bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Extensive Intratumor Proteogenomic Heterogeneity Revealed by Multiregion Sampling in a High-Grade Serous Ovarian Tumor Specimen

Allison L. Hunt[1,2^], Nicholas W. Bateman[2-4^], Brian L. Hood[2,3], Kelly A. Conrads[2,3], Ming Zhou[1,2], Tracy J. Litzi[2,3], Julie Oliver[2,3], Dave Mitchell[2,3], Glenn Gist[2,3], Brian Blanton[2,4], Kunle Odunsi[5], Anil K. Sood[6], Yovanni Casablanca[2-4], Kathleen M. Darcy[2-4], Craig D. Shriver[4], Uma N. M. Rao[2,3], G. Larry Maxwell[1,2,4] and Thomas P. Conrads[1,2,4]

[1]Department of Obstetrics and Gynecology, Inova Fairfax Medical Campus, 3300 Gallows Rd. Falls Church, VA 22042, USA.

[2]Gynecologic Cancer Center of Excellence, Department of Obstetrics and Gynecology, Uniformed Services University and Walter Reed National Military Medical Center, 8901 Wisconsin Avenue, Bethesda 20889, MD, USA.

[3]The Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., 6720A Rockledge Drive, Suite 100, Bethesda, MD 20817.

[4] The John P. Murtha Cancer Center Research Program, Department of Surgery, Uniformed Services University, 8901 Wisconsin Avenue, Bethesda 20889, MD, USA.

[5]Department of Gynecologic Oncology, Roswell Park Comprehensive Cancer Center, Buffalo, NY, 14263, USA.

[6]Department of Gynecologic Oncology and Reproductive Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

^Author's contributed equally to this work

Address correspondence to: Dr. Thomas P. Conrads (thomas.conrads@inova.org), 3289 Woodburn Rd, Suite 375, Annandale, VA, 22003, or Dr G. Larry Maxwell (george.maxwell@inova.org), 3300 Gallows Road, Falls Church, VA, 22042

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

A series of 200 consecutive thin sections were generated from a high-grade serous ovarian tumor and laser microdissected four spatially separated "core" regions of tumor epithelium, along with tumor epithelium, stroma or whole tissue harvests at 200 µm intervals. These distinct tissue collections were analyzed by quantitative proteomics and RNA-Seq. Unsupervised analyses revealed co-clustering of tumor cores with enriched tumor epithelium, which were distinct from the enriched stroma and whole tumor collections. Strong correlations in protein and transcript abundance in tumor epithelium and stromal collections from neighboring thin sections were decreased in samples harvested just hundreds of microns apart. Stroma (mesenchymal) and tumor epithelium (differentiated) displayed a distinct association with ovarian cancer prognostic molecular sub-types with a 2-year difference in median survival. These data reveal substantial tumor microenvironment protein and transcript expression heterogeneity that directly bear on prognostic signatures and underscore the need to enrich cellular subpopulations for expression profiling.

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Ovarian cancer is the fifth leading cause of death in women in the U.S. [1], with more than 22,000 new cases and nearly 14,000 deaths projected to occur in 2019. Most ovarian cancer cases are diagnosed at an advanced stage after accumulation of numerous genomic and molecular signaling alterations resulting in heterogeneous populations and subpopulations of tumor cells that confound successful therapeutic intervention. The 5-year survival rate of patients diagnosed with metastatic ovarian cancer is less than 30%.

Numerous high-throughput sequencing studies aimed at broadly characterizing the genomic landscape of specific cancer types have been or are being conducted. Although profoundly key in adding to our molecular understanding of cancer, these studies have resulted in limited clinical translation, due in part to an ever increasing body of evidence that points to previously underappreciated levels of heterogeneity in the tumor microenvironment (TME) [2].

Substantial molecular and pathologic differences between HGSOC primary and metastatic tumors have also been described [3, 4]. One study demonstrated that stable and/or regressing tumors lacked common neoepitopes and mutations compared to progressing tumors in the same patient [3], implicating non-somatic factors within the TME as critical determinants of immune response and overall tumor fate. Multiregion sampling has revealed extensive variation between subpopulations of cells within a single tumor [5-7], allowing individual tumor samples to have multiple subtype signatures present with differing levels of activation [8]. Single-cell RNA-Seq of HGSOC samples has revealed grade-specific and cell type-specific transcriptional profiles present within individual tumor specimens [5]. The presence of subclonal cell populations within primary and/or metastatic tumors has been demonstrated to influence the state of immune infiltration and activation [6].

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

In this study, we investigated proteomic and transcriptomic heterogeneity in the TME from a single representative HGSOC patient tumor using laser microdissection of spatially separated tumor "core" regions throughout the depth of the tumor specimen, enriched tumor epithelium and stromal cell populations, as well as whole tumor harvests at ~200 µm intervals. These results reveal stark molecular heterogeneity in the HGSOC tumor microenvironment and underscore the need to account for compartmental heterogeneity in the TME in molecular profiling analyses.

Results

Proteogenomic analyses of discrete cellular populations within a single HGSOC tumor using locoregional multi-sampling

Consecutive thin sections (~200) were generated from a single snap frozen primary HGSOC tumor to support multi-region sampling by LMD followed by integrated quantitative proteomics and transcriptomics (Figure 1A). Tumor "cores" represent defined sub-populations of tumor epithelium that were collected as technical replicate areas from alternating tissue sections spanning the entirety of the patient tumor and combined into four independent replicate sample sets to support proteome and transcriptome analyses. Enriched tumor epithelium, stromal cells, as well as whole tumor collections representing all material on a single tissue section, were collected at ~200 µm intervals from alternating sections at nine different points throughout the tumor block (Figure 1 B). Representative post-LMD collections were imaged along with adjacent hematoxylin and eosin (H&E) stained tissue sections to enable co-registration and automated quantification of tumor and stroma cell populations collected (Supplemental Table 1) for comparison with manual assessment by a board-certified pathologist (UNMR, Supplemental Table 2). Multiregion protein samples were analyzed by quantitative proteomics using a multiplexed, isobaric tagging

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

methodology (tandem mass tags (TMT)-10) and RNA samples by targeted RNA-Seq analyses. These studies quantified a total of 6,053 proteins (Supplemental Table 3) (Supplemental Table 4) and 20,784 transcripts (Supplemental Tables 5 and 6) across all samples.

Tumor and stromal cell populations exhibit diverse proteogenomic profiles

Unsupervised hierarchical cluster analysis of protein (Supplemental Table 7) or transcript (Supplemental Table 8) abundance revealed distinct sub-clustering of tumor cores, enriched tumor epithelium and stroma, as well as whole tumor collections (Figures 2A and B). Two predominant clusters stratifying tumor cores and tumor epithelium from stroma and whole tumor collections are apparent. Notably, the cryopulverized tissue proteome exhibited an intermediate cluster between these two sample groups (Figure 2A), suggesting this sample type represents a mixture of these cellular populations.

Distribution of individual epithelial markers, i.e. CA-125 (*MUC16*), KRT19 and CDH1, and stromal markers, i.e. VCAN, and FAP in HGSOC are highlighted for (Figure 2C). As expected, CA-125 and CDH1 are most abundant in the tumor cores and LMD enriched tumor epithelium collections and were least abundant in LMD enriched stroma and whole slide harvests. VCAN and FAP expression, in contrast, were highest in LMD enriched stroma and whole slide harvests, consistent with these proteins being strong markers of stroma [9, 10]. Cell type enrichment scores using the transcript expression data was performed using xCell (Supplemental Table 9) [11]. Except for Cores 4A/B, all of the LMD tumor cores and LMD enriched tumor epithelium collections correlated most strongly with epithelial cell markers, while the LMD enriched stroma and whole slide harvests had the highest correlation with fibroblast markers, stroma and microenvironment scores (Figure 2C).

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A Spearman rank correlation matrix between 5,742 proteins co-quantified proteins and transcripts revealed concordant results for corresponding samples (Figure 3). Proteins and corresponding transcripts for a given sample largely clustered with other harvests from the same LMD collection type (e.g. Core 1A protein clustered most to its corresponding transcript, while next nearest to other core collections). Unexpectedly, stromal transcripts clustered most strongly with the proteomes from the whole slide harvests. Within the core samples, the strongest correlations, apart from intra-core comparisons, were between the "A" and "B" replicates for each core. The A and B replicates for each core at the proteome-only and transcriptome-only levels had Spearman Rho between 0.65-0.91 and 0.75-0.84, respectively. For replicate cores at the protein and corresponding transcript levels the Spearman Rho were 0.38-0.51. Cores 2A/B and 3A/B also had stronger correlations to each other (Spearman Rho = 0.35-0.47) than either relative to Cores 1 and 4 (Spearman Rho = 0.23-0.42). Due to tissue morphology variations that occurred through the depth of the block, Cores 2 and 3 became less spatially separated in some sections, likely accounting for the observed cross-core similarity.

Supervised analyses of the proteomic and transcriptomic data revealed common canonical pathways, defined by Ingenuity Pathway Analysis (IPA), that were consistently elevated- or decreased within LMD collection types (Table 1A). Aryl hydrocarbon receptor signaling, and the endocannabinoid cancer inhibition pathway were among the top 5 pathways most activated in LMD enriched tumor epithelium relative to stroma in both the proteomic and transcriptomic datasets. Conversely, IL-8 signaling and GP6 signaling were among the most highly downregulated pathways (least activated in enriched tumor epithelium relative to stroma) according to our proteomic and transcriptomic data. Diseases and biofunctions enriched included activation of organismal death, morbidity and mortality, and bleeding in LMD enriched tumor

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

epithelium relative to stroma in both proteome and transcriptome data (Table 1B). Organization of the cytoplasm and cytoskeleton, as well as cell movement and migration were among the biofunctions predicted to be inhibited in tumor relative to stroma. Genes encoding known drug targets which were differentially expressed (LIMMA adjusted p-value <0.01) between the LMD enriched tumor and LMD enriched stroma were identified in the proteomic and transcriptomic data with Log_2 fold-changes > ±1 are listed in Supplemental Table 10. Similar results were not seen using the proteomic data.

Intratumor Proteogenomic Heterogeneity Impacts Prognostic Molecular Signatures Correlating with Altered Disease Outcome in HGSOC

Several large-scale studies have aimed to categorize HGSOC patients into subtypes having distinct molecular signatures correlated with unique prognostic outcomes, primarily using genomic and/or transcriptomic data from whole tumor specimens qualified above a set tumor cellularity/purity threshold [8, 12-17]. Konecny et al applied transcriptional profiling to HGSOC data to determine gene expression signatures and classified the 174 cases into 4 transcriptional subtypes with unique prognostic significance [13]. Per their categorization of tumors into immunoreactive, differentiated, proliferative, and mesenchymal subtypes, patients whose tumors were characterized by mesenchymal signatures had the worst prognosis. We compared the abundance of 401 transcripts matching 635 subtype signature genes (Supplemental Table 4, [13]), (Figure 4, Supplemental Table 11) revealing that transcripts from LMD enriched tumor cores and tumor epithelium were strongly correlated with the differentiated subtype and inversely correlated with the prognostically-poor mesenchymal subtype. Contrary to this, the transcript profile from LMD enriched stroma were strongly correlated to the prognostically-poor mesenchymal subtype.

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

These results were largely recapitulated at the proteomic level from which we co-quantified 123 gene products from the Konecny et al. transcript prognostic signature set (Figure 4). Our protein-level data showed an equivalently negative correlation between the mesenchymal subtype and LMD enriched tumor cores and epithelium, yet a strong correlation between this molecular subtype and the proteomes from the LMD enriched stroma and whole tissue harvests.

We further framed our transcriptomic data against the Liu *et al* signature genes associated with tumor invasiveness thought to influence success of complete surgical debulking [15]. Comparing the 11 transcripts of the Liu *et al* dataset revealed that gene expression was positively correlated with our LMD enriched stroma and whole tissue collections (Figure 4).

Discussion

Ovarian cancers are typically diagnosed at an advanced stage after the accumulation of numerous molecular alterations [18]. In this study, we have demonstrated that LMD enrichment reveals extensive intratumoral proteogenomic heterogeneity, which has critical implications for clinical diagnosis and implementation of appropriate therapeutic interventions, similar to what has recently been described in hepatocellular carcinoma [19].

In an unsupervised hierarchical cluster analysis of proteins (Figure 2A) and transcripts (Figure 2B), all samples of the same LMD collection type clustered together. We found that many transcripts and/or their associated proteins were negatively correlated with the abundance trends of established diagnostic and prognostic signatures within discrete cell populations, as recently described by Zhang et al [20]. The clinical relevance of this finding is born out in correlations at the transcript and protein level for each LMD enriched collection with four HGSOC prognostic molecular subtypes described by Konecny et al [13] (Figure 4). Furthermore, we observe

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

profoundly unique transitioning of molecular subtype assignment based on sampling of the tumor microenvironment at the level of the proteome versus the transcriptome. We further note the correlation of protein and transcript abundance is higher in tumor epithelium versus stroma collections and is likely attributed to the secretory nature of stroma cells [21]. This is supported by the higher correlation of transcript abundance in stroma with protein abundance in whole tumor collections where we also collect extracellular matrix. The predictive value for estimating patient outcome through molecular subtype assignment (using the molecular prognostic signatures identified by Konecny et al. [13]) showed a median survival difference of two years depending on whether the tumor epithelium or stroma and/or mixed tissue is sampled; notably the contribution of the stromal microenvironment to the mesenchymal signature has been recently described [20, 22]. Taken together, these observations provide evidence for the existence of a pathological ovarian stroma and the proposed role of cancer-associated fibroblasts contributing to disease development and/or progression [23-26]. Assignment of the cryopulverized tissue in our study to a particular subtype was less clear due to profound signal averaging of the proteome (Figure 4).

Our results suggest that molecular signatures developed to date have reflected variations in tumor purity that differentially impact outcome; specifically, lower tumor purity contributes to the assignment of tumors to poorer prognostic outcome. Tumors with more interceding stroma have lower resectability that decreases the likelihood of achieving compete resection (R0) during surgical debulking, leading to worse outcome [27, 28]. TCGA and other large-scale tumor characterization efforts to date have applied pre-determined tumor purity cutoff levels to include predominantly high purity tumors in their studies. As a result, molecular drivers of disease progression specific to prognostically poor/low purity tumors are frequently missed [2, 29], as recently realized for colon cancer [30] and gliomas [31].

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We demonstrate the unexpected stromal distribution of many markers previously published as correlative with residual disease and/or suboptimal cytoreduction in ovarian cancer (Figures 2 and 4 and Supplemental Tables 3 and 6) [14, 15]. Among these, *FABP4* and *ADH1B* expression were associated with significantly higher levels of residual disease [14]. Expression and shedding of AKAP12 has been recently described in isogenic cell line models of paclitaxel resistance and elevation of AKAP12 transcript also correlated with decreased survival in HGSOC patients [32]. We observe AKAP12 transcript is elevated in stroma versus tumor epithelium (data not shown) suggesting that elevation of AKAP12 in drug-resistant HGSOC cells may be due to chemotherapy-induced epithelial to mesenchymal transition and association with poor disease outcome may correlate with patients harboring lower-purity, likely mesenchymal-subtype tumors. *COL11A1* transcripts were abundant in the stroma and some of the whole tumor collections, consistent with implications of COL11A1 overexpression by cancer-associated fibroblasts (CAFs) [24]. *PTEN* [33] and *TP53* [34] mutations and/or loss are well-recognized alterations in HGSOC, and in some cases the presence of *BRCA1/2* somatic mutations [12] may be predictive of outcome. Functional loss of BRCA1 has been correlated with dysregulation of several proteins involved in actin cytoskeletal and cell adhesion remodeling for cell migration in epithelial ovarian cancers [35]. *PTEN* and *BRCA1* transcripts were overall reduced in the cores and LMD enriched tumor epithelium collections, although some pure tumor epithelial fractions showed elevated levels.

Compartmentalized expression of genes encoding known anticancer drug targets [36] was found in both the LMD enriched tumor epithelium and stroma collections (Supplemental Table 10). *VEGFA* and *FLT1* (also known as *VEGFR1*) are targets of bevacizumab, an anti-angiogenic therapy often used in treating HGSOC patients [37]. Dysregulated VEGF/VEGFR expression has been demonstrated to contribute to epithelial ovarian cancer development and/or progression

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

through increased vascularization and improved survival of endothelial cells via anti-apoptotic signaling in the newly formed vessels [38, 39], though bevacizumab targeting of VEGF/VEGFR has disputed levels of improvement on HGSOC patient survival [40]. In our study, *VEGFA* transcript expression was significantly enriched in tumor epithelium whereas transcript expression of *FLT1*/*VEGFR1* was enriched in LMD enriched stroma. Several additional drugs often employed in the clinical management of HGSOC patients [36] target genes that we find to be significantly elevated in LMD enriched stroma relative to tumor epithelium, including imatinib, paclitaxel, gemcitabine, doxorubicin, and crizotinib.

In conclusion, we demonstrate a critical need to account for cellular subtype and regional TME proteogenomic heterogeneity in cancer molecular profiling efforts that will substantially enable in-depth characterization of spatially distributed subclonal cell populations that have underappreciated roles in driving carcinogenesis. Further work in this direction will aid clinical diagnosis, improve the efficacy of therapeutic intervention, and increase capabilities to identify druggable molecular markers of disease development and progression.

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Acknowledgements: This study was supported in part by the U.S. Department of Defense - Uniformed Services University of the Health Sciences (HU0001-16-2-0006 and HU0001-16-2-0014) and the Ovarian Cancer Research Program from the Congressionally Directed Medical Research Program (W81XWH-16-2-0038). The authors would also like to acknowledge Lorcan Sherry and Mark Anderson from OracleBio Limited (Scotland, UK) for histopathology image analysis support.

Author Contributions: Contributed to conception: NWB, TPC. Contributed to experimental design: ALH, NWB, TPC. Contributed to identification and acquisition of the patient specimen: GG, DM, JO, GLM, TPC. Contributed to acquisition, analysis and/or interpretation of data: ALH, NWB, BLH, MZ, GW, NP, KAC, TL, BB, TL, UNMR, TPC. Drafted and/or revised the article: ALH, NWB, YC, KO, AKS, KMD, CDS, UNMR, GLM, TPC. Acquired funding for the research: YC, GLM, TPC. All authors read and approved the final manuscript.

Declaration of Interests: TPC receives research funding from AbbVie Inc.

Disclaimer: The views expressed herein are those of the authors and do not reflect the official policy of the Department of Army/Navy/Air Force, Department of Defense, or U.S. Government.

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 1A: Top altered canonical pathways in laser microdissection-enriched tumor epithelium versus stroma identified using Ingenuity Pathway Analysis. Comparison of the most differential pathways and diseases/functions between LMD enriched tumor and stroma was determined using pairwise supervised analysis of proteomic and transcriptomic data with LIMMA adjusted p-value < 0.01 exhibiting a Log_2 fold-change cut-off ± 1.

	Ingenuity Canonical Pathway	Activation z-score	# Targets
Protein	PTEN Signaling	2.449	6
	RhoGDI Signaling	2.309	18
	Antioxidant Action of Vitamin C	2	4
	Aryl Hydrocarbon Receptor Signaling	2	5
	Endocannabinoid Cancer Inhibition Pathway	0.707	8
	Integrin Signaling	-4	20
	IL-8 Signaling	-4	17
	Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	-4.025	21
	LXR/RXR Activation	-4.849	33
	GP6 Signaling Pathway	-5.84	38
Transcript	Oxidative Phosphorylation	3.3	18
	Estrogen-mediated S-phase Entry	3.051	13
	Aryl Hydrocarbon Receptor Signaling	2.858	38
	Endocannabinoid Cancer Inhibition Pathway	2.722	39
	Cyclins and Cell Cycle Regulation	2.683	25
	Phospholipase C Signaling	-4.714	64
	Role of NFAT in Regulation of the Immune Response	-4.727	56
	GP6 Signaling Pathway	-4.811	57
	IL-8 Signaling	-5.166	62
	Colorectal Cancer Metastasis Signaling	-5.222	89

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 1B: Top altered diseases and functions in laser microdissection-enriched tumor epithelium versus stroma identified using Ingenuity Pathway Analysis. Comparison of the most differential diseases/functions between LMD enriched tumor and stroma was determined using pairwise supervised analysis of proteomic and transcriptomic data with LIMMA adjusted p-value < 0.01 exhibiting a Log_2 fold-change cut-off ± 1.

	Diseases or Functions Annotation	Activation z-score	# Targets
Protein	Organismal death	11.708	220
	Morbidity or mortality	11.411	225
	Bleeding	5.755	67
	Dystrophy	3.258	54
	Organ Degeneration	2.974	61
	Microtubule dynamics	-6.472	117
	Organization of cytoskeleton	-6.674	154
	Organization of cytoplasm	-6.674	157
	Migration of cells	-7.707	257
	Cell movement	-7.941	284
Transcript	Organismal death	9.212	1031
	Morbidity or mortality	8.988	1048
	Bleeding	5.565	202
	Edema	5.342	177
	Digestive system cancer	3.484	3417
	Binding of blood cells	-6.37	222
	Cell movement	-6.577	1085
	Organization of cytoplasm	-6.763	672
	Organization of cytoskeleton	-6.81	640
	Size of body	-6.946	388

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

1. Statistics, N.C.f.H., *Health, United States, 2016: With Chartbook on Long-term Trends in Health.* Hyattsville, MD, 2017.
2. Chen, G.M., et al., *Consensus on Molecular Subtypes of High-Grade Serous Ovarian Carcinoma.* Clinical Cancer Research, 2018. **24**(20): p. 5037-5047.
3. Jimenez-Sanchez, A., et al., *Heterogeneous Tumor-Immune Microenvironments among Differentially Growing Metastases in an Ovarian Cancer Patient.* Cell, 2017. **170**(5): p. 927-938 e20.
4. Schwarz, R.F., et al., *Spatial and Temporal Heterogeneity in High-Grade Serous Ovarian Cancer: A Phylogenetic Analysis.* PLOS Medicine, 2015. **12**(2): p. e1001789.
5. Shih, A.J., et al., *Identification of grade and origin specific cell populations in serous epithelial ovarian cancer by single cell RNA-seq.* PloS one, 2018. **13**(11): p. e0206785-e0206785.
6. Zhang, A.W., et al., *Interfaces of Malignant and Immunologic Clonal Dynamics in Ovarian Cancer.* Cell, 2018. **173**(7): p. 1755-1769.e22.
7. Bashashati, A., et al., *Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling.* The Journal of Pathology, 2013. **231**(1): p. 21-34.
8. Verhaak, R.G.W., et al., *Prognostically relevant gene signatures of high-grade serous ovarian carcinoma.* The Journal of clinical investigation, 2013. **123**(1): p. 517-525.
9. Yang, X., et al., *FAP Promotes Immunosuppression by Cancer-Associated Fibroblasts in the Tumor Microenvironment via STAT3–CCL2 Signaling.* Cancer Research, 2016. **76**(14): p. 4124-4135.
10. Ghosh, S., et al., *Up-regulation of stromal versican expression in advanced stage serous ovarian cancer.* Gynecologic oncology, 2010. **119**(1): p. 114-120.
11. Aran, D., Z. Hu, and A.J. Butte, *xCell: digitally portraying the tissue cellular heterogeneity landscape.* Genome biology, 2017. **18**(1): p. 220-220.
12. The Cancer Genome Atlas Research, N., *Integrated genomic analyses of ovarian carcinoma.* Nature, 2011. **474**: p. 609.
13. Konecny, G.E., et al., *Prognostic and therapeutic relevance of molecular subtypes in high-grade serous ovarian cancer.* Journal of the National Cancer Institute, 2014. **106**(10): p. dju249.
14. Tucker, S.L., et al., *Molecular biomarkers of residual disease after surgical debulking of high-grade serous ovarian cancer.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2014. **20**(12): p. 3280-3288.
15. Liu, Z., et al., *Suboptimal cytoreduction in ovarian carcinoma is associated with molecular pathways characteristic of increased stromal activation.* Gynecologic oncology, 2015. **139**(3): p. 394-400.
16. Leong, H.S., et al., *Efficient molecular subtype classification of high-grade serous ovarian cancer.* The Journal of Pathology, 2015. **236**(3): p. 272-277.
17. Wang, C., et al., *Pooled Clustering of High-Grade Serous Ovarian Cancer Gene Expression Leads to Novel Consensus Subtypes Associated with Survival and Surgical Outcomes.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2017. **23**(15): p. 4077-4085.

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

18. Torre, L.A., et al., *Ovarian cancer statistics, 2018.* CA: A Cancer Journal for Clinicians, 2018. **68**(4): p. 284-296.
19. Buczak, K., et al., *Spatial Tissue Proteomics Quantifies Inter- and Intratumor Heterogeneity in Hepatocellular Carcinoma (HCC).* Molecular & Cellular Proteomics, 2018. **17**(4): p. 810-825.
20. Zhang, Q., C. Wang, and W.A. Cliby, *Cancer-associated stroma significantly contributes to the mesenchymal subtype signature of serous ovarian cancer.* Gynecologic Oncology, 2019. **152**(2): p. 368-374.
21. Furuya, M., *Ovarian cancer stroma: pathophysiology and the roles in cancer development.* Cancers, 2012. **4**(3): p. 701-724.
22. Schwede, M., et al., *The impact of stroma on the discovery of molecular subtypes and prognostic gene signatures in serous ovarian cancer.* bioRxiv, 2018: p. 496406.
23. Bu, L., et al., *Biological heterogeneity and versatility of cancer-associated fibroblasts in the tumor microenvironment.* Oncogene, 2019.
24. Jia, D., et al., *A COL11A1-correlated pan-cancer gene signature of activated fibroblasts for the prioritization of therapeutic targets.* Cancer letters, 2016. **382**(2): p. 203-214.
25. Narikiyo, M., et al., *Molecular association of functioning stroma with carcinoma cells in the ovary: A preliminary study.* Oncology letters, 2019. **17**(3): p. 3562-3568.
26. Eckert, M.A., et al., *Proteomics reveals NNMT as a master metabolic regulator of cancer-associated fibroblasts.* Nature, 2019.
27. Torres, D., et al., *Factors that influence survival in high-grade serous ovarian cancer: A complex relationship between molecular subtype, disease dissemination, and operability.* Gynecologic Oncology, 2018. **150**(2): p. 227-232.
28. Hamilton, C.A., et al., *The impact of disease distribution on survival in patients with stage III epithelial ovarian cancer cytoreduced to microscopic residual: a Gynecologic Oncology Group study.* Gynecologic oncology, 2011. **122**(3): p. 521-526.
29. Aran, D., M. Sirota, and A.J. Butte, *Systematic pan-cancer analysis of tumour purity.* Nature Communications, 2015. **6**: p. 8971.
30. Mao, Y., et al., *Low tumor purity is associated with poor prognosis, heavy mutation burden, and intense immune phenotype in colon cancer.* Cancer management and research, 2018. **10**: p. 3569-3577.
31. Zhang, C., et al., *Tumor Purity as an Underlying Key Factor in Glioma.* Clinical Cancer Research, 2017. **23**(20): p. 6279-6291.
32. Bateman, N.W., et al., *Elevated AKAP12 in paclitaxel-resistant serous ovarian cancer cells is prognostic and predictive of poor survival in patients.* Journal of proteome research, 2015. **14**(4): p. 1900-1910.
33. Martins, F.C., et al., *Combined image and genomic analysis of high-grade serous ovarian cancer reveals PTEN loss as a common driver event and prognostic classifier.* Genome biology, 2014. **15**(12): p. 526-526.
34. Ahmed, A.A., et al., *Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary.* The Journal of pathology, 2010. **221**(1): p. 49-56.
35. Gau, D.M., et al., *BRCA1 deficiency in ovarian cancer is associated with alteration in expression of several key regulators of cell motility - A proteomics study.* Cell cycle (Georgetown, Tex.), 2015. **14**(12): p. 1884-1892.
36. Sun, J., et al., *A systematic analysis of FDA-approved anticancer drugs.* BMC systems biology, 2017. **11**(Suppl 5): p. 87-87.

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

37. Rossi, L., et al., *Bevacizumab in ovarian cancer: A critical review of phase III studies.* Oncotarget, 2017. **8**(7): p. 12389-12405.
38. Skirnisdottir, I., T. Seidal, and H. Åkerud, *The relationship of the angiogenesis regulators VEGF-A, VEGF-R1 and VEGF-R2 to p53 status and prognostic factors in epithelial ovarian carcinoma in FIGO-stages I-II.* International journal of oncology, 2016. **48**(3): p. 998-1006.
39. Masoumi Moghaddam, S., et al., *Significance of vascular endothelial growth factor in growth and peritoneal dissemination of ovarian cancer.* Cancer metastasis reviews, 2012. **31**(1-2): p. 143-162.
40. Tewari, K.S., et al., *Final Overall Survival of a Randomized Trial of Bevacizumab for Primary Treatment of Ovarian Cancer.* Journal of Clinical Oncology. **0**(0): p. JCO.19.01009.
41. Schumacher, R.T., et al., *An automated sample preparation solution for nucleic acid and protein extraction from cells and tissues.* American laboratory (Fairfield), 2002. **34**(16): p. 38-43.
42. Kall, L., et al., *Semi-supervised learning for peptide identification from shotgun proteomics datasets.* Nat Methods, 2007. **4**(11): p. 923-5.
43. Lazar, C., et al., *Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies.* J Proteome Res, 2016. **15**(4): p. 1116-25.
44. Aran, D., Z. Hu, and A.J. Butte, *xCell: digitally portraying the tissue cellular heterogeneity landscape.* Genome Biol, 2017. **18**(1): p. 220.
45. Bateman, N.W. and T.P. Conrads, *Recent advances and opportunities in proteomic analyses of tumour heterogeneity.* The Journal of Pathology, 2018. **244**(5): p. 628-637.

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure Legends

Figure 1. Study workflow. Illustration of histological tissue preparation, laser microdissection, proteomic analysis (TMT LC-MS/MS), and RNA-Seq (Ion Torrent) (A), with representative pre- and post-LMD images from the top and bottom levels of the tissue (B). (A) Four spatially separated core regions with adjacent replicate regions were laser microdissected (LMD) from each of 100 or 50 slides evenly distributed through the depth of the specimen for proteomics or transcriptomics, respectively. Additional sets of 9 slides were each discretely microdissected for all remaining tumor and stroma after harvesting the cores by LMD, as well as a nearest neighboring whole tissue harvest. The remainder of the specimen was cryopulverized in liquid nitrogen. Proteins and transcripts isolated from each of these distinct collections were analyzed by isobaric tagging and high-resolution liquid chromatography-tandem mass spectrometry and next generation sequencing, respectively. (B) Pre- and post-LMD images were collected for tissue sections are shown for the top level used for proteomic analysis from tissue section #7 as the representative H&E, #8 for collection of tumor cores, then LMD enriched tumor epithelium, then LMD enriched stroma, and #12 for collection of tumor cores, then whole tumor harvest. Images for the bottom level are from tissue sections #176 as the representative H&E, #177 for collection of tumor cores, then LMD enriched tumor epithelium, then LMD enriched stroma, and #178 for collection of tumor cores, then whole tumor harvest.

Figure 2. Unsupervised hierarchical cluster analysis of 1,928 differentially abundant proteins (A) and 3,861 transcripts (B) with median absolute deviation (MAD)>0.5. (C) Protein and transcript abundance of epithelial and stromal markers in HGSOC as well as cellular admixture analyses (xCell - Aran D, 2017).

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 3. Protein-RNA Spearman Correlation Matrix. Spearman correlation analysis of 5,742 genes that were co-measured as proteins and corresponding transcripts.

Figure 4. Protein and transcript abundance of epithelial and stromal markers in HGSOC as well as cellular admixture analyses (xCell - Aran D, 2017).

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Methods

Tissue Specimen

A surgically resected primary HGSOC specimen was obtained from a BRCA1$^{+/-}$ patient with a familial history of ovarian cancer under an IRB-approved research protocol. The specimen was frozen within 30 minutes following surgery after being embedded in Optimal Cutting Temperature (O.C.T.) medium (Fisher Scientific). The patient was stage 3C, grade 3 and chemotherapy-naïve at the time of surgical debulking. The tissue was sectioned into ~200 consecutive slices ranging from 165-209 mm^2, 10 µm thick (2 mm total depth), by microtome and placed on PEN membrane slides (Leica Microsystems). The tissue sections used in this study spanned a depth of 1.91 mm. Representative sections were mounted onto glass slides without the PEN membrane after every 20 sections (200 µm) and stained with hematoxylin and eosin (H&E). Tissue sections on PEN membrane slides were H&E-stained prior to LMD. For the slides microdissected for proteomics, phosphatase inhibitors (Sigma Aldrich) were added to the 70% ethanol fixative and the first DEPC water; slides for transcriptomics were stained with RNAse inhibitors (ProtectRNA; Sigma Aldrich) added into all aqueous solutions.

Laser Microdissection

Four regionally separated areas within the tissue were discretely laser microdissected for isolation of pure tumor epithelium on a LMD7 (Leica Microsystems) from each of 100 slides for proteomics or 50 slides for transcriptomics. These four microdissected “cores” were pooled from approximately 1 mm^2 per slide for a depth spanning the entire block (1 mm x 1 mm x 2 mm). An adjacent 1 mm^2 was microdissected to serve as a replicate from each core; in total each core plus their respective adjacent replicate regions covered an approximate 2 mm x 2 mm x 2 mm area. For

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

a set of 9 sections (200 µm apart) each for proteomics and transcriptomics, all remaining tumor epithelium (between 20-44 mm^2/section) and stroma (44-80 mm^2/section) were discretely collected after isolation of the 8 core regions. A second set of 9 sections was used for non-discrete whole-slide mixed tissue harvests of all remaining tissue (96-155 mm^2/section), with the exclusion of necrosis, blood, and fat. Microdissected tissue was collected into LC-grade water (Fisher Scientific) or Buffer RLT (Qiagen) for proteomic or transcriptomic analysis, respectively.

OracleBio Image Analysis

Pre- and post-LMD images were collected using the Aperio ScanScope XT slide scanner (Leica Microsystems). Image analysis was performed by OracleBio using the Indica Labs HALO platform. Post-LMD images were used to develop classification algorithms for identification of the "dissection area" and "all remaining tissue". A separate algorithm was developed which involved co-registration of post-LMD images with corresponding adjacent H&E-stained sections on glass slides for detection and quantification of the cell nuclei abundance within the LMD regions. Quantification of the size and number of cells and nuclei harvested in each LMD collection was determined using matched sets of reference glass H&E sections and PEN membrane slides following LMD enrichment of tumor (n=15) and stroma (n=6) cell populations.

Pathology Review

Reference glass H&E sections were manually reviewed by a board-certified pathologist (UNMR) for estimation of percent tumor cellularity and composition of the TME. Cellular features examined included the percent by area of necrosis, stroma, normal ovarian epithelium, lymphocytes, and polymorphonuclear leukocytes (PMN).

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Peptide Preparation for TMT LC-MS/MS

LMD tissue in LC-grade water was dried, re-suspended in 100 mM triethylammonium bicarbonate (TEAB)/10% acetonitrile (ACN), and digested using SMART trypsin (1 µg/30 mm^2 tissue; Thermo Fisher) and pressure cycling technology (PCT) [41]. The remaining tissue embedded in the O.C.T. block was washed with water and cryopulverized using liquid nitrogen in a mortar and pestle. Cryopulverized tissue was re-suspended in 100 mM TEAB/10% ACN and digested using SMART trypsin (1 µg protease/ 30 mm^2 tissue) and PCT. Peptide concentrations from trypsin digests were determined using the bicinchoninic acid assay (Pierce BCA).

Tryptic peptides were labeled using isobaric Tandem Mass Tags (TMT) according to the manufacturer's instructions using the TMT 11-Plex Kit from Thermo Fisher, excluding labeling with TMT-131C. Briefly, 30 µg peptides from the 8 cores, 9 non-discrete whole slide harvests, and cryopulverized tissue sample were individually mixed with respective TMT reagents for 1 h at room temperature, then quenched using 5% hydroxylamine. For the 9 discrete collections each of remaining LMD-enriched tumor epithelium or stroma, 10 µg peptides/sample were similarly TMT labeled.

Each TMT-10 multiplex set of samples were loaded onto C-18 trap column in 10 mM NH_4HCO_3 (pH 8.0) and fractionated by basic reversed-phase liquid chromatography (bRPLC) into 96 fractions through development of a linear gradient of acetonitrile (0.69% acetonitrile/min). Thirty-six concentrated fractions were generated by pooling the samples in a serpentine manner. Ten percent (volume) of each fraction was removed for LC-MS/MS.

Liquid Chromatography-Tandem Mass Spectrometry

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The TMT-10 sample multiplex bRPLC fractions (36 total fractions for global proteomics) were analyzed by LC-MS/MS employing a nanoflow LC system (EASY-nLC 1200, ThermoFisher Scientific, Inc.) coupled online with an Orbitrap Fusion Lumos Tribrid MS (ThermoFisher Scientific, Inc.). In brief, each sample (5 μL) was loaded on a nanoflow HPLC system outfitted with a reversed-phase trap column (Acclaim PepMap100 C18, 2 cm, nanoViper, ThermoFisher Scientific, Inc) and a heated (50 °C) reversed-phase analytical column (Acclaim PepMap RSLC C18, 2 μm, 100 Å, 75 μm × 500 mm, nanoViper, ThermoFisher Scientific, Inc) connected online with an Orbitrap mass spectrometer. Peptides were eluted by developing a linear gradient of 2% mobile phase B (95 % acetonitrile with 0.1% formic acid) to 32% mobile phase B within 120 min at a constant flow rate of 250 nL/min. High resolution (R=60,000 at *m/z* 200) broadband (*m/z* 400-1600) mass spectra (MS) were acquired from which the top 12 most intense molecular ions in each MS scan were selected for high-energy collisional dissociation (HCD, normalized collision energy of 38%) acquisition in the orbitrap at high resolution (R=50,000 at *m/z* 200). Monoisotopic precursor selection mode was set to “Peptide” and MS1 peptide molecular ions selected for HCD were restricted to z = +2, +3 and +4. The RF lens was set to 30% and both MS1 and MS2 spectra were collected in profile mode. Dynamic exclusion (t=20s at a mass tolerance=10 ppm) was enabled to minimize redundant selection of peptide molecular ions for HCD.

Quantitative Proteomic Data Processing Pipeline for Global Proteome Analysis

Peptide identifications were generated by searching the .raw data files with a publicly-available, non-redundant human proteome database [Swiss-Prot, Homo sapiens, Proteome UP000005640, 20,257 sequences, downloaded 12-01-2017; http://www.uniprot.org/] appended with porcine trypsin (Uniprot: P00761) and iRT peptide [40] sequences using Mascot (Matrix

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Science) and Proteome Discoverer (Thermo Fisher Scientific). The .raw data files corresponding to each LC-MS/MS injection of the 36 offline high pH RPLC-fractions for a given TMT-10 multiplex were searched using the following parameters: precursor mass tolerance of 10 ppm, fragment ion tolerance of 0.05 Da, a maximum of two tryptic miscleavages, static modification for TMT reporter ion tags (229.1629 Da) on N-termini and lysyl residues, and dynamic modifications for oxidation (15.9949 Da) on methionine residues. The resulting peptide spectral matches (PSMs) were filtered using a false-discovery rate (FDR) < 1.0% (q-value < 0.01), as determined by the Percolator [42] module of Proteome Discoverer (ThermoFisher Scientific, Inc.). TMT reporter ion intensities were extracted using Proteome Discoverer (ThermoFisher Scientific, Inc.) at a mass tolerance of 20 ppm and PSMs lacking a TMT reporter ion signal in TMT channel *m/z* 126 (TMT-126, the pooled study reference combined from all sample digests), PSMs lacking TMT reporter ion intensity in all TMT channels, or PSMs exhibiting an isolation interference of ≥50% were excluded from downstream analyses. Log_2-transformed TMT reporter ion ratios corresponding to individual tissue samples were calculated for each PSM against the pooled reference standard (TMT-126 channel). Log_2-transformed PSM abundance distributions were normalized by calculating the mode-centered z-score transformation for each channel in the TMT-10 multiplex as follows: normalized PSM (Log_2Ratio) = [PSM (Log_2Ratio) – ModeCenter PSM (Log_2Ratio) / σ PSM (Log_2Ratio). For global protein level abundance, the abundance of proteins identified by unique PSMs (i.e. in which a PSM maps uniquely to a single protein accession) were determined by calculating the median Log_2-transformed abundance ratios of all PSMs corresponding to a unique protein accession. The abundance of PSMs mapping to multiple proteins (i.e. "multi-mapper" PSMs) were compared to mapped, unique protein abundances using a mean squared error approach to assign these to unique proteins based on comparative abundance analyses. Briefly,

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

mean squared Log_2-transformed abundance ratios are calculated for multi-mapper PSMs with intensity data observed in ≥50% of all TMT channels for a given TMT sample plex; redundant and multiply charged versions of a given PSM were considered as unique biochemical events. Multi-mapper PSMs were assigned to the corresponding unique protein accession exhibiting the smallest difference in relative abundance levels comparatively. Multi-mapper PSMs mapping to protein accessions not identified by any unique PSMs were excluded from downstream analyses. Protein-level abundance was calculated from normalized, median Log_2-transformed TMT reporter ion ratio abundances from a minimum of two PSMs corresponding to a single protein accession. Normalized Log_2-transformed protein-level abundance for each TMT-10 multiplex were merged and protein-level abundance for proteins not quantified in all samples, but in ≥50%, were imputed using a k-nearest neighbor (k-NN) strategy using pamr prediction analysis for microarrays R-package [43].

RNA sequencing

LMD tissue in Buffer RLT was purified using the RNeasy Micro Kit (Qiagen) per the manufacturer's instructions. RNA concentrations were determined using Qubit HS and BR kits (Thermo Fisher). RNA integrity numbers (RIN) were calculated using the RNA 6000 Pico Kit 2100 Bioanalyzer (Agilent). All RNA was high-quality with RIN values of 7.2 or greater.

RNA samples were reverse transcribed from 10 ng input using the SuperScript VILO cDNA Synthesis Kit. Barcoded cDNA libraries containing 5 LMD samples plus a Universal Human Reference RNA (UHR) standard (Stratagene) were prepared on the Ion Chef System using the Ion Ampliseq Chef DL8 materials and the Ion AmpliSeq Transcriptome Human Gene Expression Panel Chef-ready Kit (Thermo Fisher). Libraries were then purified via solid phase reversible immobilization (SPRI) using AMPure XP beads (Beckman Coulter) to remove

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

fragments less than 100 bp, quantified by qPCR (TaqMan Quantitation Kit; Applied Biosystems), and 25 µl of 100 pM diluted library was used for templating, amplification via emulsion PCR, and loading onto Ion 550 chips on the Ion Chef System.

Sequencing was performed on the Ion Torrent S5 XL (Thermo Fisher) and mapped to the hg19 human reference transcriptome (hg19_Ampliseq_Transcriptome_21K_v1). Successful sequencing runs achieved ≥18M reads/sample (with one exception) and 169-234X AQ20 mean coverage depth. Per the Torrent Suite Software (Torrent Suite v5.8.0), the number of reads aligning to a given gene target represents an expression value referred to as "counts". The read count per million mapped reads (RPM) for each barcoded sample was calculated by the software as (read count) x 10^6 / total number of mapped reads (Supplemental Table 5). Normalized RPM-level transcript abundances were calculated relative to the average RPM abundance quantified across all samples for a given transcript followed by Log_2 transformation (Supplemental Table 6).

Bioinformatic and Statistical Analyses

Unsupervised analyses were performed using protein and transcript alterations exhibiting a median absolute deviation (MAD) >0.5 across all samples and clustered by Pearson correlation as heatmaps using gplots (version 3.0.1.1) in R (version 3.5.2). Differential analyses of global proteome and transcript data matrixes was performed using the LIMMA package (version 3.8, cite) in R (version 3.5.2); protein alterations passing LIMMA p-value < 0.01 exhibiting a Log_2 fold-change cut-off ± 1 were prioritized for downstream analyses. Cell type enrichment analyses was performed using RPM-level RNA-seq data in xCell (http://xcell.ucsf.edu/, [44]) xCell cell type signature score scores of interest were categorized by relative rank from highest to lowest spanning a range of 1 to -1 to enable co-visualization with similarly categorized transcript and

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

protein abundance for candidates of interest (Figure 2C). Log transformed RNA-Seq data was correlated to HGSOC molecular subtypes previously described [13]. Specifically, Spearman correlations were calculated for transcripts co-quantified by RNA-Seq analyses for 635 molecular subtype genes from Supplemental Table 4 in Konecny GE et al, 2014 [13] in R (version 3.5.2). Spearman correlations were categorized by relative rank from highest to lowest spanning a range of 1 to -1 to enable co-visualization with similarly categorized transcript and protein abundance for candidates of interest (Figure 4). Molecular subtype classifications were further compared by Spearman correlation analyses using MedCalc (Version 19.0.7). Functional pathway inference and drug targets were assessed using Ingenuity Pathway Analysis.

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Discussion

Images were evaluated digitally via OracleBio's image analysis software and manually by a pathologist for tumor purity, necrosis, immune involvement overall and in the LMD collections. There was significant discrepancy regarding estimation of tumor purity between the automated and manual evaluations. Digital analysis estimated 23-56% tumor cellularity at different depths of the block, calculated as dissection area ROI / annotated tissue area (Supplemental Table 1). The software estimated that the median cell number per area is approximately 7,482 cells/mm^2. By cell type, this represents an average of 7,527 ± 172 tumor cells (n=15 protein and RNA collections) or 3,899 ± 263 stroma cells (n=6 protein and RNA collections) harvested by LMD per mm^2 of tissue area (Supplemental Table 1). Given the average mammalian cell is ~20 µm and the tissue was cut into 10 µm sections, this equates to 267 pg protein/cell, consistent with previous estimates [45]. Comparatively, manual review estimated 75-95% tumor cellularity (Supplemental Table 2). This difference can likely be attributed to the presence of small regions of interceding stroma and other cell types in and around the tumor epithelium, which were excluded from the LMD tumor epithelium harvests.

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplemental Table Legends

Supplemental Table 1. LMD enriched tumor epithelium and stroma cell areas acquired by laser microdissection and automated analyses of collected regions of interest (ROI). ROI are denoted as millimeter squared (mm^2) areas, cell nuclei as μm^2 areas or as total cell counts derived from cellular nuclei measured in LMD enriched tumor or stroma ROIs.

Supplemental Table 2. Manual pathology assessment of tumor purity throughout the depth of the HGSOC patient specimen block. Representative H&E stained glass slides from were examined at ~200 µm intervals by a board-certified pathologist for percent by area estimation of tumor cellularity, necrosis, stroma, normal ovarian epithelium, lymphocytes, and polymorphonuclear leukocytes (PMN).

Supplemental Table 3. Global protein matrix. Log_2 transformed fold-change abundances of 6,053 proteins imputed across all samples.

Supplemental Table 4. Co-quantified proteins and transcripts. Log_2 transformed fold-change abundances of 5,742 imputed proteins that were co-measured at the transcriptome level.

Supplemental Table 5. Global transcriptome matrix. Normalized abundances of 20,784 RNA transcripts calculated relative to the average RPM abundance quantified across all samples for a given transcript, as reported by the Torrent Suite (v5.8.0) software, before Log_2 fold-change transformation.

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplemental Table 6. Global transcriptome matrix. Log_2 transformed fold-change abundances of 20,784 RNA transcripts measured across all samples.

Supplemental Table 7. Unsupervised analysis of protein abundance (median absolute deviation > 0.5).

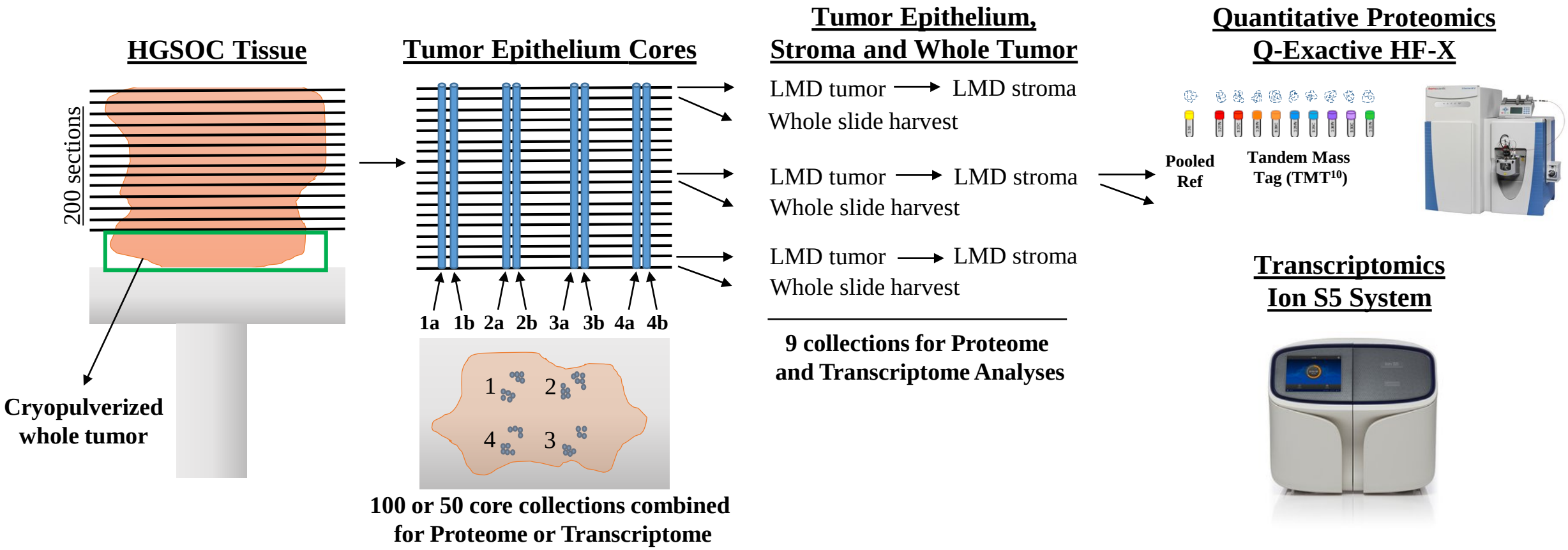
Supplemental Table 8. Unsupervised analysis of transcript abundance (Log_2 fold-change abundances with median absolute deviation > 0.5).

Supplemental Table 9. Cell type enrichment analyses as performed using RPM-level RNA-Seq data and default settings in xCell (http://xcell.ucsf.edu/, [44]).

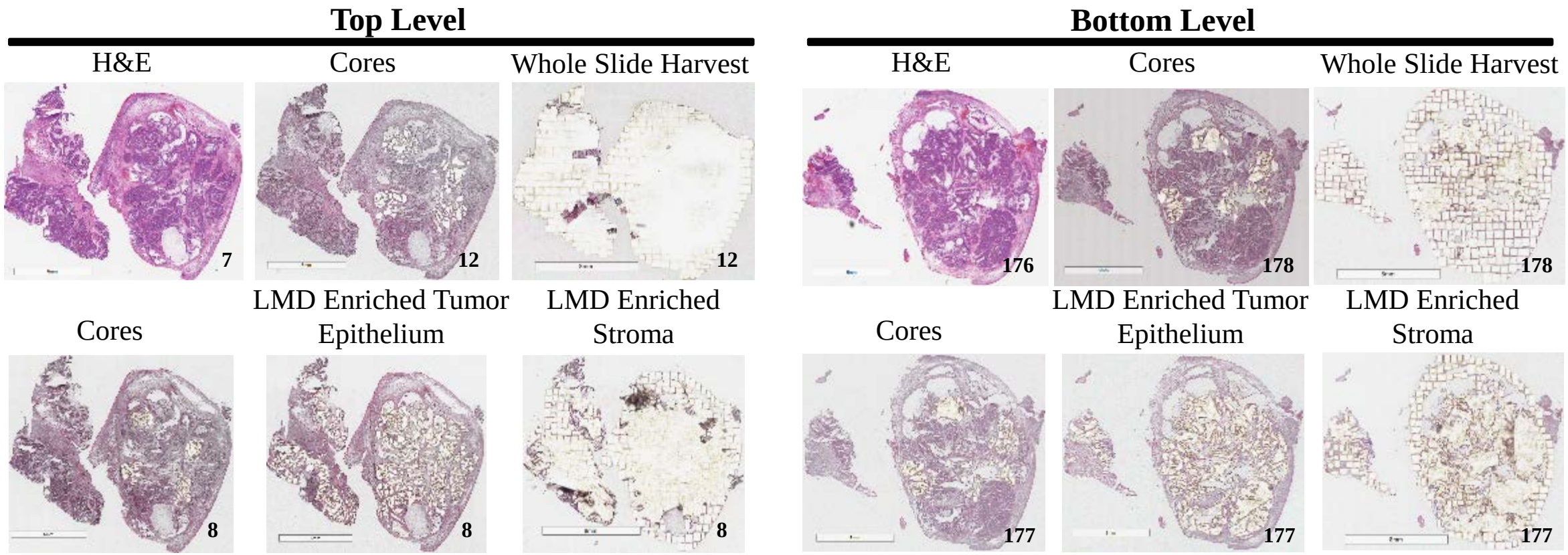
Supplemental Table 10. Significantly differentially expressed alterations between LMD enriched tumor epithelium versus stroma identified using Ingenuity Pathway Analysis (IPA) of known gene targets of FDA-approved anticancer drugs. Comparison of prevalent drug targets present in LMD enriched tumor versus stroma was determined using pairwise supervised analysis of proteomic and transcriptomic data with LIMMA adjusted p-value < 0.01 exhibiting a Log_2 fold-change cut-off ± 1. The complete list of all differentially expressed genes that are known drug targets from our transcriptomic-level and proteomic-level data was compared with a list of 150 FDA-approved anticancer drugs as of 2014 analyzed in a study by Sun et al [36].

bioRxiv preprint doi: https://doi.org/10.1101/761155; this version posted September 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

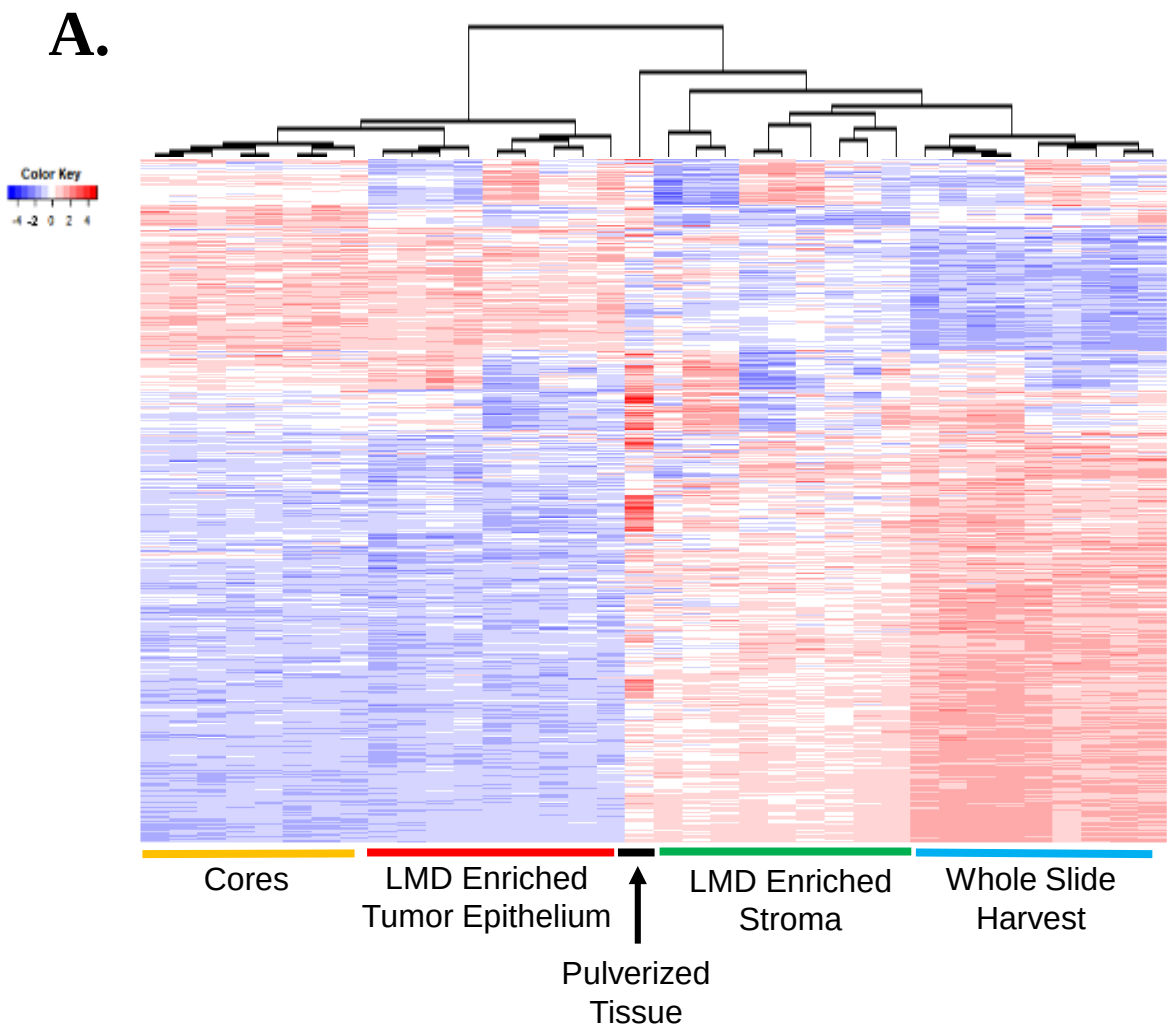
Supplemental Table 11. Spearman Rho correlations for transcripts and proteins co-quantified with molecular subtype genes from Supplemental Table 4 in Konecny GE et al, 2014 [13].

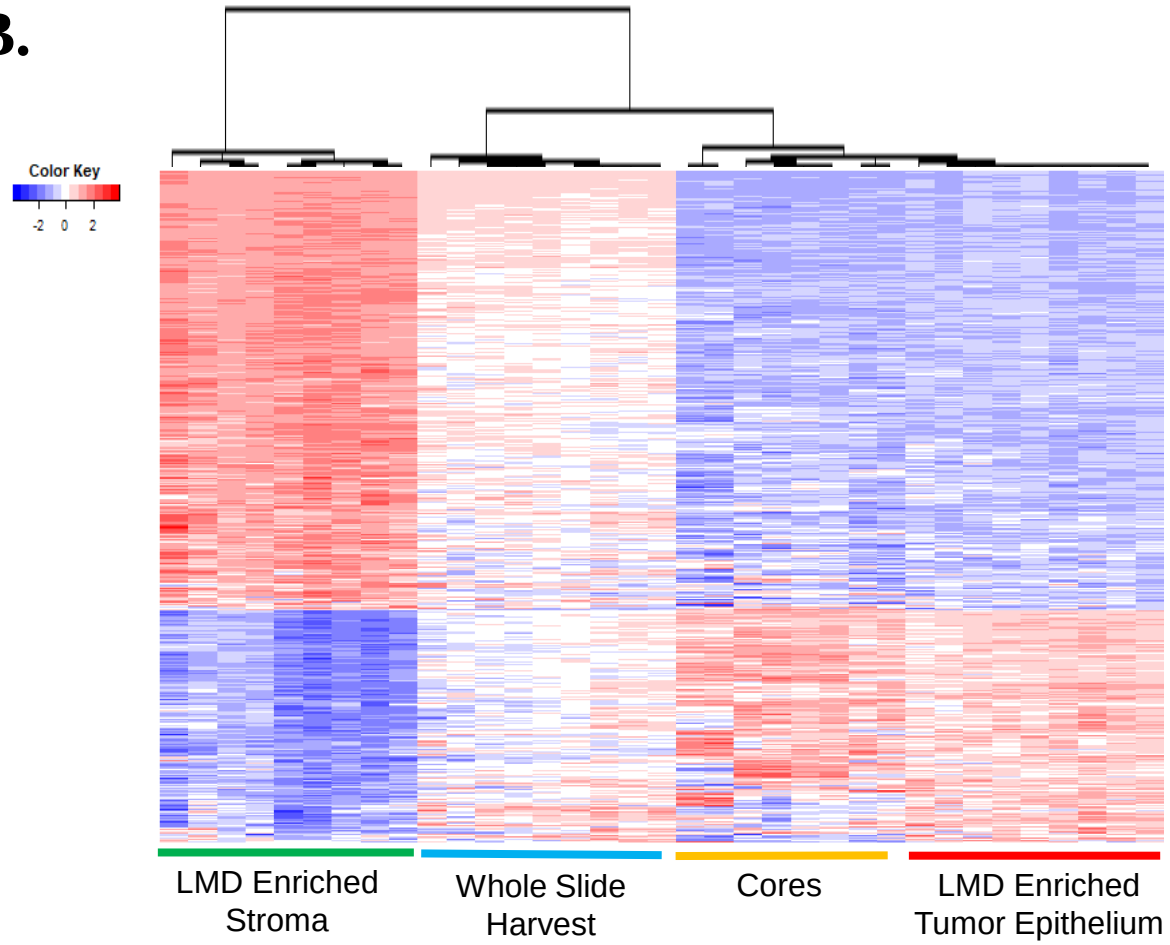


Hunt et al., Figure 1A

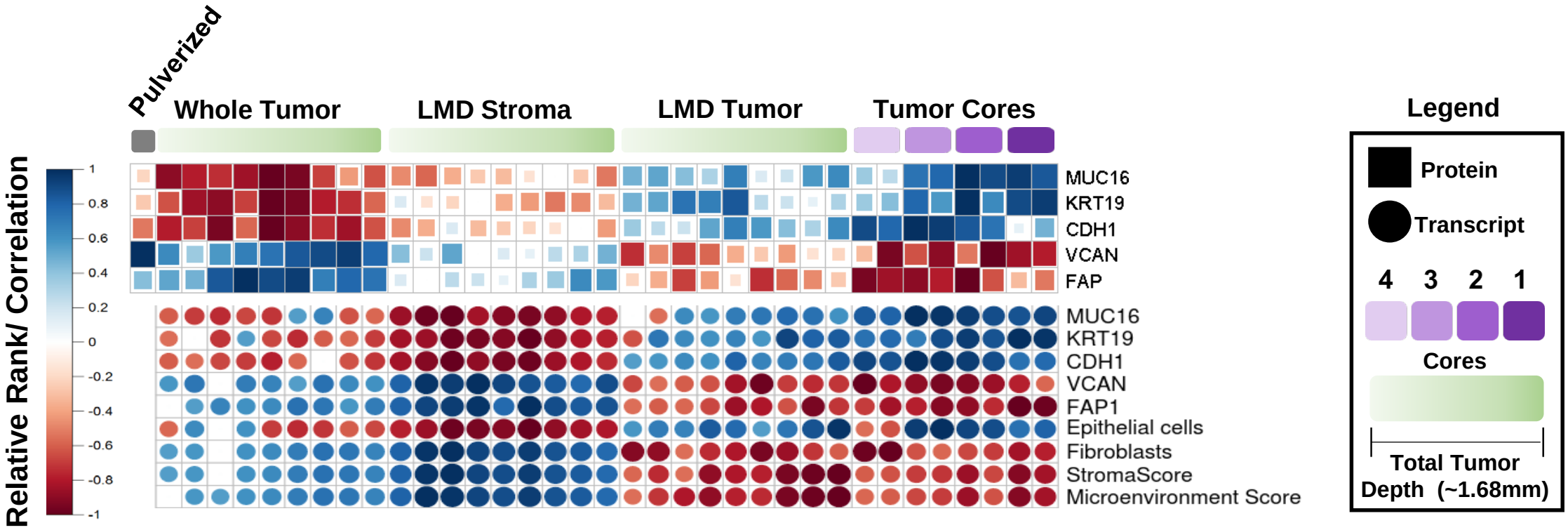


Hunt et al., Figure 1B

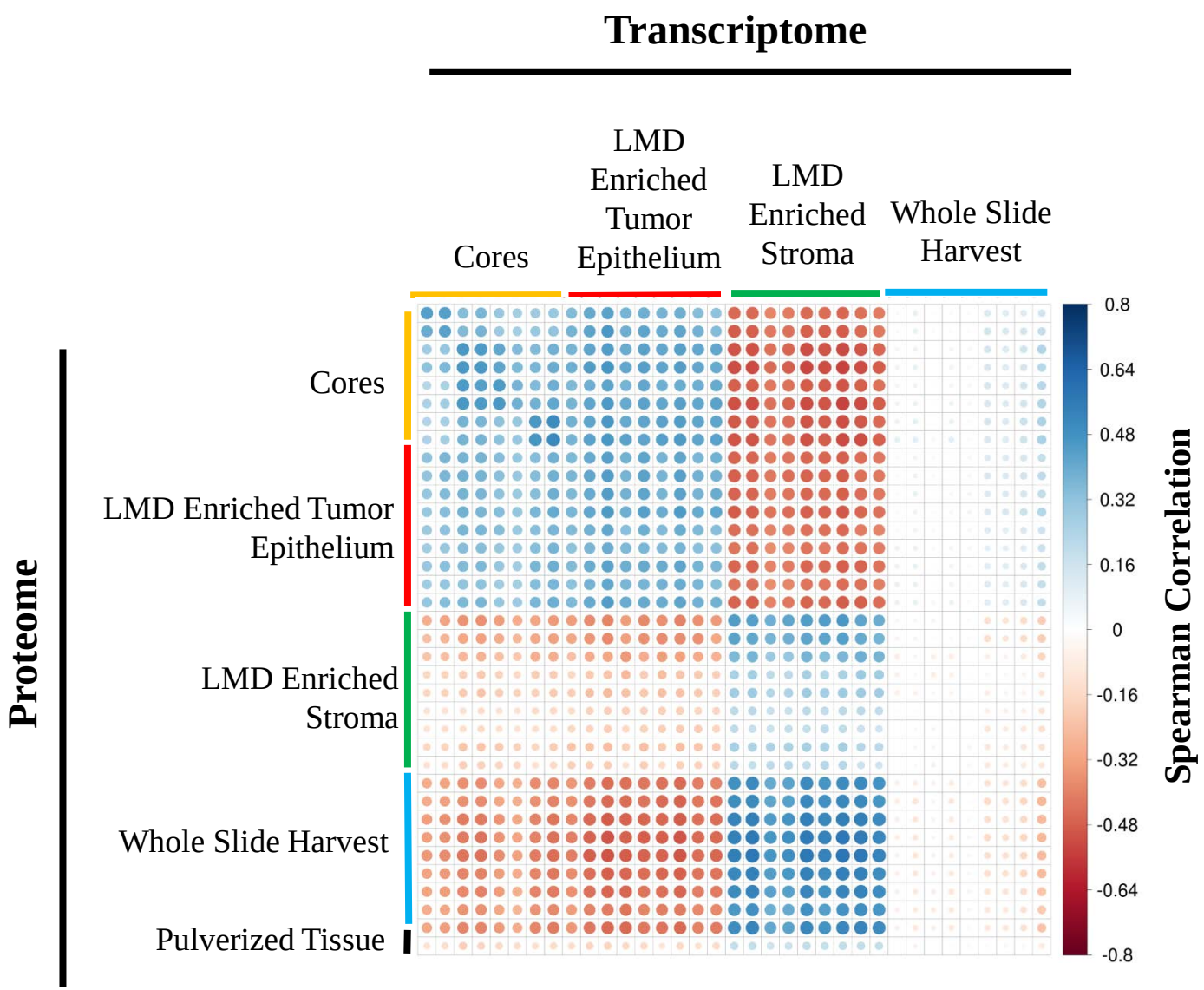




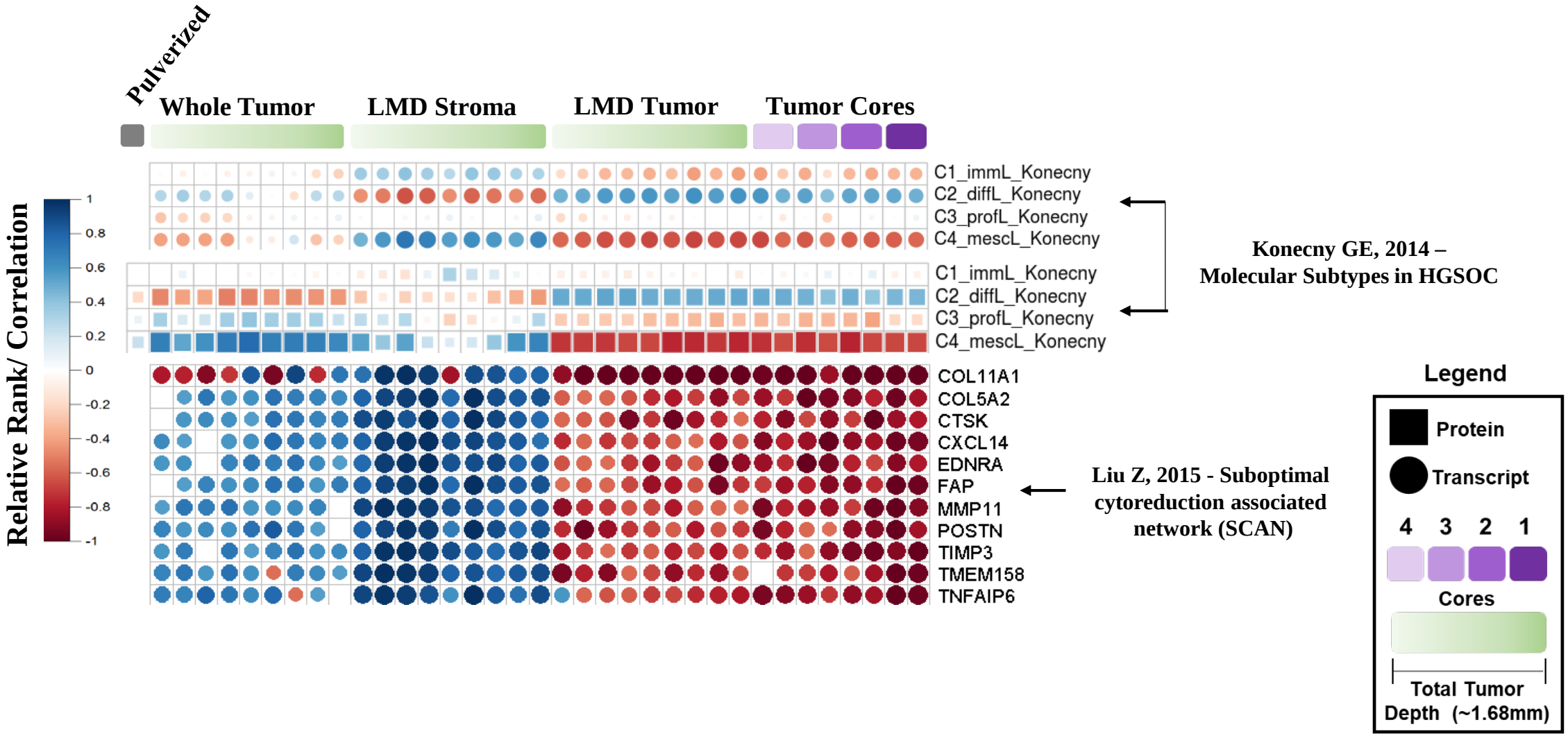
Hunt et al., Figures 2A and B



Hunt et al., Figure 2C



Hunt et al., Figure 3



Hunt et al., Figure 4