Figure S1. Major cell types identified for esophageal squamous cell carcinoma, related to Figure 1.

(A) Preprocessing of CyTOF data as indicated in the plots before pooling for formal analysis. (B) Heatmap showing the clusters computed by the FlowSOM algorithm and the cell type assignment. (C) Box plots showing the proportion of immune cells, CAFs, endothelial cells and unknown cells among normal, naïve and neo groups. *P values were derived from one-way ANOVA, Tukey’s test; *: P < 0.05, **: P < 0.01, ***: P < 0.001, ****: P < 0.0001. Error bars: Mean ± SEMs.

CAFs, cancer-associated fibroblasts.
Figure S2. Profile of CAFs subsets, related to Figure 1.
(A) t-SNE plot showing the CAFs subsets. (B) t-SNE plots of CAFs showing the expression of marker genes, CXCL8, CD73 and PDGFRα. Box plots showing the frequency of PDGFRα+ CAFs, CD73+ CAFs and unknown CAFs divided by the total CAFs number among (C) normal, naïve, neo groups, and among (D) normal, naïve, neo-c, neo-cr and neo-cri groups. $P$ values were derived from one-way ANOVA, Tukey’s test; *: $P < 0.05$, **: $P < 0.01$, ***: $P < 0.001$, ****: $P < 0.0001$. Error bars: Mean ± SEMs.
CAFs, cancer-associated fibroblasts.
Figure S3. Characterization of immune cell types, related to Figure 2.

(A) Heatmap showing the clusters computed by the FlowSOM algorithm and the immune cell type assignment. (B) Conventional gating of cDCs, macrophages and NKT cells. (C) Box plots illustrating the proportion of NKT cells, neutrophils, eosinophils, basophils and mast cells among normal, naïve and neo groups. (D) Table showing the feature genes associated with cDCs and pDCs. (E) Kaplan-Meier survival curves for overall survival of the ESCC patients in TCGA database according to the gene signatures of cDCs. *: $P < 0.05$, **: $P < 0.01$, ***: $P < 0.001$, ****: $P < 0.0001$. Error bars: Mean ± SEMs. *: $P < 0.05$, **: $P < 0.01$, ***: $P < 0.001$, ****: $P < 0.0001$. Error bars: Mean ± SEMs. $P$ values in (E) were derived from log-rank test.

Neu, neutrophils; Eos, eosinophils; cDCs, conventional dendritic cells; pDCs, plasmacytoid dendritic cells; Baso, basophils; Mast, mast cells; Mac, macrophages.
Figure S4. Comparisons of immune cell types among different neoadjuvant therapies, related to Figure 2.

(A) Box plots showing the frequency of T cells, B cells, NKT cells, macrophages, cDCs, neutrophils, eosinophils, basophils, mast cells and pDCs among normal, naïve, neo-c, neo-cr and neo-cri groups. $P$ values were derived from one-way ANOVA, Tukey’s test; *: $P < 0.05$, **: $P < 0.01$, ***: $P < 0.001$, ****: $P < 0.0001$. Error bars: Mean ± SEMs.

Neu, neutrophils; Eos, eosinophils; cDCs, conventional dendritic cells; pDCs, plasmacytoid dendritic cells; Baso, basophils; Mast, mast cells; Mac, macrophages.
Figure S5. Characterization of T cell types, related to Figure 3.

(A) t-SNE plots of T cells showing the expression of marker genes, CD103, PD-1, CCR7 and CD45RA. (B) Conventional gating that defines naïve, TCM, TEM and TEMRA status of T cells. (C) Conventional gating that defines Trm, T helper subsets and Tex. (D) Box plots showing the frequency of T helper subsets and Trm relative to the total Tconvs number, and the frequency of Tex and Trm relative to the total CD8⁺ T cells number. (E) Box plots demonstrating the mean PD-1 expression level in CD8⁺ T cells and Tregs, and mean CD73 expression level in Tregs. *: $P < 0.05$, **: $P < 0.01$, ***: $P < 0.001$, ****: $P < 0.0001$. Error bars: Mean ± SEMs.

Tconvs, conventional CD4⁺ T cells; Tregs, regulatory T cells; TCM, central memory T cells; TEM, effector memory T cells; TEMRA, terminally differentiated effector memory T cells; Trm, tissue resident memory T cells; Tex, exhausted T cells.
Figure S6. Characterization of NK cell clusters, related to Figure 4.

(A) Box plot showing the proportion of NK cells among normal, naïve, neo-c, neo-cr, and neo-cri groups by flow cytometry analysis. (B) UMAP plot of NK cells colored by NK cell clusters re-analyzed using the GSE160269 dataset. (C) UMAP plot NK cells showing the expression of FCGR3A from the GSE160269 dataset. (D) Dotplot illustrating the expression of marker genes of different NK cell clusters from the GSE160269 dataset. (E) Pathway enrichment analysis of upregulated genes in CD16⁺ NK cells using scRNA-seq data from the GSE160269 dataset. (F) Signature scores of cytotoxicity, exhaustion, G1/S and G2/M among different NK cell subsets derived from the GSE145370 dataset. (G) Scatter plots showing the frequency of CD16⁺ NK cell relative to the total NK cell number between stage I and stage II/III pathologic stages of ESCC tumors derived from the GSE160269 dataset. P values in (A) were derived from one-way ANOVA, Tukey’s test; P values in (G) were derived from two-sided unpaired Student’s t test; *: P < 0.05, **: P < 0.01, ***: P < 0.001, ****: P < 0.0001. Error bars: Mean ± SEMs.
Figure S7. Characterization of macrophage subsets, related to Figure 5.
(A) Conventional gating that defines M1/M2 status of macrophages. (B) Box plots showing the M1/M2 status of macrophages according to the expression pattern of CD86 and CD206 among normal, naïve and neo groups. (C) Box plot illustrating the frequency of CCR4⁺CCR6⁺ macrophages among normal, naïve, neo-c, neo-cr and neo-cri groups. (D) Conventional gating that defines CCR4⁺CCR6⁺ macrophages. (E) Box plot illustrating the frequency of CCR4⁺CCR6⁺ macrophages among normal, naïve, and neo groups, analyzed by flow cytometry analysis. (F) Box plot illustrating the frequency of CCR4⁺CCR6⁺ macrophages among normal, naïve, neo-c, neo-cr and neo-cri groups flow cytometry analysis. (G) Multiplex immunofluorescence was used to analyze the spatial localization of CD68⁺CCR4⁺CCR6⁺ macrophages, T cells (CD3⁺) and tumor cells (Pan-CK⁺). P values in were derived from one-way ANOVA, Tukey’s test; *: P < 0.05, **: P < 0.01, ***: P < 0.001, ****: P < 0.0001. Error bars: Mean ± SEMs.
Mac, macrophages.
Figure S8. Sources of CCR4/CCR6 ligands, related to Figure 5.

(A) UMAP plots showing the expression of CCL3, CCL5, CCL17, CCL20 and CCL22, and major clusters of immune cells from the GSE160269 dataset. Dotplots of immune clusters illustrating the expression of CCL3, CCL5, CCL17, CCL20 and CCL22 from the (B) GSE160269 and (C) GSE145370 dataset. (D) UMAP plots showing the expression of CCL3, CCL5, CCL20 and CCL22, and major clusters of non-immune cells from the GSE160269 dataset.

cDCs, conventional dendritic cells; pDCs, plasmacytoid dendritic cells; Mast, mast cells; Mono, monocytes; Mac, macrophages; Tconvs, conventional CD4+ T cells; Tregs, regulatory T cells; Plasma, plasma cells; mregDCs, mature DCs enriched in immunoregulatory molecules.
Figure S9. Comparisons of cDCs types among different neoadjuvant therapies, related to Figure 5.

(A) UMAP plots showing the expression of mregDCs markers, LAMP3, CCR7 and CD274, and myeloid clusters from the GSE160269 dataset. (C) Conventional gating of mregDCs according to the CCR7 and PD-L1 expression. Box plots illustrating the frequency of mregDCs among (C) normal, naïve and neo groups and among (D) normal, naïve, neo-c, neo-cr and neo-cri groups. (E) Conventional gating of cDC1s and cDC2s according to the FceRIa expression. Box plots illustrating the frequency of cDC1s and cDC2s among (F) normal, naïve and neo groups and among (G) normal, naïve, neo-c, neo-cr and neo-cri groups. P values were derived from one-way ANOVA, Tukey’s test; *: \( P < 0.05 \), **: \( P < 0.01 \), ***: \( P < 0.001 \), ****: \( P < 0.0001 \). Error bars: Mean ± SEMs.
cDCs, conventional dendritic cells; cDC1s, type 1 conventional dendritic cells; cDC2s, type 2 conventional dendritic cells; mregDCs, mature DCs enriched in immunoregulatory molecules.
Figure S10. Establishment of prognosis model based on CCR4/CCR6 signaling, related to Figure 6.

(A) IHC analyses of CCR4/CCR6 chemokine system among naïve and neo groups of ESCC patients. (B) H-score of CCR4/CCR6 chemokine system among naïve, neo-c, neo-cr and neo-cri groups of ESCC patients. (C) Signature scores calculated based on the established model among responders baseline (n = 9), responder on-treatment (n = 12), non-responders baseline (n = 20), non-responder on-treatment (n = 20) groups derived from the RNA-seq data of the PERFECT trial. P values were derived from one-way ANOVA, Tukey's test; *: P < 0.05, **: P < 0.01, ***: P < 0.001, ****: P < 0.0001. Error bars: Mean ± SEMs.

bs, baseline; on, on-treatment.