**Supplementary information**

****

**Figure S1.** **Schematic showing the 4T1GFP tumor model** GFP tagged 4T1 cells were orthotopically injected into Balb/c mice. The 4T1GFP tumor thus generated was surgically resected after one-month post-surgery. The mouse was followed up until ethical endpoint to harvest the lungs. Insets showing the primary tumor and lung metastases.

****

**Figure S2.** **Tumor cells outgrew the CAFs** (a) Histogram showing the gating of dissociated primary tumor cells based on GFP, Epcam using FACS. (b) Scatter plot showing the sorting of CAFs from the GFP -ve population based on CD140a expression. (c) Phase contrast and fluorescent microscopy images showing the CAF cell culture in P2 and P3 passages. GFP +ve tumor cells outgrew the CAF cells by the third passage (Magnification 10x).



**Figure S3. Differential gene expression analysis.** Schematic showing the computational tools and filters applied for the identification of genes differentially expressed in the 4T1A cells that contribute to their aggressive phenotype when compared to the 4T1NA cells.

|  |  |  |
| --- | --- | --- |
| **Phenotype** | **4T1A** | **4T1NA** |
| Morphology | Adherent | Round |
| Proliferation | High | Low |
| Self-renewal | Low | High |
| Epithelial/Mesenchymal | Epithelial | Mesenchymal |
| Disease Outcome | Aggressive | Less aggressive |

**Table S1.** **Phenotypic characterization of the heterogeneous tumor cell populations.** Table showing the comparison of the *in vitro* and *in vivo* phenotypes of the 4T1A cells and 4T1NA cells obtained from the 4T1 primary tumor

|  |  |  |
| --- | --- | --- |
| **Gene symbol** | **Log2 FC** | **P value** |
| *Fgfbp1* | 9.613072189 | 1.05E-192 |
| *Dsg2* | 9.138150775 | 1.03E-161 |
| *Dsp* | 9.121914694 | 4.26E-226 |
| *Krt14* | 8.999645887 | 1.74E-207 |
| *Cldn4* | 8.203214657 | 1.32E-113 |
| *Krt17* | 8.186290528 | 9.47E-94 |
| *Krt19* | 8.158740817 | 1.54E-116 |
| *Moxd1* | 7.782371814 | 2.49E-108 |
| *Wnt7a* | 7.673772849 | 1.56E-161 |
| *Macc1* | 7.571940264 | 5.16E-184 |
| *Ntn1* | 7.433831168 | 7.94E-84 |
| *Lncenc1* | 7.378757975 | 6.76E-146 |
| *Esrp1* | 7.355492075 | 1.05E-94 |
| *Fam83a* | 7.332464375 | 6.30E-82 |
| *Serpinb5* | 7.236955868 | 1.49E-92 |
| *Unc13b* | 7.007964245 | 4.49E-96 |
| *Anxa8* | 7.003904283 | 6.29E-110 |
| *Wisp1* | 6.968642533 | 2.77E-232 |
| *B4galnt3* | 6.841436618 | 8.31E-98 |
| *Tacstd2* | 6.815237014 | 1.08E-119 |
| *Adamts5* | 6.687909001 | 1.21E-58 |
| *C3* | 6.663375268 | 3.46E-102 |
| *Myo5b* | 6.61220331 | 1.35E-65 |
| *Tmem185b* | 6.606416729 | 2.15E-70 |
| *Mpzl2* | 6.444142475 | 2.78E-242 |
| *Ctsw* | 6.378251279 | 1.82E-79 |
| *Epcam* | 6.288581562 | 1.13E-279 |
| *Tubb2b* | 6.146066228 | 2.47E-54 |
| *Ctgf* | 6.125384483 | 0 |
| *Mal* | 6.029244829 | 2.56E-49 |
| *Pcsk6* | 6.009392893 | 4.26E-57 |
| *Mapk13* | 5.901129217 | 4.22E-185 |
| *Itgb6* | 5.867463092 | 2.24E-48 |
| *Sorbs2* | 5.860353621 | 1.64E-249 |
| *Celsr1* | 5.850815844 | 1.85E-76 |
| *Fam167a* | 5.777816816 | 1.80E-45 |
| *Edn1* | 5.758725482 | 4.87E-54 |
| *Lcp1* | 5.75021601 | 5.63E-96 |
| *Ly6e* | 5.684214796 | 2.80E-81 |
| *Cntfr* | 5.657521771 | 7.59E-41 |
| *Rasef* | 5.629302159 | 3.37E-45 |
| *Cxcl16* | 5.611714969 | 7.36E-163 |
| *Ctsc* | 5.585199733 | 1.74E-170 |
| *Piezo2* | 5.551369461 | 1.59E-45 |
| *Lama3* | 5.530057633 | 7.15E-132 |
| *Myo5c* | 5.51074033 | 7.67E-43 |
| *Postn* | 5.509196229 | 1.74E-37 |
| *Inhba* | 5.495592626 | 0 |
| *Nbl1* | 5.460232808 | 4.43E-56 |
| *AU018091* | 5.450749912 | 8.12E-61 |

**Table S2.** **50 candidate genes upregulated in the 4T1A compared to 4T1NA cells** The DEGs obtained after DeSeq2 analysis was filtered based on q value < 0.01 and Log2 FC >2. The top 50 genes upregulated in the 4T1A population compared to 4T1NA cells were identified and could be mediating the aggressive phenotype of 4T1A cells

|  |  |  |
| --- | --- | --- |
| **Symbol** | **Gene Name** | **RFS pvalue** |
| MACC1 | Metastasis Associated Colon Cancer protein 1 | 0.034 |
| NTN1 | Netrin 1 | 0.045 |
| B4GALNT3 | Beta-1,4-N-acetylgalactosaminyltransferase 3 | 0.014 |
| ITGB6 | Integrin beta 6 | 0.011 |
| POSTN | Periostin | 0.014 |

**Table S3.** **Top 5 candidate genes upregulated in the 4T1A associated with poor survival in TNBC patients** List of top 5 genes upregulated in that showed significant correlation with poor survival in TNBC along with the p values for relapse free survival (RFS).