protective effect of each full-term pregnancy (FTP) on the development of breast cancer (BC) later in life, a phenomenon that could be related to both hormonal and immunological changes during pregnancies. In this work, we studied the pregnancy-associated differences in peripheral blood gene expression profiles between healthy women and women diagnosed with BC in a prospective design.

Methods: Using an integrated system epidemiology approach, we modeled BC incidence as a function of parity in the Norwegian Women and Cancer (NOWAC) cohort (165,000 women) and then tested the resulting mathematical model using gene expression profiles in blood in a nested case–control study (460 invasive case–control pairs) of women from the NOWAC postgenome cohort. Lastly, we undertook a gene set enrichment analysis for immunogenous gene sets.

Results: A linear trend fitted the dataset precisely showing an 8% decrease in risk of BC for each FTP, independent of stratification on other risk factors and lasting for decades after a woman’s last FTP. Women with six children demonstrated 48% reduction in the incidence of BC compared to nulliparous. When we looked at gene expression, we found that 756 genes showed linear trends in cancer-free controls (false discovery rate [FDR] 5%), but this was not the case for any of the genes in BC cases. Gene set enrichment analysis of immunological gene sets (C7 collection in Molecular Signatures Database) revealed 215 significantly enriched human gene sets (FDR 5%).

Conclusion: We found marked differences in gene expression and enrichment profiles of immunological gene sets between BC cases and healthy controls, suggesting an important protective effect of the immune system on BC risk.

Keywords: breast cancer, Norwegian Women and Cancer Study, gene expression, parity, semi-allograft, hormones, pregnancy

Introduction

Each pregnancy represents a unique and remarkable challenge to the mother’s immune system, as her body must accept and carry a fetus to term without immune rejection. Immunological tolerance during pregnancy is necessary for women’s bodies to accept the fetus.¹,² This obligate tolerance has been observed for many years, and because of it, each fetus can be seen as a different semi-allograft. Interestingly, “immune evasion”, which consists of some similar immunological changes, was included as another hallmark of cancer,³ and researchers have recognized some common characteristics between these two highly different biological processes.⁴–⁷ Looking at this, it seems paradoxical that increasing number of full-term pregnancies reduces the risk of breast cancer (BC) in women.⁸,⁹ Some studies have even shown a protective effect up to eight
or nine children. However, neither spontaneous nor induced abortions confer the same risk reduction.

The role of hormones is well established in both pregnancy and tumorigenesis; however, the influence of hormonal fluctuations associated with pregnancy on the cancer development in a long-term perspective remains unclear. High, changing levels of hormones such as estrogens, gestagens, and human chorionic hormone initiate and support immunological tolerance during pregnancy, whereas the high levels of estrogens and progestins found in oral contraceptives and hormone replacement therapy are classified as carcinogens.

We have previously shown that blood gene expression profiles differ between women with BC and BC-free controls, both at time of diagnosis and during the years prior to diagnosis. Additionally, in a recent study, we compared gene expression in blood and breast tumor tissue. We revealed some similarities in expression profiles mainly for immunogenic tumors; however, we demonstrated that there is no obligate common gene expression pattern.

Here, we aim to explore changes in blood gene expression in women with BC and cancer-free controls according to number of full-term pregnancies (FTPs).

**Methods**

**Norwegian Women and Cancer (NOWAC) study**

The NOWAC study is a national population-based study in which a random sample of women living in Norway was recruited starting in 1991. Among the 172,748 women invited to participate, 165,227 women replied to at least one questionnaire and make up the study population. Death and emigration status of these women were updated through linkage to the Cause of Death Registry at Statistics Norway and Central Population Registry, respectively. Information on parity was taken from the NOWAC questionnaires. The date of entry into the present study was set as the date of the first questionnaire and make up the study population. Death and emigration status of these women were updated through linkage to the Cause of Death Registry at Statistics Norway and Central Population Registry, respectively. Information on parity was taken from the NOWAC questionnaires.

Altogether 546 cases were identified prospectively, ie, developed BC after blood sample donation.

For each BC case, a control from the NOWAC post-genome cohort was assigned, matched by time of blood sampling and year of birth, to be analyzed together with the case. The controls were used to establish the average (mean) gene expression in individuals without cancer and to serve as the basis for exposure-adjusted analyses. The expression level of a gene not involved in the carcinogenic process varied depending on day-to-day changes in exposures such as environment and nutrition, resulting in random fluctuations of the difference in gene expression between cases and matched controls around a population-average constant over time. On the other hand, the difference in the expression of genes related to different stages of the carcinogenic process varied over time in a nonrandom way, thus exhibiting a non-random trend. Moreover, the changes in the expression of genes related to the carcinogenesis could be complicated by other effects caused by exposure to carcinogens.

**Laboratory procedures**

All extraction and microarray services were provided by the Genomics Core Facility, Norwegian University of Science and Technology, Trondheim, Norway. To control for technical variability such as different batches of reagents and kits, day-to-day variations, microarray production batches, and effects related to different laboratory operators, each case–control pair was kept together throughout all extraction, amplification, and hybridization procedures. RNA extraction was performed using the PAXgene Blood mRNA Isolation kit (Preanalytix/Qiagen) according to the manufacturer’s instructions. RNA quality and purity were assessed using the NanoDrop ND 8000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), respectively. RNA amplification was performed on 96-well plates using 300 ng of total RNA and the Illumina TotalPrep-96 RNA Amplification Kit (Ambio, Inc., Austin, TX, USA). The amplification procedure consisted of reverse transcription with a T7 promoter and ArrayScript, followed by a second-strand synthesis. In vitro transcription with T7 RNA polymerase using a biotin–nucleoside triphosphates mix produced biotinylated cRNA. All case–control pairs were run on either the IlluminaHumanWG-6 version 3 expression bead chips or the HumanHT-12 version 4 expression bead chips. Outliers were excluded after visual examination of dendrograms, principal component analysis (PCA) plots and density plots. Individuals who were considered borderline outliers were
excluded if their laboratory quality measures were below given thresholds (RNA integrity number value <7, 260/280 ratio <2, 260/230 ratio <1.7, and 50 < RNA < 500).

Statistical methods
The first stage of the analysis was to explore the relationship between full-term pregnancies and BC in the whole NOWAC cohort. Based on person-years, the percentage change in incidence rates of BC for each additional child was calculated by using a Poisson linear regression. Each covariate was analyzed separately in a stratified analysis with test for interactions using the Wald test. A Cox proportional semiparametric hazard analysis with attained age as follow-up was used for estimating the overall effect of potential confounders.

Then, gene expression profiles were analyzed for BC cases and controls using microarray technology. The dataset with 30,046 probes for each individual was preprocessed as previously described. The dataset was background corrected using negative control probes, log2 transformed using a variance stabilizing technique, and quantile normalized. Data from the two Illumina chips (HumanWG-6 v3 and HumanHT-12 v4) were combined on identical nucleotide universal identifiers. We retained probes present in at least 70% of the individuals. If a gene was represented with more than one probe, the average expression level of the probes was used as the expression for the gene, resulting in a dataset with 8155 genes. The probes were translated to genes using the lumiHumanIDMapping database. Finally, the differences in the log2 gene expression for each case–control pair were computed using the Bioconductor R-package Limma in the same way as we did for individual genes by using enrichment scores for gene sets instead of differences in the log2 gene expressions as responses in the linear model. The enrichment scores for gene sets were obtained from the differences in log2 gene expressions using the Bioconductor R-package gene set variation analysis (GSVA).

Ethics approval and consent to participate
The NOWAC study was approved by the Norwegian Data Inspectorate and the Regional Ethical Committee of North Norway (REK). The study was conducted in compliance with the Declaration of Helsinki, and all participants gave written informed consent. The linkages of the NOWAC database to national registries such as the Cancer Registry of Norway and registries on death and emigration were approved by the Directorate of Health. The women were informed about these linkages. Furthermore, the collection and storing of human biological material was approved by the REK in accordance with the Norwegian Biobank Act. Women were informed in the letter of introduction that the blood samples would be used for gene expression analyses.

Results
Reduction of BC risk after full-term pregnancies
Our first step was to estimate the mathematical relationship between parity and BC incidence in the complete NOWAC cohort, which consisted of 165,227 participants and 6536 registered incident BC cases identified through linkage to the Cancer Registry of Norway. Increasing parity was related to less BC in mothers, less use of hormone replacement therapy, larger body mass index (BMI), earlier first birth, more breast feeding, and less use of oral contraceptives (Table 1). Among both premenopausal and postmenopausal women, a regular decrease in incidence rates was found (Figure 1 and Table S2), and in linear regression analyses of BC incidence in the complete NOWAC cohort, each new pregnancy conferred a strong linear reduction in the absolute risk (8.0%; Figure 1). This led to an observed reduction of 48% for a woman with six children.

Differences in BC incidence rates were of the same order between nulliparous and uniparous women and for women with five or six children, illustrating an additive effect. The linearity of BC incidence was equally clear regardless of menopausal status, showing the persistent effect of age,
but postmenopausal women had higher BC incidence rates (Figure 1). Stratification on all risk factors (current use of hormone replacement therapy, BMI, age at first birth, breastfeeding, ever use of oral contraceptives, and maternal history of BC) revealed the same linear patterns (Table S3). There were no significant interactions between different strata of the risk factors (none in the Wald test). Using the Cox’s multivariate hazard model adjusted for the same risk factors (taken from Table S3), the crude estimated relative risk (RR) in women with six children versus nulliparous women was 0.50 (95% CI 0.37–0.66) and adjusted RR was 0.49 (95% CI 0.36–0.68), similar to the ones derived from Poisson regression analysis (Table S4).

Differential pregnancy-associated blood gene expression between cases and controls

This linear relationship independent of other covariates was then used to test the hypothesis that blood gene expression of some genes increased or decreased linearly according to parity.

An overall analysis without parity status showed no significant differences in expressed genes between cases and controls. When we added one parity variable for the cases and one for the controls in the model, we found a significant linear relationship between log gene expression and parity for 756 genes in BC-free controls, but we found no such relationship in BC cases (false discovery rate [FDR] 5%). Among the controls, 96 of 756 significant genes had at least 5% change per child (91 were upregulated and five were downregulated) and four genes had at least 10% change per child (all were upregulated). A description of potential carcinogenic effects in BC of the ten top significant genes showed (Table 2) that four of them are tumor promoter genes and one is a tumor suppressor gene; however, the function of five significant genes was unknown.

Gene set enrichment analysis

Table 3 demonstrates the results of gene set enrichment analysis. We identified 588 gene sets from the C7 collection in Molecular Signatures Database (MSigDB; http://software.broadinstitute.org/gsea/index.jsp)27 that were significantly enriched when the parity of the controls varied (FDR 5%).

Table 1 Description of the entire study population according to parity

| Characteristics                       | Parity (number of children) |
|---------------------------------------|-----------------------------|
|                                       | 0   | 1–2 | 3–4 | 5–6 |
| Maternal history of BC                | 5.9 | 5.2 | 5.2 | 4.5 |
| Current use of hormone replacement therapy | 17.4 | 17.1 | 14.7 | 12.7 |
| BMI <25 kg/m²                         | 58.7 | 58.5 | 53.7 | 43.8 |
| Age at first birth <25 years          | NA  | 53.2 | 70.8 | 82.2 |
| Years of lactation                    |     |     |     |     |
| Never                                 | 100.0 | 7.8 | 4.1 | 3.1 |
| <2 years                              | NA  | 88.3 | 69.5 | 47.8 |
| 2+                                    | NA  | 3.8 | 26.4 | 49.1 |
| Ever use of oral contraceptives       | 47.6 | 59.6 | 52.7 | 37.5 |

Note: Numbers represent a percentage of women within the corresponding parity category.

Abbreviations: BC, breast cancer; BMI, body mass index; NA, data not available.

Figure 1 BC incidence rates in the NOWAC study, 1991–2013.

Note: 95% Cls and linear trends according to parity (A) and parity and menopausal status (B; n = 165,227, BC cases = 6,536, PY = 2,390,385).

Abbreviations: BC, breast cancer; CI, confidence interval; NOWAC, Norwegian Women and Cancer; PY, person-years.
Experimentally produced gene sets were submitted to MSigDB from researchers using both in vivo and in vitro materials, as well as both human cells or tissues and animal models. Of our 588 enriched gene sets, 215 were derived from human data and 373 were derived mostly from mouse data.

Detailed characterization of the top 10 gene sets from each species showed a discrepancy in the tissues analyzed: blood and cord blood in humans versus bone marrow and spleen in mice (Tables 4 and S5). We had information on gender for only one of the top 10 human gene sets and four of the mouse gene sets (after contacting main authors of the associated publications). None of the MSigDB gene sets from either humans or mice included information on the number of full-term pregnancies (Table 4).

### Discussion

We suggest that our unexpected finding of no changes in gene expression levels from circulating immune-competent cells in BC cases as opposed to parity-dependent changes in healthy women might be linked to the long-term immune-related protection against BC conferred by semi-allograft experience of each pregnancy. Our findings demonstrating the 8% decrease in BC risk for each birth correspond well to a number of previously published studies.8,9

The systems epidemiology approach we used is based on an integrated analysis in the same cohort; first, hypotheses describing the mathematical relationship between parity and BC incidence in the whole NOWAC study are explored. The findings were that the linear relationship between parity and BC could be described as independent of other risk factors. Then, this hypothesis was tested using gene expression profiles obtained from the postgenome biobank of NOWAC. Consequently, we reduced the number of multiple testing in the gene expression analysis and the potential for false-positive findings.

### Table 2 Top 10 single genes ranked according to q-value with logFC and their role in BC

| Gene ID  | logFCa | q-valueb | Role in BC | Function                                                                 |
|---------|--------|----------|------------|-------------------------------------------------------------------------|
| CNOT8   | 0.064  | 0.013    | TP         | Involved in translation and transcription regulation                     |
| ZRANB2  | 0.140  | 0.013    | NA         | Protein regulating alternative splicing; exact function is unknown       |
| TFRC    | 0.080  | 0.013    | TP         | Required for transfer of iron into cells from transferrin                |
| CD47    | 0.105  | 0.013    | TP         | Involved in many cellular processes including proliferation, migration, and apoptosis |
| PSORS1C1| -0.039 | 0.013    | NA         | Function is not known. Mutations are associated with psoriasis, systemic sclerosis, and rheumatoid arthritis |
| EML4    | 0.073  | 0.013    | NA         | Protein involved in microtubule formation                                 |
| FKBPL   | -0.027 | 0.013    | TS         | A wide range of functions including cell cycle regulation, trafficking, and angiogenesis inhibition |
| HACL1   | 0.044  | 0.013    | NA         | Exact role is unclear. Catalyzes a carbon–carbon cleavage reaction. Involved in lipid metabolism |
| PPP1R2  | 0.093  | 0.017    | NA         | Serine/threonine phosphatase. Function is not known                      |
| LPIN1   | 0.075  | 0.017    | TP         | Protein controlling the metabolism of fatty acids. Transcriptional coactivator |

**Notes:** a logFC is the estimated log-fold change in gene expression when the parity increases with one. b q-value is an FDR-adjusted p-value.

**Abbreviations:** BC, breast cancer; FDR, false discovery rate; NA, data not available; TP, tumor promoter; TS, tumor suppressor.

### Table 3 Gene set enrichment analyses for all gene sets included in MSigDB

| Collection name                      | All gene sets | Significant gene sets (FDR 5%) |
|--------------------------------------|---------------|--------------------------------|
|                                      | Total | Human | Other | Total | Human | Other |
| C1: positional gene sets             | 326   | 326   | 0     | 75    | 75    | 0     |
| C2: curated gene sets                | 4729  | 3750  | 979   | 280   | 229   | 51    |
| C3: motif gene sets                  | 836   | 835   | 1     | 22    | 0     | 22    |
| C4: computational gene sets          | 858   | 857   | 1     | 217   | 217   | 0     |
| C5: gene ontology gene sets          | 6166  | 6162  | 4     | 0     | 0     | 0     |
| C6: oncogenic signatures             | 189   | 144   | 45    | 42    | 35    | 7     |
| C7: immunologic signatures           | 4872  | 1844  | 3028  | 588   | 215   | 373   |

**Note:** In non-human gene sets, mice are the dominant species (http://software.broadinstitute.org/gsea/msigdb/index.jsp MSigDB v5.2). Copyright ©2018. Dove Medical Press. Reproduced from Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences. 2005;102(43):15545–50.27.27

**Abbreviations:** FDR, false discovery rate; MSigDB, Molecular Signatures Database.
for genes related to cancer hallmarks, while in blood, the
genes represented either general cellular or immune-related
processes.

Currently, there is no unifying concept explaining
mechanism of pregnancy-related BC protection. Yet, four
main theories may be underlined: decreased number of
mammary stem cells, changes in circulating hormone levels,
differentiation of breast epithelial cells, and altered estrogen
responsiveness.28 Notably, most of these theories are based
on studies focusing on the local changes in the mammary
gland, while we used blood samples as the source of data,
thus looking at systemic processes.

The single-gene analysis revealed a set of genes linearly
changing expression after pregnancies in controls. Almost all
of these genes were upregulated. This parity-dependent effect
was nonsignificant when we compared all cases and controls
without taking parity status into account. A possible explana-
tion could be the low parity in the study population with two
children as the median value. Of the top ten significant genes,
four upregulated genes were known to be related to breast
carcinogenesis.29–32 Similar to our results, one study found
lower levels of TFRC in plasma of BC cases compared to
controls.33 However, the overall meaning of these expression
changes in healthy women is uncertain.

Previously, Rotunno et al34 demonstrated the significant
enrichment of inflammatory and immune response genes in par-
ous women compared to nulliparous. In our gene set enrichment
analysis, we observed significant enrichment of numerous gene
sets related to the immune system (C7 collection in MSigDB). It
is nearly impossible to make an unambiguous conclusion from
these results due to the complexity of gene sets’ data and lack
of essential information in gene sets’ description (eg, we were
unable to ascertain the gender of blood donors for the major-
ity of human experiments even after contacting the authors of
publications). Nevertheless, we observed a clear interspecific
difference between components of the immune system: the
gene sets were related to (in mice, most of the gene sets were
of innate immunity origin, while in humans, they were of
adaptive origin) and between the sources of the cells for the
experiments. While the latter can be explained by technical and
ethical considerations, the former raises yet another concern on
the validity of results obtained in animal models.35–39

### Table 4 Top 10 gene sets for humans and mice with information on experimental design

| geneSetID  | scoreDiff* | FDR (q-value) | Source     | Immune system | Parity | Cells              | Gender |
|------------|------------|---------------|------------|---------------|--------|--------------------|--------|
| Human      |            |               |            |               |        |                    |        |
| GSE3982    | 0.023      | 0.027         | Cord blood | I             | NA     | Macrophages        | NA     |
| GSE2790    | -0.026     | 0.027         | Cord blood | A             | NA     | CD4+ T cells       | NA     |
| GSE16385   | 0.047      | 0.027         | Blood      | I             | NA     | Monocytes          | NA     |
| GSE4160    | 0.034      | 0.027         | Cord blood | A             | NA     | CD4+ T cells       | NA     |
| GSE13411   | -0.035     | 0.027         | Spleen     | A             | NA     | B cells            | M, F   |
| GSE2790    | 0.027      | 0.027         | Cord blood | A             | NA     | CD4+ T cells       | NA     |
| GSE29618   | -0.029     | 0.027         | Blood      | I             | NA     | DC                 | NA     |
| GSE17974   | 0.028      | 0.027         | Cord blood | A             | NA     | CD4+ T cells       | NA     |
| GSE2790    | -0.028     | 0.027         | Cord blood | A             | NA     | CD4+ T cells       | NA     |
| GSE29615   | 0.038      | 0.027         | Blood      | I, A          | NA     | PBMCs              | NA     |
| Mouse      |            |               |            |               |        |                    |        |
| GSE17721   | 0.018      | 0.027         | Bone marrow| I             | NA     | DC                 | F      |
| GSE14769   | 0.032      | 0.027         | Bone marrow| I             | NA     | Macrophages        | NA     |
| GSE3691    | 0.037      | 0.027         | Various tissues | I | NA     | DC                 | F      |
| GSE37301   | 0.035      | 0.027         | Bone marrow| I, A          | NA     | HSCs, CLPCs        | NA     |
| GSE32034   | 0.028      | 0.027         | Various tissues | I | NA     | Monocytes          | M      |
| GSE17721   | 0.026      | 0.027         | Bone marrow| I             | NA     | DC                 | F      |
| GSE21063   | 0.036      | 0.027         | Spleen     | A             | NA     | B cells            | NA     |
| GSE11924   | -0.029     | 0.027         | Spleen     | A             | NA     | CD4+ T cells       | NA     |
| GSE28237   | 0.035      | 0.027         | Spleen     | A             | NA     | B cells            | NA     |
| GSE13547   | 0.034      | 0.027         | Spleen     | A             | NA     | B cells            | NA     |

**Note:** scoreDiff is the estimated change in enrichment score when the parity increases with one.

**Abbreviations:** A, adaptive immune system; CLPCs, common lymphoid progenitor cells; DC, dendritic cells; F, female; FDR, false discovery rate; HSCs, hematopoietic stem cells; I, innate immune system; M, male; NA, data not available; PBMCs, peripheral blood mononuclear cells.
modulation of local and systemic immune response. Indeed, in our study, we observed that women with a high parity have a higher expression of tumor promoter genes and a lower expression of tumor suppressor genes; therefore, we propose that the body might experience each pregnancy as a “pseudo cancer”. It is possible to suggest that some pregnancy-related processes may have common genetic signatures and protein expression profile with initial stages of breast tumor formation. Being securely controlled by a variety of local and systemic mechanisms during pregnancy, the immune system does not respond to the changes in breast tissue, while storing memory and effector cells. Later, these cells can initiate rapid and extensive response against cancer cells during the early stages of carcinogenesis. This phenomenon of long-term storage of various subsets of antigen-experienced immune cells after pregnancies has previously been described. Slight expression differences in each pregnancy expand the pool of antigen-specific immune cells, thus increasing protective effect with the number of pregnancies. Recently, Krause et al published an experimental study confirming our hypothesis.

Our findings of an impaired immune surveillance in BC patients opposed to healthy women may generate novel approaches to immunotherapy based on the natural mechanisms involved in amplified protection against BC with repeated pregnancies. Deciphering of these mechanisms could significantly improve response rate of BC patients to immunotherapy, which at present remains poor.

The major weakness of the study is the lack of statistical power for the gene expression analyses. However, this was compensated by the strict hypothesis testing approach. The strength of the design is the complete follow-up on cancer, emigration, and mortality based on a unique national birth number. The external validity of the parity data in NOWAC was found to be excellent in a linkage study based on information from the fertility register in Statistics Norway.

**Conclusion**

To our knowledge, this is the first study revealing the involvement of the immune system in parity-associated BC protection with the use of systems epidemiology approach. Owing to descriptive study design, it is not possible to hypothesize the exact mechanisms responsible for the pregnancy-associated protection from BC using data obtained in this work. Our focus on systemic rather than on local changes may provide the scientific community with a fresh view on the biology of BC and may become a foundation for future research.

**Data sharing statement**

The datasets are stored at the European Genome-phenome Archive with accession number EGAD00010001062.

**Acknowledgments**

We are thankful to and impressed by the women who donated blood to the NOWAC postgenome cohort. Bente Augdal, Merete Albertsen, and Knut Hansen were responsible for all infrastructure and administrative issues. The funders had no role in the design of the study; collection, analyses, and interpretation of the data; writing of the manuscript; or decision to submit for publication. This study was supported by a grant from the European Research Council (ERC-AdG 232997 TICE). Some of the data in this article are from the Cancer Registry of Norway. The Cancer Registry of Norway is not responsible for the analysis or interpretation of the data presented. Microarray service was provided by the Genomics Core Facility, Norwegian University of Science and Technology, and NMC—a national technology platform supported by the functional genomics program (FUGE) of the Research Council of Norway.

**Author contributions**

EL, JT, and MH carried out conception and design of the study. EL, LH, and MH were involved in the development of methodology. EL, LH, and MH contributed toward acquisition of data. EL, AN, IS, JT, KSO, LH, and MH performed analysis and interpretation of data (eg, statistical analysis, biostatistics, and computational analysis). EL, IS, JT, KSO, LH, and MH carried out writing, review, and/or revision of the manuscript. EL, IS, KSO, LH, and MH provided administrative, technical, or material support. EL provided study supervision. All authors contributed toward data analysis, drafting and critically revising the paper, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**

1. Erlebacher A. Mechanisms of T cell tolerance towards the allogeneic fetus. Nat Rev Immunol. 2012;13(1):23–33.
2. Williams Z. Inducing tolerance to pregnancy. N Engl J Med. 2012;367(12):1159–1161.
3. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646–674.
4. Holtan SG, Creedon DJ, Haluska P, Markovic SN. Cancer and pregnancy: parallels in growth, invasion, and immune modulation and implications for cancer therapeutic agents. Mayo Clin Proc. 2009;84(11):985–1000.
5. Curigliano G. Immunity and autoimmunity: revising the concepts of response to breast cancer. Breast. 2011;20:S71–S74.

6. Elinav E, Nowarski R, Thaiss CA, Hu B, Jin C, Flavell RA. Inflammation-induced cancer: crosstalk between tumours, immune cells and microorganisms. Nat Rev Cancer. 2013;13(11):759–771.

7. Zhao H, Ozen M, Wong RJ, Stevenson DK. Heme oxygenase-1 in pregnancy and cancer: similarities in cellular invasion, cytoprotection, angiogenesis, and immunomodulation. Front Pharmacol. 2014;5:295.

8. Lund E. Childbearing in marriage and mortality from breast cancer in Norway. Int J Epidemiol. 1990;19(3):527–531.

9. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer risk in relation to abor- tion: results from the EPIC study. Int J Cancer. 2006;119(7):1741–1745.

10. Reeves GK, Kan SW, Key T, et al. Breast cancer risk in relation to abortion: results from the EPIC study. Int J Cancer. 2006;119(7):1741–1745.

11. Guo J, Huang Y, Yang L, et al. Association between abortion and breast cancer: an updated systematic review and meta-analysis based on prospective studies. Cancer Causes Control. 2015;26(6):811–819.

12. Tal R, Taylor HS, Barney RO, Mooney SB, Giudice LC. Endocrinology of Pregnancy. South Dartmouth, MA: MDText.com; 2000.

13. Henderson BE, Ponder BAJ, Ross RK. Hormones, Genes, and Cancer. Oxford; New York: Oxford University Press; 2003.

14. Robinson DP, Klein SL. Pregnancy and pregnancy-associated hormones alter immune responses and disease pathogenesis. Horm Behav. 2012;62(3):263–271.

15. IARC. Hormonal Contraception and Post-menopausal Hormonal Therapy. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Volume 72. Lyon: IARC; 1999.

16. Lund E, Holden L, Bøvelstad H, et al. A new statistical method for curve fitting in microarray experiments. Stat Appl Genet Mol Biol. 2004;3(1):1–25.

17. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. BMC Bioinformatics. 2013;14(1):7.

18. Anisimov VN, Ukraintseva SV, Yashin AI. Cancer in rodents: does it tell us about cancer in humans? Crit Rev Oncol Hematol. 2012;79(3):104981–104991.

19. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. BMC Bioinformatics. 2013;14(1):7.

20. Vonderheide RH, Domchek SM, Clark AS. Immunotherapy for breast cancer: what are we missing? Nat Rev Immunol. 2017;17(8):469–482.

21. Krause AL, Schuetz F, Boudewijns M, et al. Parity-related molecular signatures and breast cancer subtypes by estrogen receptor status. Breast Cancer Res. 2014;16(4):R74.

22. Beura LK, Hamilton SE, Bi K, et al. Normalizing the environment recapitulates adult human immune traits in laboratory mice. Nature. 2016;532(7600):512–516.

23. Krause AL, Schuetz F, Boudewijns M, et al. Parity-related molecular signatures and breast cancer subtypes by estrogen receptor status. Breast Cancer Res. 2014;16(4):R74.

24. Beura LK, Hamilton SE, Bi K, et al. Normalizing the environment recapitulates adult human immune traits in laboratory mice. Nature. 2016;532(7600):512–516.

25. Salatino M, Girotti MR, Rabinovich GA. Glycans pave the way for immunotherapy in triple-negative breast cancer. Nat Rev Cancer. 2014;14(4):371–380.

26. Buas MF, Rho JH, Chai X, Zhang Y, Lampe PJ, Li CI. Candidate early detection protein biomarkers for ER+PR+ invasive ductal breast carcinoma identified using pre-clinical plasma from the WHI observational study. Breast Cancer Res Treat. 2015;153(2):445–454.

27. Robinson DP, Klein SL. Pregnancy and pregnancy-associated hormones alter immune responses and disease pathogenesis. Horm Behav. 2012;62(3):263–271.

28. Buas MF, Rho JH, Chai X, Zhang Y, Lampe PJ, Li CI. Candidate early detection protein biomarkers for ER+PR+ invasive ductal breast carcinoma identified using pre-clinical plasma from the WHI observational study. Breast Cancer Res Treat. 2015;153(2):445–454.

29. Mor G, Aldo P, Alvero AB. The unique immunological and microbial aspects of pregnancy. Nat Rev Immunol. 2017;17(8):469–482.

30. Krause AL, Schuetz F, Boudewijns M, et al. Parity improves anti-tumor immunity in breast cancer patients. Oncotarget. 2017;8(62):104981–104991.

31. Daniels TR, Delgado T, Rodriguez JA, Helguera G, Penichet ML. The transferrin receptor part I: biology and targeting with cytotoxic antibodies for the treatment of cancer. Clin Immunol. 2006;121(2):144–158.

32. Zhang H, Lu H, Xiang L, et al. HIF-1 regulates CD47 expression in breast cancer cells to promote evasion of phagocytosis and maintenance of cancer stem cells. Proc Natl Acad Sci USA. 2015;112(45):E6215–E6223.

33. Buas MF, Rho JH, Chai X, Zhang Y, Lampe PJ, Li CI. Candidate early detection protein biomarkers for ER+PR+ invasive ductal breast carcinoma identified using pre-clinical plasma from the WHI observational study. Breast Cancer Res Treat. 2015;153(2):445–454.

34. Vonderheide RH, Domchek SM, Clark AS. Immunotherapy for breast cancer: what are we missing? Nat Rev Immunol. 2017;17(8):469–482.
Supplementary materials

Table S1 Distribution of parities for the cases and controls

| Parities | 0  | 1  | 2  | 3  | 4  | 5  | 6  | Sum |
|----------|----|----|----|----|----|----|----|-----|
| Cases    | 69 | 67 | 208| 83 | 29 | 4  | 0  | 460 |
| Controls | 74 | 62 | 174| 116| 26 | 5  | 3  | 460 |

Table S2 BC cases, PY, and incidence rates per 100,000 PY, according to menopausal status and parity

| Status                  | Parity | 0  | 1  | 2  | 3  | 4  | 5  | 6  |        |
|-------------------------|--------|----|----|----|----|----|----|----|---------|
|                         | Cases  | 168| 219| 688| 358| 86 | 13 | 2  |         |
|                         | PY     | 7,6189| 101,736| 356,276| 212,648| 55,512| 12,065| 2744   |
|                         | Rates  | 220.50| 215.26| 193.11| 168.35| 154.92| 107.75| 72.88  |
| Premenopause            | Cases  | 535| 618| 2082| 1254| 372 | 95 | 27 |        |
|                         | PY     | 132,262| 172,437| 636,545| 420,991| 144,901| 43,189| 13,662 |
|                         | Rates  | 404.50| 358.39| 327.08| 297.87| 256.73| 219.97| 197.62 |
| Postmenopause           | Cases  | 703| 837| 2770| 1612| 458 | 108| 29 |        |
|                         | PY     | 20,8451| 274,173| 992,821| 633,639| 200,413| 55,253| 16,407 |
|                         | Rates  | 337.25| 305.28| 279.00| 254.40| 228.53| 195.46| 176.76 |

Note: Data from the entire NOWAC cohort.

Abbreviations: BC, breast cancer; NOWAC, Norwegian Women and Cancer; PY, person-years.

Table S3 Tests for linear trend between strata of risk factors

| Stratification            | % decrease per child | Intercept | $\beta$ | Pr > |t|/$\beta$ |
|---------------------------|----------------------|-----------|---------|-------|---------|
| Percentage change per child | Overall | 8 | 334.34 | −26.84 | <0.001 |
| Stratified according to:  |          |   |         |        |         |
| Premenopause              | 11       | 236.4 | −24.86 | <0.001 |
| Postmenopause             | 9        | 398.29| −34.57 | <0.001 |
| With mother history of BC | 10       | 631.76| −61.68 | 0.01   |
| Without mother history of BC | 8      | 317.59| −24.56 | <0.001 |
| Age at first birth <25 years | 7     | 305.73| −22.4 | 0.002  |
| Age at first birth ≥25 years | 6    | 341.39| −20.93| 0.06   |
| BMI <25 kg/m²              | 9        | 337.08| −28.81| <0.001 |
| BMI ≥25 kg/m²              | 8        | 333.51| −25.09| <0.001 |
| Use of OC                 | 6        | 323.64| −20.76| <0.001 |
| No use of OC              | 9        | 340.42| −30.21| <0.001 |
| No lactation              | 11       | 340.34| −35.78| 0.022  |
| Years of lactation: 1–2   | 9        | 340.61| −30.33| 0.002  |
| Years of lactation: 3+    | 14       | 488.44| −68.55| 0.017  |
| HRT (only postmenopause)  | 9        | 530.49| −49.27| 0.009  |
| No HRT (only postmenopause) | 9     | 377.31| −32.77| <0.001 |

Note: Data from the entire NOWAC cohort.

Abbreviations: BC, breast cancer; BMI, body mass index; HRT, hormone replacement therapy; NOWAC, Norwegian Women and Cancer; OC, oral contraceptives; Pr, probability.
Table S4 Age-adjusted and multivariate-adjusted Cox proportional hazard ratios and 95% CIs for BC according to parity in the entire NOWAC cohort

| Parity | BC | Crude | Adjusted* |
|--------|----|-------|-----------|
| 1      | 0.91 (0.82–1.00) | 0.86 (0.77–0.96) |
| 2      | 0.81 (0.75–0.88) | 0.75 (0.68–0.83) |
| 3      | 0.73 (0.68–0.80) | 0.68 (0.61–0.76) |
| 4      | 0.63 (0.56–0.71) | 0.59 (0.51–0.67) |
| 5      | 0.52 (0.43–0.64) | 0.49 (0.39–0.62) |
| 6+     | 0.50 (0.37–0.66) | 0.49 (0.36–0.68) |

Note: *Adjustment for maternal history of BC (yes, no), BMI (<25, 25+ kg/m²), current use of hormone replacement therapy (yes, no), years of lactation (0, 1–2, 2+), and use of oral contraceptives (ever used oral contraceptives, never used contraceptives).

Abbreviations: BC, breast cancer; BMI, body mass index; CI, confidence interval; NOWAC, Norwegian Women and Cancer.

Table S5 Full names of gene sets described in Table 4

Human

- GSE3982_CTRL_VS_LPS_4H_MAC_UP
- GSE2770_UNTREATED_VS_ACT_CD4_TCELL_4H_DN
- GSE16385_MONOCYTE_VS_12H_ROSIGLITAZONE_TREATED_MACROPHAGE_UP
- GSE1460_CORD_VS_ADULT_BLOOD_NAIVE_CD4_TCELL_DN
- GSE13411_IGM_VS_SWITCHED_MEMORY_BCELL_UP
- GSE2770_UNTREATED_VS_TGFbeta_AND_IL12_TREATED_ACT_CD4_TCELL_2H_DN
- GSE29618_LAIV_VS_TIV_Flu_VACCINE_DAY7_PDC_DN
- GSE17974_2.5H_VS_72H_IL4_ANDANTI_IL12_ACT_CD4_TCELL_UP
- GSE2770_TGFbeta_AND_IL12_VS_TGFbeta_AND_IL12_TREATED_ACT_CD4_TCELL_2H_DN
- GSE29615_CTRL_VS_DAY3_LAIV_Flu_VACCINE_PBMC_DN

Mouse

- GSE17721_LPS_VS_PAM3CSK4_4H_BMDM_DN
- GSE4769_UNSTIM_VS_240MIN_LPS_BMDM_UP
- GSE3691_IFN_PRODUCING_KILLER_DC_VS_PLASMACYTOID_DC_SPLEEN_DN
- GSE37301HEMEATopoietic STEM_CELL_VS_COMMON_LYMPHoid_PROGENitor_DN
- GSE32034_UNTREATED_VS_ROSIGLITAZONE_TREATED_LY6C_HIGH_MONOCYTE_DN
- GSE17721_LPS_VS_PAM3CSK4_1H_BMDM_DN
- GSE21063_WT_VS_NFATC1_KO_BCELL_DN
- GSE11924_TFH_VS_TH17_CD4_TCELL_UP
- GSE28237_EARLY_VS_LATE_GC_BCELL_UP
- GSE13547_CTRL_VS_ANTI_IGM_STIM_ZFX_KO_BCELL_2H_DN