**Preliminary Analysis:** We have done a case study to run the pipeline for Aim1 with the RNA-seq data from cell lines MDA-MB-231 and MCF10A and ChIP-seq data and mutational profiles from MDA-MB-231.We did not include CNV and methylation profiles nor the proteomic data. As such, the biological results are not final but it allows us to see that the pipeline returns biologically meaningful results. All the source code for this preliminary analysis has already been deposited in Vera-Licona’s Github public repository https://github.com/VeraLiconaResearchGroup/CancerReversion.

**Preliminarystatic network of CL TNBC tumorigenesis for cell line MDA-MB-231**:The preliminary network we have built consists of 200 nodes, and 425 edges (fig1). The first layer of functionally enriched DEGs contains 194 genes. MATCH and IPA algorithms identified 4 TFs, which were confirmed to be in open chromatin regions by CHIP-Seq data [59]. For the final network layer, 3 upstream MRs were identified. The network has 6 genes known to be associated with TNBC tumorigenesis: TP53, STAT3, SP1, EGR1, SMAD3, and SMAD4 [60–63]. Additionally, the network captures genes that are known to be highly expressed in claudin low tumors, including EEF1A1, ACTB, HSP90AA1, and VIM [64]. CASP3, a gene associated with uncontrolled proliferation, and VIM, a gene associated with EMT, are included in the network [58,65]. We identified 28 readout nodes. Heatmaps with hierarchical clustering show that network nodes separate between MCF10A and MDA-MB231 using normalized expression values from RNA-seq data (Fig3).

**Attractor Analysis with SFA.** Using SFA we estimated the attractors from the MCF10A and MDA-MB-231 initial states. The direction of activity change (DAC) of the readout nodes was calculated between the MDA-MB231 and MCF10A attractors. For each readout node, the DAC was compared to the corresponding readout’s log2 FC values from our RNA-seq data. The readout nodes’ DAC matched the RNA-seq data, giving evidence that SFA is yielding reliable results (Fig4).

**FC Control Analysis.** There were 12 FVS, each containing 8 nodes. Because the network has 4 source nodes, each of the 12 FC control sets contains 1 FVS of 8 nodes and 4 source nodes, for a total of 12 nodes. FC sets were prioritized to select one with the least mutated genes known in MDA-MB-231 cell lines [64]24304894]. One FC control set, *FC1*, was chosen for further analysis. FC1 contains known TNBC genes including SMAD3 and STAT3 [60,61,63] and uncontrolled proliferation marker CASP3 [58].

***In Silico* Screening on FC1 Set.** SFA was performed taking as input our MD-MBA-231 network nodes, the RNA-seq normalized expression values of the MDA-MB-231 and a set of perturbations (activations or inhibitions) on FC1. SFA ran until reaching an attractor. A perturbation pattern is deemed successful if, we observe a change in the direction of expression of as many readout nodes from those expression values in the MDA-MB-231 attractor. After 100 simulations, one perturbation pattern was found to successfully change 14 of the readout nodes’ attractor activity. The 14 controlled readout nodes cover various hallmarks of cancer. CD44 is a common cancer stem cell marker, associated with malignancy in MDA-MB-231 cells. CRK, CRKL, and FGFR1 are associated with metastasis of MDA-MB-231 cells [66–69]. Other switched readout nodes, such as HMGA2 and KPNB1 are associated with EMT and malignant proliferation in MDA-MB-231 breast cancer cells, respectively [70,71]. Once we identify a satisfactory set of interventions on the 8 nodes of FC1, the last step will be to reduce the size of FC1 applying our software, OCSANA and repeat the same process for the other 12 sets.