Supplemental Methods

Flow Cytometry

48 h after lentiviral infection, Ewing sarcoma cells (A673) were trypsinized and washed once with PBS then permeabilized and fixed in 70% ethanol. Cell were washed with PBS and resuspended in PBS with propidium iodide and RNase. Cells were analyzed (CyAn) and the cell cycle profile was quantified (ModFit LT, Verity House).

Supplemental Figure Legends

Figure S1. Cell cycle profile of Ewing Sarcoma cells. The percentage of cells in G0/G1, S, G2-M, or sub-G0 (Sub) based on propidium iodine staining and flow cytometry were calculated for uninfected A673 cells ("control") and cells in which endogenous EWS-FLI was silenced with concurrent HA-EWS-FLI ("EF") or HA-FLI1 ("FLI1") expression.

Figure S2. Two genomic loci demonstrating the range of factor- and cell-type-specific binding. EWS-FLI (black) and FLI1 (red) ChIP and input control signal in both Ewing Sarcoma cells and HUVECs at (**A**) *NR0B1* and (**B**) *EPHA2***.** Viewing range is cut at 50 reads. Horizontal bars represent bound sites as identified by ZINBA. Scale bar and schema of each gene are depicted at the top and bottom of the represented tracks.

Figure S3. Differentially expressed genes in endothelial and Ewing cells. **A.** Number of up- and down-regulated RefSeq genes identified in each cell type for each transcription factor. **B.** Fold-change of genes commonly differentially expressed in EWS502 (blue) and HUVEC (red) cells by EWS-FLI or FLI1. Numbers shown represent gene counts per quadrant.

Figure S4. EWS-FLI and FLI1 differentially expressed genes have distinct biological functions. **A.** Over-enriched biological functions and **B.** Over-enriched biological pathways identified by Ingenuity Pathway Analysis indicate that EWS-FLI activates cancer-related genes and pathways, while FLI1 activates genes involved in normal endothelial growth in both cell types. Significance line ($p < 0.05$) is drawn in red.

Figure S5. Annotation of clusters of binding sites using GREAT shows biologically relevant associations with ontologies. Only those terms with FDR-corrected q-values more significant than 10^{-5} and in the top 5 significant terms are shown. Bars are color-coded by the ontology from which they were derived (MSigDB Perturbation, green; Mouse Phenotype, blue; MGI expression: Detected, red; GO Biological Process, yellow) and statistical significance is expressed as -log10(q-value).

Figure S6. Control immunoglobulin (IgG) ChIP-qPCR on chromatin isolated from EWS502 cells expressing the various Ewing Sarcoma fusions. Results are shown as a percent of input control. Overall, greater binding is identified at EWS-FLI bound regions near differentially expressed genes that contained GGAA repeats (*NR0B1, CAV1, GSTM4, JAK1, IGF1*) compared to those that bound EWS-FLI but did not harbor a repeat (*NKX2-2, KIF14, JAK1, CDKN1A, MDM2*). Five control

repeat-containing regions are included, and error bars represent standard error of three replicates.

Figure S7. Upregulated genes are closer to EWS-FLI binding sites. Median distance from TSS of all genes, all differentially expressed, all upregulated, and all downregulated RefSeq genes to the nearest EWS-FLI-bound GGAA repeat with length ≥ 5.

Figure S8. FLI1 binding sites are closer to FLI1 differentially expressed genes. Distance from the TSS of a gene differentially expressed by FLI1 (blue) or EWS-FLI to the nearest FLI1 or EWS-FLI binding site, respectively. The fraction of genes containing at least one site within the denoted distance is presented.

Figure S9. EWS-FLI-bound tetranucleotide repeats demonstrate repeat length polymorphism. **A.** Length of repeat found within *IGF1* across 7 Ewing Sarcoma cell lines (EWS502, EWS894, A673, MHH-ES-1, RD-ES, SK-ES, SK-N-MC) and compared to endothelial cells (HUVEC). Lengths determined by PCR using primers flanking repetitive region and resolved on an 8% acrylamide gel. **B.** Sequence of repeat region from EWS502 cells compared to reference genomic sequence (hg18). Multiple sequence alignment was performed using ClustalX with default parameters. Exact sequence matches denoted by "*", regions of difference highlighted in yellow.

Figure S10. EWS-FLI occupied sites similar to that of FLI1 in a normal cellular context. Heatmap showing normalized ChIP-seq signals of EWS-FLI or FLI1 in both Ewing Sarcoma cells and HUVECs around computationally predicted transcription factor binding sites of ETS, Max, NFkB, STAT, PPAR, HNF4, and CREB. Sequence logos corresponding to the computationally predicted motif are shown on the left. Color was assigned on a $\log_2$ scale from 0.5 to 9.

Figure S11. Permutation abolished EWS-FLI and FLI1 signal enrichment at all identified sites. One permutation of GGAA repeats, ETS, ETS-AP-1, AP-1, GATA, as well as the 7 computationally predicted motifs presented in Figure S10 shows the observed pattern is non-random. Color was assigned on a $\log_2$ scale from 0.5 to 9.

Figure S12. UCSC Genome Browser snapshots of epigenetic patterns in EWS502 (black), HUVEC (orange), and H1hESC (green). **A.** Three of the four *HOX* clusters show activation and aberrant H3K27me3 patterns relative to normal cell types. Values are presented on a scale of 0 to 50 reads. **B.** Normal H3K27me3 at *PAX2* (left) and *WNT3A* (right), consistent with normal cell types. Values are presented on a scale of 0 to 50 reads.

Figure S13. Chromatin accessibility and modifications of GGAA microsatellites. Percentage of GGAA repeats overlapping areas of significant enrichment for H3K4me1, H3K4me2, H3K4me3, H3K27me3, and FAIRE are presented for EWS502, HUVEC, H1hESC, K562, NHEK, and GM12878 cells.

Figure S14. Predicted nucleosome occupancy of EWS-FLI and FLI1 binding sites. Average nucleosome occupancy predicted on DNA sequence (Kaplan et al. 2009) surrounding the summits of EWS-FLI and FLI1 binding sites in EWS502.

Supplementary Table Legends

Supplementary Table 1. Summary of sequencing quality. For each experiment, the number of raw sequencing reads, approximate basecall error rate, number of uniquely aligned and filtered sequencing reads, the number of regions exhibiting significant enrichment (peaks), and the cross-replicate correlation (where applicable) are presented.

Supplementary Table 2. Binding sites of EWS-FLI and FLI1 in EWS502 and HUVEC. The chromosomal coordinates of identified binding sites for each transcription factor in each cell type are listed.

Supplementary Table 3. Significantly differentially expressed genes. Gene symbols and RefSeq IDs, as well as p-value and fold-change are presented for all genes determined to be differentially expressed in each cell type and manipulation.

Supplementary Table 4. Closest differentially expressed genes to identified binding sites. The peak coordinates as well as the gene symbol, RefSeq ID, chromosome, position, and distance to the nearest transcriptional start site are listed for each cell type and transcription factor. Distances are reported in base pairs.

Supplementary Table 5. Complete list of ChIP cluster annotations. Ontologies, along with their term name, q-value, and –log10(q-value), determined to be over-represented in each of the clusters identified in Fig. 3A and Sup. Fig. S5 are presented.

Supplementary Table 6. Locations of H3K4me1, H3K4me2, H3K4me3, H3K27me3, and FAIRE enrichment in EWS502. The chromosomal coordinates for each are presented.

Supplemental References

Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes TR, Lieb JD, Widom J et al. 2009. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**(7236): 362-366.