

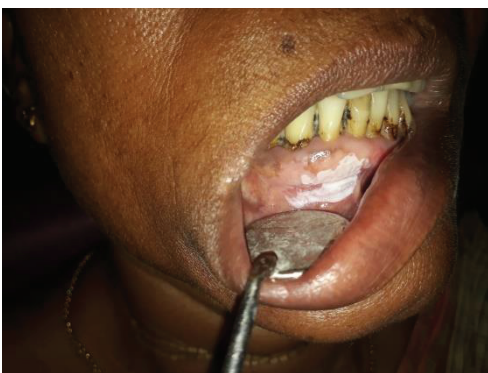
a**b****c****d**

Figure S1: Clinical photographs from randomly selected patients with oral squamous cell carcinoma of the gingivo-buccal region (OSCC-GB) illustrating (a) OSCC involving the buccal mucosa, (b) OSCC of the alveolus, (c) homogeneous leukoplakia, and (d) normal buccal mucosa. The clinician (SG) had obtained consent of the patients for use of these photographs in scientific publications.

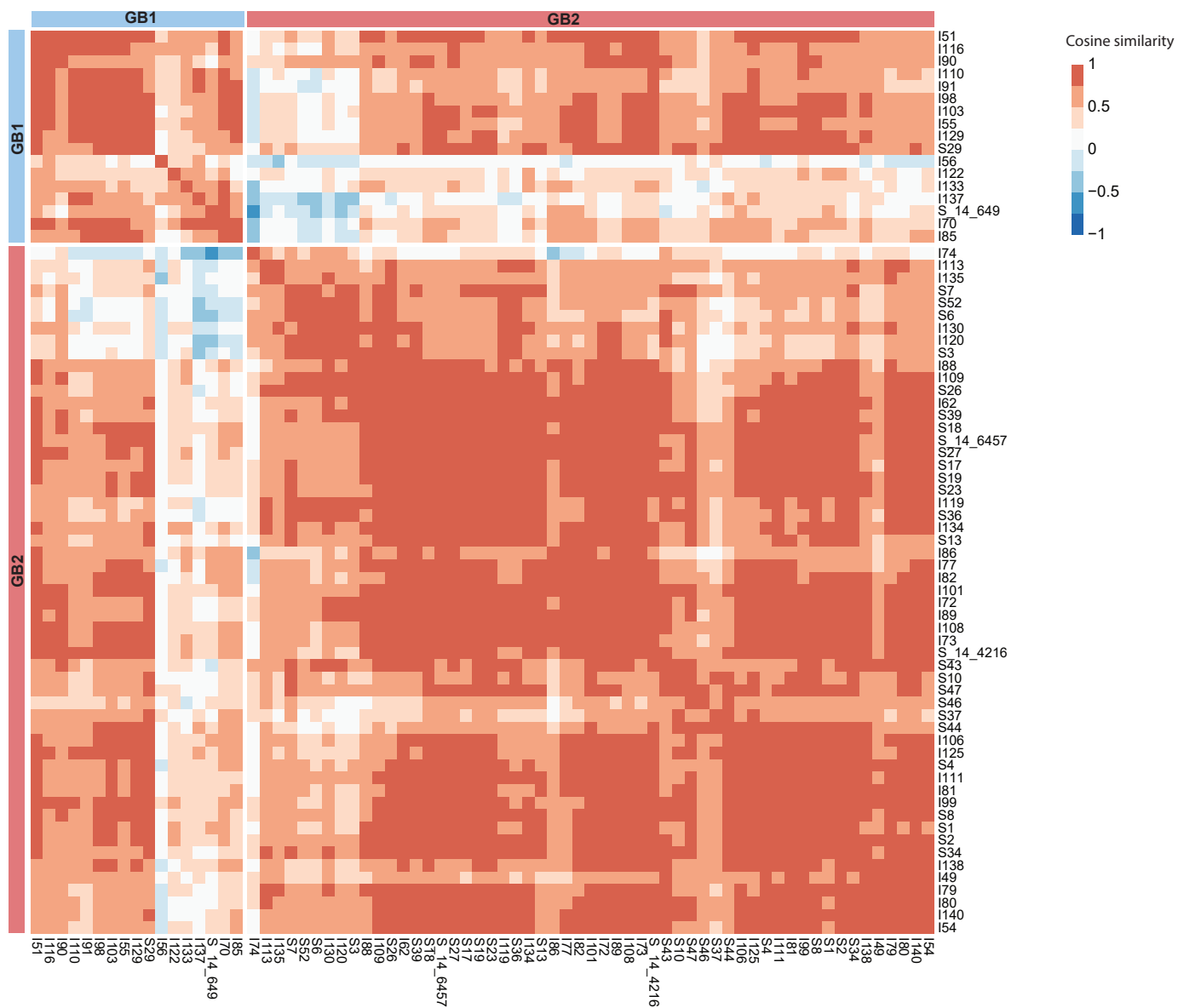


Figure S2: Comparison of GB1 (n=17) and GB2 (n=55) clusters of the OSCC-GB patients. The cosine distance metric was calculated using enrichment scores of 50 hallmark gene sets.



Figure S3: Normalized enrichment scores (NES) of immune-related hallmark gene sets between GB1 and GB2 sets of OSCC-GB patients. p-values were obtained using a 2-tailed Mann-Whitney U test.

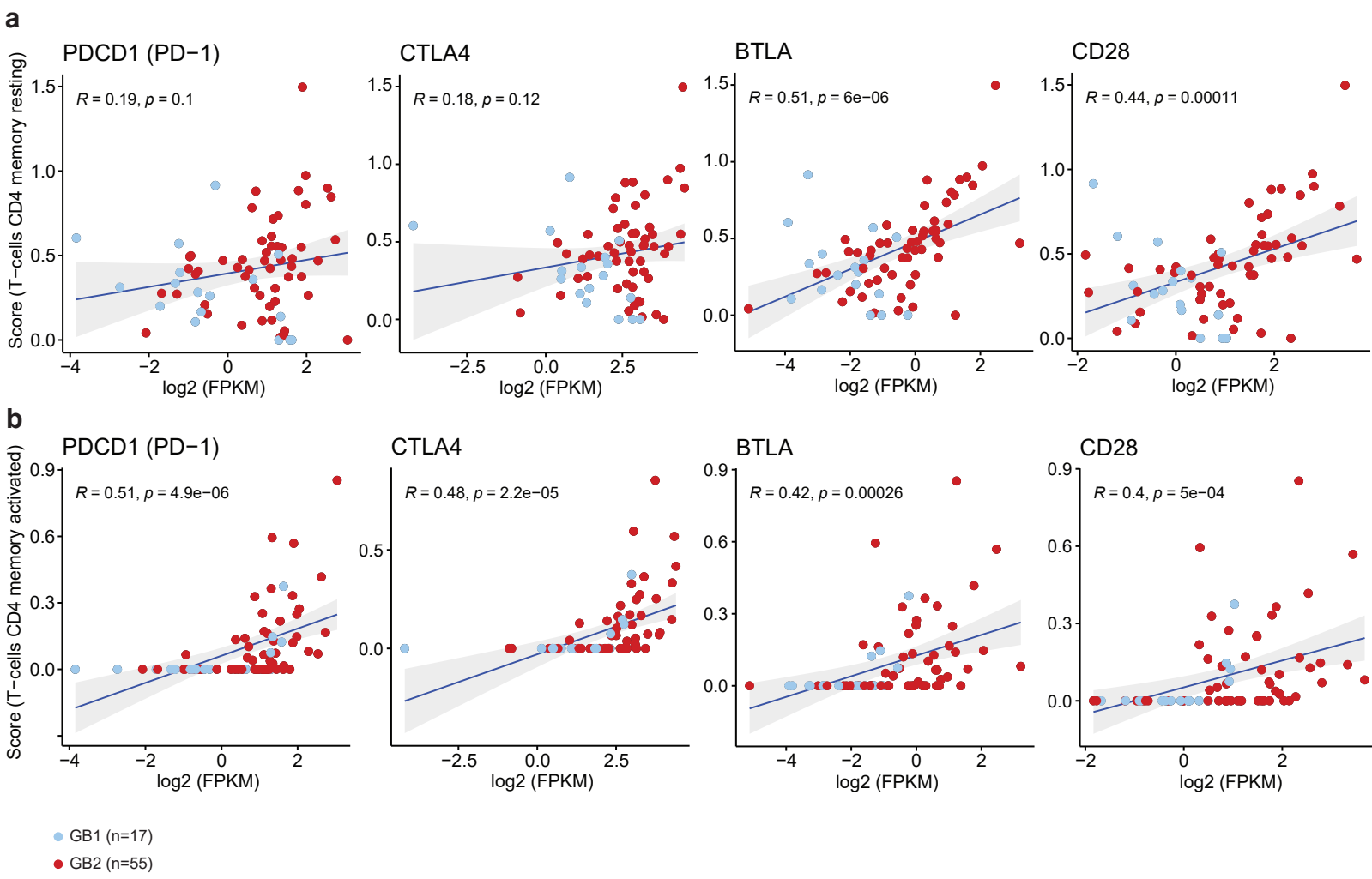
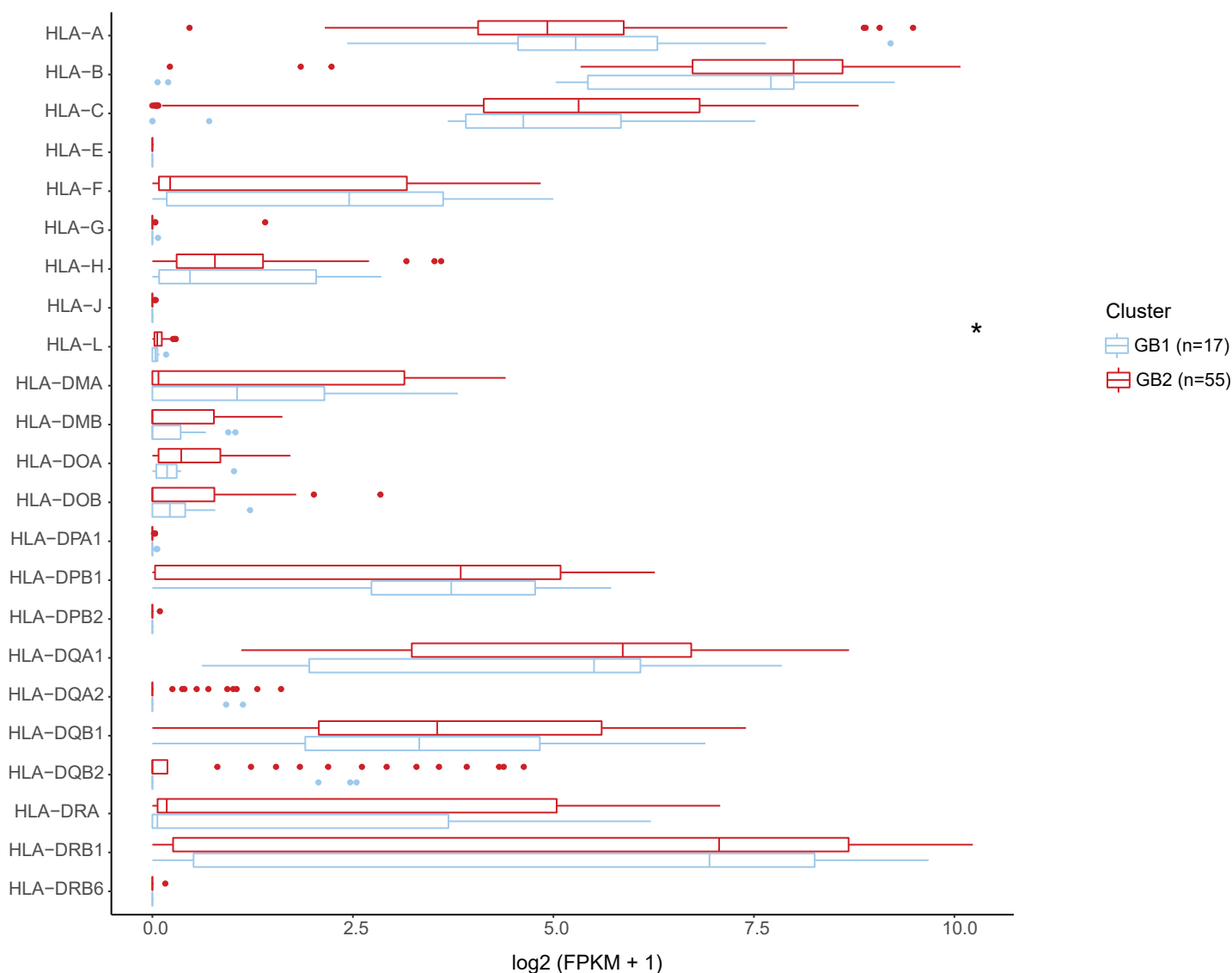


Figure S4: Correlations between abundances of memory resting or activated CD4+ T-cells and expression (log2 FPKM) of selected cell surface marker genes in tumors of OSCC-GB (n=72) patients. R values indicate Pearson's correlation coefficients.



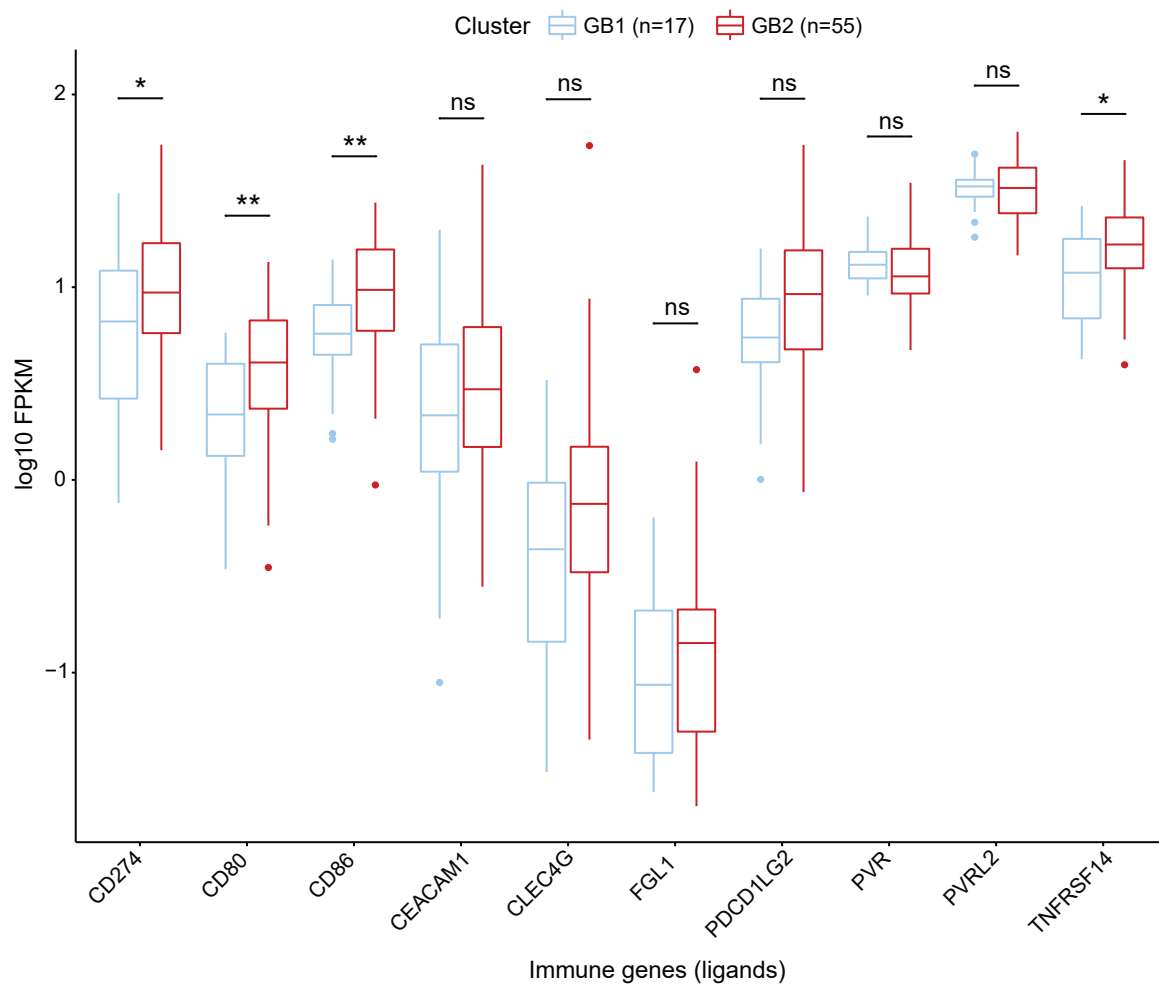


Figure S6: Expression differences of immune checkpoint ligands between the GB clusters. Mann-Whitney U test was performed to find whether the mean expression levels between the clusters were equal. * $p < 0.05$, ** $p < 0.01$.

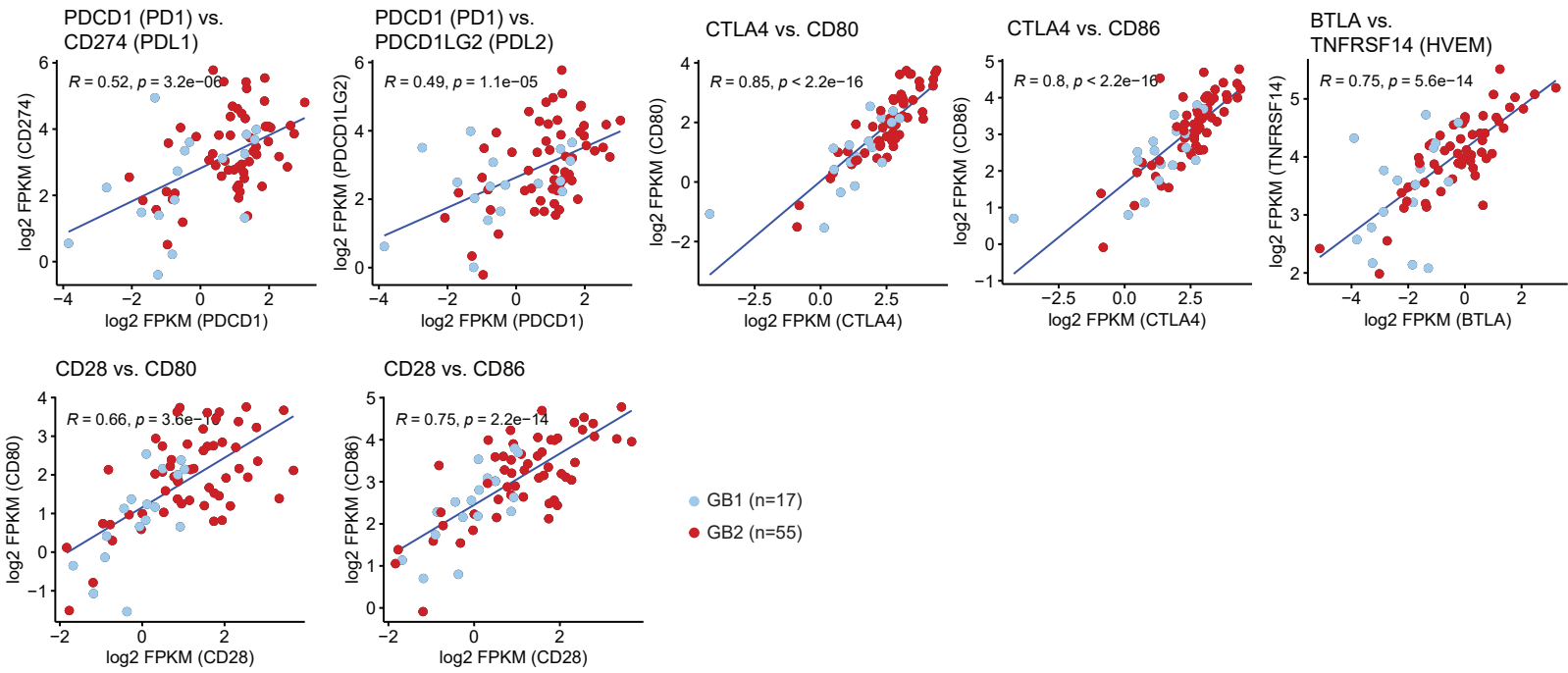


Figure S7: Correlations between expression of immune checkpoint molecules and their respective ligands in OSCC-GB tumors (n=72). The x- and y-axes indicate expressions (log2 FPKM) of an immune checkpoint receptor and its corresponding ligand. R values are Pearson's correlation coefficients.

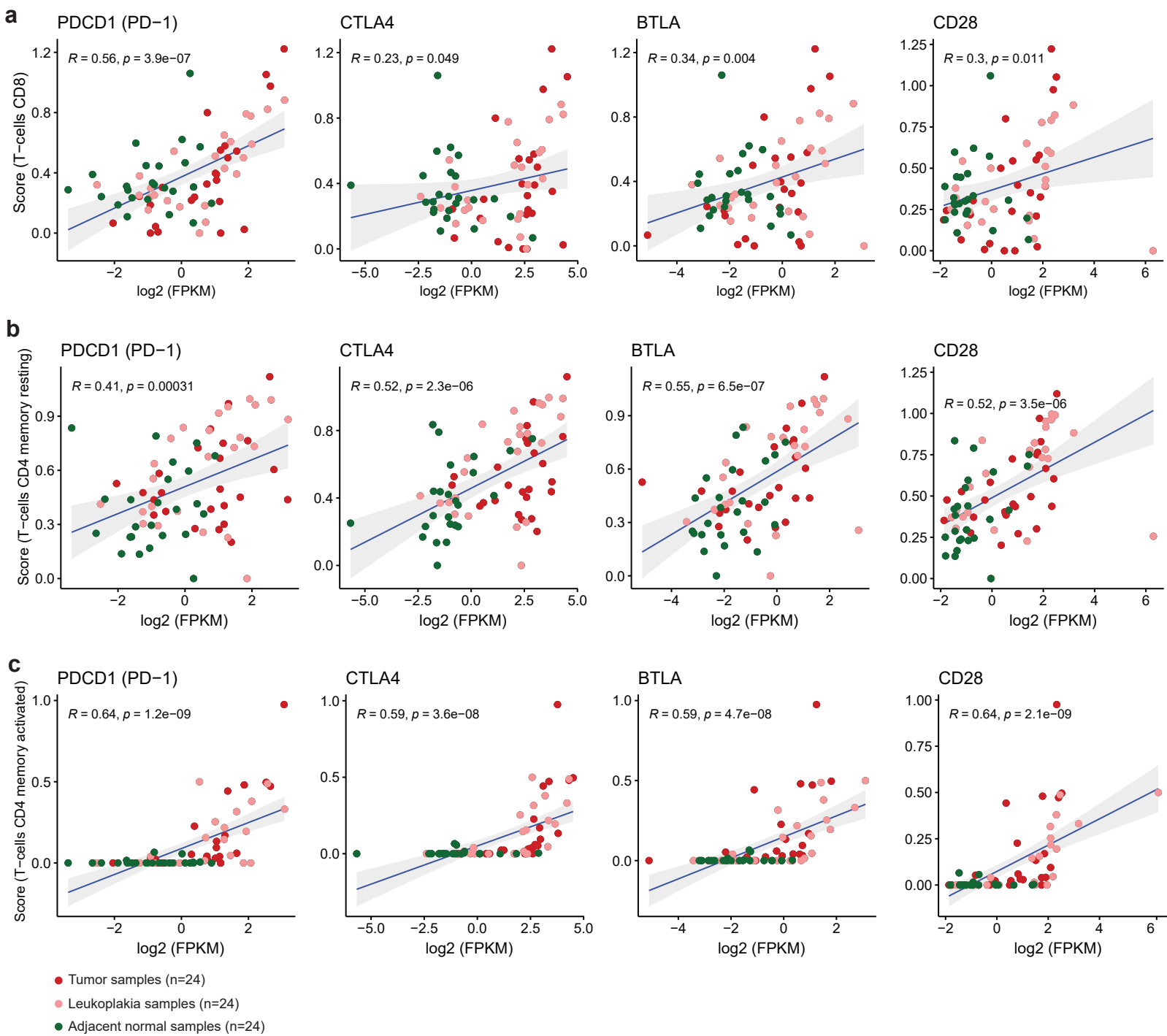


Figure S8: Correlations between abundances of CD8+ T-cells; CD4+ T-cells (memory resting or memory activated) and expression ($\log_2$ FPKM) of selected cell surface marker genes in tumor-leukoplakia-normal triad samples of the OSCC-GB (n=24) cohort. R values are Pearson's correlation coefficients.

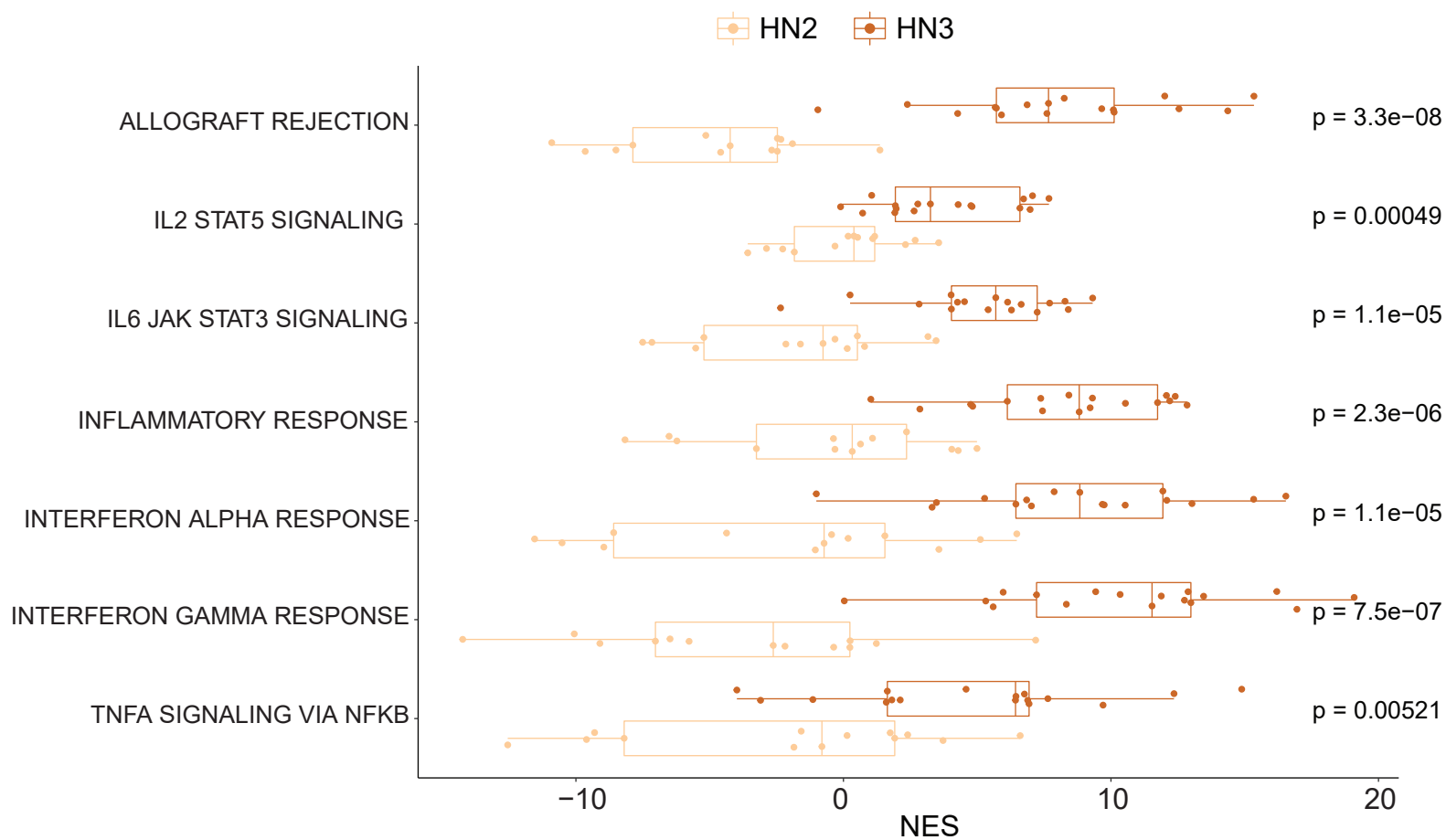


Figure S9: NES of immune-related hallmark gene sets between HN2 (n=13) and HN3 (n=17) sets of TCGA-HNSC patients. p-values were obtained using a 2-tailed Mann-Whitney U test.

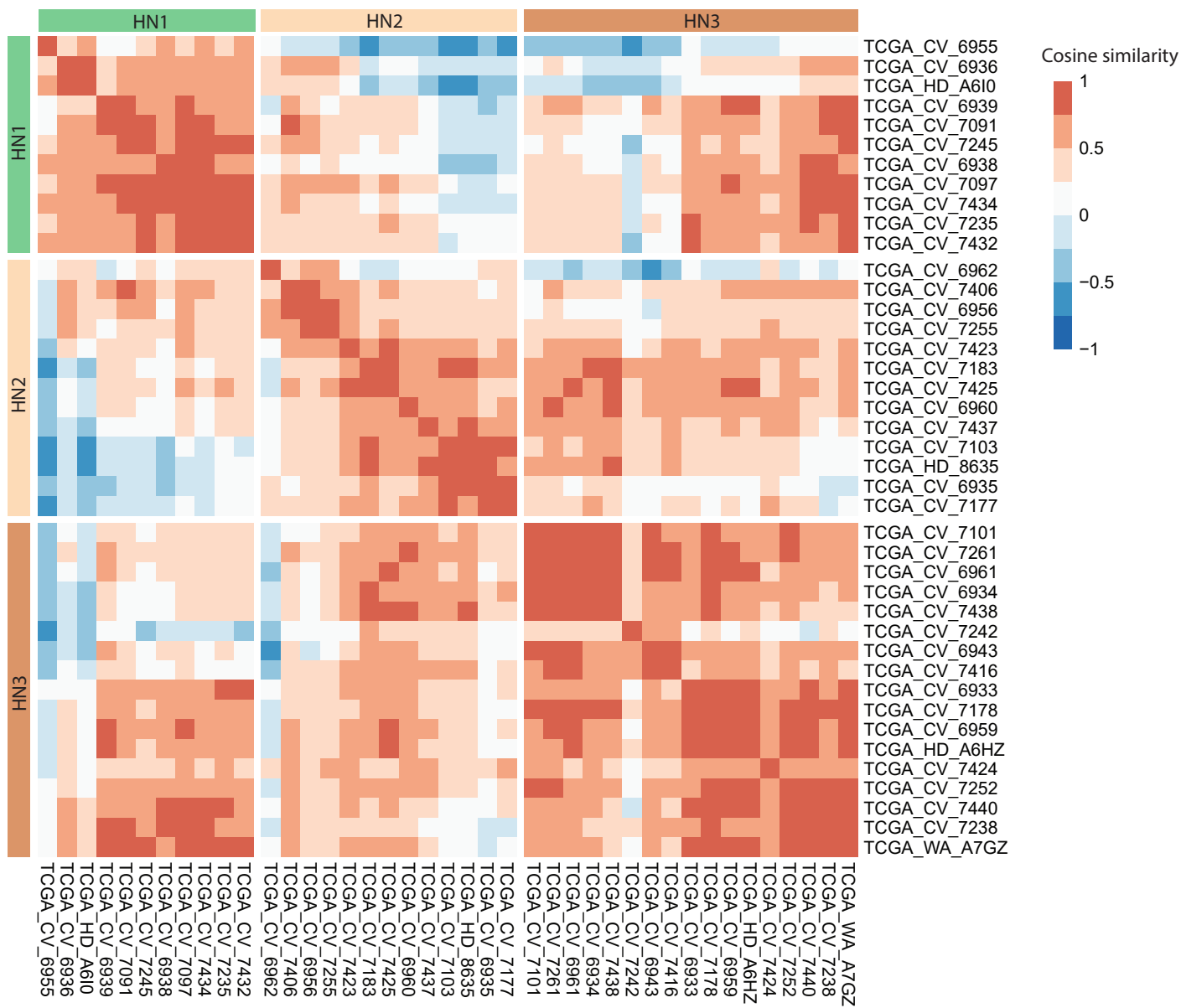


Figure S10: Comparison of HN1 (n=11), HN2 (n=13), and HN3 (n=17) clusters of the TCGA-HNSC cohort. Cosine distance metric was calculated using enrichment scores of 50 hallmark gene sets.

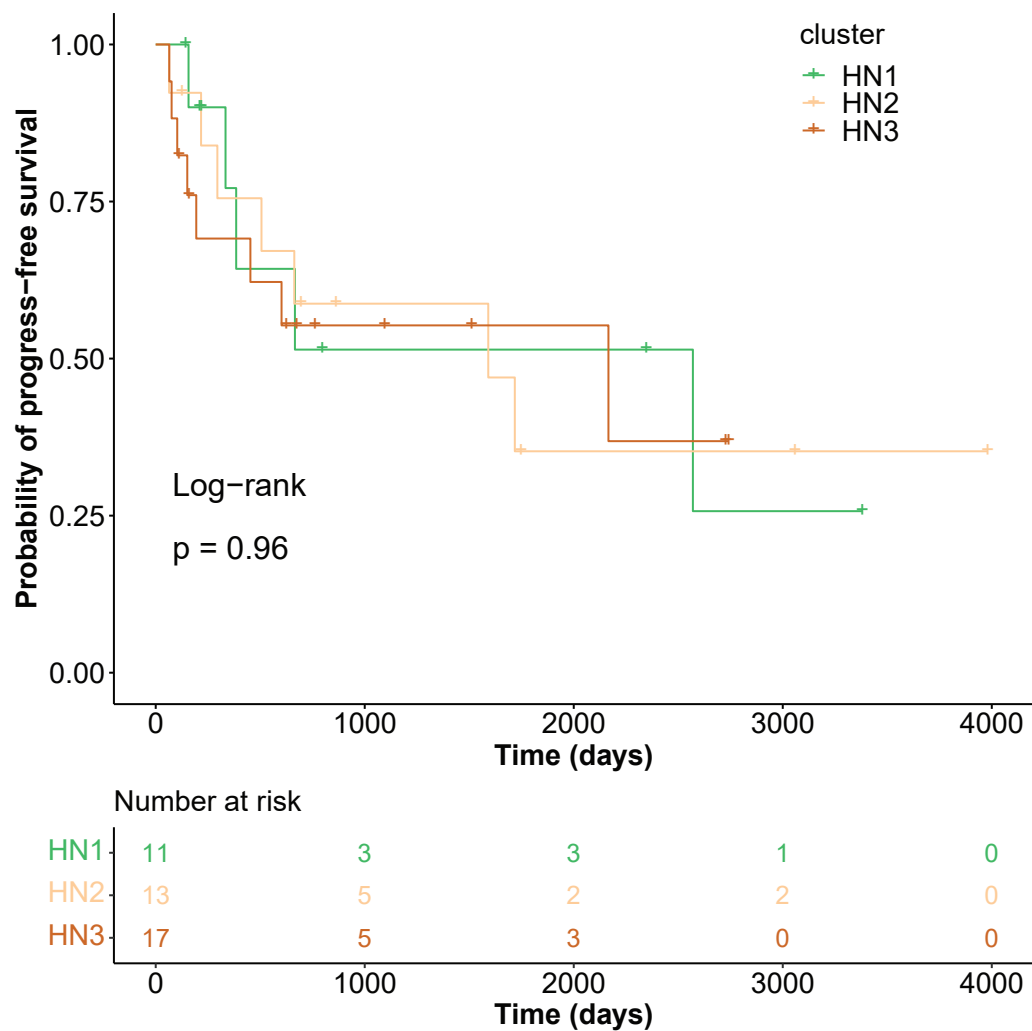


Figure S11: Kaplan-Meier curve showing progression-free survival in three TCGA-HNSC patient clusters.

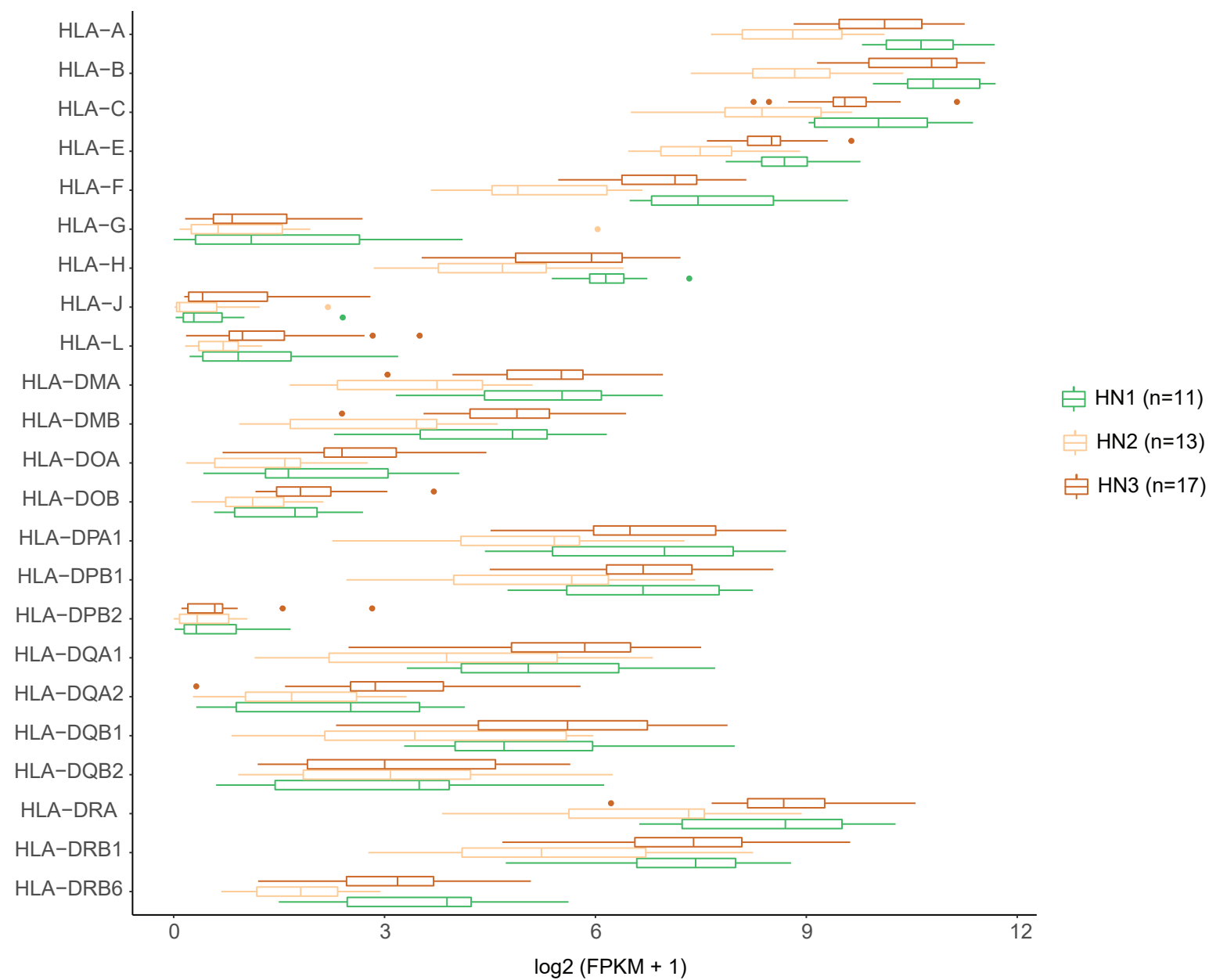


Figure S13: HLA gene expression (log2 FPKM+1) levels among the three clusters of the TCGA-HNSC cohort.

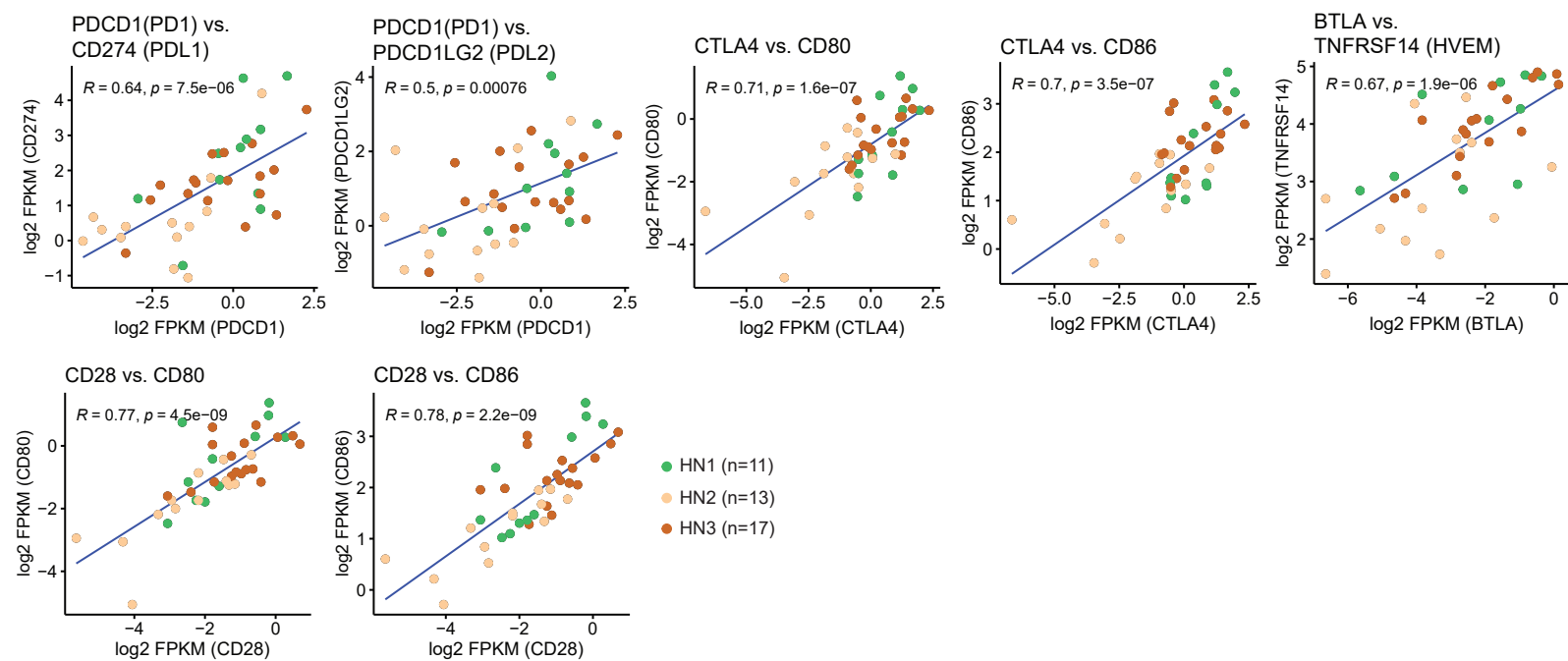


Figure S15: Correlation between immune checkpoint molecules and their respective ligands in the TCGA-HNSC tumors (n=41). The x- and y-axes indicate expressions (log2 FPKM) of immune checkpoint receptors and corresponding ligands. R values are Pearson's correlation coefficients.

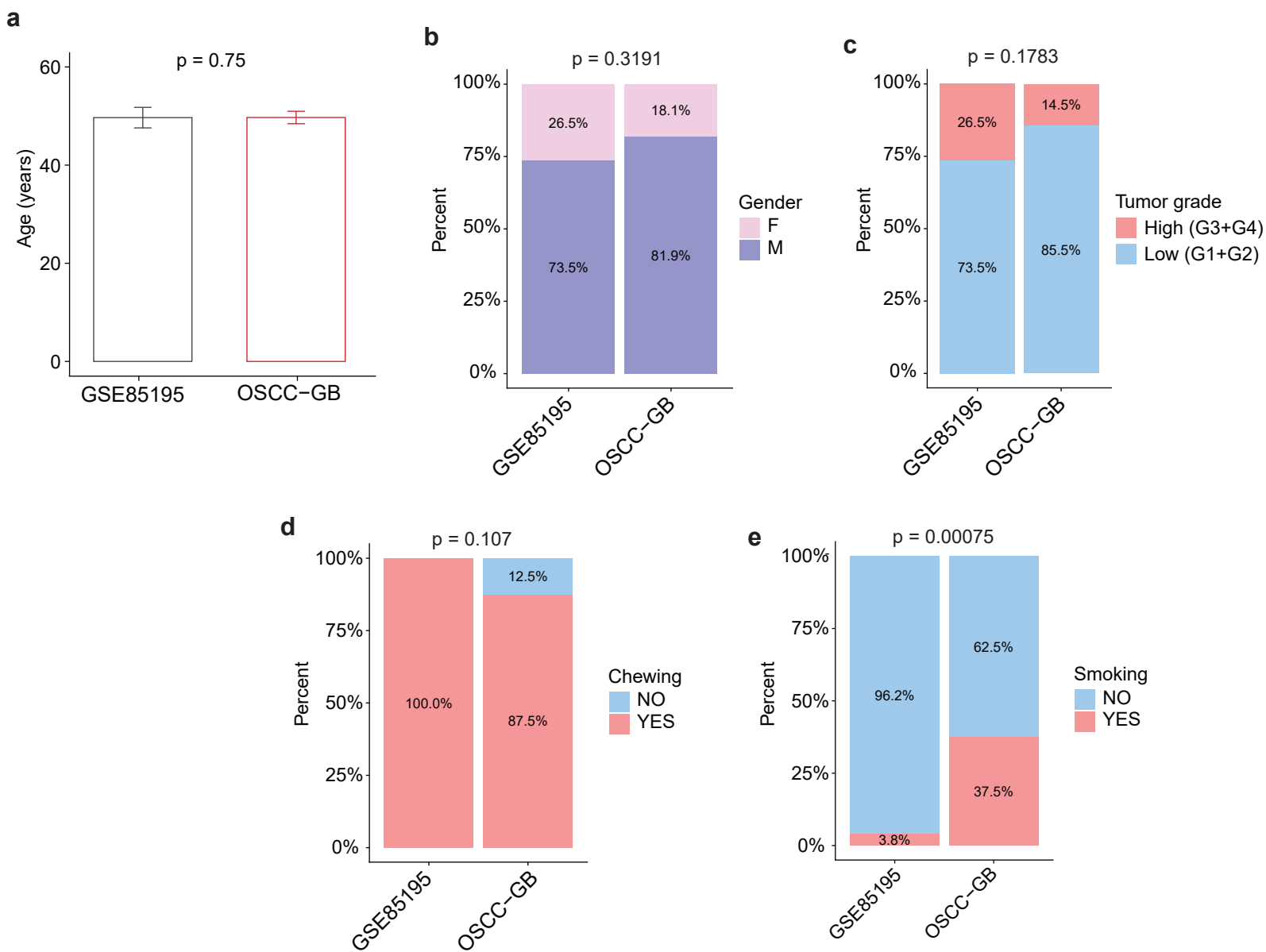


Figure S16: Comparison of clinicopathological characteristics between OSCC-GB and GSE85195 cohorts. Bar plots depict the distribution of (a) age, (b) gender, (c) tumor grade, (d) tobacco-chewing status, and (e) smoking status across the two cohorts. Percentages for categorical variables (panels b–e) are shown within the bars. The p-value in panel (a) was calculated using the Wilcoxon rank-sum test, while Fisher’s exact test was used to assess differences in categorical variables between cohorts.

Hazard ratio

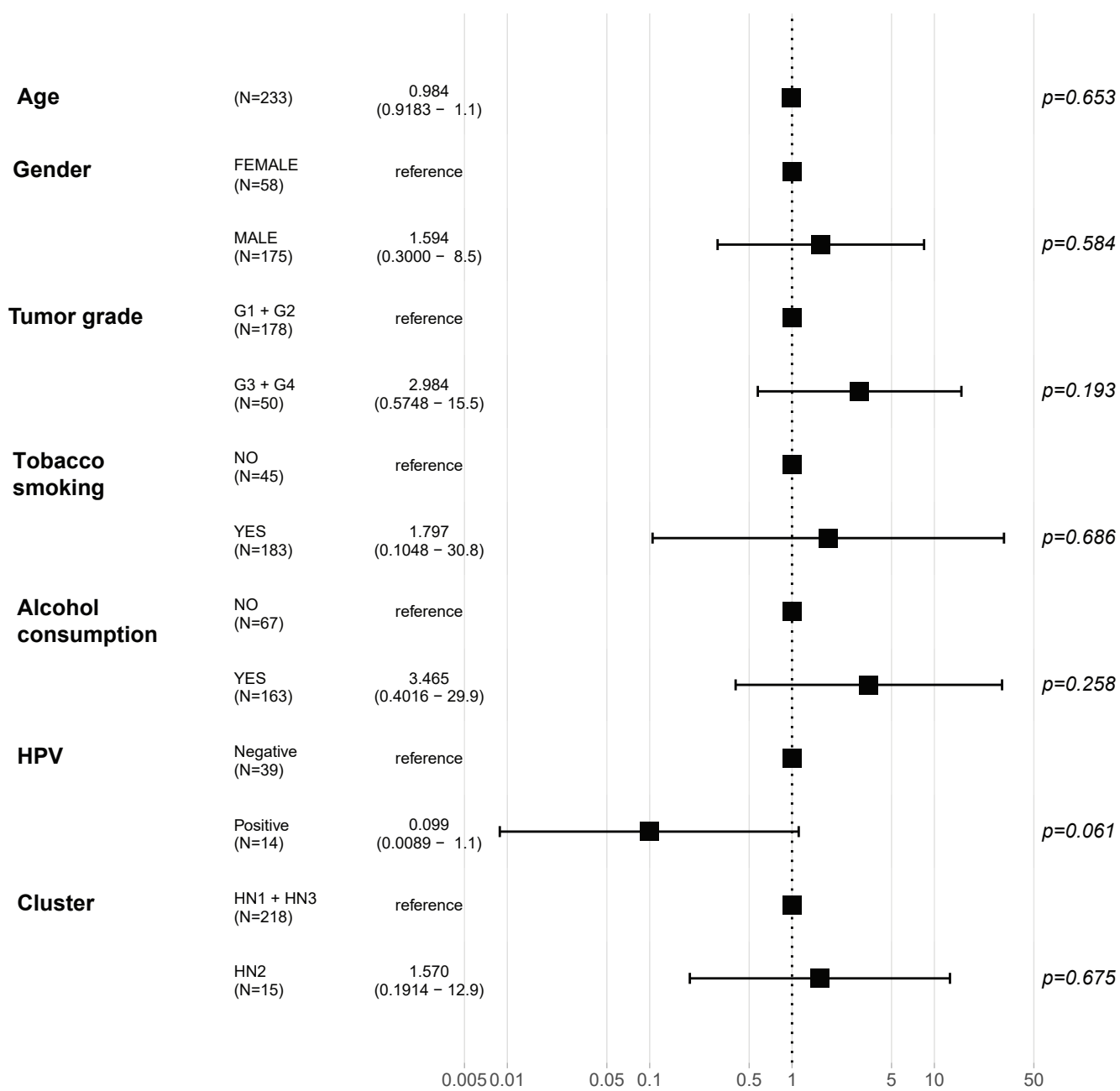


Figure S17: Multivariable Cox regression analysis of progress-free survival in the TCGA-HNSC (n=234) patients after controlling for other clinicopathologic factors.

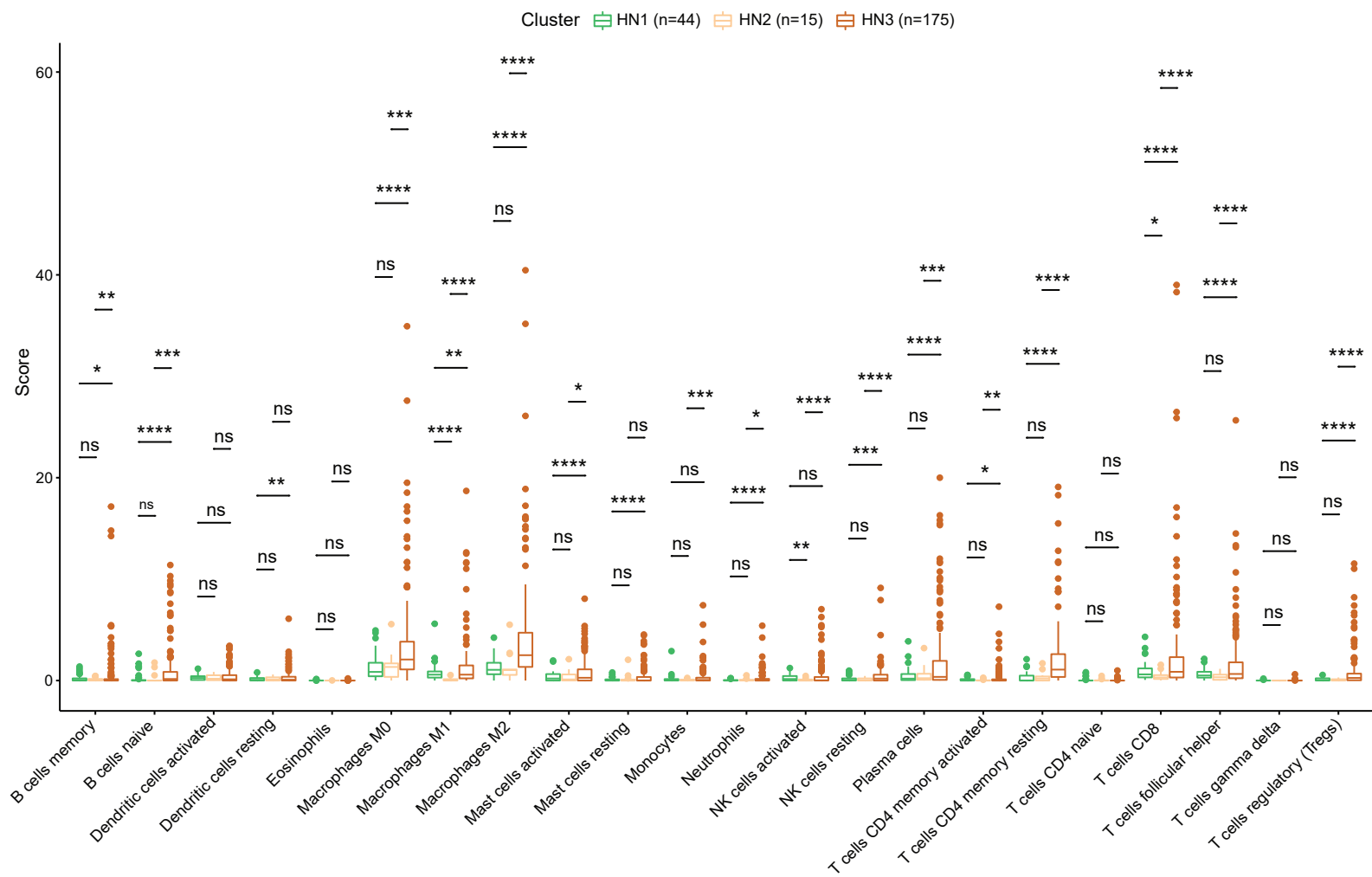


Figure S18: Abundances of immune cell types (n=22) in three predicted clusters of HN tumors in the TCGA-HNSC (n=234) cohort. The CIBERSORT absolute algorithm from RNA-Seq data estimated the infiltration of the immune cell types. Welch's t-test determined the significance of the difference between the means of the two abundances. *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.

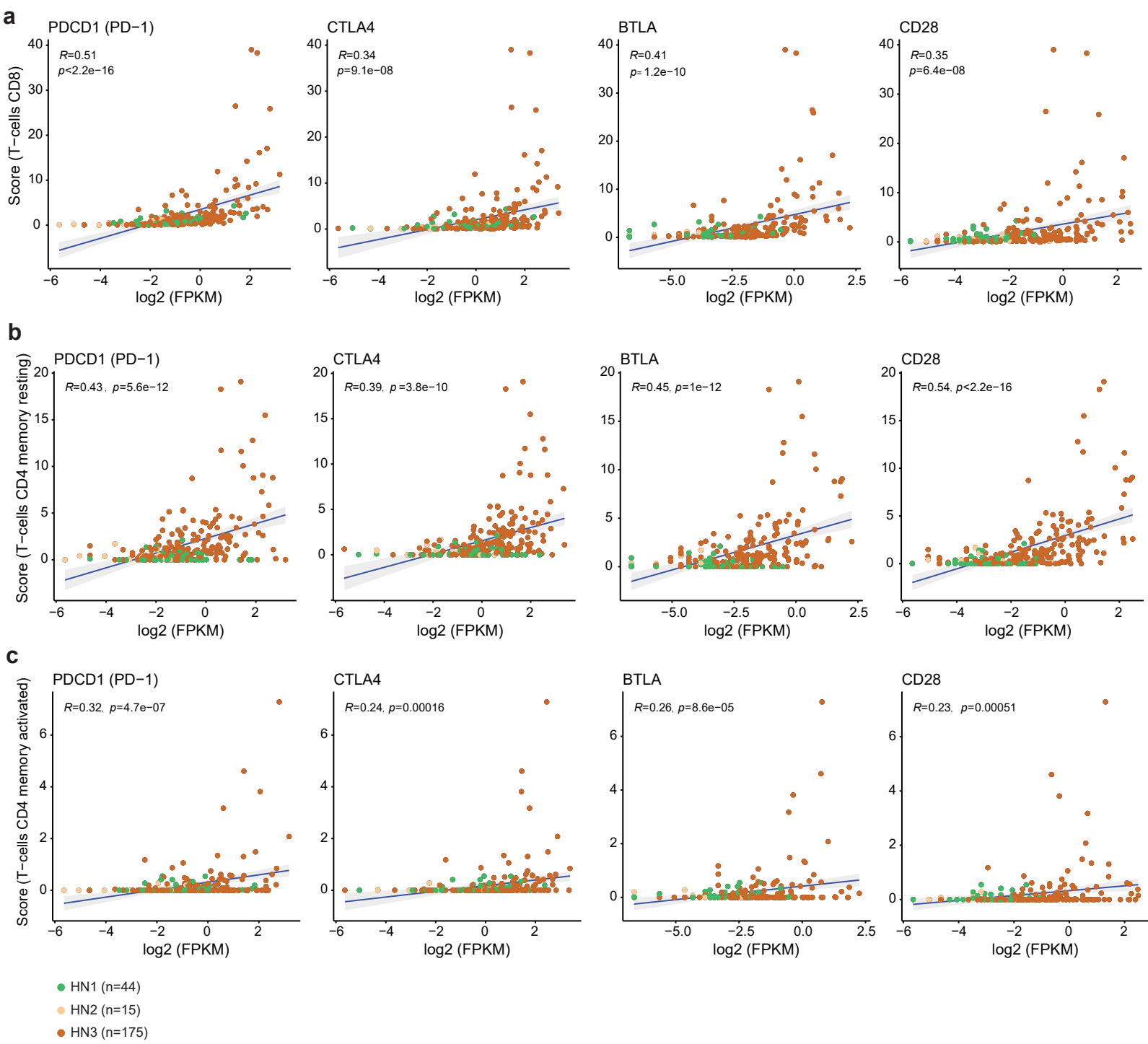


Figure S19: Correlation between abundances of CD8+ T-cells; CD4+ T-cells (memory resting), or CD4+ T-cells (memory activated) and expression (log2 FPKM) of selected cell surface marker genes in TCGA-HNSC (n=234) tumor samples from predicted HN patient clusters. R values indicate Pearson's correlation coefficients.

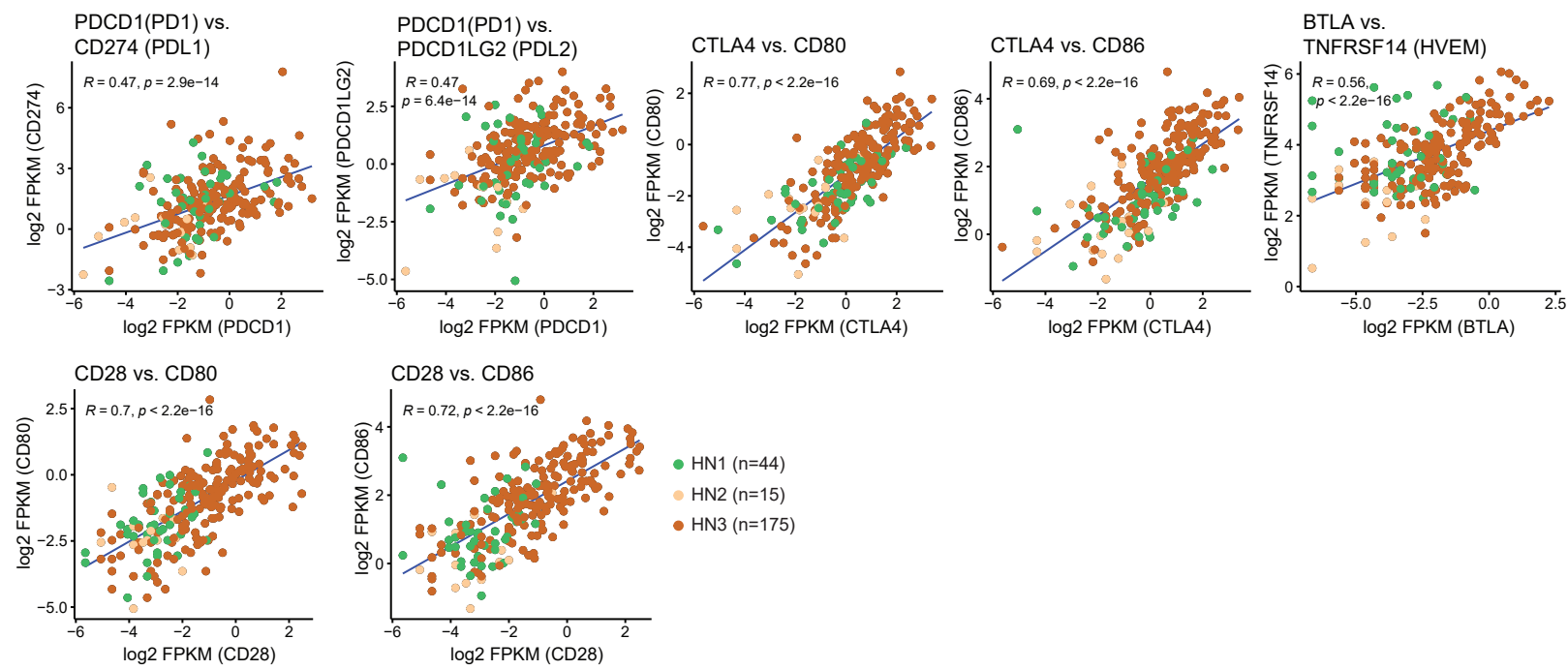


Figure S20: Correlation between immune checkpoint molecules and their respective ligands in TCGA-HNSC (n=234) cohort. The x- and y-axes indicate log2 FPKM values of immune checkpoint receptors and corresponding ligands. R values indicate Pearson's correlation coefficients.