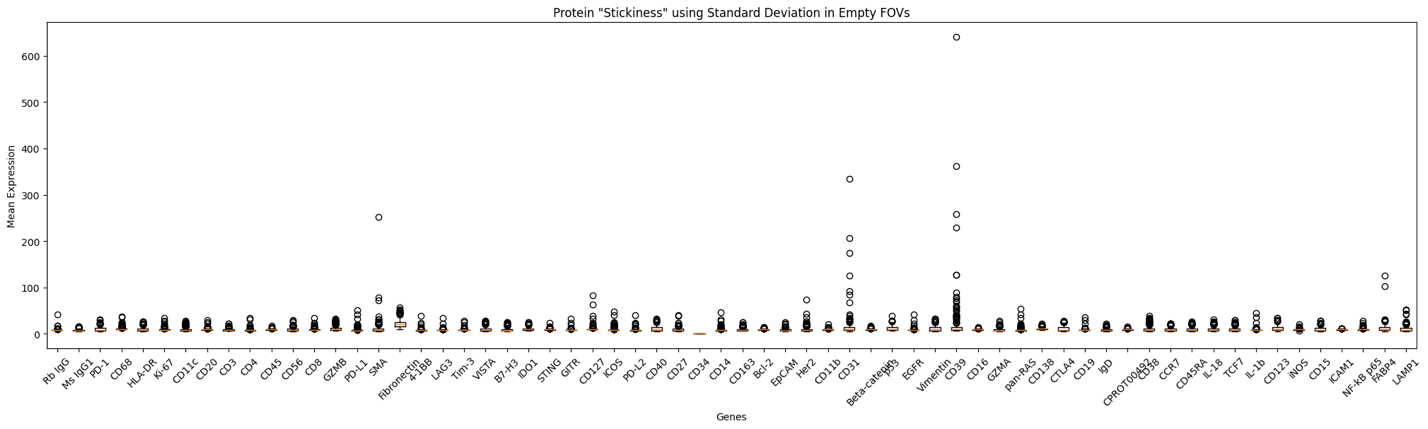
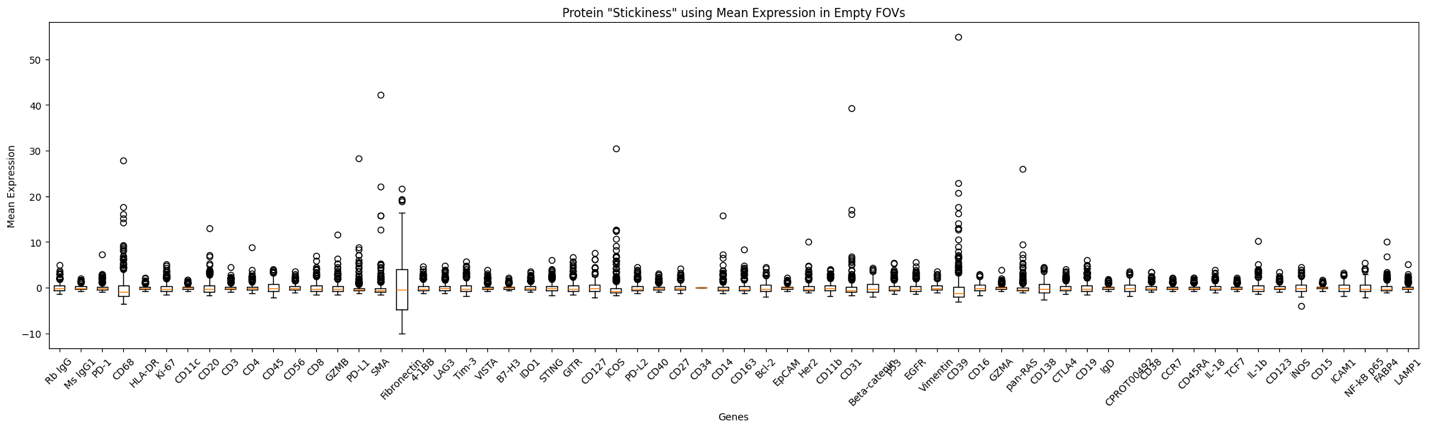
CosMx Protein Stickiness Analysis

**Overview:** One of the issues with image-based proteomic analysis is that we are unsure if the proteins are selectively binding onto tissues. For example, CosMx Protein users select FOVs of regions of interest and other areas on the slide will be ignored for imaging. Thus, it is difficult to accurately gauge protein noise. Luckily, we have a few TMA samples for which all regions on the slide are included as FOVs. This script is an analysis of those non-tissue regions to examine the noisiness of proteins. The code is commented.

**Tutorial:** This tutorial will describe the outputs of the script. Please refer to the comments on the script for more detail.

What the script does is that users define all non-tissue region FOVs. This must be done manually by visiting AtoMx and viewing the tissue and writing down the FOVs. What the script does is then it pulls the raw protein images and only look at the non-tissue FOVs. It will calculate the average fluorescence level for each of the FOVs as well as the standard deviation for the image as well. These will be stored as numpy matrices so that downstream analysis will be faster.

Boxplots of the average expression and the standard deviation are show below. These plots help us narrow down the genes of interest which are potentially problematic genes.



On the plot, you can see that Fibronectin is especially problematic with a large variance. Whereas a gene like SMA is also of interest since its outlier have incredibly high values. On the tissue map, we can see these differences more clearly.

A diagram of a graph

Description automatically generatedA diagram of a diagram of a variety of colors

Description automatically generated with medium confidence

As you can see, in the SMA example, high expression in non-tissue area is restricted in those two FOVs and expression is near 0 for rest of the FOVs. When we investigated the actual tissue, we found bits of tissue in that area which means that the protein staining is working as intended. However, in the Fibronectin example, we can see that there is low expression across the non-tissue regions which means that the staining agent is binding to areas where no tissue is present. Thus, we can identify Fibronectin as a problematic gene.