Multi-omics characterization of Microtubule-actin crosslinking factor (MACF1) using the ISB-Cancer Genomics Cloud

Dondra Bailey[1], Kawther Abdilleh[2,4], Boris Aguilar[3,4], Alexis McClary[1]

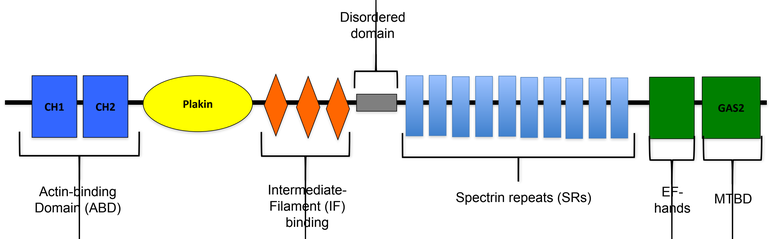
[1]Department of Natural Sciences, Coppin State University [2]General Dynamics Information Technology, [3]Institute for Systems Biology,

[4]Institute for Systems Biology- Cancer Genomics Cloud

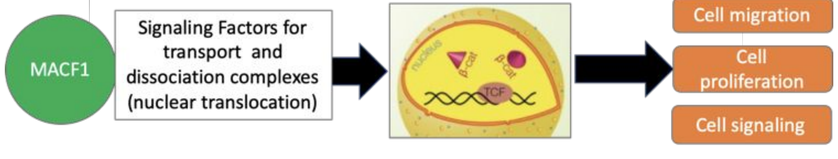
Background

Establishment of cell polarity across cell types and organisms involves distinct mechanisms that follow a common pattern: first a polarity cue arises, followed by asymmetric organization executed by polarity proteins. Loss of cell polarity has a key role in cancer development. The MACF1 gene, Microtubule actin cross-linking factor 1, or MACF1, a cytoskeletal protein is involved in oocyte development, cell proliferation, and cell migration. In addition to these roles, MACF1 is linked to metastatic invasion leading to tumor progression in numerous human cancers including gynecological cancers of endometrial and ovarian cancer. Given the functional importance of cell polarity, here we provide computational evidence of MACF1 in gynecological cancers. The comparison of multi-omic data for patient tumor and normal cells facilitates the understanding of the molecular mechanisms that contribute to tumor cell proliferation, abnormal cell adhesion and cell migration. Levering the rich datasets hosted by the NCI-funded ISB-Cancer Genomics Cloud, we performed a cloud-based patient cohort analysis across diverse multi-omics datasets. We quantified differential gene expression profiles from patients in the cohort as well as identified somatic mutations differences. The most common genomic alteration for MACF1 was the in-frame mutation. Genomic alterations and mutations were aligned to functional domains of the MACF1 protein to determine both frequency and spatial distribution. Gene-gene expression correlation analyses identified statistically significant correlations between MACF1 and other well-known cancer driver genes. Together, using a data driven cloud-computing approach we gain novel insights into the role of MACF1 regulation of cell polarity in the progression of cancer.

MACF1 Functional Domains



- Macf1 is a cytoskeleton crosslinker protein with seven main types of functional domains
- Expressed in oocyte, lung, skin cell and other tissues
- Roles in oogenesis, cell proliferation, and cell migration
- MACF1 has been implicated in signaling pathways associated with cancer



Methods

- All The Cancer Genome Atlas (TCGA) data were accessed on the Google Cloud using the ISB-CGC platform
- A cohort of patients with gynecological cancers and MACF1 mutations was identified using the ISB-CGC web-based cohort builder and explorer
- Somatic mutation, RNAseq and clinical TCGA data for patients in the cohort were explored in ISB-CGC BigQuery tables and analyzed using standard SQL

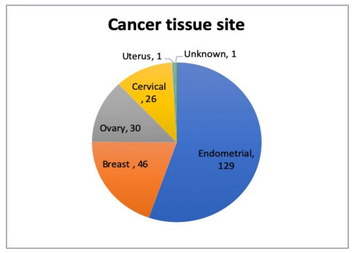
Statistical tests (t-test analyses) to compare gene expression between cases with mutations in key domains were also conducted in BigQuery

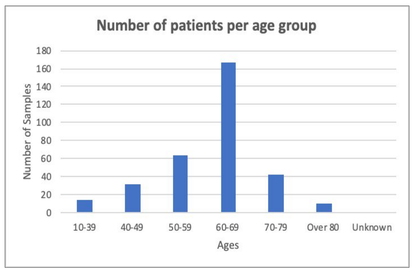
Patient Cohort Building & Data Exploration

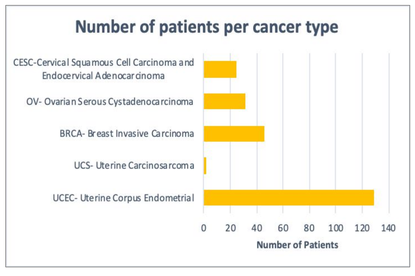
Using the ISB-CGC cohort builder and data explorer, we identified patients who possess the MACF1 mutation across five gynecological cancers. Patient demographic information is detailed below.

Gender	
Female	232
Unknown	1

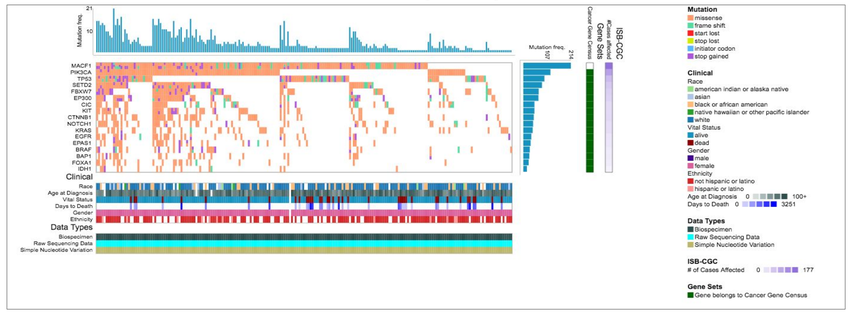
Type of sample	
Primary solid tumor	232
Metastatic	1







Mutation Frequency of MACF1



- ***MACF1*** **mutation frequency is compared to other well-known cancer driver genes including Tumor Protein p53 (*TP53*), B-Raf Proto-Oncogene (*BRAF*) and Epidermal Growth Factor Receptor (*EGFR*)**
- **MACF1 mutations in patients of this cohort include: missense, frameshift and stop-gain mutations.**

Results

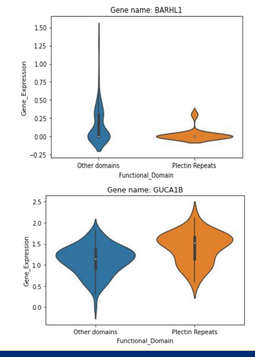
ISB-CGC BigQuery tables of patient cancer somatic mutation data were interrogated to identify both spatial distribution and frequency of mutations across MACF1 functional domains.

The table below shows the distribution of mutations across the functional domains of the five gynecological cancers.

TCGA Cancer Type	Within functional Domains of MACF1	Outside of functional domains of MACF1	Not in coding sequence	Total Num of Mutations
TCGA-BRCA (N=58)	12	44	15	71
TCGA-CESC (N=33)	6	28	23	57
TCGA-OV (n=47)	7	35	10	52
TCGA-UCEC (n=156)	85	331	153	569
TCGA-UCS (n=2)	1	3	1	5

Gene name	# mutations in Plectin repeat region	# of mutations in regions outside of the Plectin repeats	pvalue
BARHL1	25	103	1.09E-05
CTXN3	25	103	1.86E-05
GPR142	25	103	2.78E-05
KRT77	25	103	1.40E-04
GUCA1B	25	103	3.37E-04

- **The plectin repeat region is of particular interest because it is a cytoskeletal linker.**
- Joining ISB-CGC BigQuery tables of RNAseq and somatic mutation data, we observed significant gene expression differences potentially driven by mutations in the plectin repeat region when compared to mutations in other MACF1 functional domains in TCGA-UCEC (Uterine Corpus Endometrial Carcinoma).
- A total of 88 genes showed this pattern and the top 5 are listed in the table above.
- These genes are enriched for the following pathways:
  - Regulation of gene expression in endocrine-committed progenitor cells
  - Formation of the cornified envelope
  - TNF receptor superfamily (TNFSF) members mediating non-canonical NF-kB pathway



Conclusion

This work suggests that MACF1 may be a potential driver gene for gynecological cancers. We initiated a pilot study that uncovered mutations and genomic alterations linking MACF1 to progression of certain cancers. The cloud enabled rapid access to large datasets and facilitated data discovery to help elucidate new biological insights.

Future Directions

- Investigate the roles of MACF1 functional domain in cancer initiation and progression
- Investigate how MACF1 is regulated in cancer signaling pathways
- Analyze gene expression for additional domains starting with the Spectrin Repeats

Acknowledgements

- ISB-CGC has been funded in whole with Federal funds from the National Cancer Institute, National Institutes of Health, under contract No. HHSN261200800001E.
- Mary Mullins Lab, University of Pennsylvania Perelman School of Medicine