The Differentiated Signature The Differentiated expression signature correlated with gene sets such as EPITHELIAL_CELL_DIFFERENTIATION, suggestive of a terminally differentiated phenotype, (Supplemental Figure 5; Supplemental Table 8), and MYC activation associated gene sets such as ODONNELL_TARGETS_OF_MYC_AND_TFRC and LEE_LIVER_CANCER_MYC_TGFA_UP. Expression profiles of eight normal whole Fallopian tube samples showed distinctive similarity to the Differentiated signature (Supplemental Figure 5). We used estimates of purity and ploidy provided by the ABSOLUT algorithm (Carter SL et al, submitted) and found that 58 of 100 samples with highest Differentiated signature score to have a diploid genome, versus 28 of the 100 samples with lowest Differentiated gene set score.

Ploidy, age, background mutation rate and expression of the Differentiated signature in the Fallopian tube suggests an alternative path of development for patients with the Differentiated phenotype, possible related to cell of origin that is found in the Fallopian tube. The Fallopian tube has been suggested as cell of origin for a significant fraction of HGS-OVCA (Levanon et al., 2008).

The Immunoreactive Subtype. The Immunoreactive expression signature was significantly higher expressed in samples with an increased presence of intratumoral T-lymphocytes. This is exemplified by marker MSigDB gene sets such as REACTOME_SIGNALING_IN_IMMUNE_SYSTEM and KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY (Supplemental Figure 5; Supplemental Table 8). Inflammatory response is associated with pathogen infection in several cancers, but a survey of the whole sequence ovarian carcinoma data available for 13 TCGA cases found no association between specific non-human reads and subtype that might suggest a infectious origin for the Immunoreactive group (data not shown) (Kostic et al., 2011).

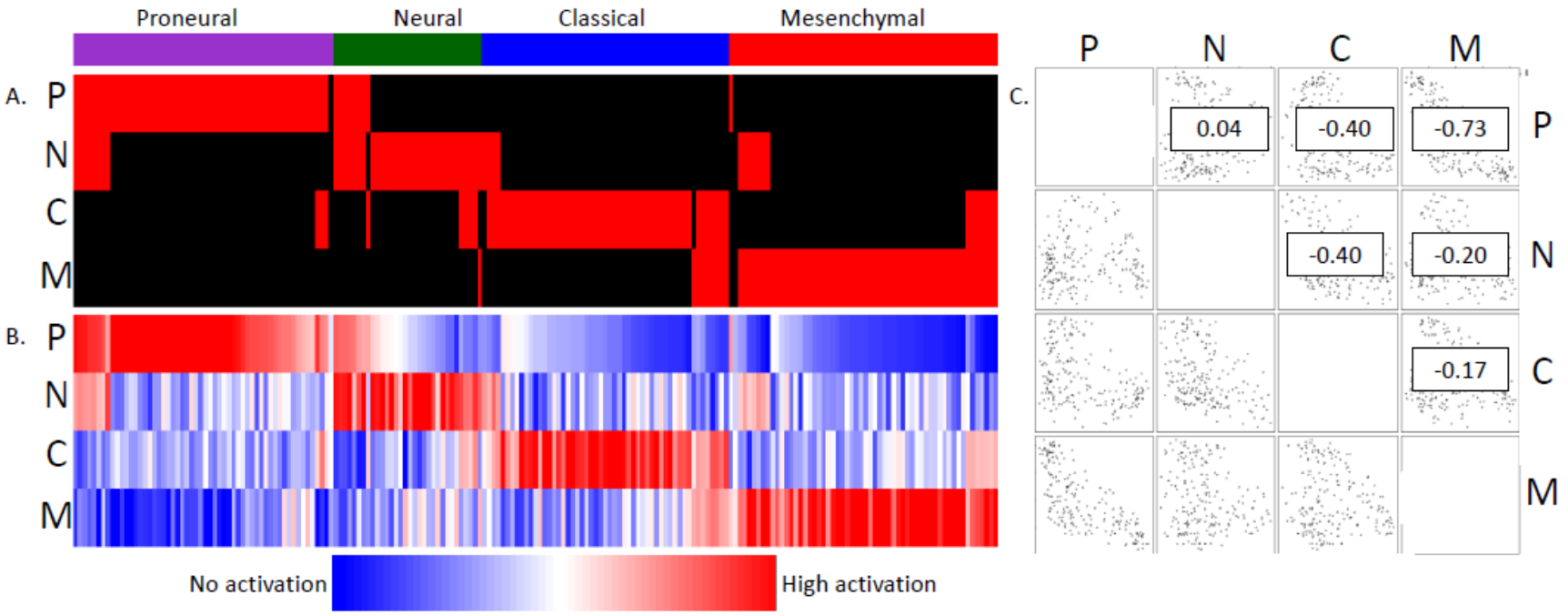
The Mesenchymal Subtype Expression of genes associated with extracellular matrix and proteolysis, and MSigDb gene sets such as INTEGRIN_BINDING and EXTRACELLULAR_MATRIX characterized the Mesenchymal signature (Supplemental Figure 5; Supplemental Table 8). Samples with high expression of

the Mesenchymal signature showed significantly lower tumor purity (Supplemental Figure 5). To investigate whether the high activation of the Mesenchymal signature was predominantly the result of stromal infiltration we calculated ovarian cancer signature scores for previously reported expression profiles of tumor samples before and after laser capture microdissection (LCM) (Tothill et al., 2008) (Supplemental Figure 6). All five profiles showed a significant decrease in Immunoreactive and Mesenchymal signature scores after LCM, and a borderline significant increase in Differentiated and Proliferative gene set scores (Supplemental Figure 7). The Mesenchymal signature resembled the expression profile of normal whole ovary (Supplemental Figure 7) which contains a high percentage of stromal cells.

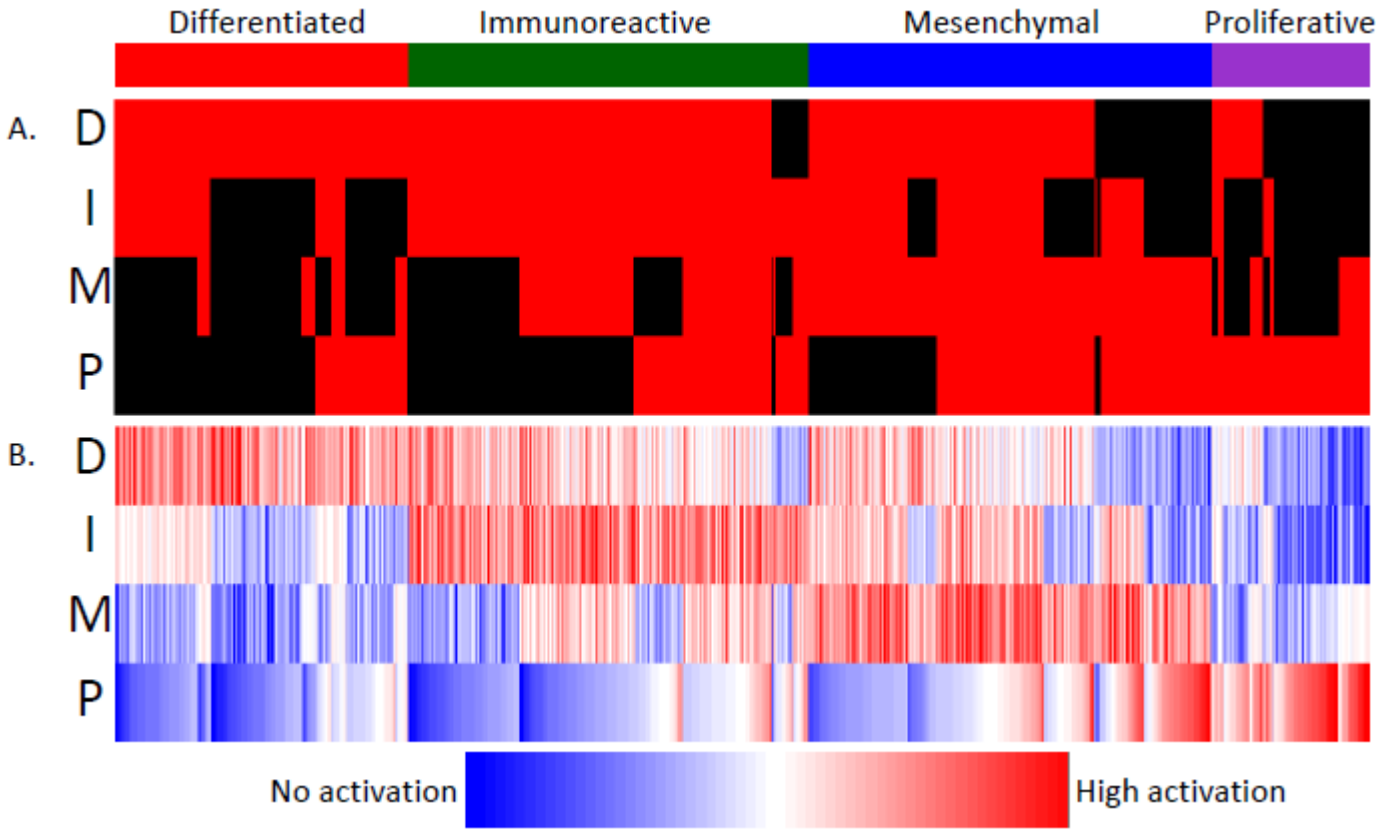
The Proliferative Subtype Tumors classified as Proliferative were less likely to harbor *BRCA1* or *BRCA2* germline mutations, but not *BRCA1* or *BRCA2* somatic mutations (Supplemental Table 9). A decrease in the frequency of *BRCA1* methylated samples was found in the Proliferative subtype (7%, n=138, versus 13%, n=351, Fisher's Exact test p-value = 0.08). Interestingly, the Proliferative subtype was associated with high expression of *FANC* genes. FANC proteins co-localize with BRCA1 to repair DNA double-strand breaks through homologous recombination (Folias et al., 2002). Gene sets such as CHROMATIN_REMODELING, CHROMOSOME_ORGANIZATION_AND_BIOGENESIS and E2F3_UP show increased activation in the Proliferative tumors, altogether suggesting that the Proliferative class undergoes a high degree of homologous recombination (Supplemental Figure 5; Supplemental Table 8)(Turner et al., 2004).

Ovarian Cancer Cell lines We calculated gene set scores for the four ovarian cancer signatures and twenty eight expression profiles from ovarian carcinoma cell lines available through the Cancer Cell Line project (http://www.broadinstitute.org/ccle and Barretina J et al, submitted). Interestingly, no cell lines showed high activation of the Immunoreactive gene set, adding to the point that the Immunoreactive signature is strongly associated with the tumor microenvironment and therefore no cancer cell lines are

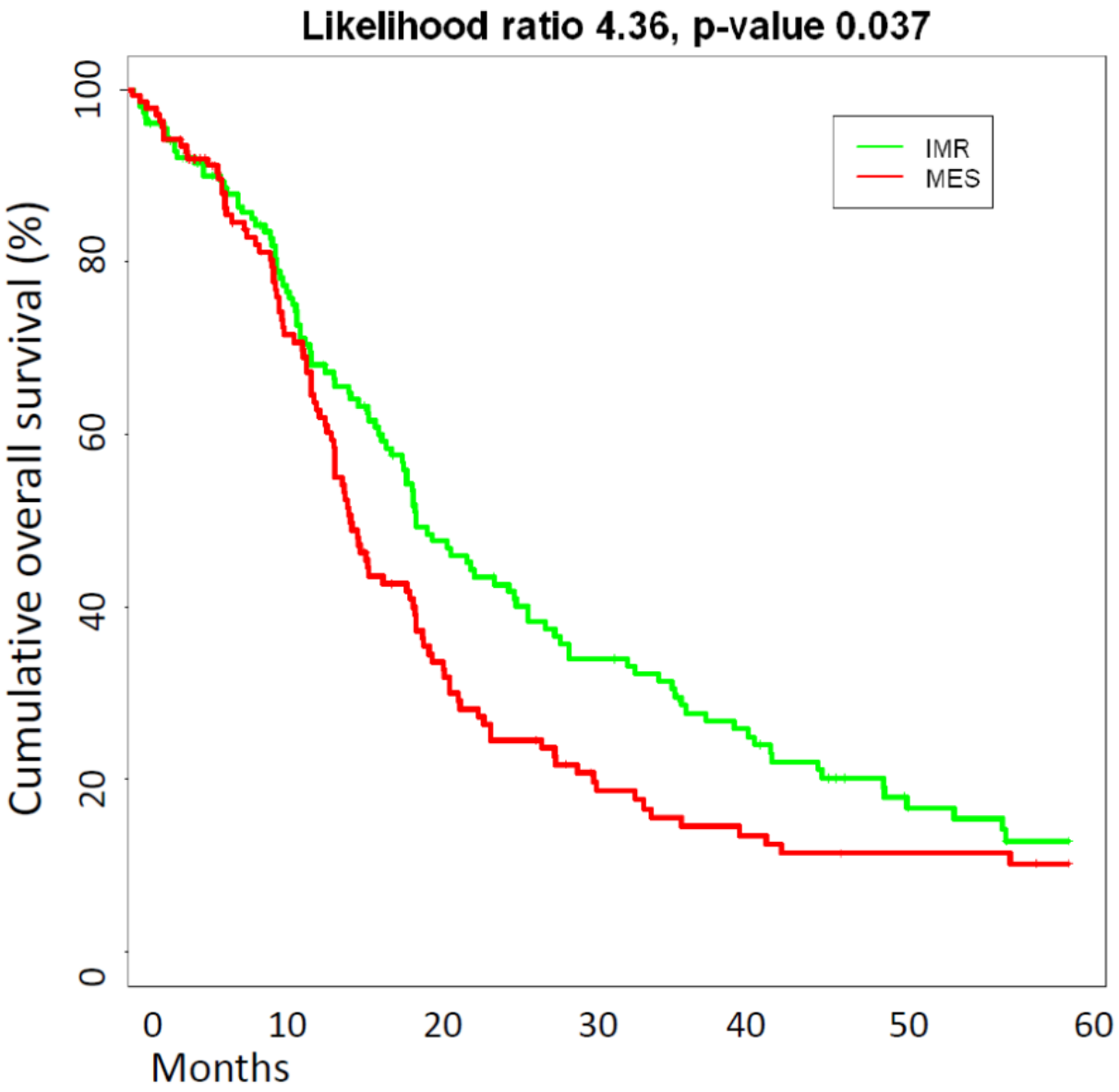
found expressing this gene set (Supplemental Table 10). Two cell lines were found to activate the Mesenchymal signature, a gene set that was also strongly expressed by the infiltrating stroma portion of these tumors (Supplemental Figure 7), suggesting that a mesenchymal signal can be the result of processes in tumor cells as well as infiltrating normal tissue. None of the ovarian cancer cell lines were found to express multiple gene sets.



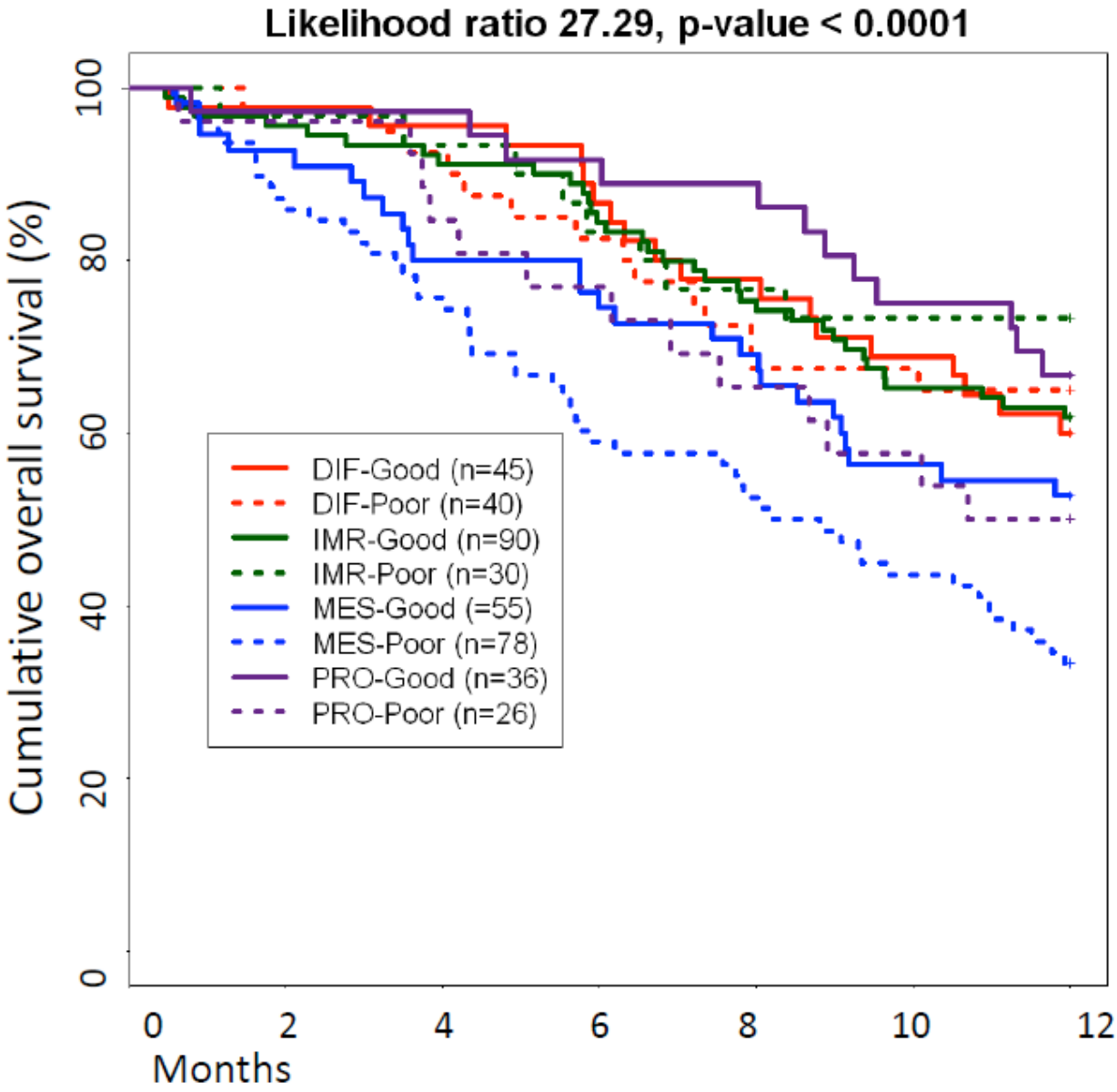
Supplemental Figure 1A. **Glioblastoma signature ssGSEA scores**. Gene set enrichment scores for 199 glioblastoma tumor samples were generated using four gene expression signatures (Verhaak et al., 2010). A. Binary scores indicating whether a tumor sample activates the gene signature. Each column represents one sample; each row represents a gene signature. Activated = red, not activated = black. B. Raw gene set activation scores. Each column represents one sample; each row represents a gene signature. C. Scatter plots for all gene set combinations.



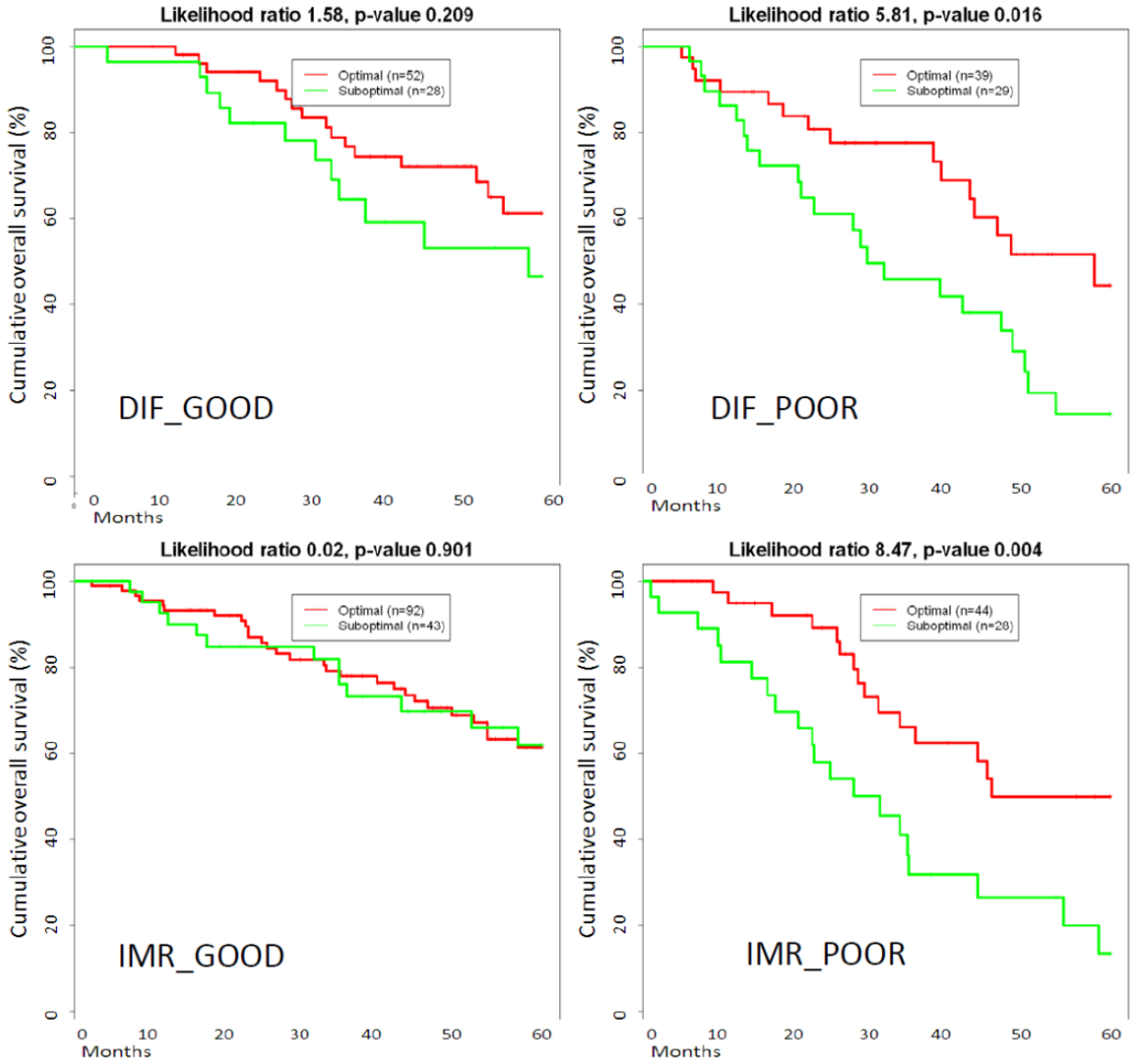
Supplemental Figure 1B. **Serous ovarian carcinoma signature ssGSEA scores**. Gene set enrichment scores for 879 validation set ovarian cancer tumor samples were generated using four gene expression signatures. A. Binary scores indicating whether a tumor sample activates the gene signature. Each column represents one sample; each row represents a gene signature. Activated = red, not activated = black. B. Raw gene set activation scores. Each column represents one sample; each row represents a gene signature. C. Scatter plots for all gene set combinations.



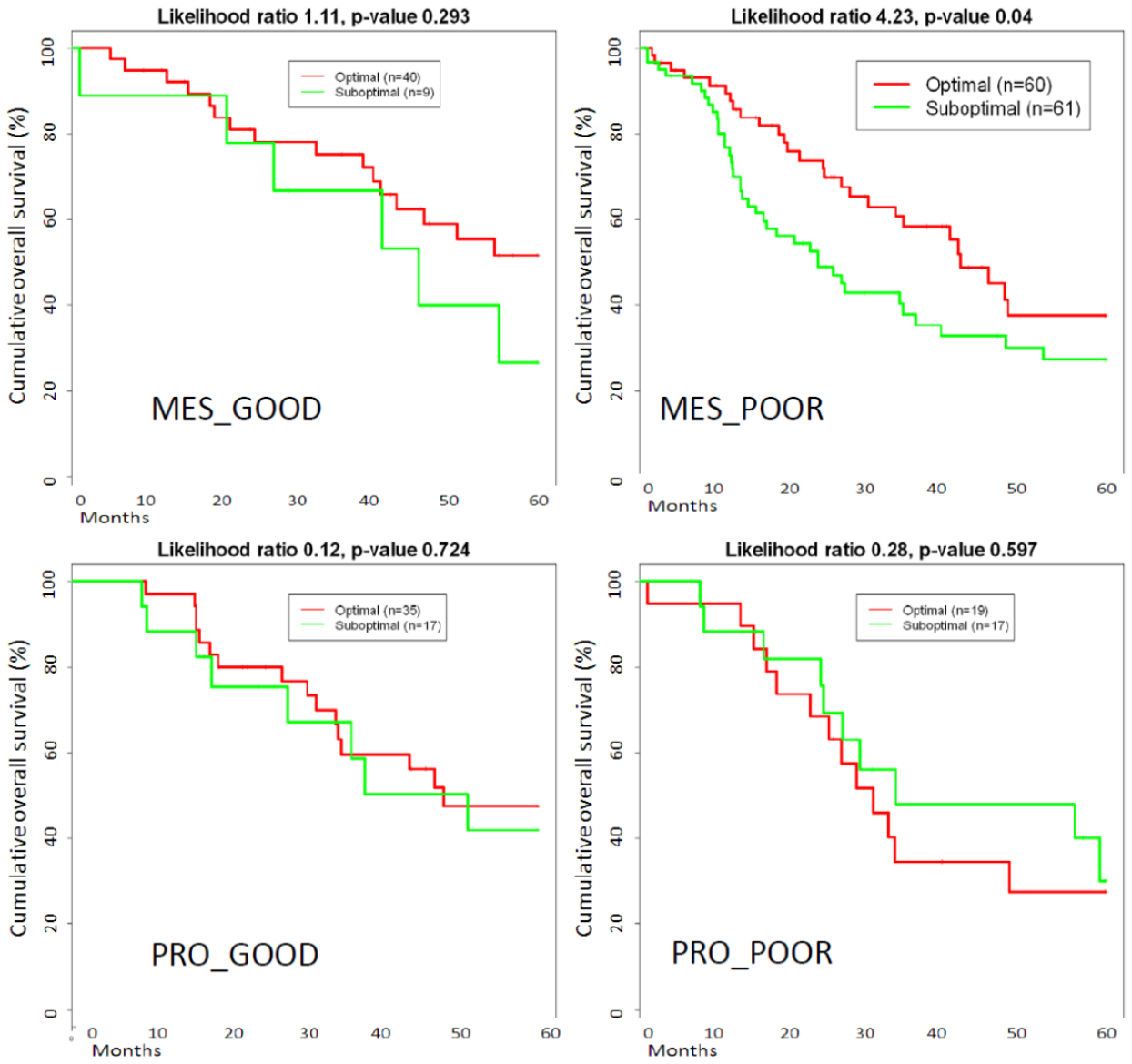
Supplemental Figure 2. **Recurrence free survival per subtype.** Samples in the TCGA cohort were assigned to the Immunoreactive class or the Mesenchymal class based on gene set activation score. The cutoff for including a sample into a group was determined by the minimum score of all samples that clustered with Immunoreactive/Mesenchymal clusters with a positive silhouette width. Time in months between date of diagnosis and date of tumor recurrence.



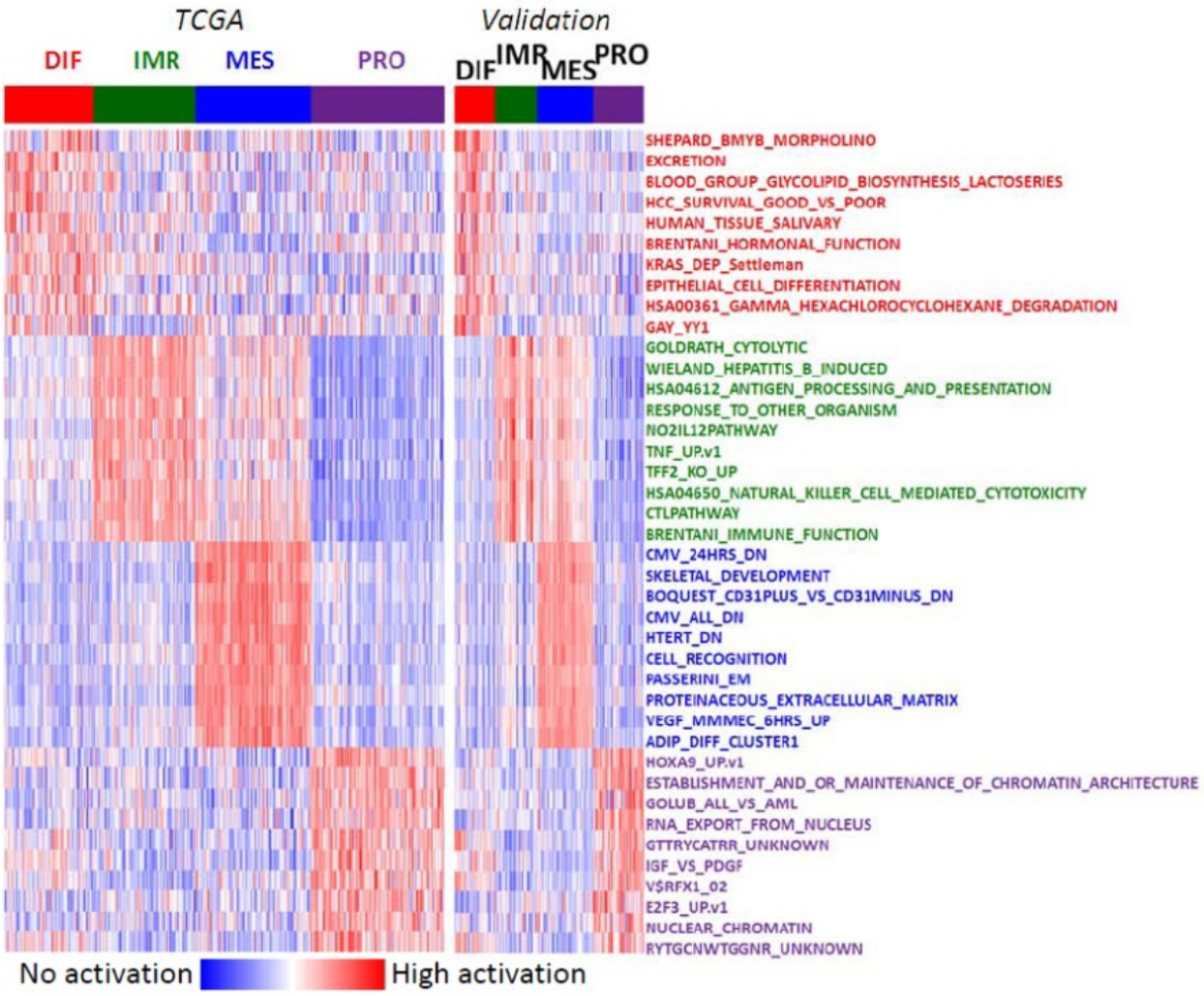
Supplemental Figure 3. **Recurrence free survival per subtype.** Samples from the Tothill et al, Bonome et al and TCGA-validation cohorts (n=400) were assigned according to the CLOVAR classification. Time to recurrence was not available for other samples in the validation cohort. Time in months between date of diagnosis and date of tumor recurrence.



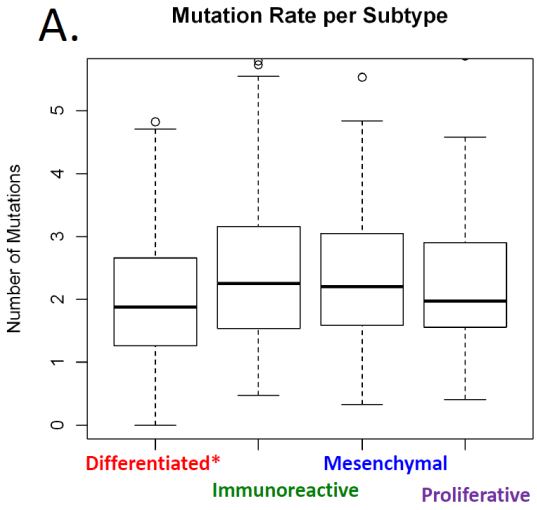
Supplemental Figure 4A. **Debulking status predicts prognosis in some but not all subtypes.** Patients from the validation set were classified into two groups, for each of the Differentiated/Immunoreactive prognostic groups, based on the size of tumor residue after surgery. Less than 1cm is optimally debulked; more than 1cm is suboptimally debulked. Time in months between date of diagnosis and date of last follow up.

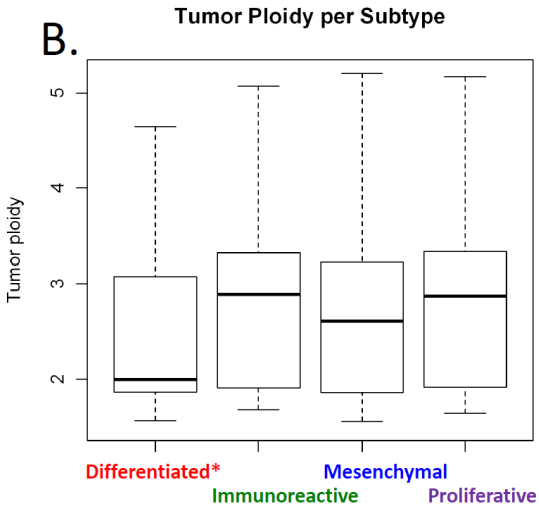


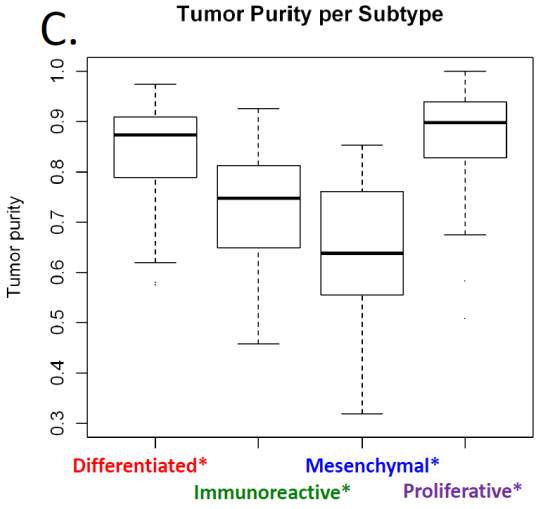
Supplemental Figure 4B. **Debulking status predicts prognosis in some but not all subtypes.** Patients from the validation set were classified into two groups, for each of the Mesenchymal/Proliferative prognostic groups, based on the size of tumor residue after surgery. Less than 1cm is optimally debulked; more than 1cm is suboptimally debulked. Time in months between date of diagnosis and date of last follow up.



Supplemental Figure 5. **Representative gene sets for four gene expression subtypes.** Gene sets were collected from MSigDb (broadinstitute.org/msigdb) and activation scores were calculated using single sample GSEA. Each column represents a sample; each row represents a gene set. The single sample GSEA scores for the ten most discriminating gene sets per subtype are shown for TCGA core samples (n=413, left panel) and Tothill et al core samples (n=206, middle panel).

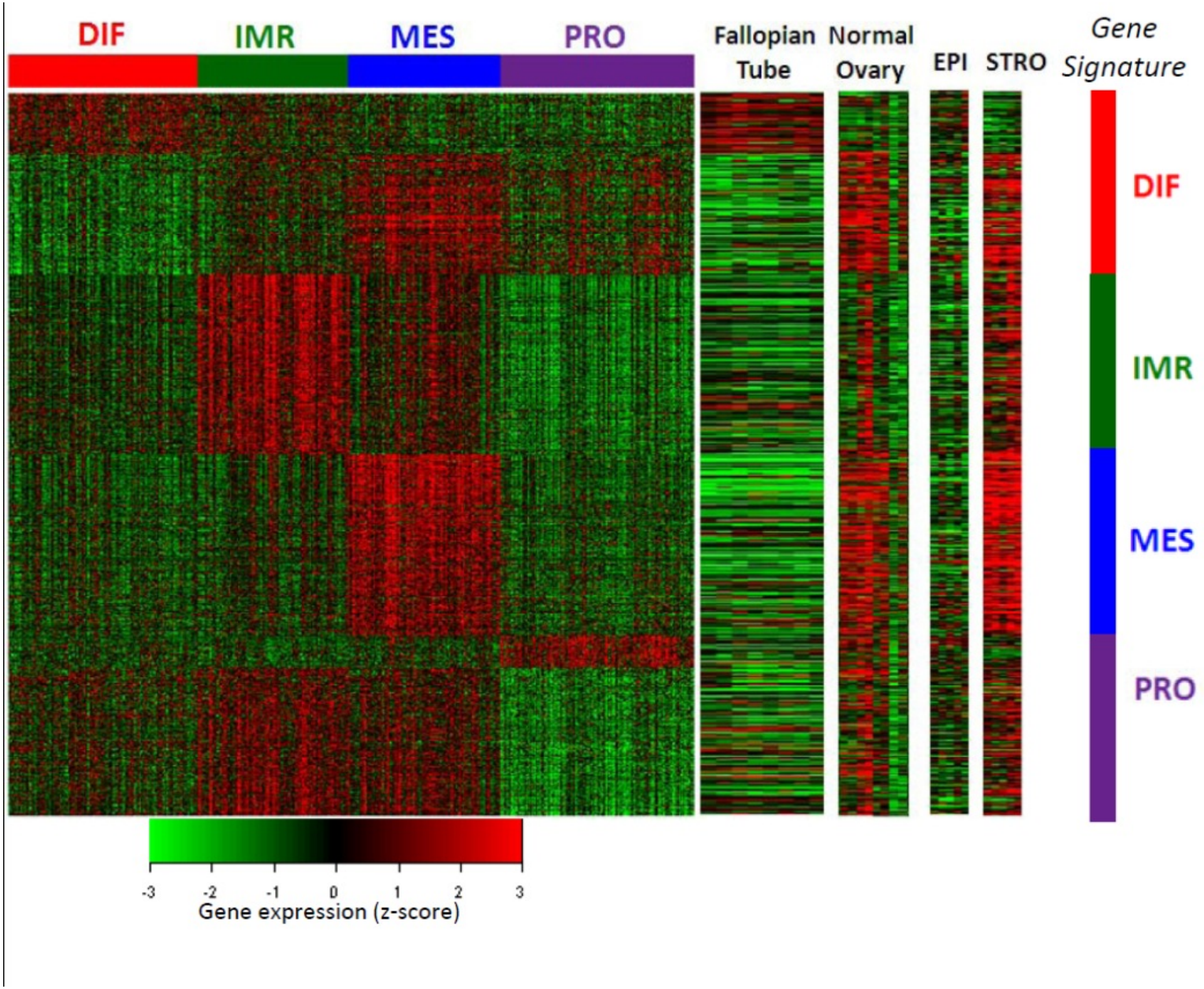

A.
Mutation Rate per Subtype
Number of Mutations
0
1
2
3
4
5
Differentiated*
Immunoreactive
Mesenchymal
Proliferative


B.
Tumor Ploidy per Subtype
Tumor ploidy
2
3
4
5
Differentiated*
Immunoreactive
Mesenchymal
Proliferative


C.
Tumor Purity per Subtype
Tumor purity
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Differentiated*
Immunoreactive*
Mesenchymal*
Proliferative*

* indicates a significant difference between the subtype and other subtypes

Supplemental Figure 6. **Mutation rate, purity and ploidy per subtype.** Mutation rate data was available for core TCGA samples: 93 Differentiated samples, 78 Immunoreactive samples, 63 Mesenchymal samples, 82 Proliferative samples. Purity and ploidy estimates from copy number arrays were available for 70 TCGA Differentiated samples; 79 TCGA Immunoreactive samples; 92 TCGA Mesenchymal samples; 104 TCGA Proliferative samples.


DIF
IMR
MES
PRO
Fallopian Tube
Normal Ovary
EPI
STRO
Gene Signature
DIF
IMR
MES
PRO
-3
-2
-1
0
1
2
3
Gene expression (z-score)

Supplemental Figure 7. **Similarity between normal tissues and Ovarian cancer subtypes.** Using the 200 gene class signatures, samples and genes were ordered based on subtype assignments. Each column represents a sample; each row represents a gene set. Data from nine normal fallopian tube samples, six normal ovary samples, five microdissected tumor cell samples (EPI) and five microdissected stromal cell samples (STRO) was similarly ordered. Expression patterns of Fallopian tube samples were most similar to the Differentiated signature where normal ovary samples were most similar to the Mesenchymal signature. Epithelial cells from the selected Mesenchymal ovarian tumors are somewhat similar to the Proliferative subtype, while the expression pattern of stromal cells from the same tumors strongly match the Mesenchymal signature.