In situ Imaging for Multiplex Tumor Microbiome Interactions via Imaging Mass Cytometry on Single-Cell Level

Zijian Feng
Shanghai Jiaotong University: Shanghai Jiao Tong University
Xianting Ding (dingxianting@sjtu.edu.cn)
Shanghai Jiao Tong University https://orcid.org/0000-0002-1549-3499
Guangxia Shen
Shanghai Jiaotong University: Shanghai Jiao Tong University
Yuli Hu
First People's Hospital of Wenling
Xin Wang
Shanghai Jiaotong University: Shanghai Jiao Tong University
Yiyang Li
Shanghai Jiaotong University: Shanghai Jiao Tong University
Lulu Zhang
Shanghai Jiaotong University: Shanghai Jiao Tong University
Jie He
Shanghai Jiaotong University: Shanghai Jiao Tong University
Hongxia Li
Shanghai Jiaotong University: Shanghai Jiao Tong University
Ting Zhang
Shanghai Jiaotong University: Shanghai Jiao Tong University
Youyi Yu
Shanghai Jiaotong University: Shanghai Jiao Tong University

Research

Keywords: Imaging mass cytometry, Bacteria, Bioinformatics, Breast cancer, Mass tags, Hybrid probes

Posted Date: June 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-643815/v1

Version of Record: A version of this preprint was published at Cytometry Part A on March 17th, 2022. See the published version at https://doi.org/10.1002/cyto.a.24550.

Abstract

Background: Co-detection of multiplex cancer subtypes and bacteria subtypes in situ is crucial for understanding tumor microbiome interactions in tumor microenvironment. Current standard techniques such as immunohistochemical staining and immunofluorescence staining are limited for their multiplicity. Simultaneously visualizing detailed cell subtypes and bacteria distribution across the same pathological section remains a major technical challenge.

Results: Herein, we developed a rapid semi-quantitative method for in situ imaging of bacteria and multiplex cell phenotypes on the same solid tumor tissue sections. We designed a panel of antibody probes labeled with mass tags, namely **p**rokaryotic and **e**ukaryotic cell **h**ybrid **p**robes for in **s**itu **i**maging (**PEHPSI**). For application demonstration, PEHPSI stained two bacteria subtypes (lipopolysaccharides (LPS) for Gram-negative bacteria and lipoteichoic acid (LTA) for Gram-positive bacteria) simultaneously with four types of immune cells (leukocytes, CD8+T-cells, B-cells and macrophages) and four breast cancer subtypes (classified by a panel of 12 human proteins) on the same tissue section.

Conclusions: We unveiled that breast cancer cells are commonly enriched with Gram-negative bacteria and almost absent of Gram-positive bacteria, regardless of the cancer subtypes (triple-negative breast cancer (TNBC), HER2+, Luminal A and Luminal B). Further analysis revealed that on the single-cell level, Gram-negative bacteria have a significant correlation with CD8+T-cells only in HER2+ breast cancer, while PKCD, ER, PR and Ki67 are correlated with Gram-negative bacteria in the other three subtypes of breast cancers. On the cell population level, in TNBC, CD19 expression intensity is up-regulated by approximately 25% in bacteria-enriched cells, while for HER2+, Luminal A and Luminal B breast cancers, the intensity of biomarkers associated with the malignancy, metastasis and proliferation of cancer cells (PKCD, ISG15 and IFI6) is down-regulated by 29-38%. The flexible and expandable PEHPSI system permits intuitive multiplex co-visualization of bacteria and mammalian cells, which facilitates future research on tumor microbiome and tumor pathogenesis.

Background

For millions of years, bacteria have parasitized and co-evolved with human beings. In this long process, the coexistence and interaction of bacteria and their host cells have played an important role in human physiology[1–4]. Bacteria that are located at the tumor site were also discovered more than a hundred years ago[5]. Afterwards, many studies have also confirmed that bacteria involve in the occurrence, development, and metastasis of tumors[6–8]. A recent report showed that different tumor types have their own unique microbial composition, among which breast cancer has the most abundant bacterial pattern[9]. Compared with the normal control group, the genus *Alistipes* was significantly up-regulated and a genus from the *Sphingomonadaceae* family was relatively decreased in the breast tissues of women with a history of breast cancer[10]. Other studies have found that compared with healthy controls, the microbiota of breast and skin tissues of breast cancer patients are relatively rich in certain specific microorganisms, including Fusobacterium, Atopobium, Gluconacetobacter, Hydrogenophaga, Bacillus,

Enterobacteriaceae, Staphylococcus, Comamonadaceae, and Bacteroidetes[11, 12]. However, scientific issues such as the cell subgroups enriched with bacteria in various tumors, the interaction between bacteria and host cells, and the further influence on the occurrence and development of cancer have not been fully studied[13]. Therefore, the investigation of breast cancer tumor microecology can facilitate to understand the complex interactions between bacteria and cells at the micro and macro levels, and further clarify the role of bacteria in the pathogenesis of cancer.

With the continuous development of immunology, biochemistry, and molecular biology, various bacterial detection techniques have been developed. The most widely used technique is immunohistochemistry (IHC) staining, which is based on the principle of the specific binding of antigen and antibody in immunology. Similar to IHC, immunofluorescence (IF) staining is applied with fluorescein-labelled antibodies and visualized by fluorescence microscopy[14]. These two visual detection methods have been used to locate bacteria on tumor tissue slices[9, 15–17]. However, both techniques only detect a small number of markers at one time, which makes them impractical to simultaneously examine complicated tumor microbiome interactions in situ. In addition, researchers have developed 16S rRNA sequence technology based on the development of molecular biology and molecular genetics. 16S rRNA is the most commonly used molecular clock in the systematic classification of bacteria. Due to its moderate size, 16S rRNA not only reflects the specific bacterial genera, but also can be easily sequenced[18]. As a result, more and more bacteria have been correctly classified or re-classified based on 16S rRNA. Combining 16S rRNA with PCR technology, the type and abundance of bacteria in tumor microenvironment can be quantitatively detected. However, due to the extremely low bacterial biomass in tumor samples, there usually exists background signal interference, and this method does not really visualize the presence of bacteria in situ. When Combining 16S rRNA technology with fluorescence in situ hybridization (FISH) technology, the bacteria distribution on tumor tissue sections can be observed in situ[19–21], but it is difficult to detect other cell-related protein markers. Meanwhile, the detection result of FISH will be affected by the processing of the tested samples, the incubation of the protease before probe hybridization, and the washing after hybridization[22].

By combining IHC, immunocytochemistry and high-resolution laser ablation techniques, imaging mass cytometry (IMC) provides a comprehensive view of the tissue microenvironment[23]. Through in situ-imaging at the single-cell resolution, IMC facilitates in-depth study of protein biomarkers, cell composition and signal pathways[24]. For example, Hartland et al. used IMC to quantify 35 biomarkers simultaneously, obtained high-dimensional pathological images of breast cancer tissue, and characterized intratumoral phenotypic heterogeneity based on spatial analysis and single cell analysis[25]. Daniel et al. further expanded the application of IMC[26]. In addition to the detection of protein markers, they also modified the mRNA probes with metal tags to achieve the simultaneous detection of three mRNA target species and 16 protein markers[26]. Compared with IHC, IF and FISH technologies, IMC mitigates the technique issues such as dislocation from serial sections, fluorescence spectra overlapping, and background fluorescence interference while achieving high-throughput multi-index imaging[27]. These unique strengths make IMC possible to simultaneously achieve high-dimensional in situ imaging of bacteria and cells in the tumor microenvironment and in-depth exploration

of the interaction between bacteria and different cell subtypes based on multi-index detection. Yet, current reports of IMC mainly focused merely on mammalian cell staining and lack a clearly pipeline for prokaryotic and eukaryotic cell co-detection.

Herein, we presented a multiplex rapid semi-quantitative approach for in situ imaging of bacteria and mammalian cells in solid tumor tissue sections with IMC. This method is based on our design of a panel of antibody probes labeled with mass tags, namely **p**rokaryotic and **e**ukaryotic cell **h**ybrid **p**robes for in **s**itu **i**maging (**PEHPSI**), which can simultaneously bind to bacterial and mammalian cell targets in situ on the same tissue sections. Detailed protocols for simultaneous staining of both prokaryotic cells and eukaryotic cells are provided in Materials and Methods session. The specificity and feasibility of PEHPSI were confirmed by cross-validation with IHC imaging results. For application demonstration, PEHPSI stained two types of bacteria (LPS for Gram-negative bacteria, LTA for Gram-positive bacteria), four subtypes of immune cells (CD45 for leukocyte, CD8a for CD8 + T-cell, CD19 for B-cell and CD68 for macrophage), and four subtypes of breast cancer cells (as distinguished by ER, PR, Ki67, HER2, ZC3HV1, ISG15, PKCD and IFI6 staining) simultaneously on the same breast cancer patient specimen. We further characterized the expression of Gram-negative and Gram-positive bacteria in samples from 20 breast cancer patients, using the PEHPSI approach. We unveiled that Gram-negative bacteria commonly existed in all four types of breast cancers (TNBC, Luminal A, Luminal B and HER2+), while no obvious Gram-positive bacteria signal was detected on any tissue sections. Furthermore, at the single-cell level, LPS and CD8a have a significant positive correlation, only in HER2 + breast cancer. For the other three types of breast cancers, namely TNBC, Luminal A, and Luminal B, LPS have a significant correlation with the relevant functional biomarkers such as ER, PR, Ki67, PKCD. Interestingly, the interaction between bacteria-enriched cell clusters and other cell subtypes showed distinct profiles among the four types of breast cancers. For TNBC, the average expression intensity of the B-cell marker CD19 in the area of bacteria-enriched cells was up-regulated by 25% compared to the distant bacteria-negative area. For HER2 + and Luminal A breast cancers, the expression intensity of PKCD and ISG15 in bacteria-enriched cells was down-regulated by 38%. For Luminal B breast cancer, the expression intensity of IFI6 and ISG15 in bacteria-enriched cells was down-regulated by 29%.

Methods

Clinical data

This study is approved by the Ethical Committee and Institutional Review Board of Wenling First People's Hospital in China. The formalin-fixed and paraffin-embedded (FFPE) tissue sections from patient specimens were obtained from the Wenling First People's Hospital. The cohort includes randomly selected 20 patients, who were classified into four subtypes of breast cancers (TNBC, HER2+, Luminal A and Luminal B), with 5 patients in each subtype. Informed consent is available for all the patients.

IHC

IHC staining of bacteria lipopolysaccharide (LPS) and lipoteichoic acid (LTA) expressing was performed according to standard procedures[28]. The formalin-fixed, paraffin-embedded (FFPE) tissue section was dewaxed in xylene (20 minutes × 2) and rehydrated in a gradient series of alcohol (100% ethanol, 95% ethanol, 80% ethanol, 70% ethanol, 0% ethanol; 5 min each). The sections were then subjected to antigen repair in a water bath of citrate buffer (pH 6.0) at 96°C for 30 minutes[29]. In order to detect the bacteria in the cells, 0.2% Triton X-100 was used to help the antibody enter the cell smoothly[9]. The tissue sections were then incubated with 3% hydrogen peroxide in PBS for 15 minutes at room temperature to quench the endogenous peroxidase. After rinsing with PBS, the tissue was blocked with 3% BSA in DPBS for 20 min at 37 ℃. Antibodies to bacterial LPS (mAb WN1 222-5, HycultBiotech #HM6011, 1:1000 dilution) or LTA (mAb 55, HycultBiotech #HM2048, 1:400 dilution) were then added and incubated overnight at 4℃. The readouts of Gram-negative and Gram-positive bacteria expression were obtained using the horseradish peroxidase (HRP) conjugated compact polymer system. Diaminobenzidine (DAB) was used as the chromogen and then the section was counterstained with hematoxylin to visualize the bacteria. Then, we used a microscope to select the stained section. In order to ensure the accuracy of the selected area corresponded to the serial slices scanned by IMC, objective lenses with different magnifications (2×, 10× and 20×) were used for imaging and shooting, and the corresponding area photos were saved as region of interest (ROI) selection for IMC.

Pehpsi Design

For panel design demonstration of PEHPSI, we chose 14 different markers: LPS for Gram-negative bacteria, LTA for Gram-negative bacteria, CD68 for macrophages, CD45 for leukocyte, CD8a for CD8 + T-cells, CD19 for B-cells, as well as ZC3HV1, ISG15, PKCD, IFI6, ER, PR, Ki67 and HER2 for breast cancer cells. Fourteen commercial antibodies were conjugated with unique metal mass tags using Maxpar Metal Labeling Kits (Fluidigm Corp.) according to the manufacturer's instructions[30]. The staining panel was listed in **Supplementary Table 1**. IMC of tissue stained with mass tag conjugated antibodies was performed according to standard procedures[23]. Metal-labeled antibody cocktails were prepared in PBS with 0.5% BSA as per the dilution scheme presented in **Supplementary Table 2**[25]. The tumor tissue sections were dewaxed, rehydrated, and pretreated for antigen retrieval as described above. Slides were then blocked with 3% BSA in DPBS for 45 minutes at room temperature. The sections were incubated overnight at 4℃ with antibody cocktail, followed by washing with 0.1% Triton X-100/DPBS twice and DPBS twice. Then, the tissue sections were stained with DNA intercalator Ir-191 (Fluidigm, catalog No. 201192B) diluted at 1:200 for 30 minutes. After that, the tissue sections were washed with DPBS and ddH_2O and dried before IMC measurements.

Images of ROI on the tissue sections were acquired using a Hyperion Imaging System (Fluidigm). Selected ROIs on the tissue sections were ablated by an argon-based laser at 200 Hz. The raw data, including the intensity of marker expression, were then output for subsequent analysis.

Data Processing And Analysis

After acquiring the imaging mass cytometry data (MCD and text file format) from Hyperion™ Imaging System, MCD Viewer was used to create and save the merged ion image to an 8-bit or 16-bit color TIFF file to visualize the marker expression of each ROI on the tissue sections. Then the raw data was exported to the Single-page or multipage 16-bit or 32-bit grayscale OME-TIFF (.ome.tiff) and text (.txt) for additional processing and analysis using third-party software.

Here CellProfiler was used to convert the pixel data acquired from MCD Viewer into significant single-cell data since identifying cells in the image (cell segmentation) is a necessary condition for quantitative single-cell biological imaging[31].

Finally, histoCAT was used to process and analyze the data of cell segmentation acquired from CellProfiler. After loading samples into histoCAT, the channel heat map drawing function was used to visualize the enrichment sites of bacterial on the ROI. Then, according to the standard instructions, we draw the boxplot of the bacterial-related markers on each ROI, such as LPS, and acquired the heat map of the expression of each marker[32]. In order to further analyze the expression of other markers on bacteria-enriched cells, the PhenoGraph clustering function of histoCAT was applied, and combined with the tSNE function to realize data visualization. Finally, based on the results of PhenoGraph clustering, neighborhood analysis was performed on the cell subgroups with high expression of bacterial markers.

Results

PEHPSI for Bacteria Detection with IMC and Cross-Validation Using IHC

For bacteria detection on cancer tissue slices using IMC, we designed specific metal labeled probes for staining of bacteria. First, we selected two kinds of antibodies, namely anti-LPS and anti-LTA antibodies, which can specifically identify Gram-negative bacteria and Gram-positive bacteria (Fig. 1a **left**). After combining with different lanthanide metal tags, PEHPSI staining panel was prepared and used for staining of tissue sections (Fig. 1a **middle**). Of note, since bacteria possess a layer of bacteria wall that differs from mammalian cells, the staining protocol needs to be re-optimized for prokaryotic and eukaryotic cell co-staining. Detailed protocols are provided in the Methods session. Meanwhile, to confirm the sensitivity and specificity of PEHPSI staining panel, we further performed IHC staining on the serial sections sequential to the tissue sections that were scanned by IMC (Fig. 1a **right**). IMC results showed that Gram-negative bacteria were present in all four types of breast cancers, while the signal of Gram-positive bacteria was not significantly detected (Fig. 1b **middle**, Fig. 1c **middle**). This result is highly consistent with the results of IHC validation (Fig. 1b **top**, Fig. 1c, **Supplemental Fig. 1**). Moreover, multipage 16-bit grayscale OME-TIFF (.ome.tiff) and text (.txt) from MCD Viewer were exported and loaded into Cellprofiler and histoCAT. We obtained the corresponding heatmap through histoCAT to show the intensity distribution of LPS and LTA (Fig. 1b **bottom**, Fig. 1c **bottom**). We quantified the mean

intensity of LPS and LTA in the heatmap through histogram and boxplot (Fig. 1d, e). These results indicated the reliable specificity of PEHPSI and the desirable accuracy of PEHPSI-based IMC detection.

PEHPSI Enables Simultaneous Multiplexed Detection of cancer subtypes, immune cell subtypes and bacteria subtypes on one tissue specimen

IMC platform can reproduce the co-expression of multiple biomarkers in situ. However, existing reports for IMC applications have been mainly applied for mammalian cells only. Therefore, we expanded the PEHPSI channel in order to investigate the multiplex co-existence of cancer subtypes, immune cell subtypes and bacteria subtypes in the tumor microecology. We investigated CD45 (leukocyte), CD68 (Macrophages), CD8a (CD8 + T-cell), CD19 (B-cell), LPS (Gram-negative bacteria), LTA (Gram-positive bacteria) and standard breast cancer functional biomarkers (PKCD, PR, IFI6, HER2, Ki67, ISG15, ER, ZC3HV1). After chelating these biomarkers with the selected lanthanide metal tags, the PEHPSI was prepared and the tissue sections were stained (Fig. 1a). After interrogating the slices with the IMC Hyperion™ Imaging System and selecting the labeled region of interest (ROI),the metal tags on these tissues were detected and presented as a mass intensity spectrum. Then, the protein expression profile on cancer tissues can be obtained based on such metal signal spectrum. MCD Viewer is a post-acquisition data processing software that allows users to visualize, review, and export imaging mass cytometry data. Therefore, we imported the raw data into the MCD Viewer to obtain the ion images of these markers detected after hybridization with PEHPSI on four types of breast cancer tissue slices (TNBC, HER2+, Luminal A, Luminal B) (Fig. 2a-d). In addition to visualizing the images, we quantified the relative pixel intensities of each biomarker in the images (Fig. 2e). Further, we performed Pearson correlation analysis of the mean intensity of each marker on all scanned images. We did not observe a significant consistent pattern of correlation between the mean intensities of LPS and any other biomarkers, demonstrating that co-staining does not induce overlap with the expression of cancer-related markers as well as immune cell markers (Fig. 2f). The overall distribution of markers in the four types of breast cancers also confirmed the specificity of co-staining. Particularly, the expression intensities of ER, PR and HER2 were low in TNBC, and in HER2 + breast cancer, a significant HER2 signal was detected (Fig. 2g). The presentation of these images and the quantification of these biomarkers prove that PEHPSI enables to stain biomarkers in both eukaryotic cells and prokaryotic cells on the same tissue section simultaneously without interfering with each other, and thus is suitable for multiplexed in situ rapid semi-quantitative detection of bacteria and mammalian cells in solid tumor tissue sections.

Heterogeneity of LPS expression in different patients within the same type of breast cancer at the single-cell level

Based on the results of IHC, we observed that LPS was present in all four types of breast cancers and the expression of LTA was extremely low on population level. Therefore, we selected the regions including the high expression areas of LPS and their surrounding areas as scanning ROI for IMC. 3–9 ROIs were selected for each of the five patient tissues from TNBC, HER2+, Luminal A and Luminal B breast cancers. Based on PEHPSI, we simultaneously analyzed expression of CD45 (leukocyte), CD68 (Macrophages),

CD8a (CD8 + T-cell), CD19 (B-cell), LPS (Gram-negative bacteria), LTA (Gram-positive bacteria) and standard breast cancer cell functional biomarkers (PKCD, PR, IFI6, HER2, Ki67, ISG15, ER, ZC3HV1). The expression of these biomarkers is shown in the form of heatmap using histoCAT (Fig. 3a). Then we counted the intensity of LPS and LTA expression in each single cell of all samples. First, we analyzed the overall expression of LPS and LTA in all four types of breast cancers. Due to the weak signal intensity of LTA, we calculated the average value of the expression intensity of LTA in each sample to reduce the error and then performed statistics and comparison. Our results again indicate that gram-negative bacteria (LPS) commonly present in all four types of breast cancers, while gram-positive bacteria (LTA) rarely exist (Fig. 3b), which is also consistent with the result of IHC staining. Moreover, statistics of the expression of LPS and LTA on all ROIs from the four types of breast cancers suggested that the peak intensity of LPS expression in TNBC was the highest, followed by HER2+, Luminal A, and Luminal B, while there was no obvious LTA signal on the tissue sections from any of the four types of breast cancer. Next, representative IMC images of all patients were selected to demonstrate the expression of LPS (Fig. 3c) and we acquired the statistics on the detection results to compare the differences of LPS and LTA expression between patients within the same type of breast cancer (Fig. 3d). Our results indicate that there are significant differences in the expression of bacteria in tumor tissues for different patients even in the same type of breast cancer. For example, TNBC patient-1, HER2 + patient-1, Luminal A patient-1and Luminal B patient-1 have significantly higher peak LPS expression intensity in tissue samples than other patients of the same kind.

In summary, these results indicate that PEHPSI-based detection method can faithfully achieve multiplexed analysis of bacterial and cellular topological distribution on one tissue section. Combined with further analysis, we found that the expression of Gram-negative bacteria in breast cancer tissues of patients with different breast cancer types (peak intensity from high to low: TNBC, HER2+, Luminal A, Luminal B) and that the bacterial composition of the same breast cancer type is largely heterogeneous.

Populational correlation between bacteria enrichment and cell clusters in breast cancer

We then performed dimensionality reduction on all sample data and draw a t-SNE graph (**Supplementary Fig. 2**). We adopted the unsupervised clustering algorithm, PhenoGraph, which identified 17–19 cell sub-clusters from all the 20 breast cancer patients. These cell sub-clusters were characterized by either specific epitopes (e.g., CD45, cluster 19 (TNBC) and CD8a, cluster 10 (HER2+)) or combinations of markers (e.g., IFI6, cluster 15 (Luminal A) and ZC3HV1, cluster 7 (Luminal B)) (Fig. 4a). These cell phenotype clusters appear at different frequencies (Fig. 4b) and were then visualized on the t-SNE map (Fig. 4c).

After the cell subtypes are grouped, the cluster 13 (TNBC), the cluster 9 (HER2+), the cluster 13 (Luminal A), and the cluster 16 (Luminal B) are again identified as the cell subtypes with relatively high expression of LPS in each cancer subtype. In addition to the higher expression of LPS, the cluster 13 (TNBC) has a high expression of PKCD, indicating that bacteria in TNBC are mainly distributed in cancer cells with high expression of PKCD (Fig. 4d). Moreover, in HER2 + breast cancer, bacteria are mainly enriched in CD8a + T-

cells (Fig. 4e); in Luminal A breast cancer, bacteria are mainly distributed in cancer cells that highly express PKCD, PR, and ER (Fig. 4f); in Luminal B breast cancer, bacteria are enriched in cancer cells with high expression of Ki67 (Fig. 4g). These results show that for TNBC, Luminal A and Luminal B breast cancers, bacteria are mainly distributed in cancer cells rather than the other non-cancer cell types. Among them, the cancer functional biomarkers that co-expressed in the bacteria-enriched cells may be related to indicators such as breast cancer malignancy. For HER2 + breast cancer, bacteria are mainly distributed in areas with high CD8 + T cell enrichment.

Spatial correlation of bacterial enriched cells with immune cells and inhomogeneous cancer cells in different types of breast cancers

In order to understand the spatial interaction of bacterial-enriched cells in different breast cancers with other cell subtypes such as immune cells or inhomogeneous cancer cells, we then performed neighborhood analysis in histoCAT. The correlations between different cell clusters in four types of breast cancers are displayed in the form of heatmaps (Fig. 5a). Furthermore, the interaction of bacterial-enriched cells with other cell clusters in the four types of breast cancers is presented as the cell social interaction networks (Fig. 5b) and heatmaps (Fig. 5c). We identified cluster 15 in TNBC, cluster 11 in HER2+, cluster 17 in Luminal A and cluster 2 in Luminal B as the four major cell clusters that correlated most significantly with bacterial-enriched cells. We therefore further counted the signal intensity of all biomarkers particularly in these four cell clusters (Fig. 5d).

Our results showed that the bacteria-enriched cells in TNBC were mainly distributed in the vicinity of CD19-expressing B-cells, but other cell clusters with longer spatial distances did not express CD19 significantly. This shows that in TNBC, bacteria may promote the colonization and survival and reproduction of B-cells in the tumor microenvironment, thereby promoting B-cells to inhibit the survival and development of tumors. For HER2 + breast cancer, the cell clusters that are far away from the bacteria-enriched cells are mainly cancer cells that highly express PKCD. PKCD is significantly up-regulated in highly metastatic breast cancer, and participates in the invasion of breast cancer cells by promoting the excessive activation of the PKC pathway[33]. Therefore, we speculate that in HER2 + breast cancer, bacteria may inhibit the PKC pathway of cancer cells. In Luminal A breast cancer, bacterial-enriched cells have very low spatial correlation with cancer cells that significantly upregulate ISG15. ISG15 is a 15 kDa protein induced by type I interferons,which is highly expressed in various malignant tumors and may be positively correlated with tumor proliferation and metastasis[34]. Similar to Luminal A breast cancer, the expression of ISG15 was lower in cancer cells surrounding bacteria-enriched cells in Luminal B breast cancer compared to distant sites. Furthermore, we quantified the expression intensity of CD19 for TNBC, PKCD for HER2+, ISG15 for Luminal A, IFI6 and ISG15 for Luminal B in cell clusters that are interaction or avoidance to bacterial-enriched cells. Our results indicate that the expression intensity of CD19 of cell clusters highly correlating with bacteria-enriched cells in TNBC is up-regulated by 26% (**Supplemental Fig. 3a**). For HER2 + and Luminal A breast cancer, the expression intensity of PKCD and ISG15 is down-regulated by 38% (**Supplemental Fig. 3b, c**). For Luminal B breast cancer, the expression intensity of IFI6 and ISG15 is down-regulated by 29% (**Supplemental Fig. 3d**). These data also

demonstrate the ability of high-dimensional PEHPSI-based IMC detection to reveal the multiplex correlation between bacteria and cells in tumor microbiota in addition to merely in situ imaging.

Discussion

We describe and validate an approach for multiplexed in situ imaging and simultaneous semi-quantitative bacteria and mammalian cell co-detection in solid tumor tissue sections by IMC. As an emerging detection technology, IMC can achieve high-throughput in-situ detection while avoiding the interference of spectrum overlap or background autofluorescence. Yet, the application of IMC has long been constrained in mammalian cell research and has rarely been applied for tumor microbiome research domain. Herein, we designed a panel of antibody probes labeled with mass tags, namely **p**rokaryotic and **e**ukaryotic cell **h**ybrid **p**robes for in **s**itu **i**maging (PEHPSI). Two bacterial-specific markers are used for PEHPSI, namely LPS for Gram-negative bacteria and LTA for Gram-positive bacteria[35]. Through IHC staining, the accuracy of IMC with PEHPSI for bacteria detection on tumor tissue sections was confirmed. Then, to expand the panel of PEHPSI, we carefully selected a panel of specific markers including CD19, CD45, CD68 and CD8a for immune cells that may exist on tumor tissue sections. For cancer cells, markers including PKCD, PR, IFI6, HER2, KI67, ISG15, ER, ZC3HV1 were also co-examined to characterize the ability of cancer cells to proliferate, metastasize, and the degree of malignancy. Compared with other detection methods such as IHC, IF and FISH, IMC coupled with PEHPSI can detect bacteria in situ while simultaneously quantifying the expression of bacteria and cellular functional markers. Based on the expression and distribution area of different markers, we divided all cells on the breast cancer tissue slices into 17–19 cell sub-clusters by PhenoGraph. Combined with the results of neighborhood analysis, we explored and investigated the correlation between bacteria-enriched cells with other cancer cell subtypes and immune cell subtypes.

Based on IMC's high-throughput detection results, we found that Gram-negative bacteria have different expression intensity in different types of breast cancer, while gram-positive bacteria are basically absent in breast cancer. The concentration and location of bacteria in different types of breast cancer tissues are also heterogeneous. For TNBC, Luminal A and B breast cancers, bacteria are mainly enriched within cancer cells. In addition to the high enrichment of bacteria, these cancer cells express PKCD, ER, PR, KI67 corresponding to the functionality of cancer. In addition to markers related to cancer cells such as PKCD, ER and PR, we found that in HER2 + breast cancer, bacteria are enriched in T-cells that express CD8a. This may be due to the fact that bacteria promoted the up-regulation of protein connexin 43 (Cx43) in tumor cells, which reduced the indoleamine 2, 3-dioxygenase (IDO)-mediated suppression of T-cells, thereby promoting the activation of CD8a + T-cells at the sites of bacterial enrichment[36–38]. Therefore, the up-regulation of T-cell expression caused by bacteria may enhanced antigen presentation in the local tumor environment[39]. The presence of T-cells in tumors is generally associated with better patient prognosis[40]. These facts support the potential of Gram-negative bacteria as a prognostic marker for breast cancer.

The results of the neighborhood analysis further reveal the relevance of bacteria to cells. For TNBC, bacteria-enriched cells are highly correlated with B-cells; for HER2+, Luminal A and Luminal B breast cancers, bacteria-enriched cells are more related to cancer cells with lower malignancy, metastasis and proliferation capacity. In the more malignant TNBC, bacteria may promote the up-regulation of the immune response and thus inhibit cancer cells. In addition to the aforementioned mechanism of bacteria causing upregulation of T-cells, for the upregulation of B-cells where bacteria are enriched, it may be due to the peptidoglycan of Gram-negative bacteria binds to nucleotide-binding oligomerization domain containing 1 (NOD1) to activate the chemokine receptor CCR6, the CCL20-related signal transduction process, thereby promoting the up-regulation of isolated lymphoid follicles (ILFs)[41]. When ILFs mature, they will become large clusters of B-cells. In HER2+, Luminal A and Luminal B breast cancers, although no correlation between bacteria-enriched cells and immune cells was found, the cell clusters where the bacteria are enriched will express fewer markers that are positively correlated with the degree of malignancy, proliferation, metastasis and invasion. There may be multiple mechanisms of action[38, 42]. After the bacteria are targeted to enter the tumor tissue, they colonize and proliferate in the microenvironment. On one hand, they can directly inhibit the proliferation of cancer cells through TLR5 signal caused by flagellin[43] or lead to cancer cell death by causing high levels of reactive oxygen species (ROS)[44]. On the other hand, bacteria will induce innate immune cells (macrophages, dendritic cells and neutrophils) to migrate to the tumors where bacteria are enriched[45–47], and promote the up-regulation of different secretions such as interleukins[48] to inhibit the proliferation and metastasis of cancer cells.

Although this study demonstrated in situ detection and quantification of bacteria in tumors, and further explored the interaction between bacteria and cells in the tumor microenvironment, current approach still has certain limitations. First of all, this study only selected two bacterial markers, LPS and LTA, to roughly classify bacteria in breast tumors. Therefore, it is necessary to further determine the specific subtypes of bacteria in the tumor through 16S rRNA and PCR technology, and then select the corresponding bacterial marker for PEHPSI panel designing. Second, due to the relative slow acquisition speed of IMC-based raw data, we did not recruit an enormous patient cohort. Twenty tumor tissue sections, five patients from each of the four subtypes of breast cancers, were examined in our study. In the future, we plan to design a larger cohort and prospectively collect tumor tissues that are not limited to breast cancer. In addition, detection methods such as gene sequencing and proteomics can be combined to explore the influence of bacteria on the expression level of genes and proteins of host cells, and then analyze the interaction between bacteria and host cells more comprehensively.

Conclusion

In summary, we developed a new method based on IMC with PEHPSI for rapid and semi-quantitative in situ imaging of bacteria and mammalian cells in solid tumor tissue sections, which presents the distribution landscape of bacteria in tumor microenvironment. Through further analysis of the subtypes of bacteria-enriched cells, the correlation between bacteria and tumor development was discussed. Based

on this technology, our data supported the potential of bacteria as a reliable prognostic marker for tumors and the feasibility of bacteria-based tumor immunotherapy.

Abbreviations

PEHPSI
Prokaryotic and Eukaryotic cell Hybrid Probes for in Situ Imaging;
LPS
Lipopolysaccharides;
LTA
Lipoteichoic Acid;
TNBC
Triple-Negative Breast Cancer;
IHC
Immunohistochemistry;
IF
immunofluorescence;
FISH
Fluorescence In Situ Hybridization;
IMC
Imaging Mass Cytometry;
FFPE
Formalin-Fixed and Paraffin-Embedded;
HRP
Horseradish Peroxidase
DAB
Diaminobenzidine
ROI
Region Of Interest

Declarations

Acknowledgements

We gratefully thank the financial support from Shanghai Municipal Science and Technology Project (18430760500 and 2017SHZDZX01), Shanghai Municipal Education Commission Project (ZXWF082101), Shanghai Jiao Tong University Projects (YG2021ZD19, Agri-X20200101, SL2020MS026, 19X190020154, ZH2018ZDA01, YG2016QN24, YG2016MS60, 2020 SJTU-HUJI, 2019 SJTU-Usyd, SD0820016), Shanghai Municipal Health Comission Project (2019CXJQ03), Shanghai Clinical Medical Research Center Project (19MC1910800). Thanks for AEMD SJTU for the supporting.

Disclaimer

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Institute for Personalized Medicine, State Key Laboratory of Oncogenes and Related Genes, School of Biomedical Engineering, Shanghai Jiao Tong University.

Author contributions

Z.F. wrote the paper. G.S. and X.D. conceived and supervised the project. Z.F., Y.H. and X.W. performed all experiments. Y.H. and Y.L. prepared the samples and antibodies. L.Z., H.L., T.Z. and Y.Y provided technical support on the project. All authors edited the paper.

Funding

Funding for this work was provided by Shanghai Municipal Science and Technology Project (18430760500 and 2017SHZDZX01), Shanghai Municipal Education Commission Project (ZXWF082101), Shanghai Jiao Tong University Projects (YG2021ZD19, Agri-X20200101, SL2020MS026, 19X190020154, ZH2018ZDA01, YG2016QN24, YG2016MS60, 2020 SJTU-HUJI, 2019 SJTU-Usyd, SD0820016), Shanghai Municipal Health Comission Project (2019CXJQ03), Shanghai Clinical Medical Research Center Project (19MC1910800).

Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

Ethics approval and consent to participate

Ethics approval for the use of the formalin-fixed and paraffin-embedded (FFPE) tissue sections from patient specimens was granted by the Ethical Committee and Institutional Review Board of Wenling First People's Hospital in China, approval number KY-2020-2006-01. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Wenling First People's Hospital.

Consent for publication

Not applicable

Competing interests

The authors declare no competing interests.

Author details

[1]Institute for Personalized Medicine,State Key Laboratory of Oncogenes and Related Genes, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, 200030, China.

[2]Department of Pathology, Wenling First People's Hospital, Wenling City, Zhejiang Province, 317500, China.

*Correspondence and requests for materials should be addressed to G.S. (email: gxshen@sjtu.edu.cn) or to X.D. (email: dingxianting@sjtu.edu.cn)

References

Acknowledgements

We gratefully thank the financial support from Shanghai Municipal Science and Technology Project (18430760500 and 2017SHZDZX01), Shanghai Municipal Education Commission Project (ZXWF082101), Shanghai Jiao Tong University Projects (YG2021ZD19, Agri-X20200101, SL2020MS026, 19X190020154, ZH2018ZDA01, YG2016QN24, YG2016MS60, 2020 SJTU-HUJI, 2019 SJTU-Usyd, SD0820016), Shanghai Municipal Health Comission Project (2019CXJQ03), Shanghai Clinical Medical Research Center Project (19MC1910800). Thanks for AEMD SJTU for the supporting.

Disclaimer

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Institute for Personalized Medicine, State Key Laboratory of Oncogenes and Related Genes, School of Biomedical Engineering, Shanghai Jiao Tong University.

Author contributions

Z.F. wrote the paper. G.S. and X.D. conceived and supervised the project. Z.F., Y.H. and X.W. performed all experiments. Y.H. and Y.L. prepared the samples and antibodies. L.Z., H.L., T.Z. and Y.Y provided technical support on the project. All authors edited the paper.

Funding

Funding for this work was provided by Shanghai Municipal Science and Technology Project (18430760500 and 2017SHZDZX01), Shanghai Municipal Education Commission Project (ZXWF082101), Shanghai Jiao Tong University Projects (YG2021ZD19, Agri-X20200101, SL2020MS026, 19X190020154, ZH2018ZDA01, YG2016QN24, YG2016MS60, 2020 SJTU-HUJI, 2019 SJTU-Usyd, SD0820016), Shanghai Municipal Health Comission Project (2019CXJQ03), Shanghai Clinical Medical Research Center Project (19MC1910800).

Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

Ethics approval and consent to participate

Ethics approval for the use of the formalin-fixed and paraffin-embedded (FFPE) tissue sections from patient specimens was granted by the Ethical Committee and Institutional Review Board of Wenling First People's Hospital in China, approval number KY-2020-2006-01. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Wenling First People's Hospital.

Consent for publication

Not applicable

Competing interests

The authors declare no competing interests.

Author details

[1]Institute for Personalized Medicine,State Key Laboratory of Oncogenes and Related Genes, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, 200030, China.

[2]Department of Pathology, Wenling First People's Hospital, Wenling City, Zhejiang Province, 317500, China.

*Correspondence and requests for materials should be addressed to G.S. (email: gxshen@sjtu.edu.cn) or to X.D. (email: dingxianting@sjtu.edu.cn)

Figures

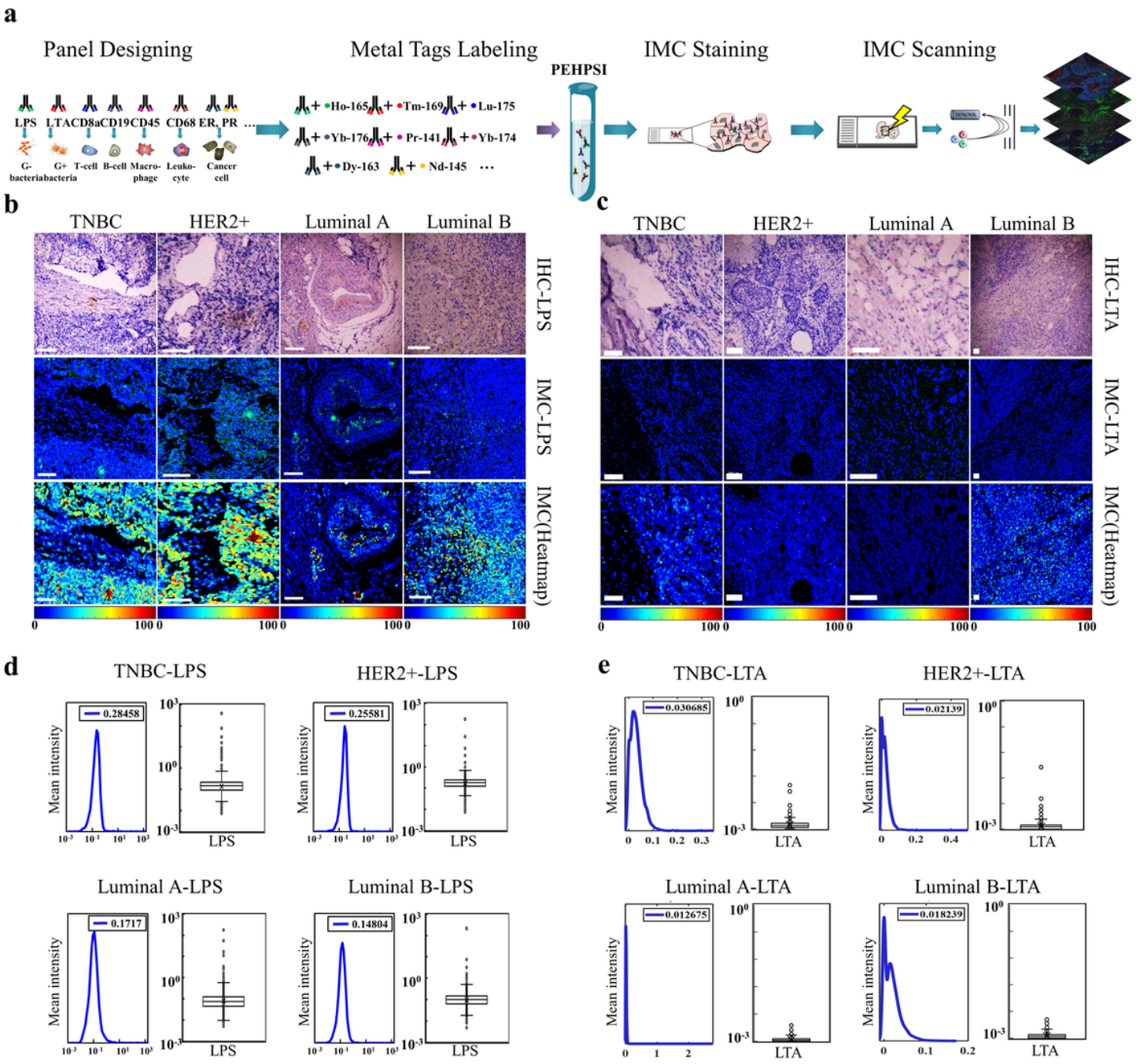

Figure 1

PEHPSI-based IMC detection and IHC cross-validation on serial sections. a PEHPSI is designed to simultaneously target bacteria, immune cells, and cancer cells that may be present on the same tissue section. Non-interfering lanthanide metal tags are selected and chelated with the PEHPSI antibody staining panel that is specifically multiplexed with in situ biomarkers. PEHPSI is prepared for slice staining. After laser ablation, the ionic metals are sent to the mass spectrometer for detection and then assembled into a high-dimension image. b IHC images (brown for LPS, violet for nuclei) and IMC images (green for LPS, blue for DNA) are acquired from serial sections. The corresponding LPS heatmaps of breast cancer tissue section depict the intensity distribution of LPS; scale bar, 100 μm. c IHC images

(brown for LTA, violet for nuclei) and IMC images (green for LTA, blue for DNA) are acquired from serial sections. The corresponding LTA heatmaps of breast cancer tissue section depict the intensity distribution of LTA; scale bar, 100 μm. d The histogram and boxplot for quantifying mean intensity of LPS. e The histogram and boxplot for quantifying mean intensity of LTA.

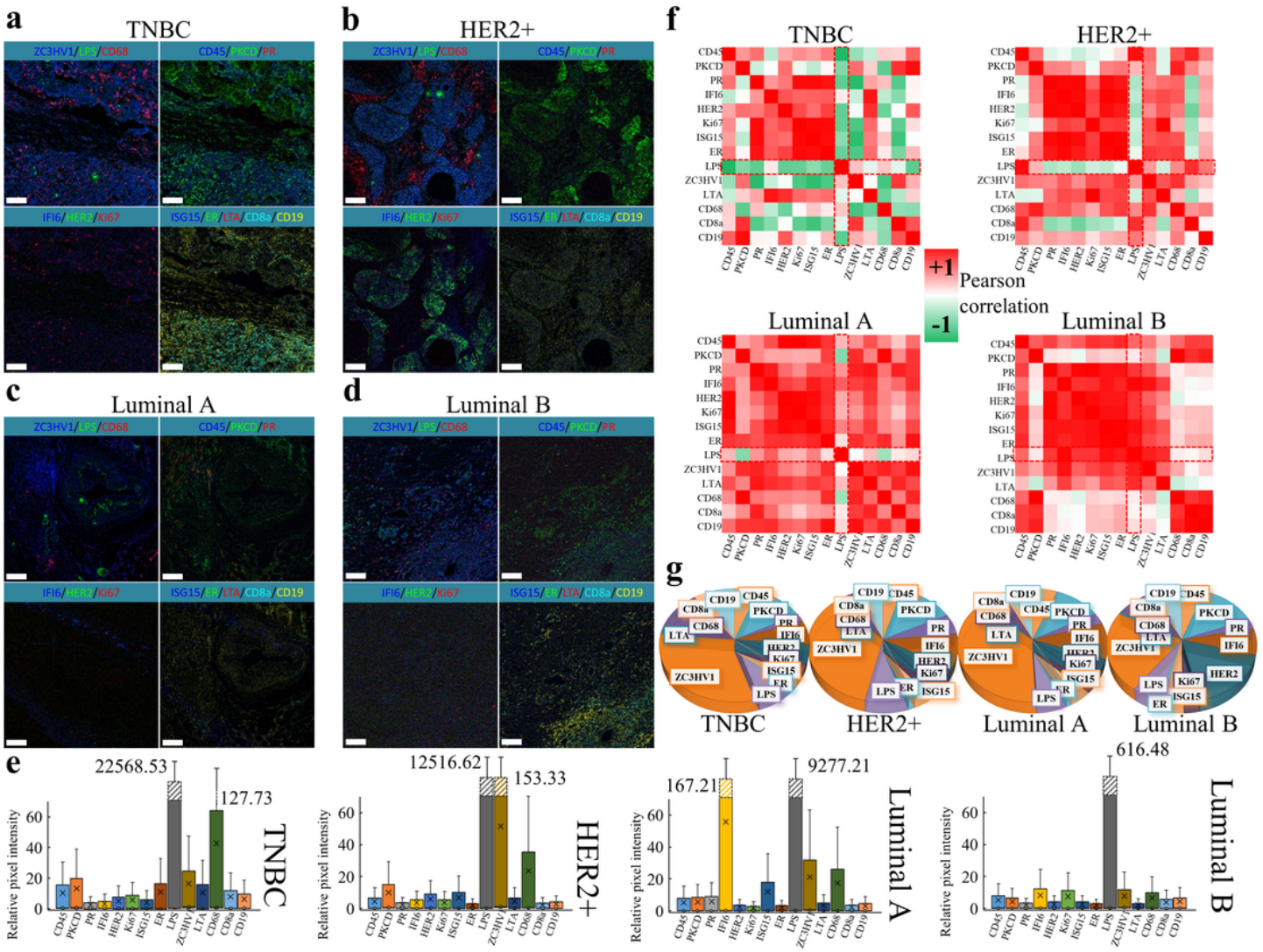

Figure 2

Multiplexed detection of bacteria and mammalian proteins using IMC based on PEHPSI. a-d Representative images of TNBC, HER2+, Luminal A and Luminal B breast cancers stained for ZC3HV1 (blue), LPS (green), CD68 (red) in channel 1; CD45 (blue), PKCD (green), PR (red) in channel 2; IFI6 (blue), HER2 (green), Ki67 (red) in channel 3; ISG15 (blue), ER (green), LTA (red), CD8a (aquamarine), CD19 (yellow) in channel 4 are shown separately. Pseudo-colored raw ion images represent the markers detected; scale bar, 100 μm. e Quantification of the relative pixel intensities of each biomarker in the four channels of representative IMC images of four types of breast cancers in (a-d). f Heatmaps indicate the Pearson correlation of markers in all measured tissue regions. The red dashed box marks the Pearson correlation of LPS with other markers in PEHPSI panel. g The pie charts show the distribution of markers in the four types of breast cancer.

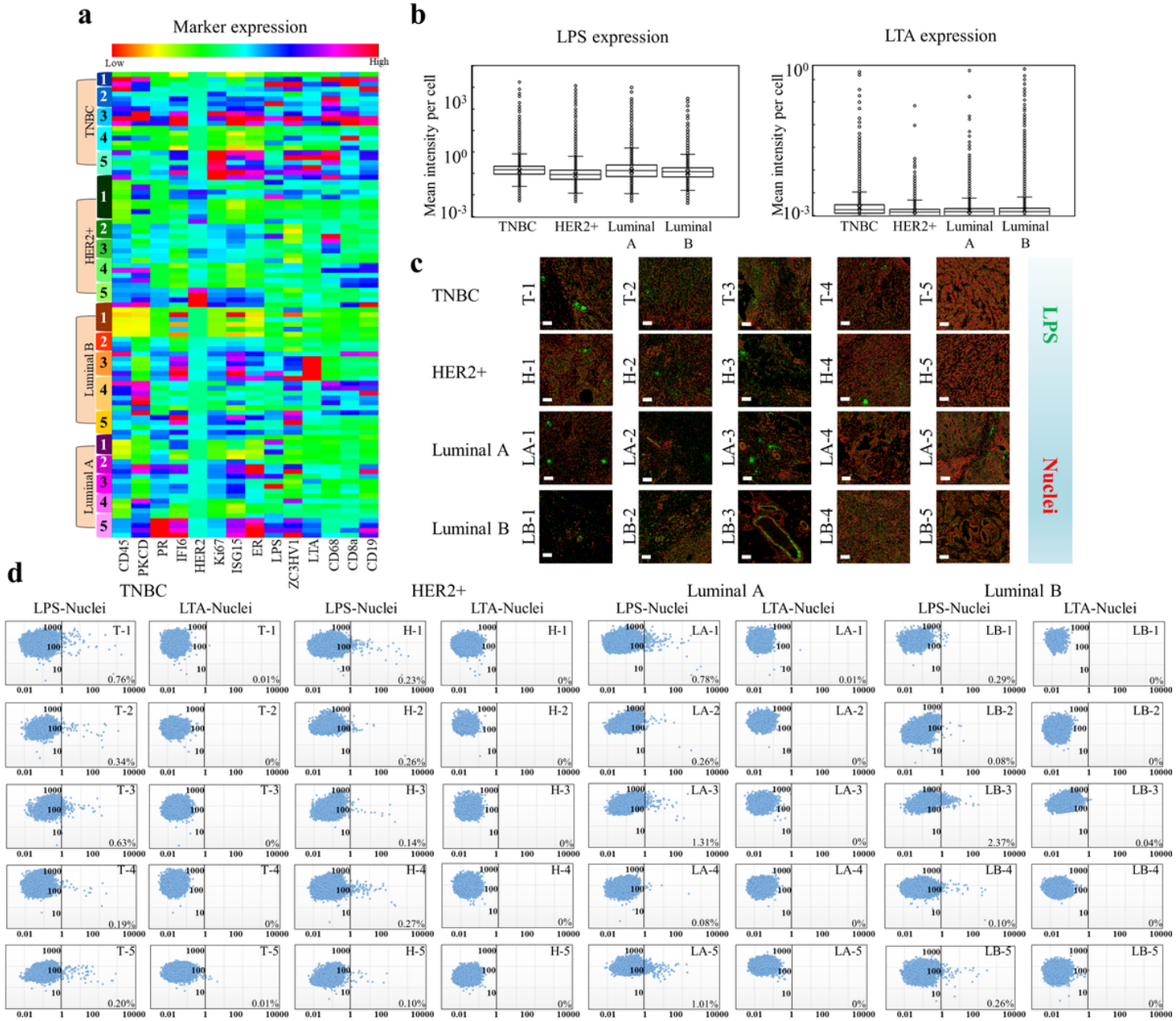

Figure 3

Expression of LPS and LTA in four types of breast cancers. a Heatmap showing the expression of markers on ROIs from 20 patients. The color in heatmap represents the level of marker expression. Five patients (1-5) for each subtype of breast cancers (TNBC, HER2+, Luminal A and Luminal B) were investigated. b Boxplots of LPS and LTA expression in four types of breast cancers (TNBC, HER2+. Luminal A and Luminal B). c Representative IMC images of 20 patients demonstrate the heterogeneous expression of LPS on tissue sections (green for LPS, red for nuclei); scale bar, 100 μm. d Comparison of the expression intensity of LPS and LTA on single cell level in each patient of the four types of breast cancers (TNBC, HER2+, Luminal A and Luminal B). The horizontal axis is the signal intensity of LPS or LTA, and the vertical axis is the signal intensity of the nuclei as a baseline. The proportion of cells with signal intensities greater than 1 is calculated and placed in the lower right corner of each graph.

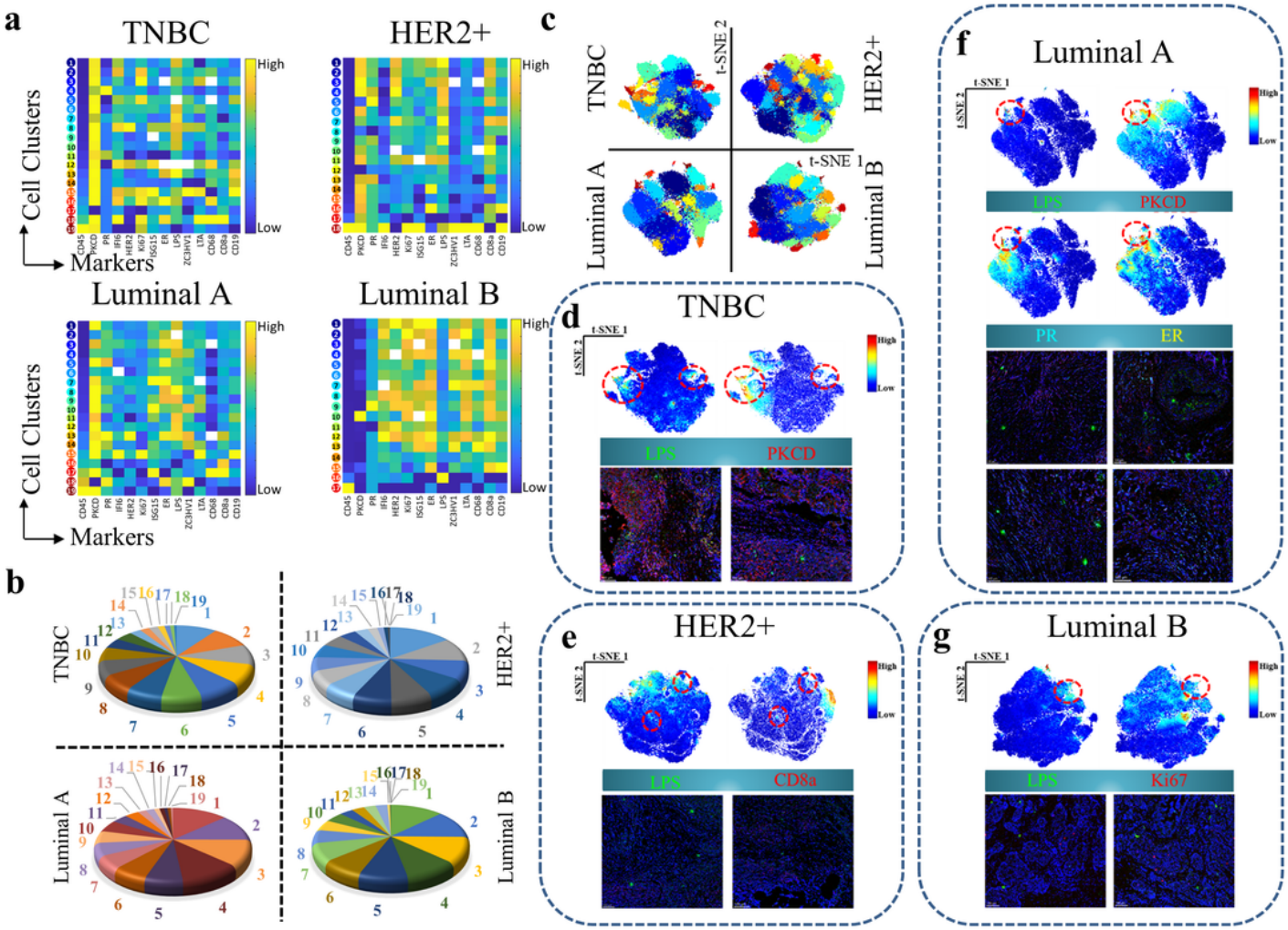

Figure 4

PhenoGraph-based cell subtyping and Single-Cell LPS-cellular protein correlation analysis. a Heatmaps showing the biomarker expression of different cell clusters in four types of breast cancer. 17-19 cell clusters in different color circles from four types of breast cancers were identified by PhenoGraph. b The pie charts indicates the relative proportion of each cell cluster in the clinical cancer subtypes. The code numbers of the cell clusters all correspond to the number in a. c The t-SNE maps depicting the cell

subtypes after PhenoGraph-based clustering in four types of breast cancer. The colors of different cell clusters in the t-SNE maps are consistent with those in a. d-g T-SNE maps and example pseudo-color images of LPS and its co-expression with other cellular markers (PKCD, PR, ER, CD8a, Ki67) in four types of breast cancers (TNBC, HER2+, Luminal A, Luminal B). The red dot ellipse region marks the cell clusters that co-expresses LPS and the marker of cancer cells or immune cells.

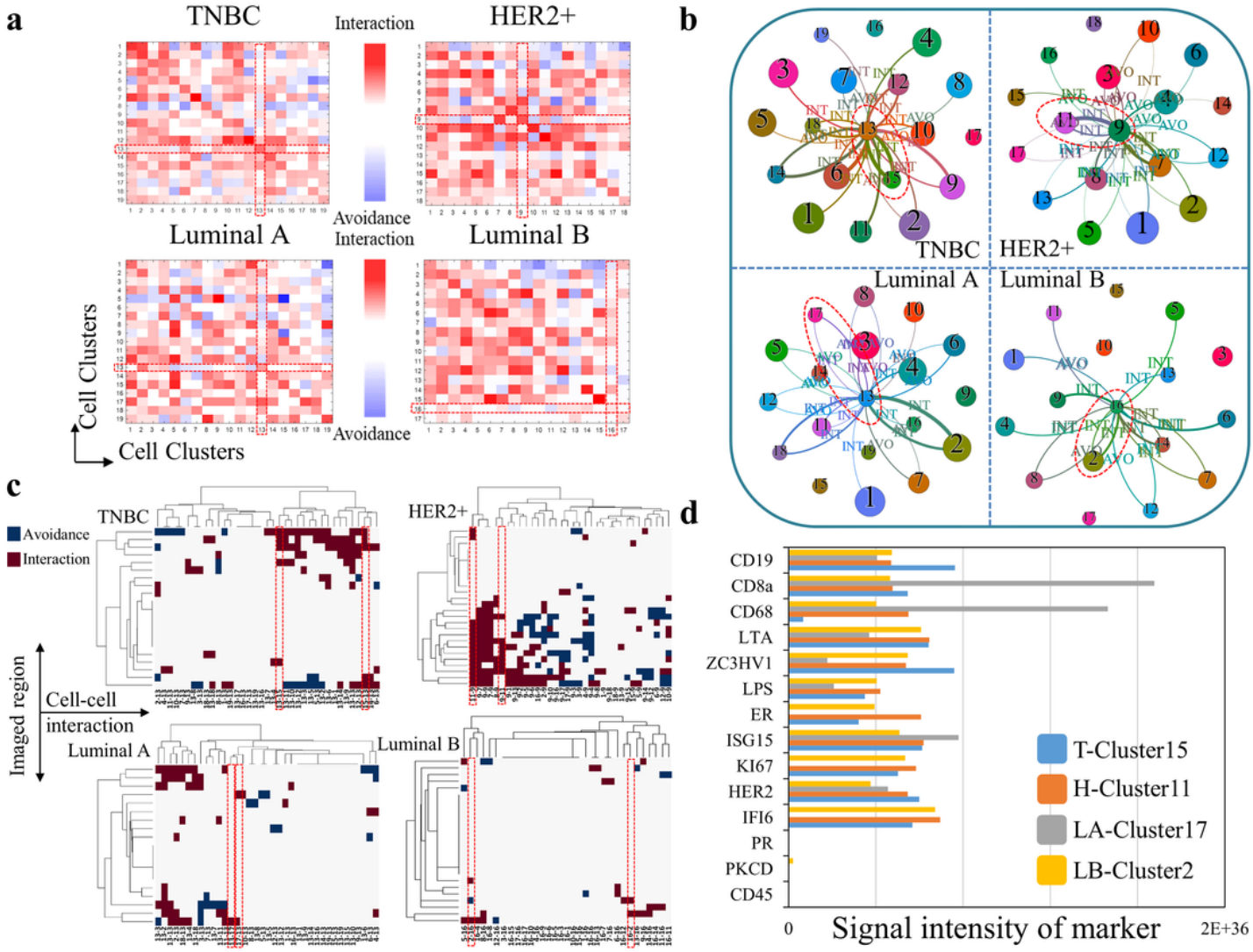

Figure 5

Neighborhood analysis of bacterial-enriched cell clusters reveals spatial correlations between bacteria and cell sub-clusters. a Interactions of all cell clusters in the four breast cancers are represented by heatmaps. All clusters of TNBC, HER2+, Luminal A and Luminal B breast cancers are listed on the left and bottom of the heatmaps. The cell cluster in the row is in close interaction (red) or avoidance (blue) to the cell cluster in the column (P□0.01). The bacterial-enriched cells of four types of breast cancer are highlighted in the red dashed boxes. b The cell social interaction network graphs representing the interaction of bacterial-enriched cells with other cell clusters in the four types of breast cancers. The circles are colored and numbered accroding to the PhenoGraph clusters, and the circle size corresponds to the proportion of the total number of cells in the phenotype. The centered circle represents the bacterial-enriched cell cluster. The lines between the different clusters represent the direction of action in a clockwise direction and the width of the line indicates significance. The types of interaction include INT (interaction) or AVO (avoidance). The red dashed circles mark the cell clusters that interact more significantly with the bacterial-enriched cells. c Heatmaps showing the interactions (red) or avoidances (blue) between the bacterial-enriched cell clusters and different cell clusters in four breast cancers. White area represents interactions that are not present or not significant. The red dashed boxes select the interrelationship between the highlighted cell clusters in b and the bacterial enriched cells on different imaged regions. d The bar graph shows the signal intensities of all biomarkers of cell clusters that interacted significantly with bacterial-enriched cells in four types of breast cancer.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Scheme1.png
- supplementalmaterials.docx