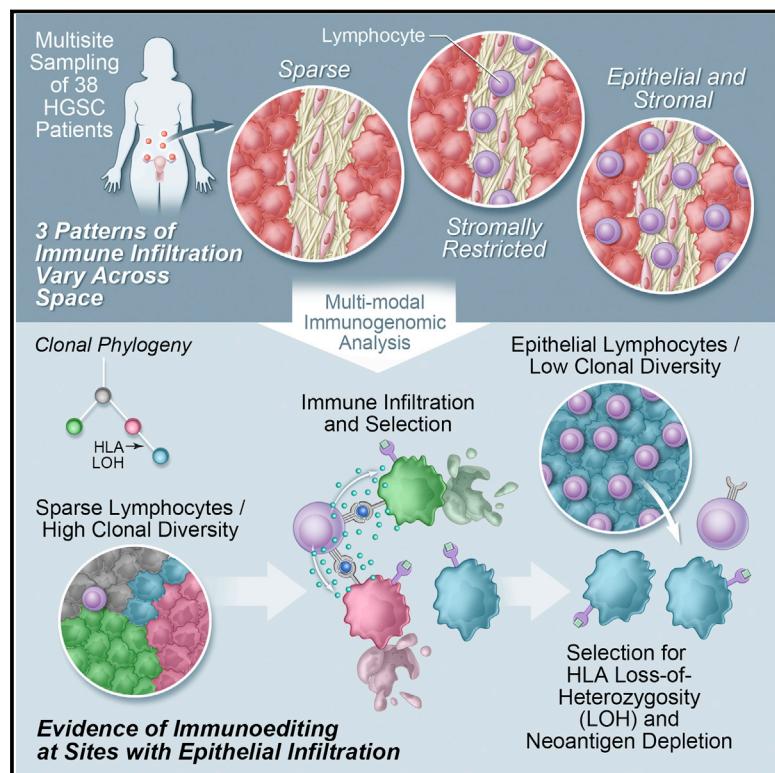


Interfaces of Malignant and Immunologic Clonal Dynamics in Ovarian Cancer

Graphical Abstract



Authors

Allen W. Zhang, Andrew McPherson,
Katy Milne, ..., Robert A. Holt,
Brad H. Nelson, Sohrab P. Shah

Correspondence

bnelson@bccrc.ca (B.H.N.),
sshah@bccrc.ca (S.P.S.)

In Brief

Integrated multi-region analysis of metastatic sites in patients with high-grade ovarian cancer highlights the connection between immune microenvironment variation and malignant spread, as well as the combinatorial prognostic value of immune and mutational features.

Highlights

- Immune infiltrates vary across space within patients at the time of diagnosis
- Immune infiltration shapes malignant cell evolutionary trajectories
- T cell clones track with tumor clones across spatial sites within patients
- Immune infiltrates and mutational processes show prognostic interactions

Interfaces of Malignant and Immunologic Clonal Dynamics in Ovarian Cancer

Allen W. Zhang,^{1,3,11} Andrew McPherson,¹ Katy Milne,⁴ David R. Kroeger,⁴ Phineas T. Hamilton,⁴ Alex Miranda,⁴ Tyler Funnell,^{1,11} Nicole Little,⁴ Camila P.E. de Souza,^{1,14} Sonya Laan,⁴ Stacey LeDoux,⁴ Dawn R. Cochrane,¹ Jamie L.P. Lim,¹ Winnie Yang,¹ Andrew Roth,^{6,7} Maia A. Smith,¹ Julie Ho,⁸ Kane Tse,² Thomas Zeng,² Inna Shlafman,² Michael R. Mayo,² Richard Moore,² Henrik Failmezger,^{9,10} Andreas Heindl,^{9,10} Yi Kan Wang,¹ Ali Bashashati,¹ Diljot S. Grewal,¹ Scott D. Brown,^{2,12} Daniel Lai,¹ Adrian N.C. Wan,¹ Cydney B. Nielsen,¹⁴ Curtis Huebner,¹ Basile Tessier-Cloutier,¹⁴ Michael S. Anglesio,^{14,15} Alexandre Bouchard-Côté,¹³ Yinyin Yuan,^{9,10} Wyeth W. Wasserman,³ C. Blake Gilks,¹⁴ Anthony N. Karnezis,¹⁴ Samuel Aparicio,^{1,14} Jessica N. McAlpine,¹⁵ David G. Huntsman,^{1,14} Robert A. Holt,² Brad H. Nelson,^{4,5,16,17,*} and Sohrab P. Shah^{1,14,17,18,*}

¹Department of Molecular Oncology, BC Cancer, Vancouver, BC V5Z 4E6, Canada

²Michael Smith Genome Sciences Centre, BC Cancer, Vancouver, BC V5Z 4E6, Canada

³BC Children's Hospital Research, Vancouver, BC V5Z 4H4, Canada

⁴Deeley Research Centre, BC Cancer, Victoria, BC V8R 6V5, Canada

⁵Department of Biochemistry and Microbiology, University of Victoria, Victoria, BC V8P 3E6, Canada

⁶Department of Statistics, University of Oxford, Oxford OX1 2JD, UK

⁷Ludwig Institute for Cancer Research, University of Oxford, Oxford OX1 2JD, UK

⁸Department of Anatomical Pathology, Vancouver General Hospital, Vancouver, BC V5Z 1M9, Canada

⁹Centre for Evolution and Cancer, The Institute of Cancer Research, London SM2 5NG, UK

¹⁰Division of Molecular Pathology, The Institute of Cancer Research, London SM2 5NG, UK

¹¹Graduate Bioinformatics Training Program, University of British Columbia, Vancouver, BC V6T 1Z3, Canada

¹²Genome Science and Technology Program, University of British Columbia, Vancouver, BC V6T 1Z3, Canada

¹³Department of Statistics, University of British Columbia, Vancouver, BC V6T 1Z3, Canada

¹⁴Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC V6T 1Z3, Canada

¹⁵Department of Gynecology and Obstetrics, University of British Columbia, Vancouver, BC V6T 1Z3, Canada

¹⁶Department of Medical Genetics, University of British Columbia, Vancouver, BC V6T 1Z3, Canada

¹⁷These authors contributed equally

¹⁸Lead Contact

*Correspondence: bnelson@bccrc.ca (B.H.N.), sshah@bccrc.ca (S.P.S.)

<https://doi.org/10.1016/j.cell.2018.03.073>

SUMMARY

High-grade serous ovarian cancer (HGSC) exhibits extensive malignant clonal diversity with widespread but non-random patterns of disease dissemination. We investigated whether local immune microenvironment factors shape tumor progression properties at the interface of tumor-infiltrating lymphocytes (TILs) and cancer cells. Through multi-region study of 212 samples from 38 patients with whole-genome sequencing, immunohistochemistry, histologic image analysis, gene expression profiling, and T and B cell receptor sequencing, we identified three immunologic subtypes across samples and extensive within-patient diversity. Epithelial CD8+ TILs negatively associated with malignant diversity, reflecting immunological pruning of tumor clones inferred by neoantigen depletion, HLA I loss of heterozygosity, and spatial tracking between T cell and tumor clones. In addition, combinatorial prognostic effects of mutational processes and immune properties were observed, illuminating how specific genomic aberration types associate with immune

response and impact survival. We conclude that within-patient spatial immune microenvironment variation shapes intraperitoneal malignant spread, provoking new evolutionary perspectives on HGSC clonal dispersion.

INTRODUCTION

High-grade serous ovarian cancer (HGSC) exhibits the highest disease mortality among gynecologic cancers. Despite recent progress with poly ADP-ribose polymerase (PARP) inhibitor-based synthetic lethal approaches exploiting homologous recombination deficiency (Mirza et al., 2016), HGSC remains incurable in most cases. Characterized by profound genomic instability and clonal diversity, HGSC often presents with widespread peritoneal dissemination. Multi-site studies have revealed genomic intratumoral heterogeneity (ITH) as a correlate to poor survival (Schwarz et al., 2015), as well as specific patterns of malignant cell spread within the peritoneal cavity (Bashashati et al., 2013). Importantly, the physical distribution of malignant clones across the peritoneal cavity is non-random, with the majority of sites exhibiting clonal homogeneity and a minority of sites harboring diverse clones (McPherson et al., 2016). This raises the hypothesis that region-specific properties, including immunologic components of the

tumor microenvironment, may modulate malignant cell invasion and expansion, thereby shaping evolutionary selection.

HGSC patients with abundant CD8+, CD4+, CD20+, and plasma cell tumor-infiltrating lymphocytes (TILs) are associated with favorable clinical outcomes (Zhang et al., 2003; Hwang et al., 2012; Nielsen et al., 2012; Kroeger et al., 2016). TILs can respond to and temporally track neoantigens (Wick et al., 2014) and mitigate resistance to platinum chemotherapy (Wang et al., 2016). However, much of our understanding of the immune response in HGSC derives from single biopsies; far less is known about spatial immunologic variation across distal tumor foci. Histologic imaging has revealed that lymphocyte abundance can vary between tumor foci in HGSC (Heindl et al., 2016). Furthermore, lymphocyte expression signatures are linked to patterns of metastasis (Auer et al., 2016). A single case report has described immunologic variation across relapse specimens (Jiménez-Sánchez et al., 2017); however, given the immunomodulatory effects of chemotherapy (Lo et al., 2017), understanding of pre-treatment spatial variation is still lacking.

Beyond immunologic features, prognostic mutational processes in HGSC through analysis of point mutation, copy number, and rearrangement features has indicated a prominent association between foldback inversions (FBIs) and poor response to platinum-based chemotherapy (Wang et al., 2017). FBI-dominated tumors, which comprise approximately 40% of HGSC, tend to be exclusive to homologous-recombination-deficient (HRD) cases and bear a distinct pattern of high-level amplifications colocalized with foldback rearrangements typical of breakage-fusion-bridge processes (Campbell et al., 2010; Wang et al., 2017). How mutational processes co-vary with immune response characteristics in HGSC remains poorly understood. This will become of central importance as clinical trials assaying synthetic lethal compounds targeting DNA repair processes combined with immune-modulation therapies read out.

We surmised that localized selective pressures imposed by immune microenvironments shape the distribution of malignant clones during disease progression. Thus, we systematically profiled the inter-relationship of clonal diversity, mutational processes, and immunologic response across a cohort of patients and multi-region samples. Genome-sequencing-based clonal decomposition, transcriptome-based T and B cell receptor sequencing, multicolor immunohistochemistry (IHC), and histologic image analyses were applied. Our results elucidate the landscape of cell-type interactions at the interface of malignant and immune cells across 212 samples from 38 patients. We show that samples robustly segregate into three distinct TIL subtypes, reflecting little or no immune infiltration, stromal infiltration, and combined epithelial and stromal infiltration. We reveal an association between these classes and malignant clone diversity properties. Regions with highest levels of epithelial immune infiltration exhibit the lowest malignant clone diversity, neoantigen depletion, and subclonal loss of heterozygosity (LOH) at human leukocyte antigen (HLA) loci as evidence of purifying selection. Moreover, T cell clonotypes, but not B cell clonotypes, spatially track with tumor clones in patients with heavily infiltrated tumors. Finally, we show combinatorial prognostic effects between mutational processes and immune infiltration

with foldback inversions exhibiting high risk even in the presence of high cytotoxicity. In aggregate, our findings illuminate molecular and evolutionary properties at the immune-malignant interface in HGSC with new insights on how tumor progression and clonal dissemination are driven by immune-related selective pressures.

RESULTS

High-Resolution Multi-site Profiling of Immune and Malignant Populations in the HGSC Tumor Microenvironment

We assembled a cohort of 212 tumor samples from 38 HGSC patients (Figure 1A and STAR Methods). Multiple samples per patient were collected via primary debulking surgery from ovary, omentum, and other distant metastatic sites (except some relapse samples from patients 7, 11, and 23; Table 1). TIL densities were measured by multicolor IHC, cell-type colocalization with 20× histologic images, clonotype diversity in T and B cell populations with T and B cell receptor sequencing (TCR-/BCR-seq), total mRNA gene expression from the 770-gene Nanostring PanCancer Immune Profiling Panel (Cesano, 2015) augmented with 39 molecular subtyping probes (Leong et al., 2015), mutational signatures and clonal diversity of malignant cells from whole-genome sequencing (WGS; mean depth: 86×), and deep amplicon sequencing (mean depth: 16,278×, median number of loci: 188, Table S1) (Figure S1). Both WGS and immune data (IHC, TCR/BCR-seq, or Nanostring) were obtained for 101 samples from 21 of 38 patients.

Tumor-Infiltrating Lymphocyte Subtypes Reveal Extensive Intrapatient Variation in Immune Responses across Peritoneal Sites

We began by profiling 188 tumor samples from 37 patients with multicolor IHC for CD8+ T cells (CD3+CD8+), CD4+ T cells (CD3+CD8-), CD20+ B cells (CD20+), and plasma cells (CD79a+CD138+). All but three patients were surveyed at multiple sites, providing an unprecedented view of intrapatient spatial variation. CD8+ T cells were the most abundant TIL type (0–1125.65 cells per high-powered field [HPF], median: 53.08), while CD20+ B cells were the rarest (0–136.77 cells per HPF, median: 2.74). Densities of all TIL types were correlated (Figure S2A), with extensive variation across the cohort (Figure S2B).

Using TIL densities as input features, we first analyzed a discovery cohort of 119 samples from 20 patients. Hierarchical clustering revealed three major TIL subtypes: N-TIL (tumors sparsely infiltrated by TILs), S-TIL (tumors dominated by stromal TILs), and ES-TIL (tumors with substantial levels of both epithelial and stromal TILs) (Figure 1B and Table S2). Based on orthogonal Nanostring probe counts, gene expression values for immune-associated pathways, including cytotoxicity, cytokines, and T cell- and B cell-associated genes, were comparable between S-TIL and ES-TIL but lower in N-TIL (Figure 1B). The three TIL subtypes mapped to previously described gene expression subtypes (C1, C2, C4, and C5) of HGSC (Leong et al., 2015) (STAR Methods). N-TIL was enriched for C4 and C5 tumors ($p < 10^{-5}$, Fisher's exact test), while S-TIL was overrepresented for C1 tumors ($p < 0.01$, Fisher's exact test) and ES-TIL for C2 tumors,

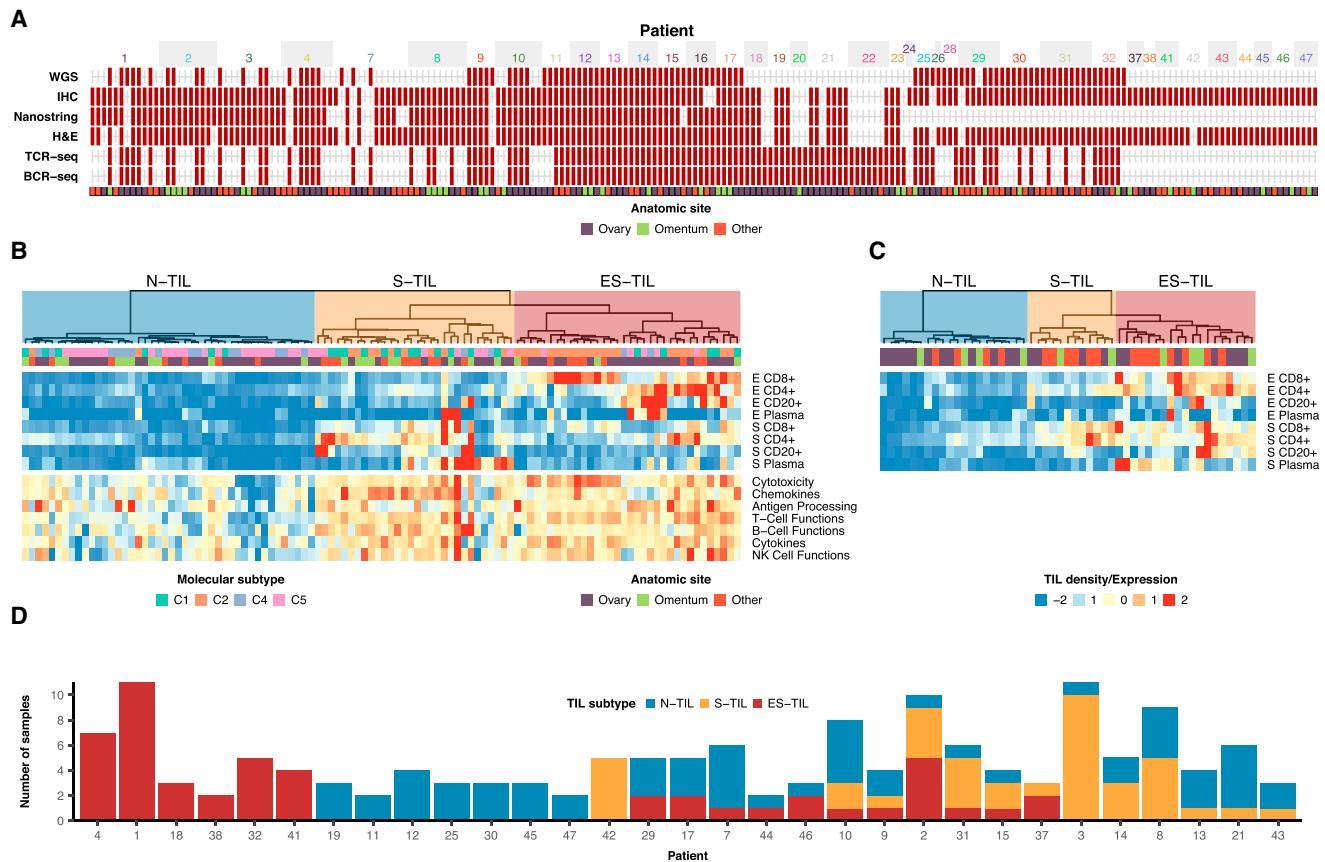


Figure 1. Multisite Profiling of HGSC Reveals Three Distinct TIL Subtypes with Extensive Intrapatient Variation

(A and B) (A) Experiments conducted on each tumor sample. Hierarchical clustering (Ward's method on L2-distances) of TIL densities from (B) discovery cohort of 119 samples from 20 patients.

(C) Additional cohort of 69 samples from 17 patients. Median expression of select immune pathways also shown in (B). Heatmap values standardized and clipped between -2 and 2. Samples with zero epithelial/stromal areas were removed (STAR Methods).

(D) Distribution of TIL subtypes by patient.

See also Figures S1, S2, and S5 and Table S2.

respectively ($p < 10^{-5}$, Fisher's exact test; Figure 1B and Table S2), suggesting previously reported HGSC gene expression subtypes (Tothill et al., 2008; Bell et al., 2011) largely reflect immune cell content. We analyzed IHC data from an additional cohort of 69 samples from 17 patients and observed a similar N-TIL, S-TIL, and ES-TIL distribution (Figure 1C), indicating reproducibility of the TIL subtypes. Among patients with ≥ 2 treatment-naive samples, 14 of 31 patients harbored only one TIL subtype: seven were N-TIL only, six were ES-TIL only, and one was S-TIL only. The remaining 17 of 31 patients harbored tumors from more than one TIL subtype (Figure 1D), and five patients harbored samples from all three subtypes, indicating extensive variation in immune response within patients.

While the ES-TIL pattern suggests active cytolytic TIL response against tumor cells, the presence of TILs in an epithelial region does not necessarily indicate active engagement with malignant cells. We therefore used histologic image analysis to profile microscopic spatial relationships between cancer cells and TILs. For each sample, we leveraged hematoxylin and eosin

(H&E) images to identify cancer cell and lymphocyte "hotspots" within the tumor epithelium—i.e., regions of local aggregation relative to epithelial cellular density (Figures 2A–2C and STAR Methods). We computed three measures of cancer-lymphocyte hotspot colocalization (Nawaz et al., 2015): f_C (the fraction of cancer cell hotspots that are lymphocyte hotspots); f_I (the fraction of lymphocyte hotspots that are cancer cell hotspots), and f_{CI} (fractional tissue area occupied by colocalized cancer-lymphocyte hotspots) (Table S2). ES-TIL tumors exhibited high levels of overlap between cancer and lymphocyte hotspots, while S-TIL samples contained relatively low overlap (all $p < 0.05$, Kruskal-Wallis test, Figures 2A and 2D). Thus, in S-TIL tumors, the rare immune cells that enter epithelial compartments appear to fail to engage with tumor cells, possibly due to lack of recognition. Although N-TIL tumors have negligible levels of TIL, they nonetheless showed occasional immune cells that could be evaluated by hotspot analysis. Where measurable, N-TIL tumors showed similar levels of colocalization as ES-TIL (Figure 2D).

Table 1. Studied Patients

Patient	Age at Diagnosis (Years)	Stage	Recurrence	RFS (Months)	Status	OS (Months)	BRCA Status (Clinical)
1	72	IIIC	no	N/A	NED	71	screen negative
2	76	IIIC	yes	12	DOD	45	screen negative
3	69	IIIC	yes	25	AWD	73	screen negative
4	53	IIIA	yes	50	AWD	71	screen negative
7 ^a	47	IIIC	yes	8	DOD	52	screen negative
8	62	IIIC	no	N/A	NED	65	BRCA1 mutation and unclassified BRCA2 variant
9	53	IIIB	yes	5	DOD	32	unknown
10	74	IIIC	no	N/A	NED	59	unknown
11 ^b	53	IIIB	yes	32	AWD	174	BRCA2 mutation
12	62	IIIC	yes	15	DOD	44	screen negative
13	80	IV	no	N/A	NED	40	screen negative
14	58	IIIC	yes	7	DOD	36	screen negative
15	61	IIIC	no	N/A	NED	38	BRCA1 VUS
16	72	IIIC	yes	23	AWD	35	screen negative
17	56	IIIC	yes	19	AWD	32	BRCA2 and MUTYH variant
18	56	IIIC	yes	19	DOD	34	unknown
19	59	IIIA	no	N/A	NED	32	screen negative
20	64	IIIA	no	N/A	NED	10	unknown
21	79	IIIC	yes	4	DOD	45	screen negative
22	73	IIIC	yes	22	AWD	75	rare BRCA2 variant (2680G>A), likely benign
23 ^c	65	IIIC	yes	9	DOD	75	screen negative
24	40	IIIB	yes	22	DOD	66	screen negative
25	46	IIIC	yes	6	AWD	23	screen negative
26	55	IB	no	N/A	NED	14	unknown
28	83	IIIC	yes	4	AWD	16	unknown
29	19	IV	yes	5	AWD	16	BRCA1 mutation
30	38	IIIC	no	N/A	NED	13	screen negative
31	38	IIIC	yes	10	AWD	16	BRCA1 mutation
32	46	IIIC	yes	1	AWD	14	screen negative
37	81	IIIB	no	N/A	NED	12	unknown
38	80	IIC	no	N/A	NED	6	unknown
41	68	IIIC	no	N/A	NED	8	screen negative
42	54	IIIC	no	N/A	NED	4	unknown
43	70	IIIC	yes	5	AWD	20	screen negative
44	35	IIIC	no	N/A	NED	6	unknown
45	79	IIIC	no	N/A	NED	5	unknown
46	77	IIIC	no	N/A	NED	6	unknown
47	45	IIIC	no	N/A	NED	4	unknown

Current disease status: NED, no evidence of disease; AWD, alive with disease; DOD, dead of disease. RFS, relapse-free survival; OS, overall survival.^aBrnM and BrnMA1 at 14 months and RPvM and BwllmA6 at 33 months post diagnosis

^bPv1, Rct1, Rct2 at 139 months post diagnosis

^cLOv1 at 14 months post diagnosis

Evidence for Purifying Malignant Clonal Selection at Tumor Sites with High Epithelial Lymphocyte Infiltration

We next evaluated whether regional variation in TIL subtypes provided insight into the evolutionary trajectories and dissemination patterns of malignant clones. Using WGS on cryopreserved

tissues (102 samples from 21 patients, of which 31 from 7 patients were previously described in McPherson et al., 2016), we profiled somatic single-nucleotide variants (SNVs), allele-specific copy number, and rearrangements (Table S2) as markers of malignant clones. In addition, we performed deep

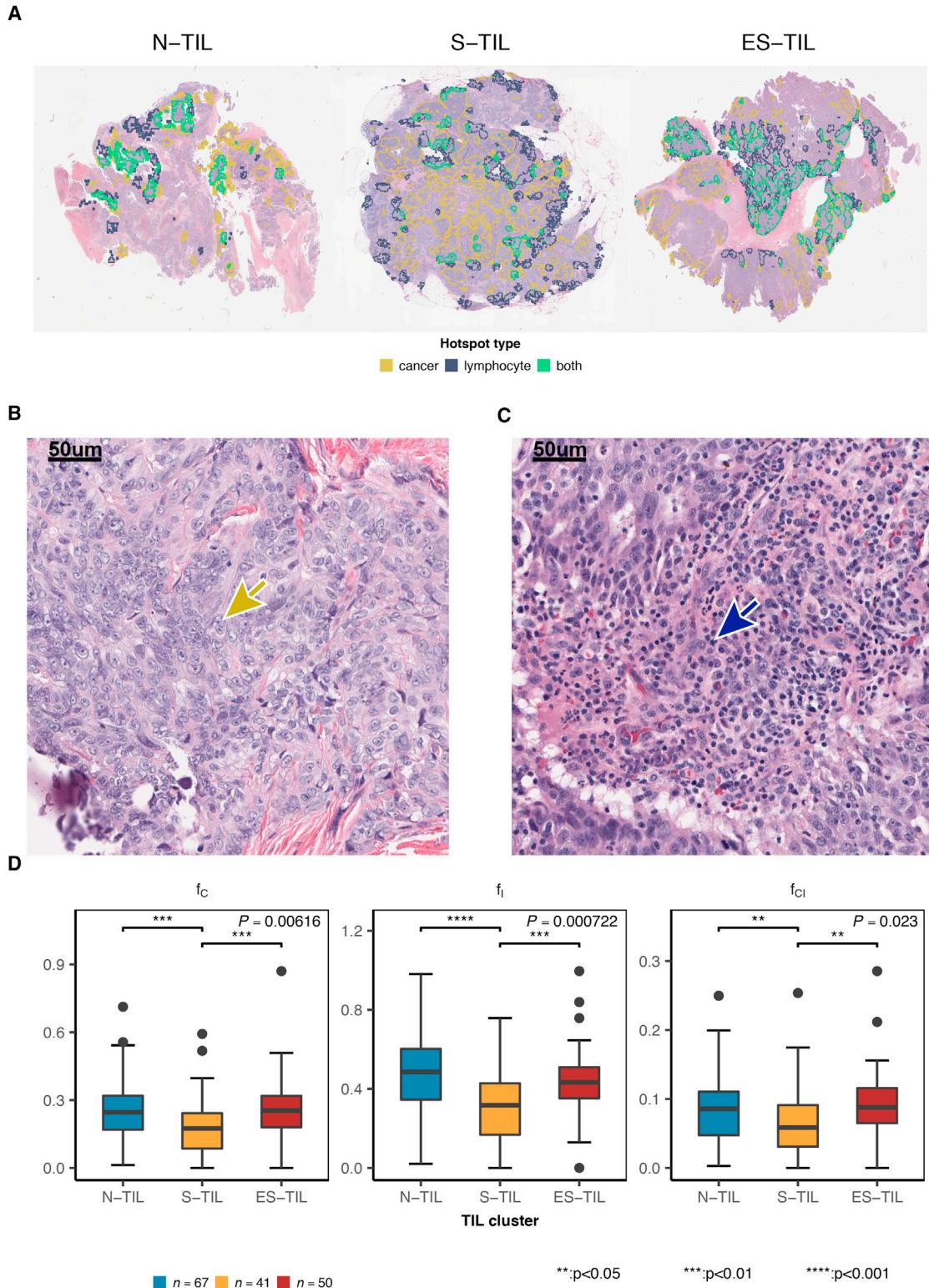


Figure 2. Differences in Cancer-Lymphocyte Hotspot Colocalization within Tumor Epithelium between TIL Subtypes

(A) Epithelial cancer cell and lymphocyte hotspots for representative N-, S-, and ES-TIL examples.

(B) Histology of a cancer cell hotspot. Yellow arrow: cancer cell.

(legend continued on next page)

amplicon sequencing on 97 samples from 14 of these patients (66 frozen and 31 formalin-fixed samples) to calculate clonal phylogenies and the clonal composition of each sample (Figure S3 and STAR Methods). We then related quantitative attributes of malignant clone composition to the N-TIL, S-TIL, and ES-TIL subtypes.

For each sample, we computed three continuous measures of malignant clone complexity: mixture entropy (the mixture distribution of clones present within a sample), clone divergence (the maximum phylogenetic distance between clones present within a sample; see McPherson et al., 2016), and heterogeneity index (the mean phylogenetic distance between a randomly selected pair of clones within a sample, weighted by abundance). We also computed an orthogonal measure from WGS directly with copy-number analysis (McPherson et al., 2017b; Table S2 and STAR Methods). All four measures of ITH were correlated (all $p < 0.1$, significance of Spearman ρ ; Figure S4A). For quality control, we confirmed entropy, clone divergence, and heterogeneity index were not correlated with tumor purity (all $p > 0.2$; Figure S4B). We evaluated the associations between measures of malignant clone complexity and the three TIL subtypes over all treatment-naïve samples. ES-TIL samples were lower for all four ITH measures relative to S-TIL and N-TIL samples (Figures 3A and 3B; accounting for tumor purity in the subclonal copy-number comparison) with mixture entropy, heterogeneity index, and subclonal copy number statistically significant. Accordingly, clonally pure tumors had the highest epithelial CD8+ TIL densities (Figure S4C). Despite the association between TIL and ITH, clonal similarity between intrapatient sites was not associated with TIL subtype ($p > 0.3$, nested ranks test; Figure S4D). For example, omentum sites 1 and 2 from patient 17 had comparable clonal composition, while ovary site 1 contained different clones (Figure S3); however, omentum site 1 was ES-TIL subtype, whereas omentum site 2 and ovary site 1 were N-TIL subtype (Table S2). Together, these data are consistent with epithelial TIL abundance as a negative determinant of regional malignant clonal complexity.

The negative association between epithelial TIL densities and malignant clone diversity could be explained by clonally complex tumors suppressing development of ES-TIL microenvironments and/or tumor clones undergoing immune-mediated purifying selection in the presence of high epithelial TIL density. In the latter scenario, subclonal (non-ancestral) neoepitopes might serve as targets of T cell recognition and hence show evidence of depletion at ES-TIL sites. To test this, we used NetMHCpan (Ternette et al., 2016) to computationally predict neoepitopes from nonsynonymous somatic SNVs (Table S3), categorizing each neoepitope as clonal or subclonal through phylogenetic analysis (STAR Methods). For each sample, we then quantified neoantigen depletion by comparing observed to expected (computed on an independent cohort of 121 primary HGSC samples) neoantigen rates (STAR Methods). Within patients, samples with higher epithelial CD8+ density exhibited

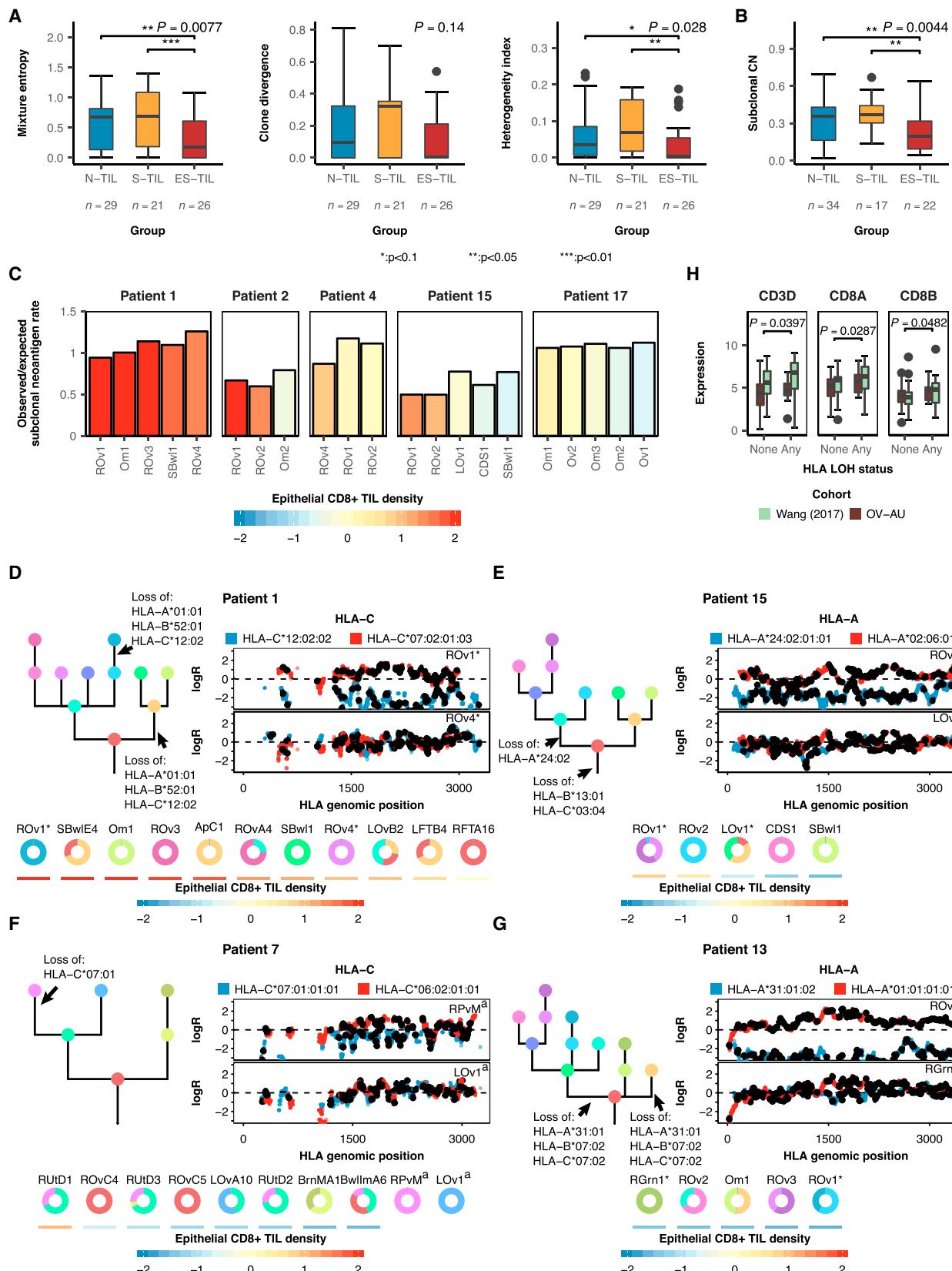
higher levels of subclonal neoantigen depletion (lower observed/expected subclonal neoantigen rate, $p = 0.09$, linear mixed model; Table S3 and STAR Methods), but not clonal neoantigen depletion ($p > 0.3$), compared to other samples from the same patient. This association was pronounced in patients containing samples with the highest epithelial CD8+ TIL densities ($p = 0.001$, linear mixed model; Figure 3C). In contrast, no significant association was observed between stromal CD8+ TIL density and clonal or subclonal neoantigen depletion (all $p > 0.2$, linear mixed model). Thus, samples with high epithelial CD8+ TILs show evidence of immune editing of subclonal neoantigens, raising the possibility that immune-driven purifying selection underlies the observed reduction in malignant cell diversity at TIL-rich sites.

In tumors with high epithelial CD8+ TIL densities, we postulated that the few remaining tumor clones might have avoided immune-related negative selection through clonal expansion of cells lacking neoantigen- or other tumor antigen-presenting HLA alleles. We used a Bayesian statistical extension (STAR Methods) of the LOHHHLA algorithm (McGranahan et al., 2017) to analyze WGS data for clone-specific HLA class I allele loss. Of 14 patients evaluated, we identified four patients harboring clonal HLA LOH and four with subclonal HLA LOH (one patient had both; Table S4). In three out of four patients with subclonal HLA LOH, the samples with the highest epithelial CD8+ TIL densities contained tumor clones with subclonal HLA LOH (Figures 3D–3G), including two of the patients (1 and 15) that demonstrated subclonal neoantigen depletion. An exception was patient 13, where subclonal HLA LOH was observed despite all samples having low epithelial CD8+ TIL density (Figure 3G; no samples were ES-TIL). Nevertheless, these findings suggest that tumor clones at ES-TIL sites have, in some cases, escaped immune clearance by somatic genomic loss of HLA haplotypes. We next examined the prevalence of HLA LOH in orthogonal WGS external cohorts (OV-AU, i.e., Patch et al., 2015; Wang et al., 2017). HLA LOH was found in 33.3% of samples (OV-AU: 34.7%, Wang: 32.1%) and was associated with significantly higher expression of lymphocyte markers (Figure 3H), establishing a link between HLA LOH and higher TIL levels.

To provide context, we also considered other known mechanisms of immune escape, including anatomic site, disruption of antigen presentation machinery (Yoshihama et al., 2016), and expression of immunosuppressive factors (Spranger et al., 2013; Rooney et al., 2015). TIL subtype was not significantly associated with any specific anatomic location (Fisher's exact test, $p > 0.05$, Table S2), and no point mutations, indels, or copy-number losses in antigen presentation machinery molecules were observed in ES-TIL samples. However, consistent with expectation from previous reports (Rooney et al., 2015), we found that inhibitory immune checkpoint molecules were generally upregulated in tumors with high epithelial CD8+ TIL density (Figure S4E).

(C) Histology of a colocalized cancer-lymphocyte hotspot. Blue arrow: lymphocyte.

(D) Comparison of epithelial cancer-lymphocyte hotspot colocalization between TIL subtypes. p values from Kruskal-Wallis test; post hoc comparisons (Benjamini-Hochberg adjusted) from Dunn's test. Whisker ends correspond to Q1 – 1.5*IQR and Q3 + 1.5*IQR.
See also Table S2.



(legend on next page)

T Cell, but Not B Cell, Clonotypes Show Evidence of Tumor Clone Tracking

We next investigated whether T and B cell clonotypes associate with tumor clones. We applied TCR β chain and BCR heavy-chain sequencing to total RNA from 116 samples (27 patients) and defined the clonotype-level composition of T and B cell populations in each sample (Figures 4A and S5A, Table S2, and STAR Methods). TCR diversity was positively correlated with IHC-based CD8+ and CD4+ TIL densities (all Spearman $p < 10^{-5}$; Figure S5B). Similarly, BCR diversity was positively correlated with CD20+ and plasma cell densities (all Spearman $p < 0.01$; Figure S5C). S-TIL and ES-TIL tumors had significantly more diverse TCR and BCR repertoires than N-TIL tumors (Figures 4B, S5D, and S5E) and a higher proportion of rare clonotypes (Figure S5F). None of the four ITH measures were significantly associated with TCR or BCR diversity across treatment-naïve samples (all Spearman $p > 0.3$; STAR Methods), indicating that diverse malignant populations do not recruit similarly diverse TIL repertoires.

We next ascertained the degree of homogeneity (similarity) between TCR and BCR repertoires across spatial samples within patients (STAR Methods). This revealed marked variation in both intrapatient TCR and BCR similarity across the cohort (Figures 4C and S5G). Considering patients with at least three samples, the extent of intrapatient TCR and BCR repertoire similarities were correlated (Spearman $p < 0.1$), but with notable exceptions (Figure S5G). Patient 15 had high TCR similarity (ranked 2nd out of 20 patients), but not BCR similarity (14th), while patients 10 and 21 had high BCR similarity (3rd and 5th), but not TCR similarity (15th and 20th). Mean intrapatient BCR similarity was not significantly correlated with IHC-based CD20+ or plasma cell density (all Spearman $p > 0.2$, Figures S5H and S5I). However, mean intrapatient TCR similarity was strongly associated with CD8+ (Spearman $p < 0.01$), but not CD4+, TIL density (Figures 4D and 4E), suggesting that CD8+ TILs were more broadly distributed (shared) across tumor sites compared to CD4+ TILs. To test this, we trained a classifier to separate TCRs as CD8+ type or CD4+ type on the basis of V/J genes and physicochemical properties of the hypervariable domain (STAR Methods). The ratio of CD8+/-CD4+ type TCRs was correlated with the ratio of CD8+/CD4+ densities by IHC (Spearman $p < 0.01$; Figure S5J). Corroborating our predictions, CD8+-type TCRs were significantly more broadly distributed than CD4+-type TCRs ($p < 0.001$; Figure 4F).

Having established that TCR-/BCR-based immune profiles vary across space, we asked how this variation is related to the spatial distribution of tumor clones. Pairwise T cell repertoire similarity was significantly correlated with malignant clone composition similarity in 7 out of 13 patients (Figure 5). Importantly, this relationship was significant in 5 of 6 patients with the highest epithelial CD8+ TIL densities (patients 1, 2, 9, 15, and 17), consistent with T cell clonotypes spatially tracking tumor clones in patients with high epithelial CD8+ TILs. This association held in the same six patients when considering only major TCR clonotypes (most abundant clonotypes constituting the top 50% of reads within each patient), but was only significant in patients 2, 9, and 12 when considering minor clonotypes (all other clonotypes), indicating that the most abundant clonotypes drove this effect. In contrast, pairwise BCR similarity was not significantly correlated with tumor clone similarity in any patient (Figure S3), suggesting an absence of spatial tracking between B cells and tumor clones.

Mutation Signatures Prognostically Associate with Patient-Level Immunologic Features

We next investigated interaction of malignant and immune infiltration from the perspective of mutational processes operating in HGSC. We previously identified two prognostically relevant mutation signature-associated subtypes: H-HRD and H-FBI (Wang et al., 2017). Here, we explored whether those subtypes could explain the observed variation in immune infiltration within and between patients. We pooled WGS data from our 21 cases with 195 additional single-site ovarian cancer cases (133 from Wang et al., 2017 and 62 from OV-AU in the International Cancer Genome Consortium [ICGC]) and applied a novel multimodal correlated topic model (MMCTM; Funnell et al., 2018), identifying six SNV and seven rearrangement signatures (Figures S6A and S6B and Table S5). Hierarchical clustering by signature proportions identified four major clusters (Figure 6A, Table S5, and STAR Methods): one subtype (HRD-DEL) dominated by the point mutation signature associated with homologous recombination deficiency (P-HRD) along with a short deletion signature (R-SDEL) associated with *BRCA2* mutations (Nik-Zainal et al., 2016), a second subtype (HRD-DUP) with P-HRD and a short tandem duplication signature (R-SDUP) associated with *BRCA1* mutations (Nik-Zainal et al., 2016), a third subtype (FBI) characterized by an FBI rearrangement signature (R-FB) associated with breakage-fusion-bridge (Wang et al., 2017), and a fourth, minor subtype distinguished by medium and large

Figure 3. Low ITH, Neoantigen Depletion, and Subclonal HLA Loss of Heterozygosity in Samples with High Epithelial TIL

(A) Clonal measures of ITH by TIL subtype. p values from Kruskal-Wallis tests; asterisks indicate post hoc significance (Benjamini-Hochberg adjusted) from Dunn's test. Whisker ends correspond to Q1 - 1.5*IQR and Q3 + 1.5*IQR.
(B) Subclonal copy-number proportion by TIL subtype. p value from ANOVA, controlling for cellularity (STAR Methods). Asterisks indicate post hoc significance from Tukey's range test. Whisker ends correspond to Q1 - 1.5*IQR and Q3 + 1.5*IQR.
(C) Ratio between observed and expected neoantigen rates for pre-treatment samples in patients with highest sample-level epithelial CD8+ densities (indicated by bar color).
(D-G) For patients with subclonal HLA class I LOH, (left) clonal phylogeny showing HLA LOH events and (right) logR values for samples with and without HLA LOH based on clonal composition. ^a:RPvM and LOv1 did not have IHC data. (Bottom) Clonal composition and epithelial CD8+ density of each sample. (D) Patient 1. (E) Patient 15. (F) Patient 7. HLA-C*07:01:01:01 was not as visually depleted in RPvM due to low cellularity (38%). (G) Patient 13. Sample labels defined in Table S2.
(H) Expression of lymphocyte markers in cases with none or any HLA LOH for Wang et al. (2017) (Nanostring) and OV-AU (RNA sequencing) cohorts. p values from nested ranks test. Whisker ends correspond to Q1 - 1.5*IQR and Q3 + 1.5*IQR.
See also Figure S4 and Tables S2, S3, and S4.

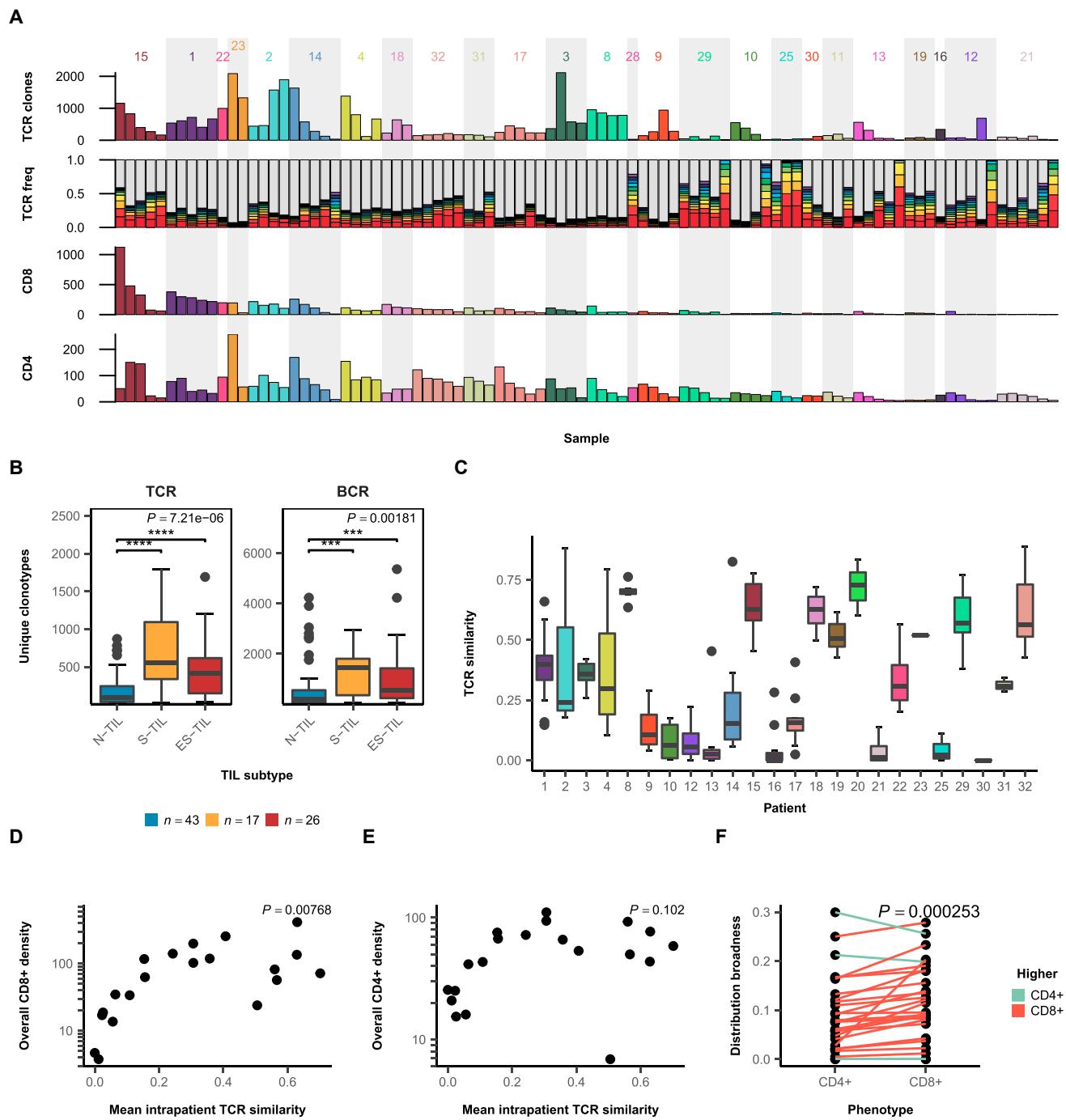


Figure 4. TCR/BCR Repertoire Diversity and Within-Patient Similarity, and Relationship to TIL Profiles

(A) Number of unique TCR clonotypes, prevalences of top 10 clonotypes (gray = all others), and CD8+ and CD4+ TIL density for each sample.

(B) Comparison of unique TCR and BCR clonotype counts between TIL subtypes. p values from Kruskal-Wallis tests; asterisks indicate post hoc significance (Benjamini-Hochberg adjusted) from Dunn's test.

(C) Distribution of pairwise TCR similarity for each patient. Whisker ends correspond to $Q_1 - 1.5 \times IQR$ and $Q_3 + 1.5 \times IQR$.

(D and E) Scatterplot of mean intrapatient TCR similarity and (D) CD8+ TIL density and (E) CD4+ TIL density. p value of Spearman ρ shown.

(F) Mean repertoire broadness for CD8+ and CD4+ type clonotypes in each patient. p value from Wilcoxon signed-rank test. Post-treatment tumors excluded in (B), (C), (D), (E), and (F).

See also Figure S5 and Table S2.

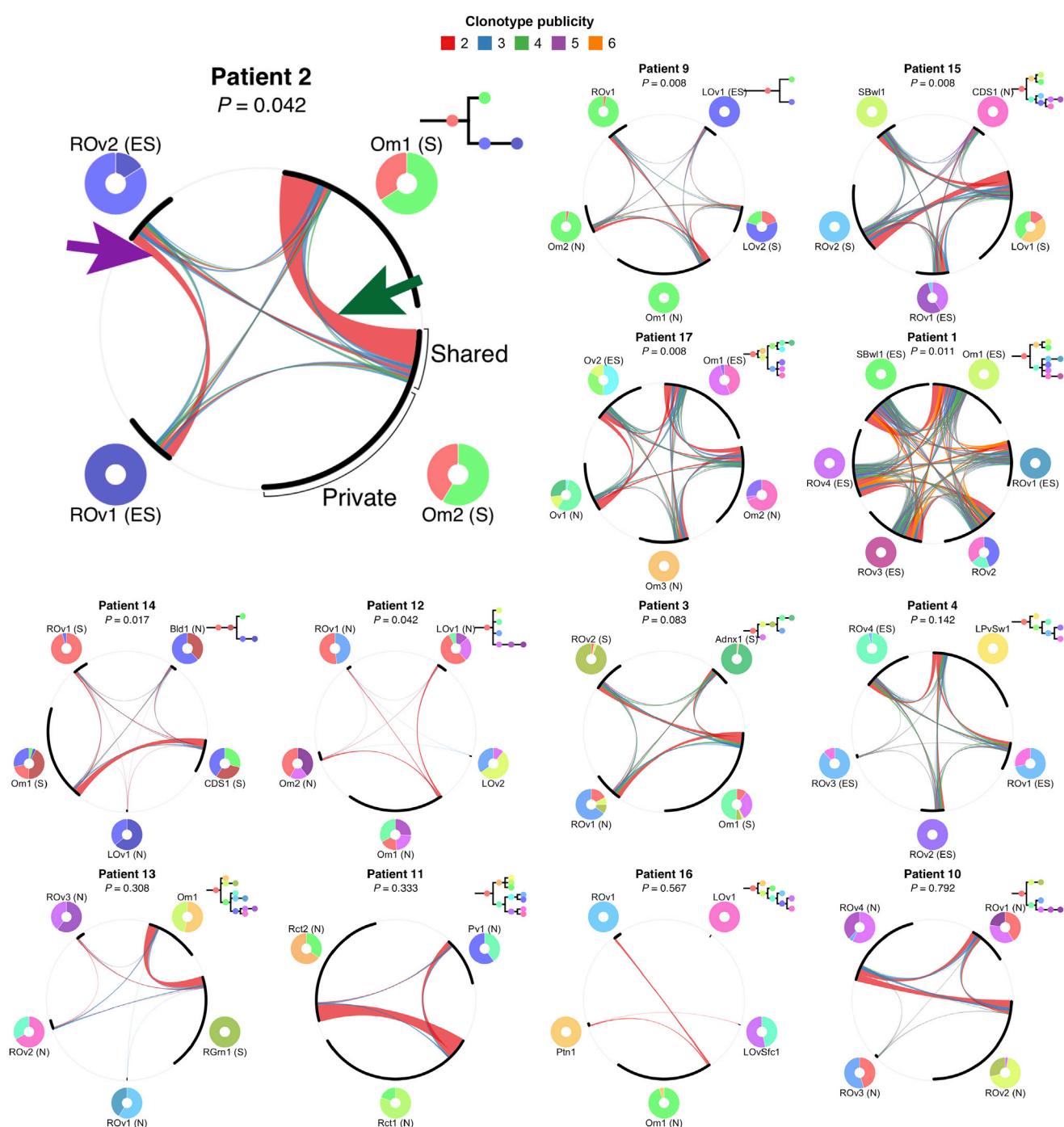
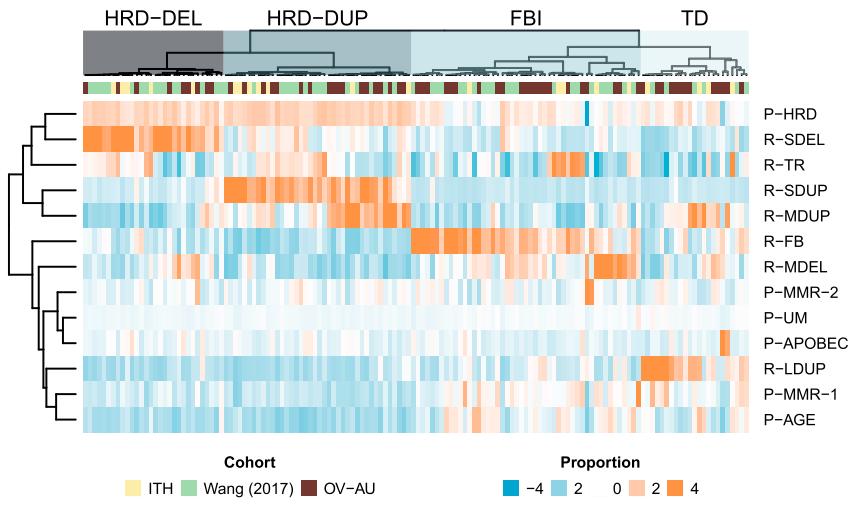


Figure 5. Relationships between Malignant Clone Composition and TCR Clonotype Repertoire

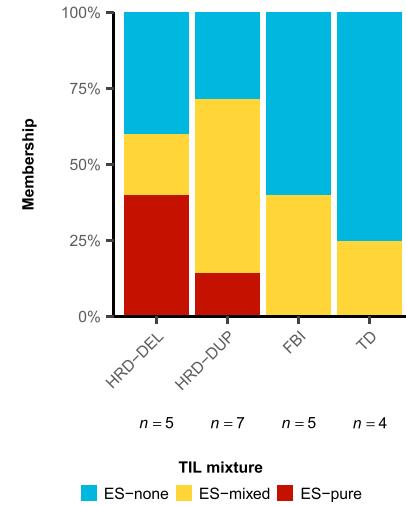
Patient 2 aside, cases ordered by significance of association between TCR repertoire and clonal composition dissimilarities (uncorrected Mantel's test p values). Chords denote shared clonotypes, width proportional to clonotype count, colored by publicity (number of samples containing a clonotype). Shared: publicity ≥ 2 ; private: publicity = 1. Purple arrow: chord denoting clonotypes shared only between right ovary sites 1 and 2. Green arrow: clonotypes shared only between omentum sites 1 and 2. Tumor clone composition and phylogenies next to each circle. TIL subtypes indicated as N (N-TIL), S (S-TIL), and ES (ES-TIL). Patient 7 excluded, as only two samples had TCR and tumor clone data.

See also Figure S3 and Table S2.

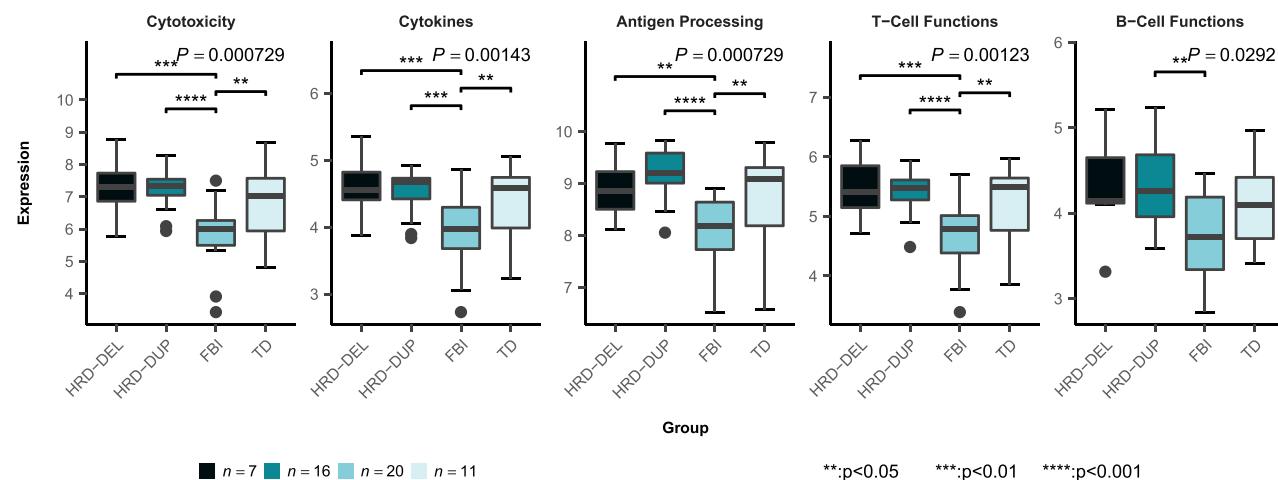
A



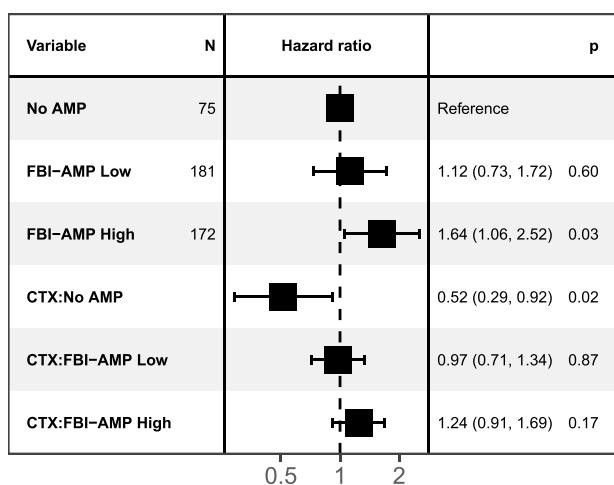
B



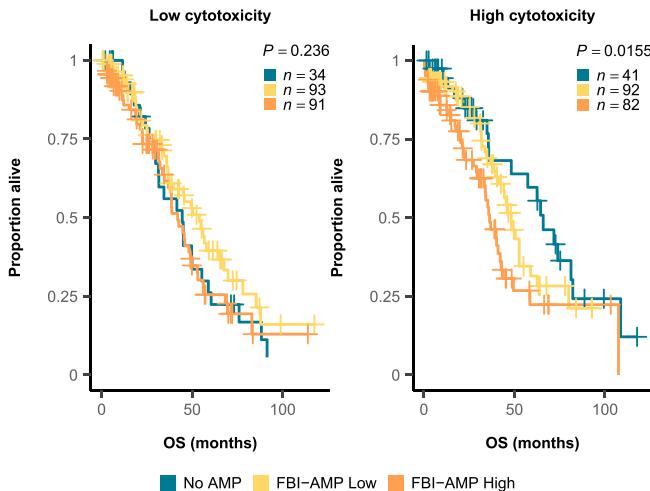
C



D



E



(legend on next page)

tandem duplications (TDs) (R-MDUP and R-LDUP, respectively) associated with CDK12 point mutations (Popova et al., 2016; Funnell et al., 2018).

Using this grouping of samples, we asked how immune response characteristics co-segregated with mutational signatures. Unlike TIL subtypes, mutational subtypes were largely invariant within patients (Figures S6C and S6D), indicating that mutational processes cannot explain intrapatient heterogeneity in TIL subtypes. We next asked whether mutational subtypes related to the mixture of TIL subtypes within each patient. Focusing on the ES-TIL subtype, we categorized patients with multi-sample IHC data as ES-none (no ES-TIL samples), ES-mixed (both ES-TIL and N-TIL/S-TIL samples), or ES-pure (all samples ES-TIL). The HRD subtypes contained the only three ES-pure patients (out of 12 HRD patients), although this did not reach significance with respect to the other mutational subtypes (Fisher's exact test, $p = 0.23$; Figure 6B). Expression values of immune-associated pathways (Cesano, 2015) for 54 OV-AU cases (STAR Methods) revealed that cytotoxicity, antigen processing, cytokine, and T cell markers were highest among HRD tumors (Figure 6C), concordant with similar findings in ER+ breast cancer (Smid et al., 2016) and among *BRCA1*-mutated tumors in HGSC (Nelson, 2015). Relative to HRD tumors, TD tumors had similar expression of immune markers, whereas FBI tumors were significantly depleted of these (Figure 6C). Corroborating these findings, differential expression analysis of OV-AU cases revealed that antigen processing, TCR/BCR signaling, cytotoxicity, and cytokine pathways were upregulated in HRD and TD relative to FBI ($Q < 0.01$), while none of these were differentially expressed between HRD and TD (Figure S7 and Table S6).

Colocalized foldback inversions and focal high-level amplifications (HLAMPs), thought to be reflective of breakage-fusion-bridge, have been associated with poor outcomes in HGSC (Wang et al., 2017). We asked whether immune activity could be used to further stratify foldback-enriched tumors into subgroups with distinct survival outcomes. Using gene expression data for 433 ovarian cystadenocarcinoma cases from the Cancer Genome Atlas (TCGA; Bell et al., 2011; Table S7), we jointly modeled the effects of colocalized foldback-HLAMP events and cytotoxicity expression with a Cox proportional hazards model, controlling for age of diagnosis and therapeutic regimen (STAR Methods). In agreement with Wang et al. (2017), high levels of colocalized foldback-HLAMP events were associated with significantly shorter overall survival (hazard ratio: 1.64, 95% CI: 1.06–2.52, $p < 0.05$; Figure 6D). The association between cytotoxicity and survival differed between FBI-HLAMP

groups ($p < 0.05$, likelihood ratio test between Cox models with and without cytotoxicity \times FBI-HLAMP interaction). In cases with no HLAMP events, cytotoxicity was significantly associated with a decreased hazard ratio (0.52, 95% CI: 0.29–0.92, $p < 0.05$; Figure 6D). However, among cases with colocalized foldback-HLAMP events, the hazard ratio for cytotoxicity was not significant (FBI-AMP low: 0.97, 95% CI: 0.71–1.34, $p > 0.3$; FBI-AMP high: 1.24, 95% CI: 0.91–1.69, $p > 0.1$; Figure 6D), suggesting that HLAMP-positive foldback-containing tumors harbor prognostic effects that are independent of immune response. We then median-stratified cases into low- and high-cytotoxicity groups. Low FBI was associated with significantly longer overall survival among tumors with high cytotoxicity (log-rank $p < 0.05$; Figure 6E), but not low cytotoxicity (log-rank $p > 0.2$; Figure 6E). Together, the covarying effects of immune activity and mutational processes suggest a combinatorial prognostic effect with high immune activity and low prevalence of FBIs leading to the best outcomes, while FBI-bearing patients have poor outcomes even in the presence of high immune activity.

DISCUSSION

Our results illuminate evolutionary properties at the malignant-immune interface of HGSC. In patients with the highest epithelial TIL densities, our data are consistent with active pruning of malignant cell diversity by TIL through subclonal neoepitope recognition, resulting in expansion of clones harboring neoantigen loss and/or HLA LOH. The underlying mechanism likely involves tracking of tumor clones across peritoneal space by T cell clones, but not B cell clones. As such, immune infiltrates impose selective constraints, shaping patterns of malignant spread and clonal diversity in HGSC. Our findings do not exclude the possibility that T cells can also recognize clonal neoepitopes (McGranahan et al., 2016); however, subclonal neoepitopes, which have been reported to have higher predicted immunogenicity than clonal neoepitopes (Jiménez-Sánchez et al., 2017), may be under stronger negative selection. Moreover, depletion of clonal neoantigens could result in complete tumor elimination and therefore go clinically undetected.

The presence of extensive intrapatient immune variation prior to treatment highlights potential shortcomings of prognostic stratification and study of the immune microenvironment from single biopsies. The widespread multi-site variation we observed suggests that even a single site harboring relative immune privilege may be sufficient to engender resistant disease, regardless of active immune responses in distal intraperitoneal regions. We suggest immunologically sheltered havens may plausibly act as

Figure 6. Mutational Subtypes Prognostically Associate with Immune Patterns in HGSC

- (A) Signature proportions in HGSC cases standardized and clipped from –4 to 4. Dendrogram computed with Ward's method on Pearson correlation dissimilarities. ITH: multisite cohort from this study.
- (B) Fractions of ES-none, ES-mixed, and ES-pure patients across mutational subtypes.
- (C) Expression of select immune-associated pathways across mutational subtypes in OV-AU. p values (Benjamini-Hochberg adjusted) from Kruskal-Wallis test; asterisks indicate post hoc significance (Benjamini-Hochberg adjusted) from Dunn's test. Survival analysis of 433 TCGA patients. Whisker ends correspond to Q1 –1.5*IQR and Q3 +1.5*IQR.
- (D) Hazard ratios, 95% confidence intervals, and p values from Cox regression of overall survival. Interaction terms indicated by colons; e.g., CTX:No AMP: effect of cytotoxicity in No AMP subtype.
- (E) Differences in overall survival between FBI-HLAMP subgroups for tumors with low/high cytotoxicity. p values from log-rank test.
- See also Figures S6 and S7 and Tables S5, S6, and S7.

reservoirs of clonal diversity from which malignant clones impacting disease relapse might emerge. As a preliminary illustrative example, ES-pure patients had better outcomes (5 of 6 no evidence of disease [NED] or alive with disease [AWD], 5 of 6 platinum sensitive, median progression-free survival [PFS] for relapsed patients was 19 months) than ES-mixed and ES-none patients (8 of 11 and 11 of 14 NED or AWD, 7/9 and 7/11 platinum sensitive, median PFS for relapsed patients was 9.3 and 7.1 months, respectively).

Our data show for the first time a prognostically relevant interaction between mutational processes and immune response in HGSC. Notably, foldback inversions associate with poor outcomes, even in highly cytotoxic tumor microenvironments. Thus, in contrast to point mutations resulting from mismatch repair deficiency (Le et al., 2015), FBIs likely represent a class of non-immunogenic genomic aberrations. Conversely, our findings also provide context for explaining superior outcomes observed in *BRCA1*- and *BRCA2*-mutated HGSC (Nelson, 2015). In contrast to previous reports that *BRCA1* disruption, but not *BRCA2* disruption, is associated with elevated TILs (Nelson, 2015; Goode et al., 2017), we observe comparably high immune activity between *BRCA1*-associated (HRD-DUP), *BRCA2*-associated (HRD-DEL), and TD subtypes. Shared deficiencies in homologous recombination between HRD and TD subtypes (Joshi et al., 2014) may result in patterns of rearrangements or point mutations (Wang et al., 2017) responsible for eliciting these immune responses (Smid et al., 2016).

Our study provides context for clinical trials investigating various classes of immunotherapy in ovarian cancer (e.g., immune checkpoint blockade, adoptive T cell transfer, neoepitope vaccination, combination immunotherapy with PARP inhibition). A recent case study tracking immune response over time in a HGSC patient with a remarkable clinical trajectory (Jiménez-Sánchez et al., 2017) demonstrated that spatiotemporal variation of the immune microenvironment relates specifically to treatment sensitivity of malignant clones. We reveal that immune-microenvironment spatial variation exists prior to treatment and is prevalent in the HGSC patient population. Given that efficacy of PD-1 axis blockade hinges on pre-existing adaptive immunity (Herbst et al., 2014), immunologically privileged sites on an otherwise highly infiltrated background may explain the limited success of immunotherapy in HGSC to date (Homicsko and Coukos, 2015; Gaillard et al., 2016). While some tumors contain abundant TILs, lack of cancer cell-lymphocyte colocalization and reduced tumor-immune engagement in S-TIL sites may result from a failure of immune recognition or region-specific barriers to infiltration. Consequently, TIL abundance alone is an insufficient predictor of active immune response. Even at sites patterned by extensive epithelial TILs, neoantigen depletion and apparent positive selection of clones harboring HLA LOH may render checkpoint blockade ineffective.

Despite these challenges, our findings inform on several potential therapeutic solutions. While FBI cases exhibit poor prognostic profiles independently of immune properties, HRD cases, typically associated with fewer foldback inversions, likely represent optimal candidates for immunotherapy approaches. Thus, mutational processes considered in conjunction with immune properties will aid in interpretation of newly initiated clinical

trials examining combination PARP inhibition with checkpoint-blockade approaches. Furthermore, if obstacles to infiltration at immunologically privileged sites can be surmounted, our findings hint at the tantalizing potential that such tumor sites may represent targetable cancer cell populations, owing to their limited neoantigen and HLA depletion at baseline.

As the cancer evolution field progresses toward a more rigorous understanding of the fitness of heterogeneous clones within disease spectra and over temporal dimensions (Lipinski et al., 2016), external selective pressures imposed by the immune system must be considered as highly relevant factors. Here we show that high-resolution measurement of the immune microenvironment together with clonal decomposition analysis is tractable and yields novel insight into forces shaping malignant cell diversity and intraperitoneal spread. Broadly disseminated intraperitoneal disease at diagnosis in HGSC remains a formidable clinical problem. Our study informs on how regional variation at the interface of immunological and cancer cells controls dissemination and diversification of clones and simultaneously identifies microenvironmental and malignant cell properties to exploit in future immuno-oncologic therapeutic strategies for HGSC.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
 - Sample acquisition, consent, & surgery
 - Sample preservation & histologic evaluation
 - Distribution and availability of samples
- **METHOD DETAILS**
 - WGSS library construction & sequencing
 - Targeted bulk sequencing analysis
 - Immunohistochemistry
 - Nanostring gene expression
 - TCR & BCR sequencing
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - WGSS analysis
 - Clonal analysis
 - RNA-seq analysis
 - Mutation signature analysis
 - Immunohistochemistry analysis
 - Nanostring analysis
 - TCR/BCR-seq analysis
 - TCR clonotype classification
 - Neoantigen analysis
 - Histologic image analysis
 - General statistical methods
- **DATA AND SOFTWARE AVAILABILITY**

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and seven tables and can be found with this article online at <https://doi.org/10.1016/j.cell.2018.03.073>.

ACKNOWLEDGMENTS

We thank Christine Kenney for illustrating the graphical abstract. Funding was provided by the BC Cancer Foundation (1NSRG033 to S.P.S., B.H.N., and R.A.H.), Canadian Cancer Society Research Institute (702497; S.P.S., B.H.N., and R.A.H.), Canada's Networks of Centres of Excellence (B.H.N. and R.A.H.), Canadian Institutes of Health Research (FDN-143246 to S.P.S.; B.H.N., P.T.H.), Terry Fox Research Institute (B.H.N.; 1021 to S.P.S.), Cancer Research Society (B.H.N.), Genome Canada and Genome British Columbia (9114 to S.P.S.; R.A.H.), Canada Research Chairs Program (S.P.S. and D.G.H.), Michael Smith Foundation for Health Research (CI-SCH-02987(11-1) to S.P.S., A. McPherson), CANARIE Foundation (RS-219; S.P.S.), and Vanier Canada Graduate Scholarship (A.W.Z.). The results shown here are in whole or part based upon data generated by the TCGA Research Network: <https://cancergenome.nih.gov/>. Finally, we acknowledge cloud computing support from Microsoft (S.P.S.).

AUTHOR CONTRIBUTIONS

Study design, A.W.Z., R.A.H., B.H.N., and S.P.S.; Writing, A.W.Z., B.H.N., and S.P.S.; Manuscript review, A.W.Z., B.H.N., S.P.S., R.A.H., S.A., A.B.-C., Y.Y., Y.K.W., A.B., W.W.W., C.B.G., C.P.E.d.S., A. McPherson, P.T.H., A. Miranda, and T.F.; Data interpretation, A.W.Z., P.T.H., D.R.K., A. Miranda, T.F., A. McPherson, N.L., J.N.M., R.A.H., B.H.N., W.W.W., and S.P.S.; Curation, D.R.C., D.L., and A.N.C.W.; Formal analysis, A.W.Z., A. McPherson, A.R., T.F., Y.K.W., A.B., S.D.B., H.F., A.H., Y.Y., D.S.G., and C.H.; DNA/RNA extraction, J.L.P.L. and W.Y.; IHC, K.M., D.R.K., S. Laan, and S. LeDoux; TCR-BCR-seq, K.T., T.Z., I.S., M.R.M., and R.M.; Nanostring, J.H. and M.S.A.; Pathologic evaluation, A.N.K. C.B.G., and B.T.-C.; Supervision, S.P.S., B.H.N., R.A.H., D.G.H., and W.W.W.; Visualization, A.W.Z., M.A.S., and C.B.N.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 25, 2017

Revised: February 22, 2018

Accepted: March 27, 2018

Published: May 10, 2018

REFERENCES

- Auer, K., Bachmayr-Heyda, A., Sukhbaatar, N., Aust, S., Schmetterer, K.G., Meier, S.M., Gerner, C., Grimm, C., Horvat, R., and Pils, D. (2016). Role of the immune system in the peritoneal tumor spread of high grade serous ovarian cancer. *Oncotarget* 7, 61336–61354.
- Bashashati, A., Ha, G., Tone, A., Ding, J., Prentice, L.M., Roth, A., Rosner, J., Shumansky, K., Kaloger, S., Senz, J., et al. (2013). Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling. *J. Pathol.* 231, 21–34.
- Bell, D., Berchuck, A., Birrer, M., Chien, J., Cramer, D.W., Dao, F., Dhir, R., DiSaia, P., Gabra, H., Glenn, P., et al.; Cancer Genome Atlas Research Network (2011). Integrated genomic analyses of ovarian carcinoma. *Nature* 474, 609–615.
- Bolotin, D.A., Poslavsky, S., Mitrophanov, I., Shugay, M., Mamedov, I.Z., Putintseva, E.V., and Chudakov, D.M. (2015). MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods* 12, 380–381.
- Campbell, P.J., Yachida, S., Mudie, L.J., Stephens, P.J., Pleasance, E.D., Stebbings, L.A., Morsberger, L.A., Latimer, C., McLaren, S., Lin, M.-L., et al. (2010). The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* 467, 1109–1113.
- Cesano, A. (2015). nCounter® PanCancer Immune Profiling Panel (NanoString Technologies, Inc., Seattle, WA). *J. Immunother. Cancer* 3, 42.
- Ding, J., Bashashati, A., Roth, A., Oloumi, A., Tse, K., Zeng, T., Haffari, G., Hirst, M., Marra, M.A., Condon, A., et al. (2012). Feature-based classifiers for somatic mutation detection in tumour-normal paired sequencing data. *Bioinformatics* 28, 167–175.
- Emerson, R., Sherwood, A., Desmarais, C., Malhotra, S., Phippard, D., and Robins, H. (2013). Estimating the ratio of CD4+ to CD8+ T cells using high-throughput sequence data. *J. Immunol. Methods* 397, 14–21.
- Funnell, T., Zhang, A., Shiah, Y.-J., Grewal, D., Lesurf, R., McKinney, S., Bashashati, A., Wang, Y.K., Boutros, P., and Shah, S. (2018). Integrated single-nucleotide and structural variation signatures of DNA-repair deficient human cancers. *bioRxiv*. <https://doi.org/10.1101/267500>.
- Gaillard, S.L., Secord, A.A., and Monk, B. (2016). The role of immune checkpoint inhibition in the treatment of ovarian cancer. *Gynecol. Oncol. Res. Pract.* 3, 11.
- Getis, A., and Ord, J.K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geogr. Anal.* 24, 189–206.
- Goode, E.L., Block, M.S., Kalli, K.R., Vierkant, R.A., Chen, W., Fogarty, Z.C., Gentry-Maharaj, A., Toloczo, A., Hein, A., Bouligny, A.L., et al.; Ovarian Tumor Tissue Analysis (OTTA) Consortium (2017). Dose-Response Association of CD8+ Tumor-Infiltrating Lymphocytes and Survival Time in High-Grade Serous Ovarian Cancer. *JAMA Oncol.* 3, e173290.
- Ha, G., Roth, A., Khattra, J., Ho, J., Yap, D., Prentice, L.M., Melnyk, N., McPherson, A., Bashashati, A., Laks, E., et al. (2014). TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. *Genome Res.* 24, 1881–1893.
- Heindl, A., Lan, C., Rodrigues, D.N., Koelble, K., and Yuan, Y. (2016). Similarity and diversity of the tumor microenvironment in multiple metastases: critical implications for overall and progression-free survival of high-grade serous ovarian cancer. *Oncotarget* 7, 71123–71135.
- Herbst, R.S., Soria, J.-C., Kowanetz, M., Fine, G.D., Hamid, O., Gordon, M.S., Sosman, J.A., McDermott, D.F., Powderly, J.D., Gettinger, S.N., et al. (2014). Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* 515, 563–567.
- Homicsko, K., and Coukos, G. (2015). Targeting Programmed Cell Death 1 in Ovarian Cancer. *J. Clin. Oncol.* 33, 3987–3989.
- Hwang, W.-T., Adams, S.F., Tahirovic, E., Hagemann, I.S., and Coukos, G. (2012). Prognostic significance of tumor-infiltrating T cells in ovarian cancer: a meta-analysis. *Gynecol. Oncol.* 124, 192–198.
- Jiménez-Sánchez, A., Memon, D., Pourre, S., Veeraraghavan, H., Li, Y., Vargas, H.A., Gill, M.B., Park, K.J., Zivanovic, O., Konner, J., et al. (2017). Heterogeneous Tumor-Immune Microenvironments among Differentially Growing Metastases in an Ovarian Cancer Patient. *Cell* 170, 927–938.
- Joshi, P.M., Sutor, S.L., Huntoon, C.J., and Karnitz, L.M. (2014). Ovarian cancer-associated mutations disable catalytic activity of CDK12, a kinase that promotes homologous recombination repair and resistance to cisplatin and poly(ADP-ribose) polymerase inhibitors. *J. Biol. Chem.* 289, 9247–9253.
- Klarenbeek, P.L., Doorenspleet, M.E., Esveldt, R.E.E., van Schaik, B.D.C., Lardy, N., van Kampen, A.H.C., Tak, P.P., Plenge, R.M., Baas, F., de Bakker, P.I.W., and de Vries, N. (2015). Somatic Variation of T-Cell Receptor Genes Strongly Associate with HLA Class Restriction. *PLoS ONE* 10, e0140815.
- Kroeger, D.R., Milne, K., and Nelson, B.H. (2016). Tumor-Infiltrating Plasma Cells Are Associated with Tertiary Lymphoid Structures, Cytolytic T-Cell Responses, and Superior Prognosis in Ovarian Cancer. *Clin. Cancer Res.* 22, 3005–3015.
- Le, D.T., Uram, J.N., Wang, H., Bartlett, B.R., Kemberling, H., Eyring, A.D., Skora, A.D., Luber, B.S., Azad, N.S., Laheru, D., et al. (2015). PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* 372, 2509–2520.
- Leong, H.S., Galletta, L., Etemadmoghadam, D., George, J., Köbel, M., Ramus, S.J., and Bowtell, D.; Australian Ovarian Cancer Study (2015). Efficient molecular subtype classification of high-grade serous ovarian cancer. *J. Pathol.* 236, 272–277.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760.
- Li, H.M., Hiroi, T., Zhang, Y., Shi, A., Chen, G., De, S., Metter, E.J., Wood, W.H., 3rd, Sharov, A., Milner, J.D., et al. (2016). TCR β repertoire of CD4+ and

- CD8+ T cells is distinct in richness, distribution, and CDR3 amino acid composition. *J. Leukoc. Biol.* 99, 505–513.
- Lipinski, K.A., Barber, L.J., Davies, M.N., Ashenden, M., Sottoriva, A., and Gerlinger, M. (2016). Cancer Evolution and the Limits of Predictability in Precision Cancer Medicine. *Trends in cancer* 2, 49–63.
- Lo, C.S., Sanii, S., Kroeger, D.R., Milne, K., Talhouk, A., Chiu, D.S., Rahimi, K., Shaw, P.A., Clarke, B.A., and Nelson, B.H. (2017). Neoadjuvant Chemotherapy of Ovarian Cancer Results in Three Patterns of Tumor-Infiltrating Lymphocyte Response with Distinct Implications for Immunotherapy. *Clin. Cancer Res.* 23, 925–934.
- McGranahan, N., Furness, A.J.S., Rosenthal, R., Ramskov, S., Lyngaa, R., Saini, S.K., Jamal-Hanjani, M., Wilson, G.A., Birkbak, N.J., Hiley, C.T., et al. (2016). Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* (80-) 351, 1463–1469.
- McGranahan, N., Rosenthal, R., Hiley, C.T., Rowan, A.J., Watkins, T.B.K., Wilson, G.A., Birkbak, N.J., Veeriah, S., Van Loo, P., Herrero, J., and Swanton, C.; TRACERx Consortium (2017). Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell* 171, 1259–1271.
- McPherson, A., Roth, A., Laks, E., Masud, T., Bashashati, A., Zhang, A.W., Ha, G., Biele, J., Yap, D., Wan, A., et al. (2016). Divergent modes of clonal spread and intraperitoneal mixing in high-grade serous ovarian cancer. *Nat. Genet.* 48, 758–767.
- McPherson, A., Shah, S.P., and Sahinalp, S.C. (2017a). deStruct: Accurate Rearrangement Detection using Breakpoint Specific Realignment. *bioRxiv*. <https://doi.org/10.1101/117523>.
- McPherson, A.W., Roth, A., Ha, G., Chauve, C., Steif, A., de Souza, C.P.E., Eirew, P., Bouchard-Côté, A., Aparicio, S., Sahinalp, S.C., and Shah, S.P. (2017b). ReMixT: clone-specific genomic structure estimation in cancer. *Genome Biol.* 18, 140.
- Mirza, M.R., Monk, B.J., Herrstedt, J., Oza, A.M., Mahner, S., Redondo, A., Fabbro, M., Ledermann, J.A., Lorusso, D., Vergote, I., et al.; ENGOT-OV16/NOVA Investigators (2016). Niraparib Maintenance Therapy in Platinum-Sensitive, Recurrent Ovarian Cancer. *N. Engl. J. Med.* 375, 2154–2164.
- Nawaz, S., Heindl, A., Koelble, K., and Yuan, Y. (2015). Beyond immune density: critical role of spatial heterogeneity in estrogen receptor-negative breast cancer. *Mod. Pathol.* 28, 766–777.
- Nelson, B.H. (2015). New insights into tumor immunity revealed by the unique genetic and genomic aspects of ovarian cancer. *Curr. Opin. Immunol.* 33, 93–100.
- Nielsen, J.S., Sahota, R.A., Milne, K., Kost, S.E., Nesslinger, N.J., Watson, P.H., and Nelson, B.H. (2012). CD20+ tumor-infiltrating lymphocytes have an atypical CD27- memory phenotype and together with CD8+ T cells promote favorable prognosis in ovarian cancer. *Clin. Cancer Res.* 18, 3281–3292.
- Nik-Zainal, S., Davies, H., Staaf, J., Ramakrishna, M., Glodzik, D., Zou, X., Martincorena, I., Alexandrov, L.B., Martin, S., Wedge, D.C., et al. (2016). Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 534, 47–54.
- Patch, A.-M., Christie, E.L., Etemadmoghadam, D., Garsed, D.W., George, J., Fereday, S., Nones, K., Cowin, P., Alsop, K., Bailey, P.J., et al.; Australian Ovarian Cancer Study Group (2015). Whole-genome characterization of chemoresistant ovarian cancer. *Nature* 521, 489–494.
- Popova, T., Manié, E., Boeva, V., Battistella, A., Goundiam, O., Smith, N.K., Mueller, C.R., Raynal, V., Mariani, O., Sastre-Garau, X., and Stern, M.H. (2016). Ovarian Cancers Harboring Inactivating Mutations in CDK12 Display a Distinct Genomic Instability Pattern Characterized by Large Tandem Duplications. *Cancer Res.* 76, 1882–1891.
- Rooney, M.S., Shukla, S.A., Wu, C.J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 160, 48–61.
- Roth, A., Khattri, J., Yap, D., Wan, A., Laks, E., Biele, J., Ha, G., Aparicio, S., Bouchard-Côté, A., and Shah, S.P. (2014). PyClone: statistical inference of clonal population structure in cancer. *Nat. Methods* 11, 396–398.
- Saunders, C.T., Wong, W.S.W., Swamy, S., Becq, J., Murray, L.J., and Cheetham, R.K. (2012). Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* 28, 1811–1817.
- Schwarz, R.F., Ng, C.K.Y., Cooke, S.L., Newman, S., Temple, J., Piskorz, A.M., Gale, D., Sayal, K., Murtaza, M., Baldwin, P.J., et al. (2015). Spatial and temporal heterogeneity in high-grade serous ovarian cancer: a phylogenetic analysis. *PLoS Med.* 12, e1001789.
- Smid, M., Rodríguez-González, F.G., Sieuwerts, A.M., Salgado, R., Prager-Van der Smissen, W.J.C., Vlugt-Daane, M.V., van Galen, A., Nik-Zainal, S., Staaf, J., Brinkman, A.B., et al. (2016). Breast cancer genome and transcriptome integration implicates specific mutational signatures with immune cell infiltration. *Nat. Commun.* 7, 12910.
- Spranger, S., Spaepen, R.M., Zha, Y., Williams, J., Meng, Y., Ha, T.T., and Gajewski, T.F. (2013). Up-Regulation of PD-L1, IDO, and Tregs in the Melanoma Tumor Microenvironment Is Driven by CD8+ T Cells. *Sci. Transl. Med.* 5, 200ra116.
- Szolek, A., Schubert, B., Mohr, C., Sturm, M., Feldhahn, M., and Kohlbacher, O. (2014). OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics* 30, 3310–3316.
- Ternette, N., Yang, H., Partridge, T., Llano, A., Cedeño, S., Fischer, R., Charles, P.D., Dudek, N.L., Mothe, B., Crespo, M., et al. (2016). Defining the HLA class I-associated viral antigen repertoire from HIV-1-infected human cells. *Eur. J. Immunol.* 46, 60–69.
- Tothill, R.W., Tinker, A.V., George, J., Brown, R., Fox, S.B., Lade, S., Johnson, D.S., Trivett, M.K., Etemadmoghadam, D., Locandro, B., et al.; Australian Ovarian Cancer Study Group (2008). Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. *Clin. Cancer Res.* 14, 5198–5208.
- Wang, W., Kryczek, I., Dostál, L., Lin, H., Tan, L., Zhao, L., Lu, F., Wei, S., Maj, T., Peng, D., et al. (2016). Effector T Cells Abrogate Stroma-Mediated Chemo-resistance in Ovarian Cancer. *Cell* 165, 1092–1105.
- Wang, Y.K., Bashashati, A., Anglesio, M.S., Cochrane, D.R., Grewal, D.S., Ha, G., McPherson, A., Horlings, H.M., Senz, J., Prentice, L.M., et al. (2017). Genomic consequences of aberrant DNA repair mechanisms stratify ovarian cancer histotypes. *Nat. Genet.* 49, 856–865.
- Wick, D.A., Webb, J.R., Nielsen, J.S., Martin, S.D., Kroeger, D.R., Milne, K., Castellarin, M., Twumasi-Boateng, K., Watson, P.H., Holt, R.A., and Nelson, B.H. (2014). Surveillance of the tumor mutanome by T cells during progression from primary to recurrent ovarian cancer. *Clin. Cancer Res.* 20, 1125–1134.
- Yoshihama, S., Roszik, J., Downs, I., Meissner, T.B., Vijayan, S., Chapuy, B., Sidiq, T., Shipp, M.A., Lizee, G.A., and Kobayashi, K.S. (2016). NLRC5/MHC class I transactivator is a target for immune evasion in cancer. *Proc. Natl. Acad. Sci. USA* 113, 5999–6004.
- Zhang, L., Conejo-Garcia, J.R., Katsaros, D., Gimotty, P.A., Massobrio, M., Regnani, G., Makrigiannakis, A., Gray, H., Schlienger, K., Lieberman, M.N., et al. (2003). Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N. Engl. J. Med.* 348, 203–213.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-CD8, clone C8/144B	Cell Marque, Rocklin, CA	Cat#108M-94
Rabbit anti-CD3, clone SP7	Spring Biosciences, Pleasanton, CA	Cat#M3074, RRID: AB_1660772
Rabbit anti-CD79a, clone SP18	Spring Biosciences, Pleasanton, CA	Cat#M3182, RRID: AB_11217923
Mouse anti-CD138, clone B-A38	Biocare Medical, Pacheco, CA	Cat#CM167
Mouse anti-CD20, clone L26	Biocare Medical, Pacheco, CA	Cat#CM004
Biological Samples		
Multisite HGSC tissue samples, patient IDs 1-47	BC Cancer Agency OvCaRe tissue bank	N/A
Deposited Data		
Raw WGS sequence data for patients 11-17, 25, 26, 28-32	This paper	EGA: EGAS00001002839
Raw WGS sequence data, patients 1-10	McPherson et al., 2016	EGA: EGAS00001000547
Raw WGS sequence data, patients 18-23	Bashashati et al., 2013	EGA: EGAS00001000547
Raw TCR/BCR-seq data, patients 1-23, 25, 28-32	This paper	EGA: EGAS00001002839
Raw Nanostring data, patients 1-23	This paper	EGA: EGAS00001002839
Immunohistochemistry data, patients 1-23	This paper	Table S1
Oligonucleotides		
Primers for targeted deep sequencing, patients 11-17	This paper	Table S6
Primers for targeted deep sequencing, patients 1-10	McPherson et al., 2016	Supplemental Tables 10, 14, 15
Software and Algorithms		
BWA v0.7.12	BWA	http://bio-bwa.sourceforge.net/
samtools v1.4.1	SamTools	http://samtools.sourceforge.net/
MutationSeq 4.2.0	Shah Lab	http://shahlab.ca/
Strelka 1.0.14	Illumina	https://sites.google.com/site/strelkasomaticvariantcaller/home
PyClone 0.13.0	Shah Lab	http://shahlab.ca/
ReMixT	Shah Lab	https://bitbucket.org/dranew/remixt
deStruct	Shah Lab	https://bitbucket.org/dranew/destruct
TITAN	Shah Lab	http://shahlab.ca/
primer3 v2.3.7	Primer3	http://primer3.sourceforge.net/
UCSC <i>in silico</i> PCR tool	UCSC, Kent Informatics	http://www.kentinformatics.com/
LICHeE v1.0	GitHub	https://github.com/viq854/lichee
MiXCR 2.0	MiLaboratory, GitHub	https://github.com/milaboratory/mixcr/releases
OptiType v1.2	GitHub	https://github.com/FRED-2/OptiType
NetMHCpan 3.0	Center for Biological Sequence Analysis, Technical University of Denmark	http://www.cbs.dtu.dk/services/NetMHCpan/
inForm Cell Analysis	PerkinElmer	http://www.perkinelmer.com/category/image-analysis-software

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
LOHHLA	McGranahan et al., 2017	https://bitbucket.org/mcgranahanlab/lohlla
Clone-level HLA LOH model	This paper	Available from authors
QuPath v0.1.2	GitHub	https://github.com/qupath/qupath
QuPath-trained algorithms	This paper	Available from authors

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Sohrab P. Shah (sshah@bccrc.ca).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Sample acquisition, consent, & surgery

Ethical approval for this study was obtained from the University of British Columbia (UBC) Research Ethics Board. Women (biological sex: XX) undergoing debulking surgery (primary or recurrent) for carcinoma of ovarian/peritoneal/fallopian tube origin were approached for informed consent to bank tumor tissue. Cases of high-grade serous carcinoma where more than one sample was collected were chosen for this analysis. Clinicopathologic and outcome data were collected by chart review. Consistent with the practice at UBC and BC Cancer, all patients with high-grade serous ovarian cancer (HGSC) are referred to the hereditary cancer clinic and offered genetic testing for BRCA1 and BRCA2 mutations (http://www.bccancer.bc.ca/screening/Documents/HCP_GuidelinesManuals-HBOCCriteria.pdf).

For consented patients, when multiple tumor sites were encountered intraoperatively, effort was made to bank as many sites as possible. Samples were flash frozen and stored according to conditions outlined below. For cases where multiple tumor sites were encountered but not all anatomic sites could be frozen (e.g., due to unavailability of trained staff), archival specimens stored within our pathology department were used. All samples were from removed structures during attempts at optimal debulking; hence the majority of samples were from omentum and ovarian sites.

Platinum sensitive is defined as no relapse within 6 months of the chemotherapy stop date.

Sample preservation & histologic evaluation

When adequate tumor volume was available, multiple tissue samplings were obtained from each tissue specimen. Up to 5 samplings were taken from a given tumor, with effort made to equally space samples while staying within grossly apparent tumor tissue. Each sampling was cut into three pieces, yielding two end-pieces for cryovials and a middle portion placed in 10% buffered formalin. End pieces were homogenized manually and with a paddle blender (Stomacher). All paraffin-embedded blocks, including formalin-fixed tumor samples and molecular-fixed fallopian tubes, were sectioned and stained with hematoxylin and eosin prior to expert histopathological review to confirm the presence of high-grade serous carcinoma. Pieces from the same sampling were given the same sample identifier for the analysis steps described below.

Distribution and availability of samples

Investigators interested in the use of samples from this study should contact the Lead Contact for further information.

METHOD DETAILS

WGSS library construction & sequencing

Frozen tumor samples from 14 patients (patients 11-17, 25, 26, 28-32, total 71 samples) were submitted for library construction and sequencing. Sample size was determined by availability of resectable, cryopreserved tissue, and DNA quality. For all tumor and normal samples, DNA extraction was followed by library construction and sequencing using Illumina HiSeq2500 whole genome shotgun v4 chemistry with paired-end 125bp reads. Samples were sequenced to an average of 96X coverage. Patients 1-4, 7, 9, and 10 were previously sequenced according to specifications described in McPherson et al. (2016).

Targeted bulk sequencing analysis

Target selection

For each patient we performed targeted sequencing on (11-17), a total of 192 positions were deeply sequenced, including 4 experimental controls, a TP53 variant, heterozygous germline SNPs lost in dominant loss of heterozygosity (LOH) events, lost SNVs that

could and could not be explained by copy number events, and SNVs inferred to originate at each node of the sample phylogeny obtained by applying the stochastic Dollo approach (infinite sites with loss model) (McPherson et al., 2016) (Table S1). SNVs were sampled as evenly as possible across nodes.

Data for patients 1-4, 7, 9, and 10 was obtained from McPherson et al. (2016), and used as input for section Clonal analysis onward.

Primer design

Primers targeting the positions described above were designed using primer3. The full list of primers is included in Table S1. Optimal primer length was 27nt (18-30nt) and products were designed to be 150-250nt long with 53-61°C melting temperature. Breslauer thermodynamic correction and Schildkraut and Lifson salt correction settings in primer3 were used. Additionally, primers targeting SNVs were required to pass the following preliminary filters: minimum of 5 alignments to the genome as given by BLAT for each primer, and each primer position at least 30nt away from the target SNV.

Primers were additionally tested using a combination of UCSC's *in silico* PCR tool (<http://genome.ucsc.edu/cgi-bin/hgPcr>) aligned against the reference hg19 genome and custom in-house code (Canada's Michael Smith Genome Sciences Centre) to verify a unique hit and check that the variant was located within 150bp of the nearest end of the amplicon to ensure coverage in an Illumina NextSeq 150bp paired end read. The primers were tagged with Illumina adapters to enable a direct sequencing approach that precludes the need for adaptor ligation during sample preparation. The Illumina adaptor tags were: 5'-CGCTCTCCGATCTCTG-3' on the forward amplicon primer and 5'-TGCTCTTCCGATCTGAC-3' on the reverse amplicon primers.

PCR and Illumina sequencing

Genomic DNA templates were used as starting material to generate PCR products. PCR was set up using Phusion DNA polymerase (Fisher Scientific, USA) according to the manufacturer's specifications. The standard PCR conditions used were an initial denaturation at 98°C for 30 s, followed by 35 cycles of 98°C for 10 s, 60°C for 15 s and 72°C for 8 s, and a final extension at 72°C for 10 minutes.

Amplicons were pooled by template for sequencing sample preparation. Sample preparation involved a second round of amplification using Phusion DNA polymerase with 6 PCR cycles using PE primer 1.0-DS (5'-AATGATAACGGCGACCACCGAGATC TACACTTTCCCTACACGACGCTCTCCGATCTCTG-3') and a custom PCR Primer (5'-CAAGCAGAAAGACGGCATACGAGATNNNNNGTGA CTTGGAGTTCA GACGTGTGCTCTCCGATCTGAC-3') that contains a unique six-nucleotide 'index' shown as N's. PCR products were cleaned up using PCRClean DX beads (Aline Biosciences, USA). DNA quality was assessed using the Caliper LabChip GX High Sensitivity Assay (Caliper Life Sciences, USA) and DNA quantity was measured using a Qubit dsDNA HS assay kit on a Qubit fluorometer (Life Technologies, USA).

The indexed libraries were pooled together and sequenced on the Illumina NextSeq500 platform with paired-end 150bp reads using v2 chemistry reagents.

Immunohistochemistry

All reagents were from Biocare Medical (Pacheco, CA) unless otherwise stated. Slides of formalin-fixed, paraffin embedded tissue were deparaffinized and rehydrated through xylene and graded alcohols. Antigen retrieval was performed using Diva Decloaker in a Biocare decloaking chamber at 125°C for 30 s. Slides were then rinsed with water, marked with PAP pen and loaded into the Biocare Intellipath FLX autostainer. Slides were blocked with peroxidized-1 and background sniper for 5 minutes and 10 minutes respectively then a cocktail of either CD8 (1/250, clone C8/144B, Cell Marque, Rocklin, CA) and CD3 (1/500, clone SP7, Spring Biosciences, Pleasanton, CA), or CD79a (1/400, clone SP18, Spring Biosciences, Pleasanton, CA) and CD138 (1/200, clone B-A38, Biocare Medical, Pacheco, CA) in Da Vinci Green diluent was added for 30 minutes at room temperature. Following a wash step, Mach2 Doublestain #2 polymer was added for 30 minutes at room temperature and then antigens detected with IP Ferengi Blue chromogen for 7 minutes followed by IP DAB chromogen for 5 minutes. To denature the first round of antibodies, slides were removed from the autostainer and placed in pre-warmed SDS-glycine pH 2.0 solution for 45 minutes at 50°C with periodic agitation. Slides were then washed with water and replaced in the autostainer for the 2nd round of staining. CD20 (1/300, clone L26, Biocare Medical, Pacheco, CA) diluted in Da Vinci Green diluent was added to the slides and incubated for 30 minutes at room temperature. Mach2 Mouse-AP polymer or Mach2 Rabbit-AP polymer was added for 30 minutes at room temperature to detect CD20. Warp red chromogen was added to the slide for 7 minutes, hematoxylin at a 1/5 dilution was then added for 5 minutes. The slides were then washed, air-dried and coverslipped with Ecomount coverslipping medium.

Nanostring gene expression

FFPE samples were deparaffinised with xylene and washed with 100% ethanol. Tissue was then extracted using QIAGEN miRNeasy FFPE Kit, following the protocol for purification of total RNA (including miRNA) from FFPE tissue sections. RNA quality was assessed with Nanodrop. 500ng of high quality RNA (260/280 ratio of 1.7-2.3 and A260/230 ratio of 1.8-2.3) for each sample was used in the Nanostring assay (PanCancer Immune Profiling panel (Cesano, 2015) additionally containing markers for high-grade serous ovarian cancer subtypes C1, C2, C4, and C5 (Leong et al., 2015)). Data was normalized with the voom function from the R package limma and TMM normalization. Samples flagged by nSolver (Nanostring Technologies) were removed from further analysis.

TCR & BCR sequencing

In the text below, *TRB* and *IGH* refer to TCR- β chain and Ig-heavy chain, respectively.

RNA was extracted from frozen tissue using the miRNeasy Mini kit. Quality (260/280) and quantity were determined using Nanodrop. Total RNA samples were also QC checked using the Caliper HT RNA HiSens assay (Caliper Life Sciences, USA). Samples ranging from 60–255ng RNA were re-arrayed into a 96-well plate. First-strand cDNA was synthesized from the total RNA samples using the SMARTScribe Reverse Transcriptase from Clontech, BNA oligo, *TRB* and *IGH* gene specific primers at a concentration of 0.5uM. Reactions were incubated on a tetrad using the following program: 90mins at 42°C, 15mins at 70°C and 2mins at 4°C. Using cDNA as a template, first round PCR for *TRB* and *IGH* was set up using Phusion DNA polymerase (Fisher Scientific, USA) according to manufacturer's specifications. The gene specific primers used were *TRB* 5'-TCTCTGCTTCTGATGGCTCAAAC-3' and *IGH* 5'-ACACCGTCACCGGTTGG-3'. The PCR conditions used were an initial denaturation of 98°C for 30 s, followed by 35 cycles of 98°C for 10 s, 55°C for 10 s and 72°C for 20 s, and a final extension at 72°C for 5 minutes. PCR products were size selected and cleaned up using PCRClean DX beads (Aline Biosciences, USA). Using first round PCR product as a template, a nested round of PCR for *TRB* and *IGH* was set up using Phusion DNA polymerase (Fisher Scientific, USA) according to manufacturer's specifications. The gene specific primers used were *TRB* 5'-TGCTCTTCCGATCTGACAGCGACCTGGGTGGAAACA-3' and *IGH* 5'-TGCTCTTCCGATCTGACAAGACSGATGGCCCTGGT-3'. The PCR conditions used were an initial denaturation of 98°C for 30 s, followed by 10 cycles of 98°C for 10 s, 65°C for 10 s and 72°C for 20 s, and a final extension at 72°C for 5 minutes. PCR products were cleaned up using PCRClean DX beads (Aline Biosciences, USA).

TRB and *IGH* amplicons were pooled by template for sequencing sample preparation. Sample preparation involved a second round of amplification using Phusion DNA polymerase with 6 PCR cycles using PE primer 1.0-DS (5'-AATGATACGGCGACCACC GAGATCTACACTCTTCCCTACACGACGCTCTCCGATCTCTG-3') and a custom PCR Primer (5'-CAAGCAGAACGACGGCATAC GAGATNNNNNNTGACTGGAGTTCAGACGTGTGCTTCCGATCTGAC-3') that contains a unique six-nucleotide 'index' shown here as N's. Products were cleaned up using PCRClean DX beads (Aline Biosciences, USA). DNA quality was assessed using the Caliper LabChip GX High Sensitivity Assay (Caliper Life Sciences, USA) and DNA quantity was measured using a Qubit dsDNA HS assay kit on a Qubit fluorometer (Life Technologies, USA).

The indexed libraries were pooled together and sequenced on the Illumina HiSeq platform with paired-end 250bp reads using v2 chemistry reagents.

QUANTIFICATION AND STATISTICAL ANALYSIS

WGSS analysis

Alignment

Reads were aligned to the hg19 reference genome downloaded from http://www.bcgsc.ca/downloads/genomes/9606/hg19/1000genomes/bwa_ind/genome/GRCh37-lite.fa. Alignments were performed using bwa (Li and Durbin, 2009) using the aln and sampe commands. Duplicates were flagged with Picard <http://broadinstitute.github.io/picard/>.

SNV and indel calling

Somatic SNVs were called using both Strelka 1.0.14 (Saunders et al., 2012) and MutationSeq 4.2.0 (Ding et al., 2012) with default parameters. Somatic indels were additionally called with Strelka. We considered a somatic SNV high quality if it was predicted by both MutationSeq and Strelka to be present in any sample from a patient, not necessarily the same sample for each program. Germline SNVs and indels were called with samtools mpileup and bcftools call 1.4.1, with default parameters.

Gene name, predicted effect and impact of SNVs and indels were annotated using SnpEff 4.0e. Mappability scores were annotated for each position using precomputed values downloaded from UCSC (<http://hgdownload-test.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeMapability/release3/wgEncodeCrgMapabilityAlign50mer.bigWig>). For downstream analysis we only considered variants with a mappability score > 0.99.

Breakpoint calling

We used deStruct (McPherson et al., 2017a) and lumpy (<https://github.com/arq5x/lumpy-sv>) to call breakpoints from WGS data. deStruct breakpoints were filtered for those with at least 2 discordant reads, and at least 2 split reads. Additional filters removed breakpoints for which the reconstructed sequence was less than 120nt, and removed breakpoints with read data likelihood less than -20. Following this, the intersection of deStruct and lumpy predictions was taken, and events lying within poor mappability regions, with break distance ≤ 30bp, and deletions with breakpoint size < 1000bp were excluded (Wang et al., 2017). Furthermore, breakpoints overlapping germline structural variation as determined from the database of genomic variants or identification of a similar event in the matched normal sample. Classification of breakpoint and rearrangement type was performed according to Wang et al. (2017).

Copy number calling

We applied ReMixT (McPherson et al., 2017b) to predict allele and clone-specific copy number from WGS samples. ReMixT jointly infers clone and allele specific copy number of both segments and breakpoints, allowing for increased statistical strength for detecting subclonal rearrangements associated with subclonal copy number changes. Additionally, ReMixT uses haplotype blocks obtained from phased SNPs to increase the power for detecting small allelic imbalances resulting from subclonal copy number changes. ReMixT was run on each patient's full set of WGS samples with default parameters. Accurately inferred clone specific segment copy number was used to calculate the length-normalized proportion of segments predicted with divergent clonal copy number.

In order to call high-level amplification (HLAMP), we employed identical methods to Wang et al. (2017). We ran TITAN (Ha et al., 2014) on WGS data to infer logR values; HLAMP was called for segments with median logR values > 1.

Identifying BRCA variants

Point mutations and indels in *BRCA1* and *BRCA2* were called from germline and somatic WGS data, as described above. Variants with high SnpEff-annotated impact were used. Somatic BRCA status was determined from variant calls. Where available, clinical test results were used to determine germline BRCA status; germline variant calls were used for patients that did not consent to clinical testing. Clinically-determined BRCA status is shown in (Table 1).

Clonal analysis

Mutation cluster inference

We ran PyClone 0.13.0 (Roth et al., 2014) in multi-sample mode to perform initial clonal analysis. Parental copy number and tumor content estimates from ReMixT along with reference and alternative allele counts from deep sequencing data of SNVs (PCR and Illumina sequencing) were used as input for PyClone. The following SNVs were filtered out for clonal analysis: germline SNVs, SNVs absent (probability < 0.01) in all samples in a patient (probabilities computed from a binomial test, assuming a sequence error rate of 0.001), and SNVs on sex chromosomes. The MCMC chain was run for 100,000 iterations, with a burn-in of 50,000. Posterior plots were visually inspected to confirm convergence. Flat cluster assignments were produced from posterior similarity matrices using the MPEAR method described in Roth et al. (2014). SNVs with broad posterior cellular prevalence distributions (width of 95% credible interval ≥ 0.2) far from the corresponding cluster median (difference of ≤ 0.05) were excluded from further analysis. Additionally, clusters absent or present at low prevalence in all samples (median cluster prevalence across SNVs ≤ 0.05 in all samples), with only one SNV, or with $\geq 50\%$ SNVs lost were filtered out.

Archival samples without a corresponding flash frozen sample (i.e., no copy number predictions) were excluded from this initial analysis. They are reintroduced in section Clonal phylogenies & postprocessing.

Clonal phylogenies & postprocessing

Filtered PyClone results were provided as input to LICHeE, a multi-sample cancer lineage inference method (<https://github.com/viq854/lichee>), to elucidate clonal phylogenies. LICHeE was run in cellular prevalence mode (-cp), with additional options -completeNetwork -sampleProfile. Other parameters were set to the defaults. The top ranking lineage tree from LICHeE was kept.

To remove artifacts (e.g., falsely called low prevalence clones) and obtain clonal prevalences for archival samples, clonal prevalences were refined by resampling alternative and reference allele counts for deeply sequenced tumor samples and matched peripheral blood (normal) according to the following Bayesian generative model, adapted from McPherson et al. (2016). We suppress indices for samples as these can be treated independently.

We assume that the alternative allele counts of SNV n in the matched normal and tumor samples, b_{normal}^n and b_{tumor}^n , respectively, are distributed as:

$$b_{normal}^n | p_{normal}^n \sim \text{Binomial}(d_{normal}^n, p_{normal}^n)$$

and

$$b_{tumor}^n | Z_n = c, \xi_{nc}, \sigma_{tumor} \sim \text{BetaBinomial}(d_{tumor}^n, \xi_{nc}, \sigma_{tumor})$$

where d_{normal}^n and d_{tumor}^n correspond to the total read depth of SNV n in the normal and tumor sample, respectively, p_{normal}^n is the probability of observing the alternative allele of SNV n in the normal sample, σ_{tumor} is the dispersion parameter, Z_n is the cluster membership of SNV n , and ξ_{nc} , using similar notation to Roth et al. (2014), is given by:

$$\xi_{nc} \equiv \xi(\psi^n, \phi^c, t, p_{normal}^n) = \frac{(1-t)c(g_N)}{T^n} p_{normal}^n + \frac{t\phi^c\psi^n}{T^n}$$

where ψ^n is the copy number genotype of SNV n in the tumor variant population, t is tumor content, $c(g_N) = 1$ is the copy number genotype of the alternative allele in the normal population, total copy number $T^n = 2(1-t) + \psi^n t$, and ϕ^c is the cellular prevalence of PyClone cluster c , which can be expressed as the summation of clonal prevalences f_j over clones that contain PyClone cluster c . That is:

$$\phi^c = \sum_{j: G_j^c = 1} f_j$$

where G_j^c is a binary indicator of whether clone j contains PyClone cluster c . We then assume the following distributions over the parameters in Equations (1) and (2):

$$f \sim \text{Dirichlet}(\kappa)$$

$$\psi^n \sim \text{Categorical}(1)$$

$$p_{normal}^n \sim \text{Beta}(\zeta * \sigma_{normal}, (1 - \zeta) * \sigma_{normal})$$

with κ the Dirichlet parameter as defined in McPherson et al. (2016), and σ_{normal} the dispersion parameter. The value ζ corresponds to twice the mean allelic fraction of alternative alleles in the normal sample (twice because we model $c(g_N) = 1$). In essence, our model is analogous to that of McPherson et al. (2016), but we now consider the probability of sampling a variant allele from non-tumor cells to be nonzero, equal to p_{normal}^n , rather than 0.

Informally, the model can be described as follows. For each tumor sample:

1. Generate clonal prevalences
2. Compute the cellular prevalence of a mutation n by summing the prevalences of all clones containing the PyClone cluster associated with n
3. Generate the SNV-specific normal contamination fraction p_{normal}^n and allelic count data for the matched normal sample
4. Based on the contamination fraction, apply a modified PyClone likelihood model to simulate allelic count data in the tumor sample

The normal contamination fraction can be interpreted as the allelic fraction of SNV n in the matched normal, likely due to sequence errors or contamination. Samples with low tumor purity are particularly confounded by these issues; the addition of step 3 and modification of step 4 relative to McPherson et al. (2016) helps eliminate erroneously identified rare clones in these samples.

We set the following hyperparameter values: $\sigma_{tumor} = \sigma_{normal} = 200$ and κ as a repeating vector of 0.01. The effect of our setting for κ is to assume clonal purity unless there is substantial evidence for the contrary.

The Hamiltonian Monte Carlo chain was run for 10,000 iterations, with an additional burn-in of 5000. Posterior plots were visually inspected for convergence. Clones falling below a prevalence threshold (< 90% of the posterior distribution of clonal prevalence > 0.01) were removed.

Due to difficulties in lineage construction for patients with several samples composed of divergent clonal lineages (McPherson et al., 2016), results for patients 3 and 9 were taken from previously analyzed single-cell sequencing data (McPherson et al., 2016).

Clonal architecture distance

Pairwise similarity between clonal compositions (within a given patient) was computed using a modified version of the weighted UniFrac measure, to simultaneously incorporate clonal architecture and phylogeny information. First, clonal phylogenies from Clonal phylogenies & postprocessing were taken as ground truth and used to recompute cellular prevalences for all SNVs (denoted here as ψ_a and ψ_b) determined by WGS, where a and b denote the samples being compared. Clonal distance was computed as the summation of the differences in cellular prevalences across SNVs, or equivalently $\|\psi_a - \psi_b\|_1$.

Measures of intratumoral heterogeneity

Sample mixture entropy and clone divergence were defined as in McPherson et al. (2016). In order to compute divergence, SNVs from WGS data were assigned to PyClone clusters - and transitively, clones - by maximum likelihood according to the PyClone likelihood model (Roth et al., 2014). Proportion subclonality (copy number based measure) was computed as the proportion of the genome with subclonal copy number according to results from ReMixT. Heterogeneity index, a combined measure of intratumoral heterogeneity incorporating both clone prevalences and phylogenetic relationships, was computed as the sum of relative phylogenetic divergence between all pairs of distinct clones, weighted by clonal prevalence. The heterogeneity index is the mean phylogenetic divergence between a randomly selected pair of tumor cells from a sample (based on inferred clonal composition). Formally, for a sample A with clone set $C(A) = \{c_i\}$ and corresponding prevalences p_i (where $0 < p_i < 1$, $\sum_i p_i = 1$):

$$HI(A) = \sum_{c_j, c_k \in C(A)} p_j p_k D(c_j, c_k)$$

where $D(c_j, c_k)$ is the relative phylogenetic divergence between clones c_j and c_k , defined as:

$$D(c_j, c_k) = \frac{|S_{c_j} \cup S_{c_k}| - |S_{c_j} \cap S_{c_k}|}{|S_{c_j} \cup S_{c_k}|}$$

where S_{c_i} is the set of WGS SNVs assigned to clone c_i . By construction, the heterogeneity index takes values between 0 and 1. Intratumoral heterogeneity values for each sample are listed in Table S2.

Samples were also assigned to clonal mixture classes (pure, chain, branched) based on the phylogenetic relationships between constituent clones. Pure samples contained a single clone; chain samples contained clones along a single lineage (in other words, the minimal spanning tree is a line); branched samples contained at least 2 clones that were not ancestors/descendants of each other (in other words, the minimal spanning tree contains a bifurcation).

The significance of differences in the 3 clone-derived intratumoral heterogeneity measures (entropy, clone divergence, heterogeneity index) between the 3 TIL subtypes was assessed with the Kruskal-Wallis test (Figure 3A). Post hoc comparisons were made with Dunn's test (P -values were BH corrected).

To assess the significance of differences in subclonal copy number proportion between the 3 TIL subtypes, ANOVA was performed (aov function in R) with subclonal CN proportion as the dependent variable (logit-transformed, as subclonal CN proportion values lie between 0 and 1, exclusive), TIL subtype and cellularity as independent variables (to control for tumor cellularity). The residual plot did not indicate any substantial deviations from normality, with relatively constant variance across the fitted range. Post hoc comparisons were made with Tukey's range test (P -values were BH corrected).

RNA-seq analysis

RNA-seq raw counts for 54 primary HGSC tumors from the Australian Ovarian Cancer Study (OV-AU) ([Patch et al., 2015](#)) were downloaded from the International Cancer Genome Consortium (ICGC) Data Portal. Ensembl Gene IDs were mapped to gene symbols using biomaRt. Duplicate entries were summarized by taking the mean of expression values. Raw counts were normalized using voom from the R package limma with quantile normalization.

Mutation signature analysis

Data

Mutation signatures were jointly inferred for 102 multi-site HGSC tumors (21 patients), 62 primary HGSC tumors from the Australian Ovarian Cancer Study with BAM files ([Patch et al., 2015](#)), and 133 additional ovarian tumors (59 HGSC, 35 clear cell, 10 germinal cell, and 29 endometrioid) ([Wang et al., 2017](#)) ([Table S5](#)). Note that a *POLE* hypermutant (one of the endometrioid cases) was excluded from the original set of 133 cases described in [Wang et al. \(2017\)](#), and while 93 cases were available from the Australian Ovarian Cancer Study, only 62 had BAM files on the data portal. Similarly processed variant calls to WGSS analysis were obtained from [Wang et al. \(2017\)](#). In order to avoid counting the same variant more than once, the union of SNVs from all samples for each multi-site HGSC patient was analyzed together as a ‘meta-sample’.

Signature inference & clustering

Signatures and proportions were inferred from WGS SNV and rearrangement (structural variation, SV) calls (section WGSS analysis) by applying the multimodal correlated topic model method ([Funnell et al., 2018](#)). For SNVs, the pentanucleotide context of each variant is considered. Rearrangements (deletions, duplications, inversions, and foldback inversions) were binned by breakpoint distance (<10kb, 10kb-100kb, 100kb-1Mb, 1Mb-10Mb, >10Mb) and microhomology length ([Nik-Zainal et al., 2016](#); [Funnell et al., 2018](#)). The optimal number of SNV and SV clusters was determined using the elbow method on model log-likelihoods ([Funnell et al., 2018](#)). The probable identity of each point mutation signature is as follows: P-MMR-1 ↔ mismatch repair (MMR), P-HRD ↔ homologous recombination deficiency (HRD), P-UM ↔ ultramutator-associated mutation signature (present at very low levels in the HGSC samples; primarily observed because of an endometrial sample from [Wang et al. \(2017\)](#)), P-APOBEC ↔ APOBEC, P-AGE ↔ age signature, and P-MMR-2 ↔ uncertain, but with a strikingly similar T → C substitution pattern to the MMR signature. Sample-specific and non-ancestral mutation signatures were calculated by adding signature assignment weights for all constituent variants. For non-ancestral analysis ([Figure S6D](#)), non-ancestral SNVs were defined as those not present (and not called as ancestral) in all samples from that patient, and samples with fewer than 50 non-ancestral SNVs or SVs were excluded. Prior to clustering ([Figures 6A, S6C, and S6D](#)), signature proportions were scaled across the entire pooled cohort to a standard Gaussian distribution. Hierarchical clustering was performed with Ward’s method and a Pearson correlation-based distance measure ($d = (1-r)/2$, where r is the Pearson correlation coefficient). For patients in the discovery cohort with more than 2 samples, molecular subtype annotations on the heatmap correspond to the mode of subtype assignments for each patient. The 4 described subtypes (HRD-DEL, HRD-DUP, FBI, and TD) were recovered using the dynamicTreeCut R package (or equivalently, by cutting the dendrogram into 4 clusters).

Association with immune markers

RNA-seq expression data (see RNA-seq analysis, Nanostring analysis) from a set of 54 untreated primary OV-AU cases was used for the comparison depicted in [Figure 6C](#).

Differential gene expression

Differential gene expression analysis between mutation signature clusters for ICGC OV-AU cases (see RNA-seq analysis) was carried out using the limma method (R package). limma results for HRD versus FBI, TD versus FBI, and HRD versus TD contrast matrices were fed as input to the R package GAGE for gene set enrichment analysis using KEGG pathways. Pathways significantly up- or downregulated with $Q < 0.01$ were regarded as significant. Results of differential expression analysis are shown in [Figure S7](#) and [Table S6](#).

TCGA foldback inversions

A set of $n = 433$ TCGA ovarian serous cystadenocarcinoma cases with complete copy number, clinical, hg19 exome BAM files, and array-based gene expression data was selected for analysis ([Bell et al., 2011](#)). Selected TCGA cases are listed in [Table S7](#). Expression data was downloaded from the TCGA data portal and clinical data was downloaded from the TCGA Pancancer project under Synapse (ID: syn1461171).

Array gene expression data was preprocessed with the voom function from limma (R package), using quantile normalization. The median of normalized expression values for genes associated with cytotoxicity (derived from Nanostring PanCancer Immune Profiling Panel annotations ([Cesano, 2015](#))) was computed. Samples were stratified into immune-high and immune-low classifications by thresholding on median cytotoxicity score across the cohort ([Table S7](#)). To threshold on FBI status, foldback-amplification colocalization status (FBI-AMP High, FBI-AMP Low, No AMP) for all cases was retrieved from [Wang et al. \(2017\)](#). We performed a survival analysis on FBI groups after subsetting by immune cluster. The log-rank test was used to compare survival outcomes between subgroups.

A Cox proportional hazards model was also fit to the overall survival data, using foldback-amplification colocalization status as a discrete explanatory variable, interaction terms between cytotoxicity score and FBI-HLAMP status, along with control variables for age of pathologic diagnosis and treatment regimen (columns immunotherapy, additional immunotherapy, additional drug therapy, and additional chemotherapy in the Synapse table). Age of diagnosis was binned into < 50, 50-70, and > 70 categories, and along

with immunotherapy and additional chemotherapy used as stratification variables (as these originally violated the proportionality assumption). Patients without available data for age of diagnosis (5) were excluded. To assess the validity of the proportional hazards assumption, the cox.zph function the survival R package was used. None of the individual proportionality assumption tests or the global test were violated.

The R formula for the model was:

```
coxph(survival ~ mutation_signature_subgroup + cytotoxicity:mutation_signature_subgroup + strata(age_binned) + strata(immuno_therapy) + strata(additional_chemo_therapy) + additional_drug_therapy + additional_immuno_therapy, data)
```

To evaluate the significance of the model including the cytotoxicity \times FBI-HLAMP interaction term, we constructed an identical model, but with a cytotoxicity score as an explanatory variable without the interaction terms with FBI-HLAMP. A likelihood ratio test was performed on the resulting fits of the 2 models.

Immunohistochemistry analysis

Tissue segmentation & cell counting

Slides were scanned using the Vectra Multispectral Imaging System (Perkin Elmer) and 20 random 20X images (high-powered fields, HPFs) collected for each sample. The resulting multispectral images were then analyzed using Inform software (Perkin Elmer) with the resulting cell segregation data consolidated using Spotfire (Tibco). Phenotyping algorithms were created by 2 independent researchers (K.M., S.L.) and the results validated by a 3rd researcher (A.W.Z.). Briefly a training set of 10 images, selected to be histologically diverse on visual inspection, was used by each of the researchers to train Inform to recognize the different phenotypes of interest in each image. Training was run until at least 98% validation accuracy was achieved. The 2 algorithms were compared and visual inspection used to confirm the cell counts. TIL densities for each image were calculated by normalizing validated TIL counts by total area covered by tissue in the image (in units of cells/HPF). Overall TIL densities for each slide were similarly calculated, but using the summation of TIL counts and area across all constituent images. Epithelial and stromal TIL densities employed similar calculations, with counting and area restricted to epithelial/stromal regions identified by tissue segmentation (Inform). Thus, a cell was called epithelial if it fell within epithelial regions identified by Inform, and stromal if it fell within identified stromal regions.

Correlations between TIL densities

Correlations between TIL densities (epithelial and stromal CD8+, CD4+, CD20+, and plasma cell) were quantified with Spearman's correlation coefficient (Figure S5A) and *P* values of their significance were adjusted for multiple testing with the Benjamini-Hochberg method.

Clustering

Hierarchical clustering of TIL density profiles was performed using Ward's method with Euclidean distance. Heatmap values were obtained by normalizing (to a standard Gaussian distribution) across samples for each TIL type. For Figures 1B and 1C, only samples with valid epithelial and stromal TIL densities (i.e., non-zero epithelial and stromal tissue area) are shown. Additionally, for Figure 1B, only samples with both TIL density and Nanostring expression data are shown. The optimal number of clusters (3) was determined with the Dunn index.

Malignant clone similarity and TIL subtype

To compare whether samples from the same TIL subtype were more clonally similar (within patients), we used a nested ranks test (nestedRanksTest R package), treating patient as a random effect. Specifically, for each pair of samples within a patient, we (1) categorize them as belonging to the same, or different TIL subtypes (til_cluster_comparison); and (2) compute clonal composition similarity as per Clonal architecture distance (clonal_similarity). Then, we run:

```
nestedRanksTest(clonal_similarity ~ til_cluster_comparison | patient_id, data)
```

Nanostring analysis

Molecular subtyping

Ground truth molecular subtypes for a training set of 62 primary HGSC tumors from Patch et al. (2015) were obtained from the authors. Matched RNA-seq data for these tumors was obtained from the International Cancer Genome Consortium (project OV-AU) and normalized according to section RNA-seq analysis. The resulting expression profiles were pooled with Nanostring-derived expression profiles, and subjected to batch effect correction with the ComBAT R package. To confirm the effectiveness of batch correction, expression profiles from all samples were hierarchically clustered. Samples from different batches were not clearly segregated.

Following this, a *k*-nearest neighbors classifier (*k* = 5) was trained and applied to the data using the Patch et al. (2015) molecular subtypes as ground truth. Six-fold cross-validation accuracy of 85.8% on ground truth data was obtained, similar to that reported in Leong et al. (2015). As comparison, the diagonal LDA classifier attained an inferior 80.9% cross-validation accuracy and was thus not used. To further test these molecular subtypes, a subset of 62 tumors was additionally profiled with the Affymetrix U133A2 microarray platform. As described in Bashashati et al. (2013), the expression data from these tumors was normalized with RMA and quantile normalization, corrected for batch effects with ComBAT, pooled with TCGA array expression data (see TCGA foldback inversions), and subjected to another level of batch effect correction with ComBAT. Following the methods of TCGA (Bell et al., 2011), consensus non-negative matrix factorization (NMF) was applied to determine molecular subtypes (*k* = 4). NMF-derived subtypes and *k*-nearest neighbor-derived subtypes were largely concordant (mutual information: 0.74).

Overrepresentation of each molecular subtype or set of molecular subtypes within each IHC-based subgroup (N-TIL, S-TIL, ES-TIL) was computed relative to the other 2 subgroups and other molecular subtypes with Fisher's exact test.

Pathway signature analysis

Genes were grouped on the basis of pathway annotations from the Nanostring PanCancer Immune Profiling panel ([Cesano, 2015](#)). Metagene expression values were constructed by taking the median of expression values for constituent genes in each pathway.

TCR/BCR-seq analysis

Alignment and clonotype calling

Alignment to germline TCR and BCR segments was performed with mixcr align from MiXCR 2.0 ([Bolotin et al., 2015](#)), using the human IMGT reference (<https://github.com/repseqio/library-imgt/releases>, commit d993d704553c0a1e905c702ab93c99c0001b30d9). Reads mapping to the same clonotype were clustered using mixcr assemble, and the resulting *TRB* and *IGH* clonotypes were exported with mixcr export. Clonotypes were identified by V and J germline gene names and CDR3 nucleotide sequence. All other mixcr parameters were set to the defaults.

Decontamination and quality control

Clonotypes with fewer than 5 assigned reads were immediately removed. In order to filter out potential cross-sample contamination, clonotypes shared between samples from different patients were identified. Clonotypes present at an absolute prevalence (read count) in one sample > 25 times lower than in another sample from a different patient were removed (from the former sample). Consistent with contamination, samples (from different patients) arranged close by on each 96-well PCR plate contained a larger number of shared clonotypes. Finally, clonotypes that produced non-functional (frameshift or premature stop) receptor sequences were removed.

Prior to computing repertoire diversity or similarity, TCR/BCR reads were randomly downsampled (using the minimal nonzero library size across the cohort, for TCR/BCR separately) were randomly downsampled (10 times) with replacement from each sample to account for differences in library size. Mean clonotype abundances across these resamplings were used for the computations described below, and the corresponding statistics are reported in [Table S2](#).

Calculating repertoire diversity

The following indices of diversity were calculated:

- Number of unique clonotypes
- Shannon's entropy
- Efron-Thisted index
- D50 index (<https://patents.google.com/patent/WO2012097374A1/en>)

The Efron-Thisted index estimates the total repertoire diversity (by estimating the number of unseen clonotypes), and the D50 index quantifies the preponderance of rare clonotypes in a repertoire.

Correlations between repertoire diversity and ITH were computed as Spearman's rank correlation, using the first 2 measures listed above.

Repertoire similarity analysis

Pairwise similarity between TCR/BCR repertoires A and B was calculated with the Morisita-Horn index (R package vegan):

$$S(\mathbf{A}, \mathbf{B}) = \frac{2 \sum_{i=1}^N A_i B_i}{|\mathbf{A}| |\mathbf{B}| \left(\frac{\sum_{i=1}^N A_i^2}{|\mathbf{A}|^2} + \frac{\sum_{i=1}^N B_i^2}{|\mathbf{B}|^2} \right)}$$

where A_i denotes the number of reads associated with clonotype i in repertoire A , $|\mathbf{A}|$ and $|\mathbf{B}|$ are the total number of clonotype reads in A and B , respectively, and N is the number of unique clonotypes in $\mathbf{A} \cup \mathbf{B}$.

Correlation with clonal composition

TCR repertoire and clonal dissimilarity matrices were computed as described above. These dissimilarities were correlated with Mantel's test. Uncorrected P -values are reported in [Figure 5](#) and [Figure S3](#).

TCR clonotype classification

Previous studies have revealed differences in the physicochemical properties of CDR3 sequences ([Li et al., 2016](#)) and VJ ($V\beta$ - $J\beta$) gene usage ([Emerson et al., 2013](#)) between CD8+ and CD4+ T cells. We designed a binary classifier to predict the class (CD8+ or CD4+) of a T cell receptor based on both germline VJ genotype and physicochemical properties of the TCR CDR3 sequence.

Training data

To train the classifier, unprocessed TCR sequence data from flow-sorted naive CD8+ and CD4+ mononuclear cells derived from 18 unrelated healthy donors were obtained from a previous study ([Klarenbeek et al., 2015](#)). We made an effort to obtain TCR-sequence data of flow sorted CD8+ and CD4+ T cells from other sources as well ([Emerson et al., 2013; Li et al., 2016](#)), but these data were short-read or had been preprocessed (with no raw sequence files available), and thus not amenable to uniform downstream analysis. While these training data were derived from naive T cells, [Emerson et al. \(2013\)](#) have reported that there are no significant differences in $V\beta$

and $J\beta$ usage between naive and memory T cells (for both CD4+’s and CD8+’s separately). For the analysis described below, we operated under the assumptions that differences in VJ gene usage patterns and CDR3 physicochemical features between CD4+ and CD8+ T cells are similar in the training and multi-site HGSC datasets. We later assessed the validity of these assumptions by comparing predicted CD8/CD4 abundance with results from immunohistochemistry (see Classifier). Alignment and clonotype calling were carried out according to the methods described in Alignment and clonotype calling. Twenty percent of the data, stratified by class, was randomly split off for testing; 5-fold cross-validation was carried out on the remaining 80%.

Features

V and J genotypes were binarized (80 features). Additionally, Atchley factors (R package HDMD) quantifying the physicochemical properties of amino acids at each position in the CDR3 were used ($5n$ features, where n is the CDR3 amino acid length). Separate classifiers were trained for each length category between 11 and 18 amino acids (0.70 of all clonotypes). The distribution of V and J gene usage was comparable between training and test data.

Classifier

A binary gradient-boosted tree classifier was trained on the data described in section Alignment and clonotype calling. Training with 5-fold cross-validation was allowed to proceed until 100 consecutive rounds of no improvement in validation accuracy. Based on area under the receiver operating characteristic curve, the gradient-boosted tree classifier outperformed random forest, logistic regression, support vector machine (SVM), and extreme value regression classifiers. The classifier was then applied to clonotype calls from TCR-seq data of multisite HGSC samples to predict whether each clonotype was CD8-type or CD4-type. Clonotypes assigned to either class with >80% probability were kept.

Clonotype distribution broadness across tumor samples within each patient was computed with Simpson’s diversity index on the vector of per-sample relative clonotype prevalence values (R package vegan). The significance of differences in the distribution broadness between CD4+ and CD8+ associated TCRs was evaluated by computing the average of CD4+ and CD8+ TCR distribution broadness values within each patient, and applying the Wilcoxon signed-rank test for paired data between the two groups.

Neoantigen analysis

HLA typing

Four-digit HLA class I types were determined from WGS data for each multisite and background patient (see Neoantigen depletion score) using OptiType (Szolek et al., 2014). OptiType was run on the WGS bam of the normal sample.

Sample-level HLA LOH prediction

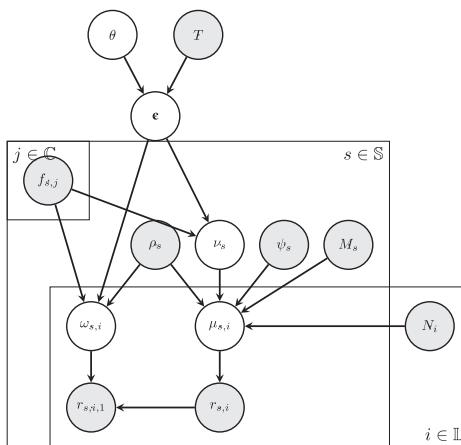
For OV-AU and Wang et al. (2017) patients, HLA class I loss-of-heterozygosity (LOH) was called from tumor and matched normal bams as well as OptiType 4-digit HLA types using LOHHHLA (McGranahan et al., 2017). HLA LOH was called for an allele if the estimated copy number (with binning and B-allele frequency settings) was < 0.5 and the significance of allelic imbalance $p < 0.1$ (paired t test, no duplicate counts). A less stringent P -value threshold (compared to McGranahan et al., 2017) was used due to the lower depth of the input bams.

Clone-level HLA LOH prediction

We devised a Bayesian statistical extension to call clone-level HLA LOH from multi-sample WGS data leveraging clonal phylogenies and clonal compositions inferred from Clonal phylogenies & postprocessing as input. Inference is done separately for each heterozygous HLA locus and patient. We define:

Variable	Description
T	Tumor clone phylogeny
$\mathbf{c} = \{c_j : j \in \mathbb{C}\}$	Set of HLA locus copy number genotypes, one for each clone
θ	“Stay” rate between copy number states
$f_{s,j}$	Prevalence of clone j in tumor sample s
$r_{s,i,1} \in \mathbb{N}_0$	Read depth at polymorphic site i for allele 1 in sample s
$r_{s,i} \in \mathbb{N}_0$	Total read depth at polymorphic site i in sample s (sum of allele 1 and 2)
ρ_s	Cellularity/tumor content of tumor sample s
ψ_s	Ploidy of tumor sample s
$\omega_{s,i}$	Allele 1 fraction at polymorphic site i in sample s
v_s	Total copy number of HLA locus in sample s
$\mu_{s,i}$	Mean parameter for total read depth at site i in sample s
M_s	Multiplicative factor between WGS library sizes of tumor sample s and the matched normal sample
$N_i \in \mathbb{N}_0$	Observed read depth at site i in the matched normal sample
L	Set of all polymorphic sites between the 2 alleles at a given HLA locus
S	Set of all tumor samples for a given patient
\mathbb{C}	Set of all clones in a given patient

Ploidy and cellularity estimates are assumed to be known and equal to the estimates from ReMixT (McPherson et al., 2017b). We present our graphical model:



We begin by defining the clone-specific copy number genotype at a given HLA locus c_j as a $(c_{j,1}, c_{j,2})$ tuple (allele 1 copy number and allele 2 copy number, respectively), where allele 1 can be arbitrarily assigned to either one of the 2 HLA alleles at a heterozygous locus without loss of generality. Given a clonal phylogeny T , we assume that the latent clone-specific copy number genotype at a given HLA locus evolves according to a Markov chain with transition rate $1 - \theta$, “stay rate” θ and the initial state distribution defined to be uniform across all possible genotypes. The transition and stay rates can be described by an n -by- n transition matrix P (n is the total number of genotype states) with diagonal entries $P_{ii} = \theta$ and non-diagonal entries satisfying $\sum_{j \neq i} P_{ij} = 1 - \theta$. In addition, the total transition probability $1 - \theta$ is divided evenly among all valid transitions (transitions from zero to non-zero allelic copy number are deemed invalid, as an allele cannot be acquired from nothing).

We use Markov chain Monte Carlo (MCMC) to sample from the posterior of c , the assignment of genotypes to clones described above. In what follows we describe our proposed distributions for the observed data given c . We assume that, given c , the observed read depth of allele 1, $r_{s,i,1}$, is distributed as:

$$r_{s,i,1}|c \sim \text{BetaBinomial}(r_{s,i}, \omega_{s,i}, \sigma),$$

where σ is the dispersion parameter and $\omega_{s,i}$, the fraction of allele 1 in tumor sample s (accounting for normal contamination), is given by:

$$\omega_{s,i} = \sum_{j \in C} \rho_s f_{s,j} c_{j,1} + (1 - \rho_s).$$

To then anchor the total copy number estimates, we use data from the matched normal bam. Given c , we assume that the total observed read depth at site i in sample s , $r_{s,i}$, follows:

$$r_{s,i}|c \sim \text{NegBinomial}(\mu_{s,i}, \alpha),$$

where α is the hyperparameter of the Gamma-distributed rate parameter in the negative binomial, and $\mu_{s,i}$, the expected read depth of polymorphic site i , can be computed as:

$$\mu_{s,i} = \frac{\nu_s}{\psi_s \rho_s + 2(1 - \rho_s)} \times M_s \times N_i,$$

with ν_s , the total copy number at the HLA locus under consideration for sample s , accounting for normal contamination, given by:

$$\nu_s = \sum_{j \in C} \rho_s f_{s,j} (c_{j,1} + c_{j,2}) + 2(1 - \rho_s).$$

The space of possible clonal genotypes c_j is restricted to those with total copy number ≤ 6 . The dispersion parameter σ for the beta binomial distribution is set to 200, and α for the negative binomial distribution is set to 0.5.

We consider the following prior distribution for the stay rate of the genotype Markov chain:

$$\theta \sim \text{TruncNormal}(\pi, \delta, 0, 1)$$

where 0 and 1 correspond to the lower and upper bounds of the truncated normal distribution, and the mean and standard deviation π and δ were set to be relatively uninformative (0.75 and 0.4, respectively).

MCMC was run for 100,000 iterations, using 50,000 additional tuning iterations. Bin median values were used to anchor total copy number. HLA LOH for a given clone j and allele a was called when $\geq 90\%$ of the posterior trace supported $c_{j,a} = 0$.

Identification of putative neoepitopes

All 8 to 11-mer peptides overlapping nonsynonymous SNVs were considered candidate epitopes. MHC-I binding affinity was computed for every mutant and corresponding wild-type allele using netMHCpan-3.0 (Ternette et al., 2016). Percentile binding scores of $\leq 2\%$, where the mutant epitope had equal or better affinity than the wild-type epitope, were considered as putative neoepitopes. In cases of HLA LOH, predicted neoepitopes associated with the lost HLA allele were excluded (for subclonal HLA LOH, a neoepitope was only excluded if all clones containing the neoepitope also exhibited loss of the corresponding HLA allele).

Neoantigen depletion score

Neoepitopes were predicted from nonsynonymous SNVs in a background set of ovarian tumors consisting of 62 primary HGSC tumors from the Australian Ovarian Cancer Study (Patch et al., 2015) and 59 additional HGSC tumors (Wang et al., 2017), following the methods described above. Following similar methods to Rooney et al. (2015), the probability of generating at least one overlapping neoepitope from each trinucleotide pattern was determined.

For each considered tumor sample (from the multi-site HGSC cohort), the expected rate of neoepitope-generating SNVs was calculated from the trinucleotide context of synonymous SNVs and the expected rate of nonsynonymous SNVs per synonymous SNV for each trinucleotide pattern. Mathematically, define \bar{N}_s to be the expected number of nonsynonymous SNVs per synonymous SNV with trinucleotide pattern s and \bar{B}_s to be the expected number of neoepitope-generating SNVs per nonsynonymous SNV with pattern s . Then, for a given sample i , define Y_i as the set of synonymous SNVs and N_i the set of nonsynonymous SNVs. We can write:

$$N_{pred,i} = \sum_m^{Y_i} \bar{N}_{s(m)}$$

$$B_{pred,i} = \sum_m^{Y_i} \bar{N}_{s(m)} \bar{B}_{s(m)}$$

where $N_{pred,i}$ and $B_{pred,i}$ are the expected number of nonsynonymous SNVs and neoepitope-generating SNVs in sample i under the null model, respectively. $s(m)$ is the trinucleotide pattern for synonymous SNV m . Denote $B_{obs,i}$ to be the observed number of neoepitope-generating SNVs in i , and $N_{obs,i} = |N_i|$ the observed number of nonsynonymous SNVs in i . We then define the neoantigen depletion score as:

$$E_i = \frac{B_{obs,i}}{\frac{N_{obs,i}}{\frac{B_{pred,i}}{N_{pred,i}}}}$$

Lower values of this score were interpreted as evidence of higher neoantigen depletion.

The within-patient relationship between the response, neoantigen score and the covariate, epithelial CD8+ TIL density was modeled with a Bayesian linear mixed model with patient-specific random intercepts. Samples with fewer than 3 nonsynonymous mutations were excluded. The corresponding R code (using the MCMCglmm R package) was:

```
MCMCglmm(log(observed_neoantigen_ratio/expected_neoantigen_ratio) ~ E_CD8_rescaled, random=~patient_id, data=data, family = "gaussian", nitt = 500000, thin = 500, burnin = 50000, prior = prior)
```

where observed_neoantigen_ratio/expected_neoantigen_ratio corresponds to E_i , epithelial CD8+ TIL density values were rescaled between 0 and 1, the residual covariance prior was set to be relatively uninformative ($V = 1$ and $nu = 0.002$ in R), and likewise for the random effect prior ($V = 1$, $nu = 1$, $alpha.mu = 0$, $alpha.V = 1000$ in R). For the fixed effect coefficient, an uninformative prior with mean 0 and variance 10^{10} was used. Lack of autocorrelation in the MCMC traces was confirmed with autocorr from the coda R package. Posterior densities of parameter estimates were checked to ensure certain assumptions of the model (e.g., fixed effect being Gaussian-distributed) were met. Reported significance values correspond to area under the (right) tail of the posterior distribution of the fixed effect coefficient.

The across-patient relationship was computed similarly, but with no patient-specific intercept term. To compute subclonal- or clonal-specific correlations, observed nonsynonymous mutations (and transitively, neoepitopes) were classified based on the clonal phylogenies inferred in Clonal phylogenies & postprocessing. Similar correlations between subclonal neoantigen depletion and epithelial CD8+ TIL densities were observed using multilevel analysis (intratpatient Spearman's correlation $p = 0.034$ across the cohort and $p = 6.1 \times 10^{-5}$ in patients containing samples with highest epithelial CD8+ TIL densities; all between-patient $p > 0.2$).

Lymphocyte marker expression and HLA LOH

CD3D, CD8A, and CD8B expression values was extracted from Nanostring expression data for HGSC cases from Wang et al. (2017) and RNA-seq expression data from OV-AU cases (see RNA-seq analysis). As expression data from few genes was available from the

Nanostring data, expression values were modeled as a function of HLA LOH using the nested ranks test (nestedRanksTest R package; gene expression as the dependent variable, HLA LOH status as the explanatory variable, and cohort as a random effect). *P*-values representing significance of the HLA LOH coefficient are shown in [Figure 3H](#).

The corresponding R code for the nested ranks test is:

```
nestedRanksTest(expression ~ loh_status | cohort, data = data)
```

Histologic image analysis

Cell classification and tissue segmentation

QuPath v.0.1.2 (<https://qupath.github.io/>) was used to detect epithelial tissue and presumptive lymphocytes on hematoxylin & eosin (H&E) pathology slides. Briefly, slides were subjected to superpixel segmentation following automated tissue detection, and intensity features calculated for superpixels. A random trees classifier was trained (by P.T.H.) to distinguish epithelial (tumor) and stromal regions from whitespace and other tissues using small sub-regions from 10 slides on the basis of 145 superpixel features to produce tissue segmentation masks. QuPath's cell detection algorithm was subsequently used to detect individual cells, and an additional random trees classifier trained to distinguish putative immune cells on the basis of 22 cellular features. Trained classifiers are available from the authors.

Hotspot identification

Cell location (coordinate) data for tumor epithelial regions from the classifier were used as input for Getis-Ord G_i^* hotspot detection ([Getis and Ord, 1992](#)). Getis-Ord G_i^* hotspots denote regions with statistically significant clustering of a variable of interest. Getis-Ord G_i^* hotspots were identified for each cell type (cancer and lymphocyte).

To identify hotspots, a grid composed of squares with side length $s = 30$ pixels was first applied to each tissue section image. Only epithelial regions of each image were considered. Grid squares devoid of cells were excluded from further analysis by applying a binary mask. Neighborhood weights were computed using a neighborhood size of $4s$ ([Nawaz et al., 2015](#)). Getis-Ord G_i^* G_i^* values for each grid square i were computed using localG from spdep. For each image, permutation testing (400 random permutations of grid point counts) was applied to compute empirical *P*-values of G_i^* . Regions with associated $p_i < 0.05$ were called as hotspots.

Samples with no identifiable epithelial regions from which to call hotspots were excluded.

Cancer-immune hotspot colocalization

Spatial colocalization between cancer and immune hotspots was computed with the following statistics ([Nawaz et al., 2015](#)):

- f_C = proportion of cancer cell hotspots that are also lymphocyte hotspots
- f_I = proportion of lymphocyte hotspots that are also cancer cell hotspots
- f_{CI} = fractional area of tumor occupied by colocalized cancer-lymphocyte hotspots

General statistical methods

Unless otherwise indicated, correlations between continuous data types were computed using Spearman's correlation coefficient and hierarchical clustering was performed with Ward's method on pairwise Euclidean distances. Sample sizes (*n*) for statistical comparisons are shown in the respective figures and supplemental figures. $p < 0.05$ was considered statistically significant (after adjusting for multiple testing with the BH method). All Dunn's test *P*-values were BH-adjusted. All boxplot whisker ends correspond to Q1 (first quartile) - $1.5 \times \text{IQR}$ (interquartile range) and Q3 + $1.5 \times \text{IQR}$. Sample size estimation was not performed.

DATA AND SOFTWARE AVAILABILITY

Raw WGSS, TCR/BCR-seq, and Nanostring data have been deposited in the European Genome-phenome Archive, which is hosted by the EBI and CRG, under accession number EGAS00001002839. Processed TCR and BCR clonotype calls and WGSS variant calls can be obtained from the Lead Contact. Requests for additional data and code should be directed to the Lead Contact.

Supplemental Figures

Cell

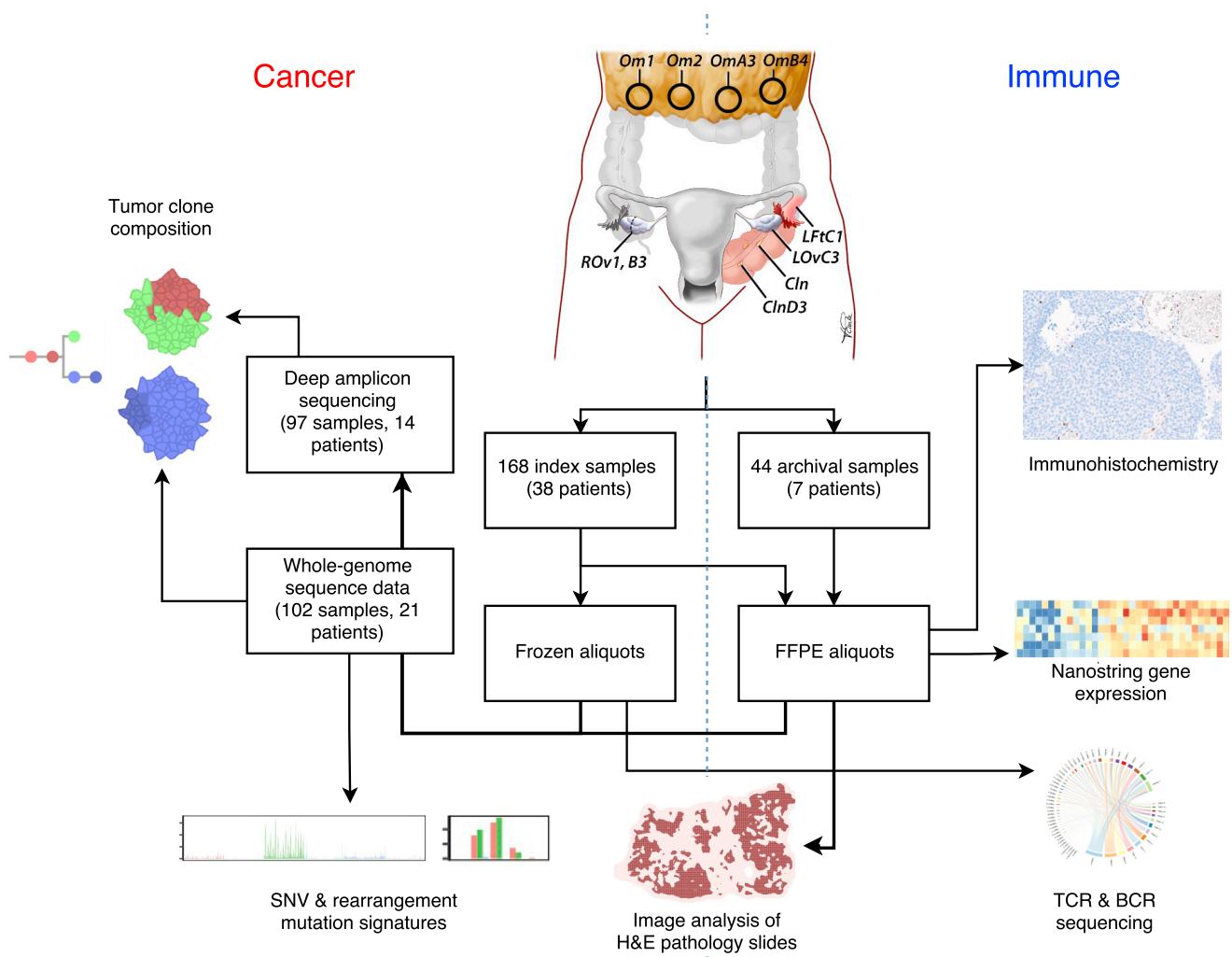


Figure S1. Schematic Diagram Depicting Sample Collection, Experimental Modalities, and Analysis Workflows Applied to the Data, Related to Figure 1

Index sample = cryopreserved; archival = formalin-fixed, paraffin-embedded (FFPE).

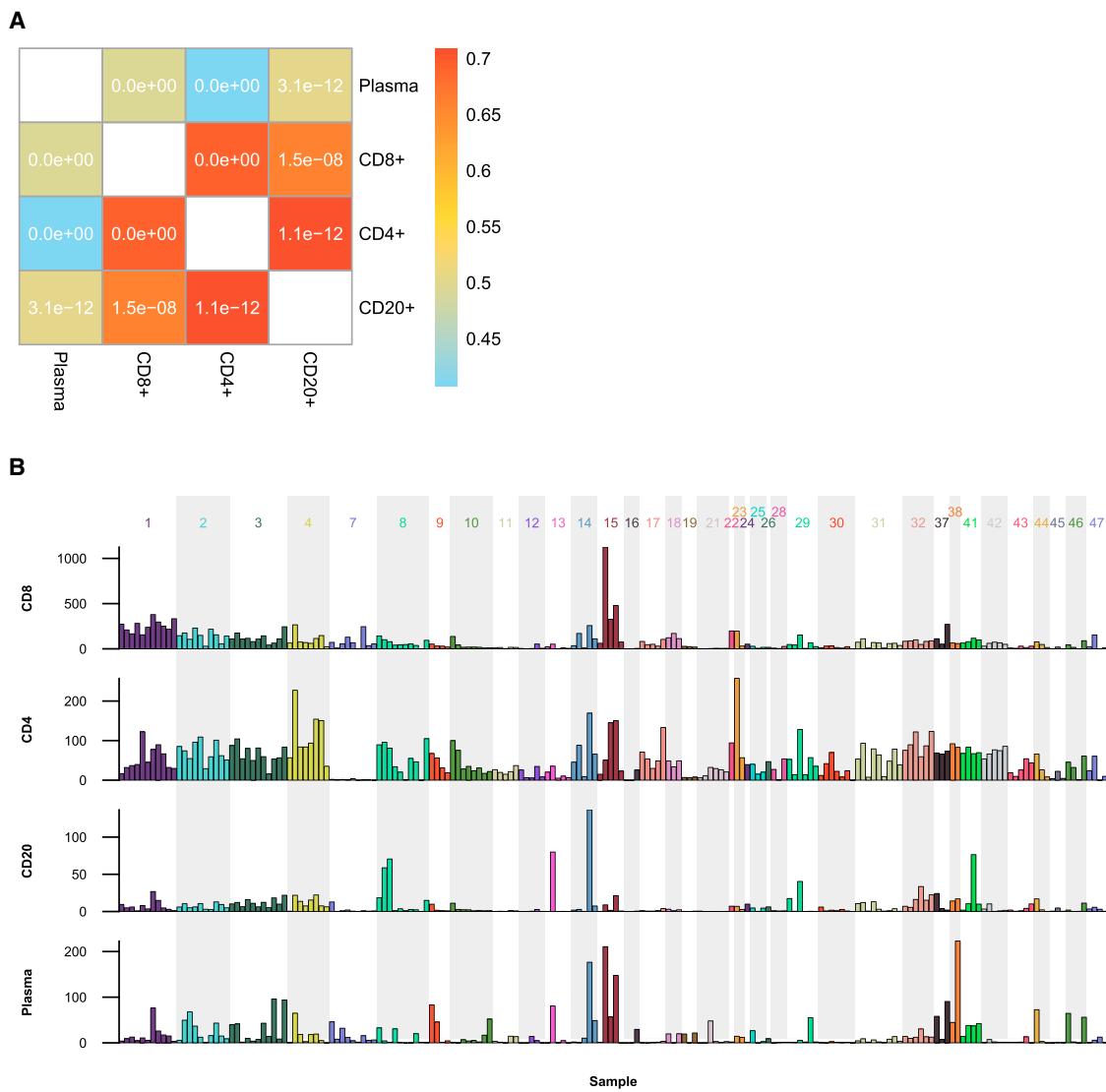


Figure S2. TIL Densities in Multisite HGSC, Related to Figure 1

(A) Correlations between overall TIL densities. Color indicates Spearman's ρ , P -values shown inside each cell.
 (B) Overall CD8+, CD4+, CD20+, and plasma cell densities across the cohort. Bars colored by patient.

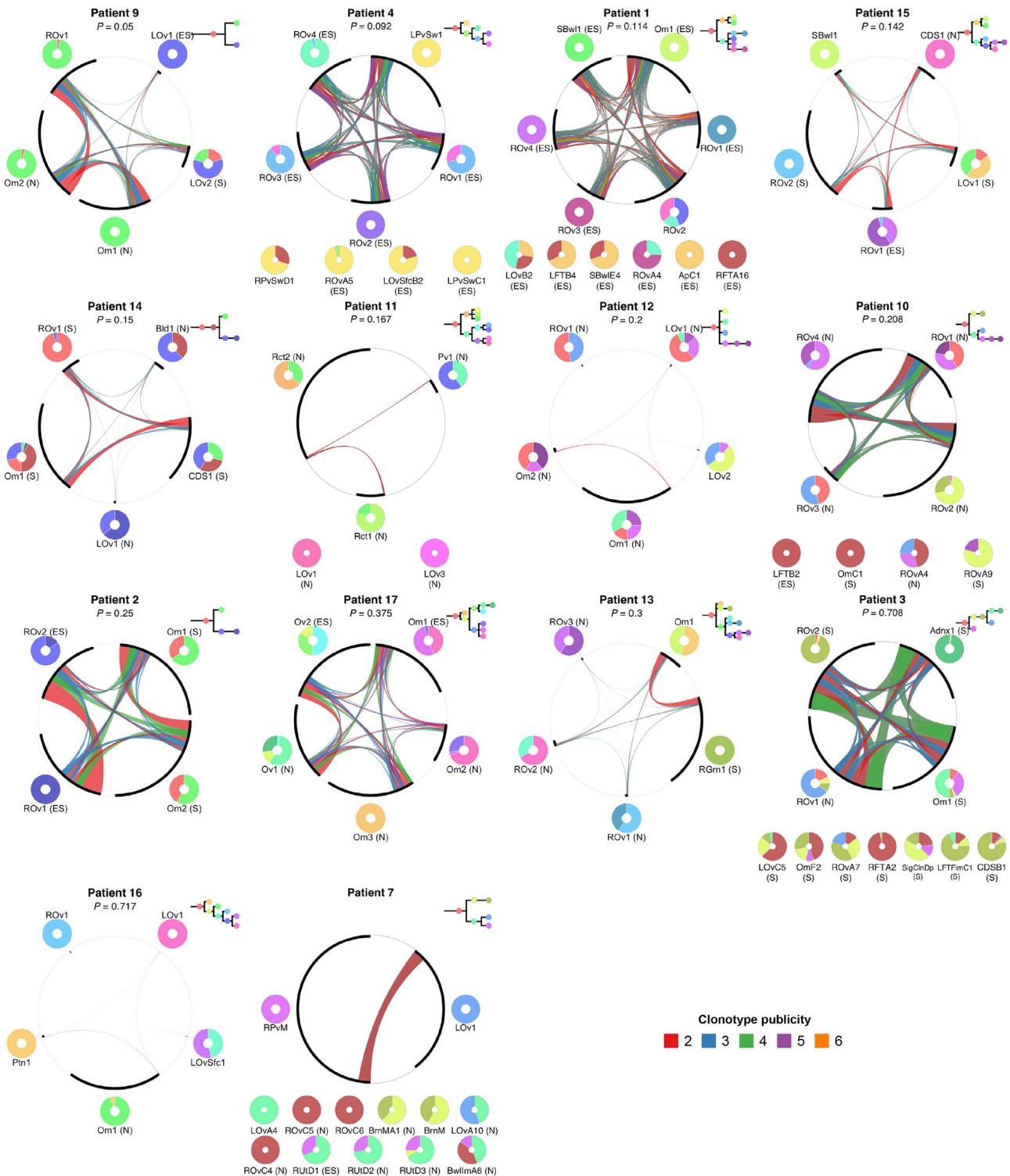


Figure S3. Relationship between Malignant Clone Composition and BCR Clonotype Repertoires, Related to Figure 5

Patients are ordered by significance of the association between BCR repertoire and clonal composition dissimilarities. Chords denote shared clonotypes, width proportional to clonotype count, colored by publicity (number of samples containing a clonotype). Shared clonotypes: publicity ≥ 2 , private clonotypes: publicity = 1. Arc length (along circumference) is proportional to total clonotype count. Tumor clone composition and phylogenies shown external to each circle. Samples without BCR-seq data shown separately below each circle. TIL subtypes indicated by N (N-TIL), S (S-TIL), and ES (ES-TIL) labels. Uncorrected Mantel's test P -values between BCR repertoire dissimilarity and clonal dissimilarity shown below patient labels.

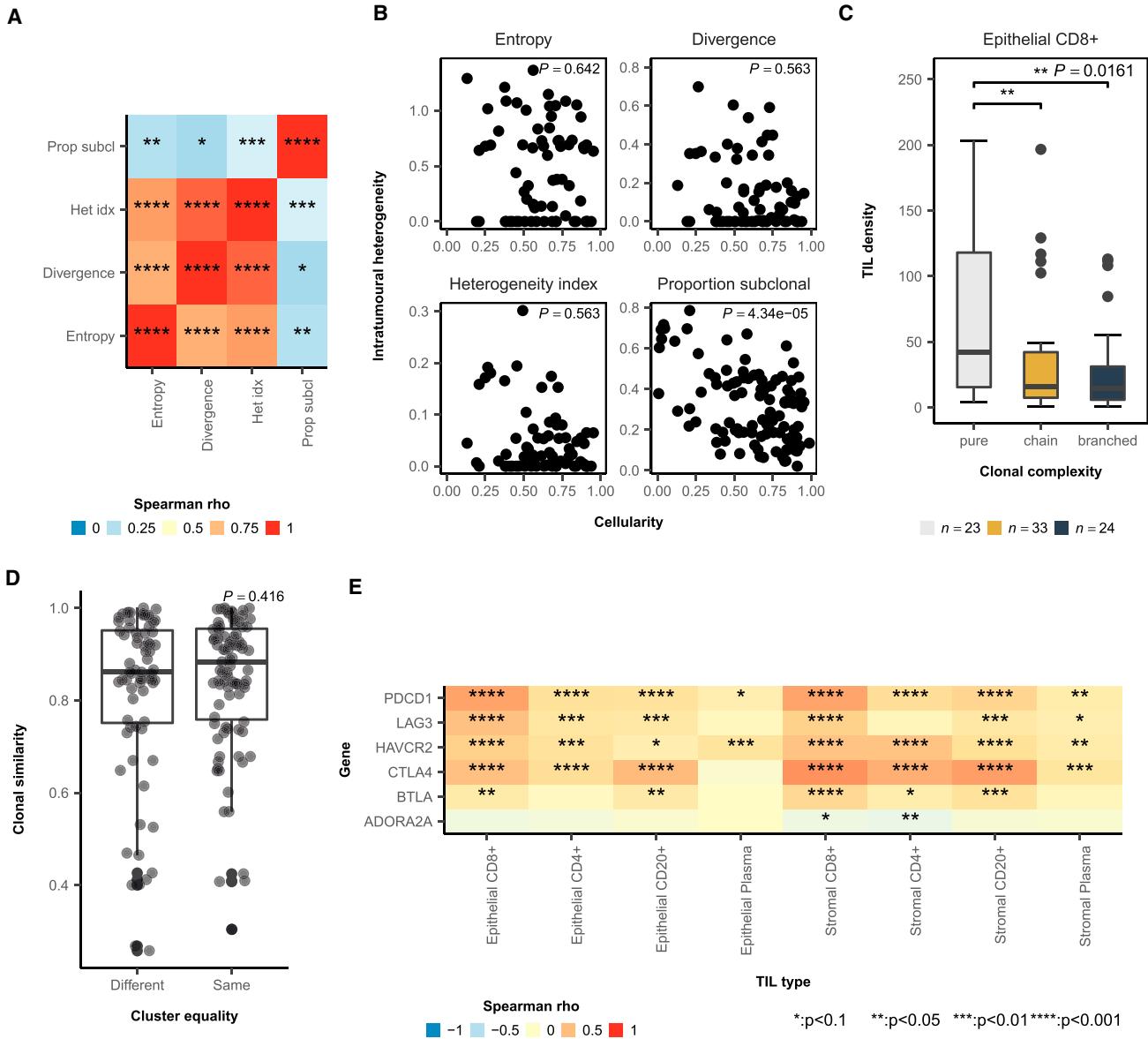
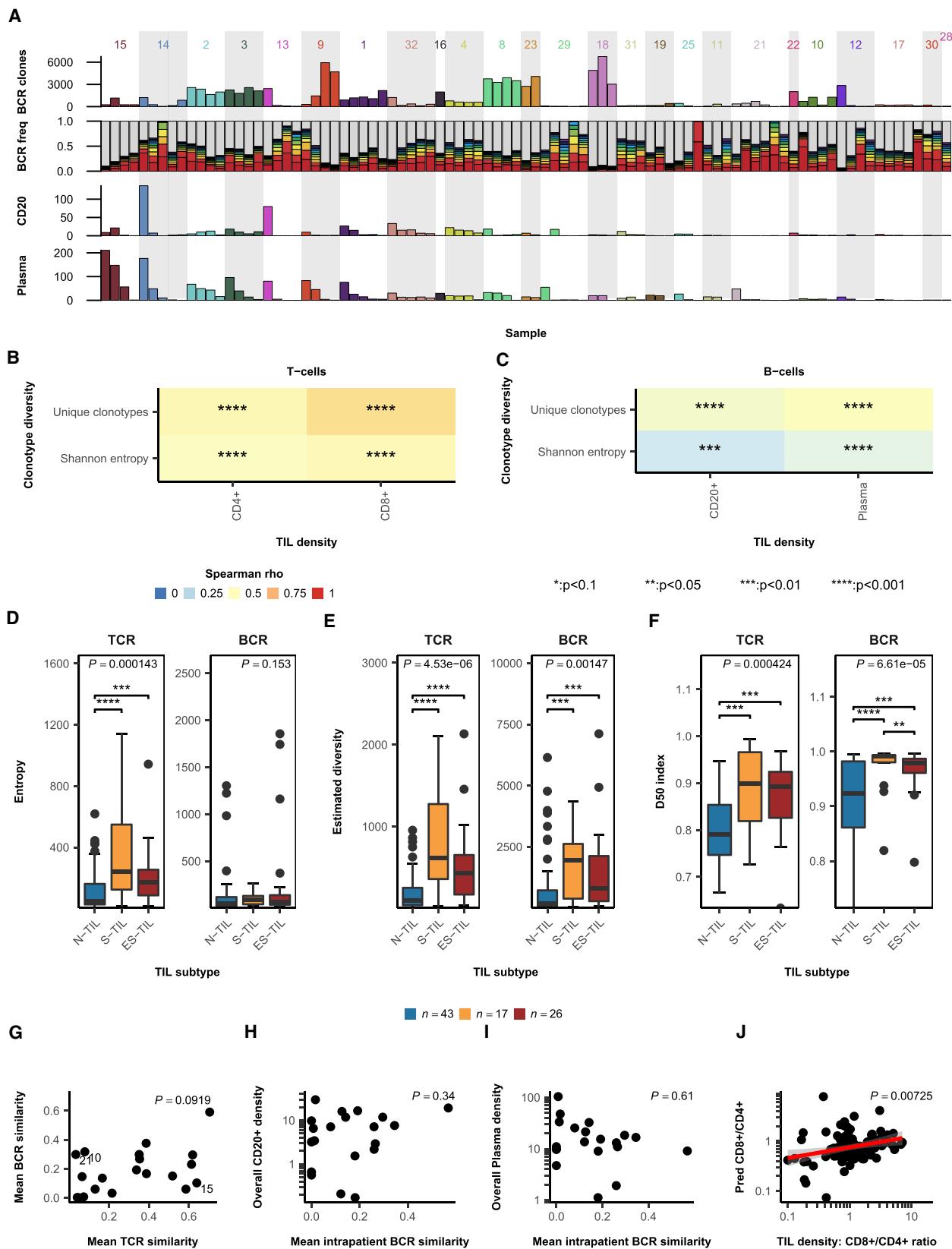


Figure S4. Patterns of Clonal Complexity, Relationship to TIL Subtypes, and Expression of Inhibitory Immune Checkpoint Molecules, Related to Figure 3

- (A) Correlations between ITH measures. Asterisks indicate significance of Spearman's correlation (legend shown in D).
- (B) Correlations between tumor cellularity and ITH. P-values of Spearman's correlation are shown.
- (C) Epithelial CD8+ TIL densities for pre-treatment samples, stratified by clonal mixture type. P-value from the Kruskal-Wallis test shown. Whisker ends correspond to Q1 - 1.5*IQR and Q3 + 1.5*IQR. Significance of post hoc Dunn's test shown (legend in D).
- (D) Degree of similarity in tumor clone composition for pre-treatment samples with different or identical TIL subtypes. Subtype comparisons were made within patients; mean similarity across all comparisons was used. Lines connect comparisons made within the same patient. Nested ranks test P-value is shown. Whisker ends correspond to Q1 - 1.5*IQR and Q3 + 1.5*IQR.
- (E) Correlation matrix between TIL densities and Nanostring-derived expression of inhibitory immune checkpoint genes. Asterisks indicate significance of Spearman's correlation.



(legend on next page)

Figure S5. Correlations between TIL Densities/Subtypes and TCR/BCR-Seq Data, Related to Figure 1 and Figure 4

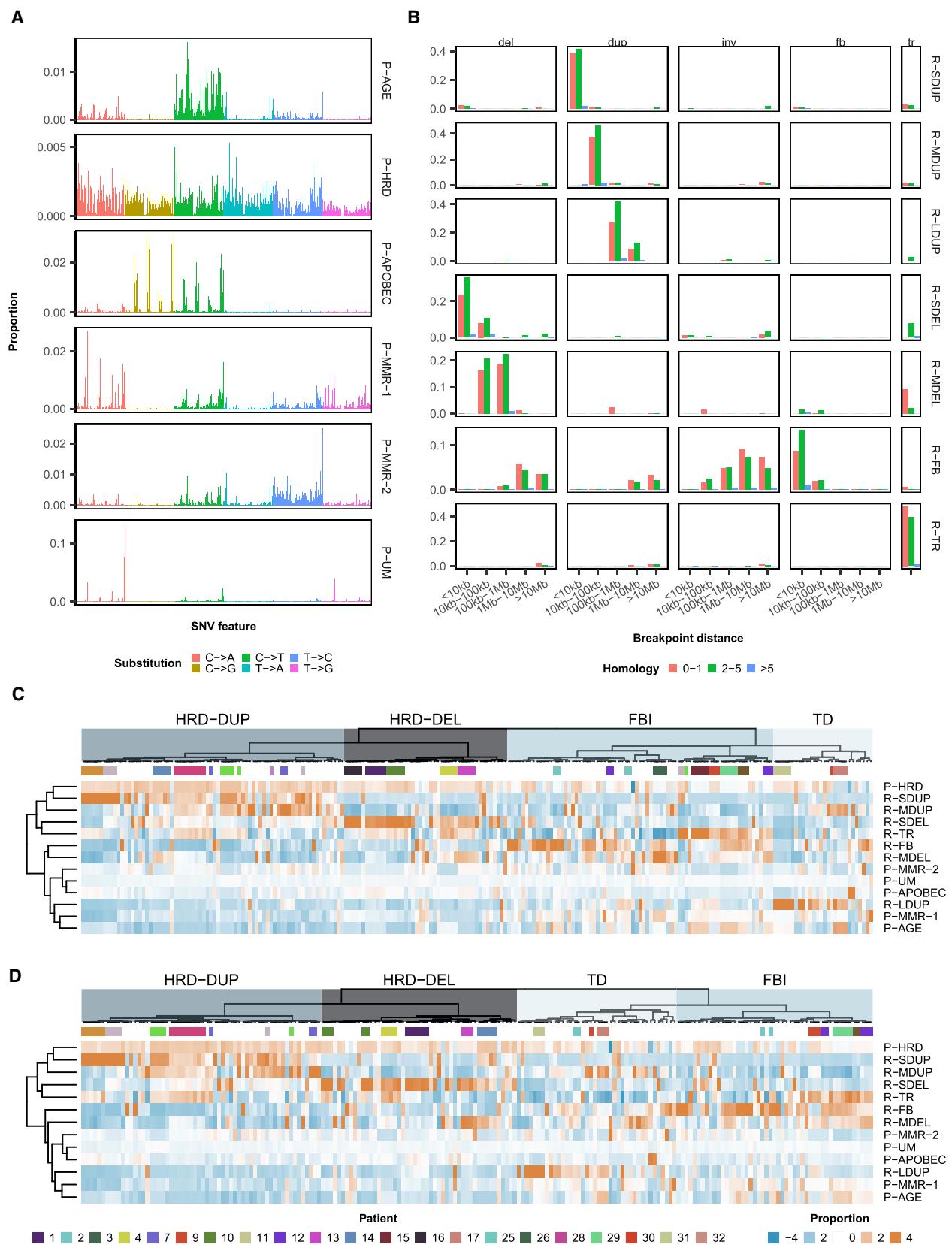
(A) Unique BCR clonotype count, relative frequencies of the top 10 BCR clonotypes (gray = all other clonotypes), CD20+ and plasma TIL density for each sample. (B and C) Correlations between (B) overall CD8+ and CD4+ densities and TCR diversity; (C) overall CD20+ and plasma densities and BCR diversity. Diversity was quantified as the (1) number of unique clonotypes and (2) the entropy of the clonotype abundance distribution. BH-adjusted *P*-values of Spearman's ρ are shown. (D–F) TCR and BCR repertoire diversity across TIL subtypes. Diversity was measured by (D) Shannon entropy of clonotype prevalences, (E) Efron-Thisted index, and (F) D50 index. Whisker ends correspond to Q1 - 1.5*IQR and Q3 + 1.5*IQR. *P*-value from Kruskal-Wallis tests shown; asterisks indicate post hoc Dunn's test significance.

(G) Correlation between mean intrapatient TCR and BCR repertoire similarity. Spearman correlation *P*-value is shown.

(H) Correlation between mean intrapatient BCR repertoire similarity and CD20+ TIL density. *P*-value of Spearman ρ is shown.

(I) Correlation between mean intrapatient BCR repertoire similarity and plasma cell density. *P*-value of Spearman ρ is shown.

(J) Consistency between CD8+/CD4+ ratios from immunohistochemistry and from TCR-based prediction. *P*-value of Spearman ρ is shown.



(legend on next page)

Figure S6. Mutation Signatures Inferred from MMCTM, Related to Figure 6

Figure S6. Mutation Signatures Inferred from MMCTM, Related to Figure 6

(A and B) Jointly inferred SNV and rearrangement signature profiles from MMCTM. Point mutation signatures: P-AGE, age associated; P-HRD, homologous recombination deficiency; P-APOBEC, APOBEC associated; P-MMR-1 + P-MMR-2, mismatch-repair associated. P-UM: ultramutator-associated (virtually absent in HGSC). Rearrangement signatures: R-TDUP, tandem duplications; R-SDUP, short duplications; R-MDUP, medium length duplications; R-LDUP, long duplications; R-SDEL, short deletions; R-MDEL, medium length deletions; R-FB, foldback inversions; R-TR, translocations. Pentanucleotide contexts are shown for each SNV signature and relative prevalences of deletions, duplications, inversions, foldback inversions, and translocations are shown for each rearrangement signature. For rearrangements, microhomology length is labeled.

(C) Standardized proportions of each mutation signature for multisite HGSC, OV-AU, and HGSC Wang et al. (2017) samples, showing clustering of samples from the same patient.

(D) Standardized proportions of each mutation signature for multisite HGSC, OV-AU, and HGSC Wang et al. (2017) samples, where only non-ancestral mutations were considered for the multisite HGSC cohort ([STAR Methods](#)). Heatmap values were clipped between -4 and 4.

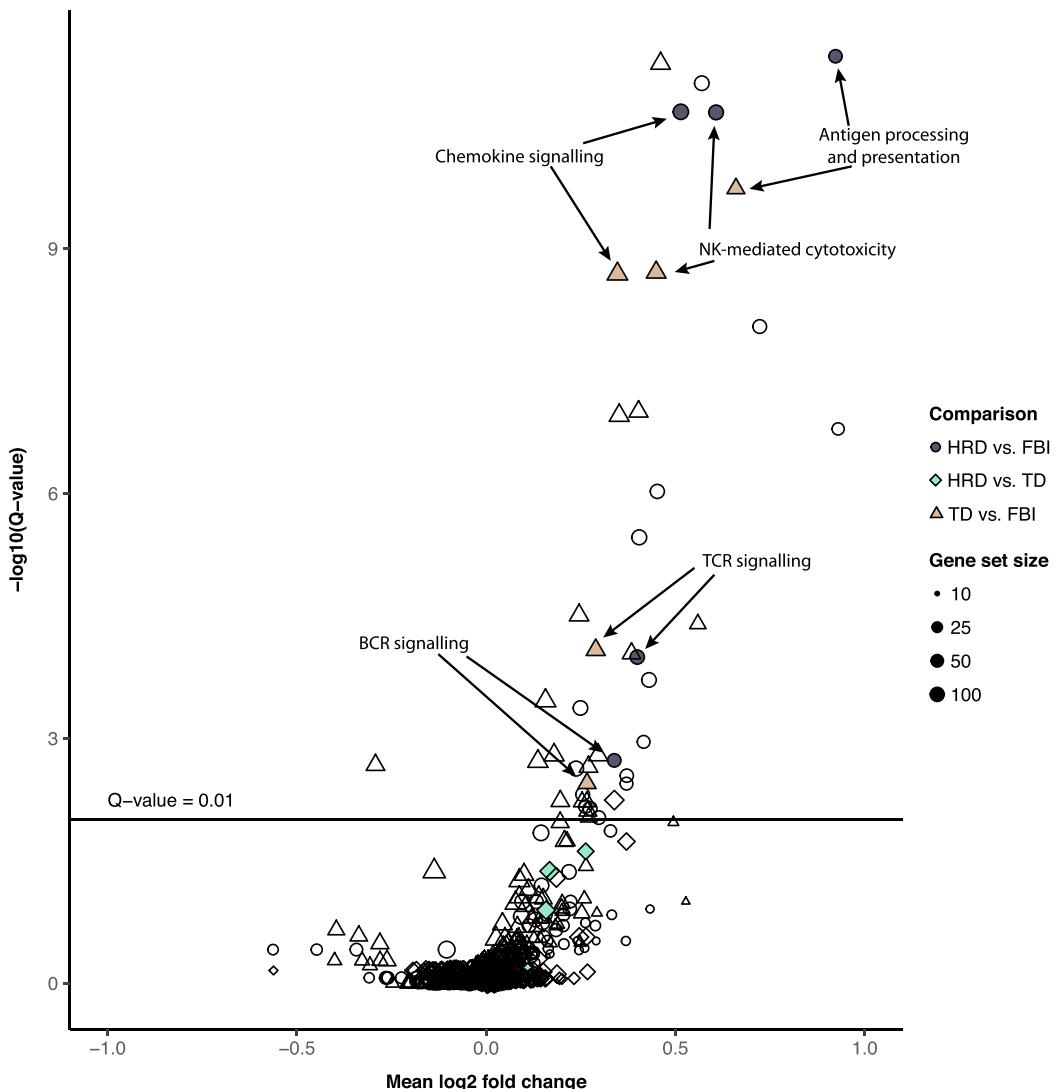


Figure S7. Differentially Expressed Pathways between Mutational Subtypes (for HRD versus FBI, TD versus FBI, and HRD versus TD Comparisons) in OV-AU Cases, Related to Figure 6

Pathway annotations derived from KEGG; enrichment computation described in [STAR Methods](#). Fold change indicated as, e.g., HRD versus TD: mean log₂ fold change in expression from HRD to TD (> 0 = higher in HRD). Q-values computed with the BH procedure; Q < 0.01 was considered significant. Selected immunologic pathways are highlighted; significant hits are additionally labeled.