Molecular subtyping of serous ovarian tumors reveals multiple connections to intrinsic breast cancer subtypes.

Jönsson, Jenny-Maria; Johansson, Ida; Dominguez, Mev; Kimbung, Siker; Jönsson, Mats; Bonde, Jesper Hansen; Kannisto, Päivi; Måsbäck, Anna; Malander, Susanne; Nilbert, Mef; Hedenfalk, Ingrid

Published in: PLoS ONE

DOI: 10.1371/journal.pone.0107643

2014

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Molecular Subtyping of Serous Ovarian Tumors Reveals Multiple Connections to Intrinsic Breast Cancer Subtypes

Jenny-Maria Jönsson1*, Ida Johansson1,2, Mev Domínguez-Valentin1,5, Siker Kimbung1,2, Mats Jönsson1, Jesper Hansen Bonde3, Päivi Kannisto4, Anna Måsbäck1, Susanne Malander1, Mef Nilbert1,5, Ingrid Hedenfalk1,2

1 Division of Oncology-Pathology, Department of Clinical Sciences, Skåne University Hospital, Lund University, Lund, Sweden, 2 CREATE Health Strategic Center for Translational Cancer Research, Lund University, Lund, Sweden, 3 Department of Pathology, Clinical Research Centre, Hvidovre Hospital, Copenhagen University, Hvidovre, Denmark, 4 Division of Obstetrics and Gynecology, Department of Clinical Sciences, Skåne University Hospital, Lund University, Lund, Sweden, 5 Clinical Research Centre, Hvidovre Hospital, Copenhagen University, Hvidovre, Denmark

Abstract

Objective: Transcriptional profiling of epithelial ovarian cancer has revealed molecular subtypes correlating to biological and clinical features. We aimed to determine gene expression differences between malignant, benign and borderline serous ovarian tumors, and investigate similarities with the well-established intrinsic molecular subtypes of breast cancer.

Methods: Global gene expression profiling using Illumina’s HT12 Bead Arrays was applied to 59 fresh-frozen serous ovarian malignant, benign and borderline tumors. Nearest centroid classification was performed applying previously published gene profiles for the ovarian and breast cancer subtypes. Correlations to gene expression modules representing key biological breast cancer features were also sought. Validation was performed using an independent, publicly available dataset.

Results: 5,944 genes were significantly differentially expressed between benign and malignant serous ovarian tumors, with cell cycle processes enriched in the malignant subgroup. Borderline tumors were split between the two clusters. Significant correlations between the malignant serous tumors and the highly aggressive ovarian cancer signatures, and the basal-like breast cancer subtype were found. The benign and borderline serous tumors together were significantly correlated to the normal-like breast cancer subtype and the ovarian cancer signature derived from borderline tumors. The borderline tumors in the study dataset, in addition, also correlated significantly to the luminal A breast cancer subtype. These findings remained when analyzed in an independent dataset, supporting links between the molecular subtypes of ovarian cancer and breast cancer beyond those recently acknowledged.

Conclusions: These data link the transcriptional profiles of serous ovarian cancer to the intrinsic molecular subtypes of breast cancer, in line with the shared clinical and molecular features between high-grade serous ovarian cancer and basal-like breast cancer, and suggest that biomarkers and targeted therapies may overlap between these tumor subsets. The link between benign and borderline serous ovarian and luminal breast cancer may indicate endocrine responsiveness in a subset of ovarian cancers.


Editor: Goli Samimi, Garvan Institute of Medical Research, Australia

Received June 5, 2014; Accepted August 12, 2014; Published September 16, 2014

Copyright: © 2014 Jönsson et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All gene expression data files are available in the National Center for Biotechnology Information Gene Expression Omnibus [GEO accession number: GSE57477].

Funding: Financial support was granted from a governmental funding of clinical research within the National Health Services (JMJ), the Swedish Research Council (MN, IH) and an unrestricted educational grant from the Swedish Society of Gynecologic Oncology sponsored by Roche (JMJ). Funding has neither influenced study design, data collection nor analysis or manuscript writing.

Competing Interests: The authors have declared that no competing interests exist.

* Email: jenny-maria.jonsson@med.lu.se

Introduction

Epithelial ovarian tumors constitute a heterogeneous group of neoplasms that differ in epidemiology, genetic risk factors, precursor lesions and clinical behavior. The different histopathologic subtypes, i.e. serous, mucinous, endometrioid, clear cell and transitional carcinomas and carcinosarcomas, likely have different origins and appear to evolve along distinct pathways [1–3]. Alongside with the standard taxane and platinum based agents used for ovarian cancer [4–6], multiple targeted agents are being evaluated, with e.g. bevacizumab recently being included in the therapeutic arsenal [7]. Personalized therapy is called for in ovarian cancer particularly since the histopathologic subtypes, as well as tumors with different malignant potential and tumor grade, can be viewed as separate diseases with differences related to both prognosis and treatment response [8–12]. Refined molecular subtyping and recognition of key genetic mechanisms constitutes
an encouraging basis for further development of subtype-specific targeted therapies.

Previous efforts to characterize ovarian cancers at the molecular level have identified distinct profiles related to the histologic subtypes and have suggested predictive gene signatures [13–17]. Totthill et al. suggested six different subtypes, referred to as C1–C6, based on serous and endometrioid ovarian, primary peritoneal and fallopian tube tumors. The C1–C2 and C4-C5 subtypes, in general, are thought to characterize high-grade serous tumors. The C1 signature is characterized by a high degree of desmoplasia, C5 by mesenchymal genes and overexpression of proliferation genes and the C2 and C4 signatures by high numbers of intra-tumoral and stroma associated CD3+ cells. The signatures are outcome predictive, with the C1 signature corresponding to a considerably worse outcome than the other signatures. The C3 signature represents low-grade serous and borderline tumors and the C6 signature low-grade, early-stage endometrioid tumors; in general they show good response to treatment and long-time survival [18]. Likewise, molecular subtyping in breast cancer is well established and recent reports have recognized similarities between high-grade serous ovarian cancer and basal-like breast cancer [19].

We performed global gene expression profiling of serous ovarian tumors, including serous cystadenomas, serous borderline tumors and serous adenocarcinomas, and applied previously described gene signatures including the well-known intrinsic breast cancer subtypes [18,20–22] to outline further possible similarities between these tumor types. Since mutations in the MAPK/ERK pathways are common in both borderline and low-grade ovarian cancer and luminal breast cancers the presence of KRAS and BRAF mutations was investigated among the ovarian tumors [23,24]. Shared common features between ovarian and breast cancer may be useful for future development of predictive biomarkers and tailored treatments in both tumor types, and in this study we present interesting connections between the molecular subtypes of ovarian and breast cancer.

Materials and Methods

Tumor samples

In total, 37 serous ovarian adenocarcinomas, 17 serous cystadenomas/adenolbromas and 5 serous borderline tumors were obtained from the Skåne University Hospital ovarian tumor biobank (table 1). A total of 13 biological replicates (6 omental metastases, 1 pelvic metastasis and 2 metastases to the contralateral ovary as well as 3 benign and 1 borderline ovarian tumors) were included to account for intra-tumor heterogeneity. All tumor samples were collected at primary surgery (2003–2011) and the patients had not received chemotherapy prior to surgery. Histologic subtype and grade were determined according to Silverberg and WHO [25,26] and all tumors were staged according to the International Federation of Gynecology and Obstetrics (FIGO) criteria. Hematoxylin & Eosin stained slides were used to assess tumor grade. This was performed by a senior pathologist (AM). Ethical approval for the study was granted from the Lund University ethics committee, Sweden, waiving the requirement for informed consent for the study.

RNA extraction and gene expression analysis

Total RNA was extracted using the Allprep kit (Qiagen, Heidelberg, Germany) according to the manufacturer’s instructions. RNA concentration was determined using a NanoDrop Spectrophotometer (NanoDrop Technologies, Wilmington, DE) and samples with ≥200 ng RNA with 260/280 ratios ≥1.8 were used for further analysis. RNA quality was assessed using a Bioanalyzer (Agilent Technologies, Santa Clara, CA), and RNA integrity numbers (RIN) >6 were regarded as sufficient.

Gene expression profiling analyses were performed at the SCIBLU Genomics Centre, Lund University, Sweden. The cDNA synthesis, labeling, and subsequent hybridization to the HumanHT-12 v4 Expression BeadChips (Illumina Inc., San Diego, CA) was performed according to the manufacturer’s instructions. The Illumina HumanHT-12 v4 Expression BeadChips allow genome-wide expression profiling of more than 47,000 gene transcripts and splice variants. The 59 samples and 13 biological replicates were randomized on the chips. The BeadChips were then scanned on an i-Scan (Illumina Inc.), during which fluorescence intensities were read and images extracted. The gene expression data are available in the National Center for Biotechnology Information Gene Expression Omnibus [ GEO accession number: GSE57477] [27].

Data analysis

Gene expression data were uploaded to the GenomeStudio software (Illumina Inc.), quantile normalized, background corrected and log2 transformed. Probes with a mean intensity <2.5 and variance <0.1 were excluded, leaving a total of 16,024 probes corresponding to 12,313 unique genes. Thereafter the data were uploaded to the MeV v4 software, an application used for identification of genes and expression patterns in microarray data [28], mean centered and a variance filter was applied to select the 20% of the probes with the greatest variation of expression across the dataset. Unsupervised hierarchical clustering was performed using complete linkage and Pearson distant metric. Two-class unpaired significance analysis of microarrays (SAM) was performed based on all 16,024 probes to identify differentially expressed probes between the different tumor subgroups (benign, borderline, malignant) at a false discovery rate (FDR)<0.01 [29]. Hierarchical clustering, supervised by the SAM analysis results and thereby identifying significantly differentially expressed genes between the tumor subgroups, was performed using the same methods as for unsupervised clustering. Gene ontology analyses based on the significant genes were performed in the gene ontology enrichment analysis and visualization tools GOrilla and ToppGene for identification of possible gene enrichment with biological or functional differences separating the subgroups [30,31].

Molecular subtyping and external data sets

Gene signatures outlining six molecular subtypes of high-grade and advanced stage serous ovarian tumors as well as endometrioid, low-grade serous and borderline ovarian tumors (referred to as the “Totthill dataset”) [18] were applied to the serous ovarian tumors in our cohort. Data were normalized and log2 transformed using the Gene Chip Operating Software Version 1.4 with Affymetrix default analysis settings. Probes with intensity values <4 and variance <0.15 were excluded. The six ovarian cancer signatures contained in total 4,732 probes of which 4,099 probes, corresponding to 2,725 unique genes, with good quality were left after filtering away probes with bad quality. 1,295/2,725 (47.5%) of these genes were identified in our dataset and used for further analyses. To validate the classifier, the 1,295 genes present in our dataset were re-applied to the 285 ovarian tumors in the original cohort, thereby re-assigning subtypes to each tumor. A gene signature for intrinsic subtyping of breast cancer was also applied to our serous ovarian tumors [22]. Each tumor in our cohort was classified into the molecular subtypes of ovarian cancer (C1–C6) as well as the intrinsic subtypes of breast cancer (luminal A, luminal
Table 1. Clinicopathologic features of malignant and borderline ovarian tumors in the study cohort.

<table>
<thead>
<tr>
<th>Id</th>
<th>Feature</th>
<th>Stage</th>
<th>Grade</th>
<th>Age at diagnosis (years)</th>
<th>Tissue type</th>
<th>Survival (years)</th>
<th>C- signature</th>
<th>BC subtype</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>69</td>
<td>Ovary</td>
<td>3</td>
<td>C1</td>
<td>Normal</td>
</tr>
<tr>
<td>70</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>82</td>
<td>Omentum</td>
<td>1</td>
<td>C1</td>
<td>Normal</td>
</tr>
<tr>
<td>84</td>
<td>M</td>
<td>IIIC</td>
<td>2</td>
<td>69</td>
<td>Ovary</td>
<td>6</td>
<td>C5</td>
<td>HER2</td>
</tr>
<tr>
<td>90</td>
<td>M</td>
<td>IIIC</td>
<td>2</td>
<td>60</td>
<td>Ovary</td>
<td>2</td>
<td>C3</td>
<td>Normal</td>
</tr>
<tr>
<td>106</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>71</td>
<td>N/A</td>
<td>4</td>
<td>C2</td>
<td>Basel</td>
</tr>
<tr>
<td>118</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>50</td>
<td>N/A</td>
<td>2</td>
<td>C5</td>
<td>Basel</td>
</tr>
<tr>
<td>125</td>
<td>M</td>
<td>IIIA</td>
<td>3</td>
<td>71</td>
<td>N/A</td>
<td>5</td>
<td>C4</td>
<td>Basel</td>
</tr>
<tr>
<td>137</td>
<td>M</td>
<td>II A</td>
<td>3</td>
<td>70</td>
<td>FT</td>
<td>Alive</td>
<td>C4</td>
<td>Basel</td>
</tr>
<tr>
<td>153</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>61</td>
<td>Ovary</td>
<td>1,5</td>
<td>C5</td>
<td>Basel</td>
</tr>
<tr>
<td>159</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>91</td>
<td>N/A</td>
<td>&lt;1</td>
<td>C1</td>
<td>Basel</td>
</tr>
<tr>
<td>186</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>82</td>
<td>Cystic fluid</td>
<td>2</td>
<td>C4</td>
<td>Basel</td>
</tr>
<tr>
<td>190</td>
<td>M</td>
<td>IIIC</td>
<td>2</td>
<td>79</td>
<td>Omentum</td>
<td>1,5</td>
<td>C5</td>
<td>Luminal A</td>
</tr>
<tr>
<td>192</td>
<td>M</td>
<td>IV1</td>
<td>2</td>
<td>55</td>
<td>Omentum</td>
<td>1</td>
<td>C1</td>
<td>Basel</td>
</tr>
<tr>
<td>207</td>
<td>M</td>
<td>IIIC</td>
<td>1</td>
<td>80</td>
<td>N/A</td>
<td>2</td>
<td>C1</td>
<td>Luminal A</td>
</tr>
<tr>
<td>219</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>67</td>
<td>FT</td>
<td>1</td>
<td>C1</td>
<td>Basel</td>
</tr>
<tr>
<td>225</td>
<td>M</td>
<td>IV1</td>
<td>2</td>
<td>53</td>
<td>N/A</td>
<td>2</td>
<td>C2</td>
<td>HER2</td>
</tr>
<tr>
<td>226</td>
<td>M</td>
<td>IV3</td>
<td>3</td>
<td>61</td>
<td>Omentum</td>
<td>Alive</td>
<td>C4</td>
<td>Basel</td>
</tr>
<tr>
<td>232</td>
<td>M</td>
<td>IV3</td>
<td>1</td>
<td>87</td>
<td>Ovary</td>
<td>&lt;1</td>
<td>C3</td>
<td>Normal</td>
</tr>
<tr>
<td>251</td>
<td>M</td>
<td>IV3</td>
<td>3</td>
<td>59</td>
<td>Omentum</td>
<td>1</td>
<td>C1</td>
<td>Basel</td>
</tr>
<tr>
<td>273</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>62</td>
<td>N/A</td>
<td>4</td>
<td>C2</td>
<td>HER2</td>
</tr>
<tr>
<td>275</td>
<td>M</td>
<td>IIIC</td>
<td>2</td>
<td>65</td>
<td>Ovary</td>
<td>2</td>
<td>C1</td>
<td>Basel</td>
</tr>
<tr>
<td>279</td>
<td>M</td>
<td>IIIB</td>
<td>2</td>
<td>69</td>
<td>Ovary</td>
<td>1</td>
<td>C1</td>
<td>Basel</td>
</tr>
<tr>
<td>284</td>
<td>M</td>
<td>IIIB</td>
<td>3</td>
<td>70</td>
<td>Omentum</td>
<td>Alive</td>
<td>C1</td>
<td>Normal</td>
</tr>
<tr>
<td>293</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>72</td>
<td>N/A</td>
<td>Alive</td>
<td>C1</td>
<td>Basel</td>
</tr>
<tr>
<td>297</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>65</td>
<td>Ovary</td>
<td>2</td>
<td>C1</td>
<td>Basel</td>
</tr>
<tr>
<td>305</td>
<td>M</td>
<td>IIIB</td>
<td>2</td>
<td>63</td>
<td>Ovary</td>
<td>2</td>
<td>C2</td>
<td>Basel</td>
</tr>
<tr>
<td>306</td>
<td>M</td>
<td>IIIC</td>
<td>1</td>
<td>40</td>
<td>N/A</td>
<td>Alve</td>
<td>C4</td>
<td>HER2</td>
</tr>
<tr>
<td>307</td>
<td>M</td>
<td>IIIC1</td>
<td>1</td>
<td>42</td>
<td>Ovary</td>
<td>Alive</td>
<td>C2</td>
<td>Basel</td>
</tr>
<tr>
<td>311</td>
<td>M</td>
<td>IIIB</td>
<td>3</td>
<td>67</td>
<td>N/A</td>
<td>Alive</td>
<td>C2</td>
<td>Basel</td>
</tr>
<tr>
<td>314</td>
<td>M</td>
<td>IIIC</td>
<td>2</td>
<td>69</td>
<td>N/A</td>
<td>Alive</td>
<td>C4</td>
<td>HER2</td>
</tr>
<tr>
<td>330</td>
<td>M</td>
<td>IC1</td>
<td>3</td>
<td>64</td>
<td>Ovary</td>
<td>Alive</td>
<td>C5</td>
<td>Luminal B</td>
</tr>
<tr>
<td>344</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>68</td>
<td>Ovary</td>
<td>Alive</td>
<td>C4</td>
<td>Basel</td>
</tr>
<tr>
<td>393</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>81</td>
<td>Ovary</td>
<td>3</td>
<td>C2</td>
<td>Basel</td>
</tr>
<tr>
<td>397</td>
<td>M</td>
<td>IIIC</td>
<td>2</td>
<td>79</td>
<td>Omentum</td>
<td>Alive</td>
<td>C3</td>
<td>Normal</td>
</tr>
<tr>
<td>402</td>
<td>M</td>
<td>IIIC</td>
<td>3</td>
<td>70</td>
<td>Ovary</td>
<td>Alive</td>
<td>C4</td>
<td>Luminal B</td>
</tr>
</tbody>
</table>
B, basal-like, normal-like, and HER2 enriched) using nearest centroid classification. The methodology for nearest centroid classification is outlined by Johansson et al.[32]. Validation was performed by classifying the 285 ovarian tumors in the Tothill dataset into the intrinsic breast cancer subtypes. Furthermore, seven gene expression modules representing key biological processes in breast cancer (AURKA/proliferation, CASP3/apoptosis, ERBB2/HER2 signaling, ESR1/ER signaling, STAT1/immune response, PLAU/tumor invasion and metastasis, VEGF/angiogenesis; referred to as the “Desmedt modules”) were applied to the serous ovarian tumors in our dataset as well as the Tothill dataset [33], and their relationship to the previously described intrinsic breast cancer subtypes was investigated. These modules, derived from 917 breast cancers in publicly available datasets and characterized by computed module scores, comprise in total 889 genes.

Mutation analysis
KRAS mutation analysis was performed using the Roche cobas K-RAS Mutation Kit (product number 05852170190) (Roche, Pleasanton, CA), a CE-IVD real-time melting curve KRAS mutations assay, detecting mutations in codons 12, 13 and 61 of the KRAS oncogene. BRAF mutation analysis was performed using the Roche cobas BRAF V600 mutational analysis (product number 05985595190), which evaluates the BRAF V600 site in exon 15 and detects wildtype or mutated V600. The analyses were performed according to the manufacturer’s instructions, and the assays were run on the z480 Lightcycler (Roche). The mutation analyses were performed at the Department of Pathology, Clinical Research Centre, Hvidovre Hospital, Denmark.

Statistical methods
Mann-Whitney U-test and Pearson correlation were used for comparison between expression profiles of the different tumor subsets using the MeV 4.6.02 software. Correlations between different subtype classifications were assessed using Fisher’s exact test and between module scores using Mann-Whitney U-test and Kruskal Wallis test in SPSS (IBM SPSS Statistics 19). P-values < 0.05 were considered statistically significant.

Results
Comparison of malignant, benign and borderline serous ovarian cancers
Unsupervised hierarchical clustering of benign and malignant tumors based on the 20% of the probes that showed the most variability revealed two distinct clusters, one containing only malignant tumors and one containing all benign and four malignant tumors (Figure S1). All but one of the biological replicates clustered together pair-wise. The clusters remained stable after removal of the biological replicates, suggesting stable transcriptional differences between the clusters (Figure S2). Of note, the four malignant tumors in the benign cluster one was grade 1, two were grade 2 and one was grade 3, but no significant differences regarding stage or mean age at diagnosis were seen between the malignant tumors in the two clusters.

Next, a SAM analysis was performed to explore transcriptional differences between benign and malignant tumors, revealing 3,944 significantly differentially expressed genes (FDR <0.01), of which 2,984 were upregulated and 2,960 were downregulated among the malignant tumors (figure 1, Table S1).

Cell cycle kinases (e.g. CDC2, CDC5, CDC7 and CDC20) as well as AURKA and S100A9, which can all broadly be linked to cell cycle regulation and mitosis, were upregulated in the

B, basal-like, normal-like, and HER2 enriched) using nearest centroid classification. The methodology for nearest centroid classification is outlined by Johansson et al.[32]. Validation was performed by classifying the 285 ovarian tumors in the Tothill dataset into the intrinsic breast cancer subtypes. Furthermore, seven gene expression modules representing key biological processes in breast cancer (AURKA/proliferation, CASP3/apoptosis, ERBB2/HER2 signaling, ESR1/ER signaling, STAT1/immune response, PLAU/tumor invasion and metastasis, VEGF/angiogenesis; referred to as the “Desmedt modules”) were applied to the serous ovarian tumors in our dataset as well as the Tothill dataset [33], and their relationship to the previously described intrinsic breast cancer subtypes was investigated. These modules, derived from 917 breast cancers in publicly available datasets and characterized by computed module scores, comprise in total 889 genes.

Mutation analysis
KRAS mutation analysis was performed using the Roche cobas K-RAS Mutation Kit (product number 05852170190) (Roche, Pleasanton, CA), a CE-IVD real-time melting curve KRAS mutations assay, detecting mutations in codons 12, 13 and 61 of the KRAS oncogene. BRAF mutation analysis was performed using the Roche cobas BRAF V600 mutational analysis (product number 05985595190), which evaluates the BRAF V600 site in exon 15 and detects wildtype or mutated V600. The analyses were performed according to the manufacturer’s instructions, and the assays were run on the z480 Lightcycler (Roche). The mutation analyses were performed at the Department of Pathology, Clinical Research Centre, Hvidovre Hospital, Denmark.

Statistical methods
Mann-Whitney U-test and Pearson correlation were used for comparison between expression profiles of the different tumor subsets using the MeV 4.6.02 software. Correlations between different subtype classifications were assessed using Fisher’s exact test and between module scores using Mann-Whitney U-test and Kruskal Wallis test in SPSS (IBM SPSS Statistics 19). P-values < 0.05 were considered statistically significant.

Results
Comparison of malignant, benign and borderline serous ovarian cancers
Unsupervised hierarchical clustering of benign and malignant tumors based on the 20% of the probes that showed the most variability revealed two distinct clusters, one containing only malignant tumors and one containing all benign and four malignant tumors (Figure S1). All but one of the biological replicates clustered together pair-wise. The clusters remained stable after removal of the biological replicates, suggesting stable transcriptional differences between the clusters (Figure S2). Of note, the four malignant tumors in the benign cluster one was grade 1, two were grade 2 and one was grade 3, but no significant differences regarding stage or mean age at diagnosis were seen between the malignant tumors in the two clusters.

Next, a SAM analysis was performed to explore transcriptional differences between benign and malignant tumors, revealing 3,944 significantly differentially expressed genes (FDR <0.01), of which 2,984 were upregulated and 2,960 were downregulated among the malignant tumors (figure 1, Table S1).

Cell cycle kinases (e.g. CDC2, CDC5, CDC7 and CDC20) as well as AURKA and S100A9, which can all broadly be linked to cell cycle regulation and mitosis, were upregulated in the
malignant tumors. Consistent with this, gene ontology analyses revealed a significant upregulation of cell cycle associated biological processes (ToppGene, FDR < 0.05; table 2).

To investigate whether borderline tumors are more closely related to benign or malignant tumors, an unsupervised clustering based on the 20% most varying probes across the dataset, and a hierarchical clustering supervised by the 5,944 significantly differentially expressed genes between benign and malignant tumors, were performed on the whole dataset. The analyses resulted in two distinct clusters, one malignant and one benign. The borderline tumors were split between the two main clusters, implying heterogeneity within this group (Figure S3).

We next applied the gene signatures described by Tothill et al. [18] to study the representation of molecular ovarian subtypes ("C-signatures") in our dataset. Using nearest centroid classification, a specific ovarian cancer C-signature was assigned to each tumor in our cohort, 52/59 (88%) of the tumors had a correlation coefficient ≥ 0.2. The centroid classifications revealed considerable heterogeneity across the tumors (p < 0.001; table 3). These differences prompted us to investigate each signature individually, and significant correlations between the malignant tumors in our cohort and the C1, C2 and C4 signatures (p = 0.020), and between the benign and borderline tumors in our cohort and the C3 signature (p < 0.001) were revealed (Table S2). 251/285 tumors in the original Tothill cohort had an assigned C-signature, and 239/251 (95.2%) were correctly re-assigned to their respective C-signatures, thereby validating the classification method (Table S7).

Exploring similarities between ovarian and breast cancer

To investigate potential similarities between ovarian cancer and the widely acknowledged intrinsic subtypes of breast cancer, beyond the similarities between high-grade serous ovarian cancer and basal-like breast cancer that have been reported, we applied the signatures representing the intrinsic breast cancer subtypes to our cohort [19,22]. 40/59 (68%) of the tumors had a correlation coefficient ≥ 0.2. Classification of the intrinsic breast cancer subtypes was applied to the ovarian tumors in our cohort, revealing considerable heterogeneity (p < 0.001; table 4). Significant correlations between the malignant ovarian tumors and the basal-like breast cancer subtype (p < 0.001), and between the non-malignant (benign and borderline) ovarian tumors and the normal-like breast cancer subtype were found (p < 0.001) (Table S3). The borderline tumors in our cohort, all of which were most highly correlated to the ovarian cancer C3 signature, also had highest correlation to the luminal A breast cancer subtype (p < 0.001) (Table S4), thus extending the links between the two tumor types.

Validation of the intrinsic breast cancer subtypes in ovarian cancer

Next, potential correlations between ovarian cancer C-signatures and the intrinsic breast cancer subtypes were explored. When the C-signatures for the tumors in our cohort were correlated to the intrinsic breast cancer subtypes for the same tumors, a significant heterogeneity within the tumor cohort was observed (p < 0.001; figure 2). The ovarian cancer C2 and C4 signatures correlated significantly with the basal-like breast cancer subtype (p = 0.019 and p = 0.001, respectively) (Table S5), while the C3 signature correlated to the normal-like breast cancer subtype (p < 0.001) (Table S6). These observations suggest commonalities between the transcriptionally based molecular classifiers of ovarian and breast cancer (Figure S4). The results were verified by classifying the ovarian tumors in the Tothill dataset [18] into the intrinsic breast cancer subtypes; similar links between the C2 signature and the basal-like subtype, and the C3 signature and normal-like subtype were observed (p < 0.001; Table S8). Notably, 16/26 patients (61.5%) whose tumors resembled the basal-like and HER2 enriched intrinsic breast cancer subtypes died from their disease within four years from diagnosis, compared to only 3/9 patients (33.3%) whose tumors displayed gene expression profiles corresponding to the luminal A and luminal B intrinsic breast cancer subtypes.

To further explore potential connections between the molecular subtypes of ovarian and breast cancer, we applied the gene expression modules representing key biological features of breast cancer described by Desmedt et al. to our tumors [33]. As for the ovarian cancer C-signatures and the intrinsic breast cancer subtypes, considerable heterogeneity within the tumor cohort was found. The malignant tumors displayed a significantly higher module score than the benign and borderline tumors for the AURKA/proliferation, STAT1/immune response, CASP3/apoptosis, VEGF/angiogenesis and ERBB2/HER2 signaling modules. The borderline and benign tumors on the other hand correlated to the ESR1/ER signaling module (Figure S5). The somewhat surprising absence of a correlation between the malignant tumors and the PLAU/tumor invasion and metastasis module led us to investigate the correlation between the Desmedt modules and the ovarian C-signatures. Again, significant differences within the cohort were observed, with highly significant correlations between the C1 signature and PLAU/invasion.
module ($p<0.001$), the C2 signature and STAT1/immune response module ($p=0.001$), and the C4 signature and VEGF/angiogenesis module ($p=0.001$). The C5 signature showed a trend towards correlation to the AURKA/proliferation module compared to non-C5 tumors, but did not reach statistical significance ($p=0.176$). The C3 signature correlated significantly to the ESR1/ER module ($p<0.001$, figure 3). Finally, we also applied the Desmedt modules to the tumors in the Tothill dataset, and could verify the correlations described between the ovarian cancer C-signatures and the functional breast cancer derived gene expression modules (Figure S6). In addition, in this larger cohort, the C5/AURKA correlation was also found to be significant ($p=0.001$).

Mutation analysis

Mutation analyses revealed four tumors with KRAS mutations and two tumors with BRAF mutations in our cohort. Four of these mutations were present among the borderline tumors, all of which corresponded to the C3 ovarian cancer signature and the luminal A breast cancer subtype, respectively. Two malignant tumors, both grade 1, harbored mutations in KRAS and corresponded to the C3/normal-like subtypes and C4/luminal B subtypes, respectively (table 5).

Discussion

Serous carcinomas account for about 50% of the malignant epithelial ovarian tumors and thereby constitute the predominant histologic subtype. Type 1 tumors, i.e. low-grade serous carcinomas, along with low-grade endometrioid, mucinous and clear cell tumors, are thought to develop step-wise from benign cystadenomas/adenofoibromas via borderline tumors and are typically slowly proliferating and frequently harbor mutations in KRAS, BRAF and PTEN. In contrast, type 2 tumors (high-grade serous and high-grade endometrioid ovarian tumors, carcinosarcomas and undifferentiated carcinomas) are suggested to develop from

Table 2. The 10 most significantly enriched biological processes in the malignant ovarian tumors in the study cohort [31].

<table>
<thead>
<tr>
<th>Biological process*</th>
<th>p-value</th>
<th>Genes from input*</th>
<th>Genes in annotationb</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. mitotic cell cycle</td>
<td>3.823E-24</td>
<td>213</td>
<td>874</td>
</tr>
<tr>
<td>2. cell cycle process</td>
<td>1.544E-23</td>
<td>265</td>
<td>1192</td>
</tr>
<tr>
<td>3. antigen processing and presentation of exogenous peptide antigen</td>
<td>2.768E-19</td>
<td>67</td>
<td>171</td>
</tr>
<tr>
<td>4. antigen processing and presentation of exogenous peptide antigen via MHC class I</td>
<td>3.943E-19</td>
<td>43</td>
<td>80</td>
</tr>
<tr>
<td>5. cellular response to stress</td>
<td>6.028E-19</td>
<td>279</td>
<td>1370</td>
</tr>
<tr>
<td>6. symbiosis, encompassing mutualism through parasitism</td>
<td>1.886E-18</td>
<td>175</td>
<td>741</td>
</tr>
<tr>
<td>7. interspecies interaction between organisms</td>
<td>1.886E-18</td>
<td>175</td>
<td>741</td>
</tr>
<tr>
<td>8. antigen processing and presentation of exogenous antigen</td>
<td>3.374E-18</td>
<td>67</td>
<td>178</td>
</tr>
<tr>
<td>9. mitotic cell cycle phase transition</td>
<td>5.793E-18</td>
<td>121</td>
<td>445</td>
</tr>
<tr>
<td>10. cell cycle phase transition</td>
<td>6.556E-18</td>
<td>122</td>
<td>451</td>
</tr>
</tbody>
</table>

*FDR<0.05 and ≥3 recognized genes/biological function were required to consider a gene ontology (GO) process significant. 731 significant. GO processes were identified.
*Number of genes in the study cohort correlating to the GO process.
*Number of genes in the GO process.
doi:10.1371/journal.pone.0107643.t002

Table 3. Ovarian cancer subtypes.

<table>
<thead>
<tr>
<th>C-signature</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malignant</td>
<td>12</td>
<td>8</td>
<td>3</td>
<td>9</td>
<td>5</td>
<td>37</td>
</tr>
<tr>
<td>(% within group)</td>
<td>(32.4)</td>
<td>(21.6)</td>
<td>(8.1)</td>
<td>(24.3)</td>
<td>(13.5)</td>
<td>(100.0)</td>
</tr>
<tr>
<td>Borderline</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>(% within group)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(100.0)</td>
<td>(0.0)</td>
<td>(0.0)</td>
<td>(100.0)</td>
</tr>
<tr>
<td>Benign</td>
<td>1</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>(% within group)</td>
<td>(5.9)</td>
<td>(0.0)</td>
<td>(88.2)</td>
<td>(0.0)</td>
<td>(5.9)</td>
<td>(100.0)</td>
</tr>
<tr>
<td>p&lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Serous ovarian tumors in the study cohort with corresponding ovarian cancer subtypes (“C-signatures”) [18]. The rows outline the tumor types with the representation in each subtype in percent within parentheses. The p-value is calculated using Fisher’s exact test.
doi:10.1371/journal.pone.0107643.t003
precursor lesions in the fallopian tube [24] and are characterized by rapid progression and frequent TP53 mutations [10,34–36]. Molecular subtyping of ovarian cancer is being increasingly recognized, with e.g. the six transcriptionally based ovarian C-signatures proposed by Tothill et al. being both descriptive (with good correlations to clinical factors) and predictive of outcome [18].

Molecular subtypes of breast cancer have been established and linked to clinical behavior and treatment response [20,21]. Although about 60% of all ovarian tumors display high expression of estrogen receptors (ER) [37], features shared between high-grade serous ovarian cancer and basal-like breast cancer (the majority of which are “triple negative”, i.e. lack expression of estrogen and progesterone receptors and HER2 amplification) have recently been reported. Both tumor groups display frequent TP53 mutations and genomic instability and are clinically aggressive. Also, BRCA mutations are more frequent in high-grade serous ovarian cancer and in basal-like breast cancers in the case of BRCA1 [19]. The diagnostic and therapeutic potential in clarifying ovarian cancer heterogeneity and identifying mechanisms shared between ovarian cancer and breast cancer constituted the basis of our study.

Table 4. Intrinsic breast cancer subtypes.

<table>
<thead>
<tr>
<th>Intrinsic breast cancer subtype</th>
<th>Luminal A</th>
<th>Luminal B</th>
<th>Basal-like</th>
<th>Normal-like</th>
<th>Her2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malignant (%) within group</td>
<td>2 (5.4)</td>
<td>3 (8.1)</td>
<td>21 (56.8)</td>
<td>6 (16.2)</td>
<td>5 (13.5)</td>
<td>37 (100.0)</td>
</tr>
<tr>
<td>Borderline (%) within group</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Benign (%) within group</td>
<td>1 (5.9)</td>
<td>0 (0.0)</td>
<td>16 (94.1)</td>
<td>0 (0.0)</td>
<td>0</td>
<td>17</td>
</tr>
</tbody>
</table>

Serous ovarian tumors in the study cohort with corresponding intrinsic breast cancer subtypes [22]. The rows outline the tumor types with the representation in each subtype in percent within parentheses. The p-value is calculated using Fisher’s exact test.

doi:10.1371/journal.pone.0107643.t004

Classification of the ovarian tumors in our cohort using the C-signatures demonstrated the presence of all but the C6 signature. Since the C6 signature is characterized mainly by low-grade endometrioid tumors, which were not present in our cohort, this finding supports the ability of gene signatures to capture histological differences, and indicates that the C-signatures are stable and widely applicable across datasets and microarray platforms. The malignant tumors correlated significantly to the C1, C2 and C4 signatures, and as anticipated from the recent data reported from the TCGA, classification into the intrinsic breast cancer subtypes also revealed a significant correlation between malignant ovarian tumors and the basal-like breast cancer subtype [19]. This link was further supported by the finding that the ovarian tumors classified as basal-like in our cohort in turn correlated to the C2 and C4 signatures. A majority of both high-grade serous ovarian and basal-like breast cancers express high levels of proliferation genes [16,19]. Proliferation, among other biological processes, is captured by the Desmedt modules, and we could correlate the C-signatures to these modules and verify the finding in an independent dataset. The C2 “high immune signature” correlated significantly to the STAT1/immune response module, for example, and the C1 “high stromal response signature” to the PLAU/invasion and metastasis module. Taken together, the breast cancer derived Desmedt modules capture the nature of the C-signatures as outlined by both Tothill et al. and the TCGA well [16,18] and provide further biological information regarding the differences in phenotype between the subgroups. Moreover, the statistical correlations shown here further support the link between the subtypes of serous ovarian and breast cancer.

The C3 signature in the original study encompassed borderline (low malignant potential, LMP) tumors and, as expected, the benign and borderline tumors in our cohort correlated significantly to the C3 signature. This signature is characterized by a relative overexpression of genes in the MAPK/ERK pathway,
Figure 3. Functional gene expression modules. Correlations between the ovarian cancer C1–C5 signatures and the functional breast cancer modules by Desmedt et al [33]. Log2 mRNA values are presented on the Y axis. p-values for the highlighted boxes vs. the rest in each plot are calculated using the Mann-Whitney U Test.
doi:10.1371/journal.pone.0107643.g003
line with the fact that type 1 ovarian tumors often harbor mutations in \(KRAS\) and \(BRAF\). Four of five borderline tumors in our cohort displayed \(KRAS\) or \(BRAF\) mutations, and the two malignant tumors harboring mutations in \(KRAS\) in turn correlated to the C3 and C4 ovarian signatures. Interestingly, a significant correlation between the few borderline tumors and the luminal A breast cancer subtype was found. Luminal breast tumors frequently display mutations in the MAPK/ERK pathway, thereby resembling the ovarian type 1 tumors (and the ovarian C3 signature). This is further supported by the independent finding of a significant correlation between the ovarian C3 signature and the \(ESR1/ER\) signaling breast cancer module by Desmedt et al. [33]. Hence, although the borderline tumors in our cohort were interspersed between the malignant and benign neighbors in the clustering analyses, upon comparison with the ovarian C3-signatures and the intrinsic breast cancer subtypes, they showed obvious similarities with benign and low-grade malignant tumors – as anticipated based on the prototypic type 1 tumors they are described as. Furthermore, the luminal A and B (estrogen receptor positive) breast cancer subtypes differ in transcriptional profiles, mutation spectra and overall survival [19]; this is in line with the clinical spectrum observed in low-grade serous ovarian cancer, with low-grade, early-stage tumors showing a favorable prognosis, while low-grade, advanced stage tumors tend to respond poorly to chemotherapy. Despite the generally high expression of ER in ovarian cancer, the response to both tamoxifen and letrozole has been limited [39,40]. The different isoforms of ER seem to vary with the malignant potential, with the beta isoform (\(ER\beta\)) reported to be less expressed in malignant ovarian tumors compared to borderline tumors and benign ovaries, but whether \(ER\beta\) or \(ER\alpha\) influence outcome is not clear [41,42]. In contrast, \(ER\alpha\) is a favorable prognostic factor in breast cancer [43]. A recent study by the Ovarian Tumor Tissue Analysis consortium (OTTA) focused on expression of \(ER\alpha\) and the progesterone receptor (\(PR\)), and in a large series only strong expression of \(PR\), but not \(ER\alpha\), was correlated to increased survival in high-grade serous ovarian cancer. No significant correlations between \(ER\alpha\) or \(PR\) expression and survival were found in multivariate analyses of low-grade serous tumors [44]. The vast majority of the low-grade tumors were however \(ER\) and/or \(PR\) positive, but the few that were negative did not have a significantly different outcome despite the fact that the majority of the low-grade tumors were stage III–IV. Likewise, \(ER\alpha\) is reported to be overexpressed in serous borderline tumors [41]. Taking these results into consideration in light of our findings of a correlation between the ovarian C3 signature and the luminal A breast cancer subtype, it would be interesting to study the response to and the potential effect of endocrine treatment specifically in advanced type I ovarian tumors.

### Conclusions

The findings in this study support that transcriptional signatures indeed capture the biology of transforming events and oncogenic mutations and also support similarities between molecular subtypes of ovarian and breast cancer beyond high-grade serous ovarian cancer and basal-like breast cancer. Though limited series are sensitive to overfitting, importantly, our findings were stable and reproducible in a large independent cohort. The similarities between molecular subtypes of ovarian and breast cancer may be of potential interest for further studies regarding targeted therapies and the use of chemotherapeutic agents in ovarian cancer, as well as biomarker studies. While the proposed similarities between low-grade serous and borderline ovarian (type 1) tumors and luminal breast cancers may in part be attributable to similarities in proliferation rates compared to high-grade ovarian (type 2) and basal-like breast cancers, other biological similarities, such as potential endocrine responsiveness, are thought-provoking and merit further investigation.

### Supporting Information

**Figure S1 Hierarchical clustering.** Unsupervised clustering of malignant and benign ovarian tumors using the 20% most varying probes and including biological replicates. \(n = 66\) tumors. (TIF)

**Figure S2 Hierarchical clustering.** Unsupervised clustering of malignant and benign ovarian tumors using the 20% most varying probes. Clustering performed without biological replicates. \(n = 54\) tumors. (TIF)

**Figure S3 Hierarchical clustering.** Supervised clustering of malignant, borderline and benign tumors based on significant probes from supervised analysis of malignant and benign tumors. Clustering performed without biological replicates. \(n = 59\) tumors. (TIF)

**Figure S4 Hierarchical clustering.** The serous ovarian tumors in the study cohort with corresponding tumor features and assigned ovarian cancer C-signatures and intrinsic breast cancer subtypes. (TIF)

**Figure S5 Functional gene expression modules.** Boxplots representing the correlations between serous ovarian malignant,
borderline and benign tumors in the study cohort and the respective gene expression modules by Desmedt et al. (Desmedt et al., Clin Cancer Res 2008). Log2 mRNA values are presented on the Y axes. p-values are calculated using the Kruskal-Wallis test.

Table 56 Functional gene expression modules. Boxplots representing the correlations between the C1-C6 ovarian cancer signatures in an independent, publicly available dataset and the respective breast cancer gene expression modules by Desmedt et al. (Desmedt et al., Clin Cancer Res 2008). Log2 mRNA values are presented on the Y axes. p-values for the highlighted boxes vs. the rest in each plot are calculated using the Mann-Whitney U Test.

Table S1 Deregulated genes. All significantly deregulated genes (n = 5,944) between malignant and benign ovarian tumors in the study cohort.

Table S2 Molecular subtypes. Significant correlations between malignant, borderline and benign ovarian tumors in the study cohort (n = 59) and the ovarian cancer C-signatures, p-values for each part of the table (separated with double lines) are calculated using Fisher’s exact test.

Table S3 Molecular subtypes. Significant correlations between malignant, borderline and benign ovarian tumors in the study cohort (n = 59) and the basal-like and normal-like breast cancer subtypes, p-values for each part of the table (separated with double lines) are calculated using Fisher’s exact test.

Table S4 Molecular subtypes. Significant correlations between malignant, borderline and benign ovarian tumors in the study cohort (n = 59) and the luminal A breast cancer subtype. The p-value is calculated using Fisher’s exact test.

Table S5 Correlations between molecular subtypes. Significant correlations between assigned C-signatures and the normal-like breast cancer subtype. The p-value is calculated using Fisher’s exact test.

Table S6 Correlations between molecular subtypes. Significant correlations between and between assigned C-signatures and the normal-like breast cancer subtype. The p-value is calculated using Fisher’s exact test.

Table S7 Validation of the centroid classifier. Cross table comparing the original kmeans groups (rows) for the tumors in an independent, publicly available dataset, consisting of malignant and borderline tumors, with the centroid classification for the same tumors (columns). The numbers represent number of tumors. Correlations between the C-signature classifications are highlighted in bold (diagonal). 251 tumors had an assigned kmeans group. 239 of these (95.2%) were correctly re-assigned using nearest centroid classification.

Table S8 Correlations between molecular subtypes. Significant correlations between ovarian cancer C-signatures and intrinsic breast cancer subtypes in an independent, publicly available dataset consisting of malignant and borderline ovarian tumors (n = 285). The C2 signature correlated significantly to the basal-like breast cancer subtype and the C3 signature to the normal-like breast cancer subtype. p-values (*) are calculated using Fisher’s exact test.

Acknowledgments
We would like to acknowledge the patients who have generously contributed to the Skåne University Hospital ovarian tumor biobank.

Author Contributions
Conceived and designed the experiments: JMJ SM MN IH MJ JJ MDV. Performed the experiments: JMJ JJ MDV SK MJ AM SM MJ MN IH JHB PK. Contributed reagents/materials/analysis tools: JMJ SM AM JHB. Contributed to the writing of the manuscript: JMJ SM MN IH JJ MDV PK AM SM MN IH.

References