Profiling CEACAM7 as a potential target in pancreatic cancer

Syngene
Putting Science to Work

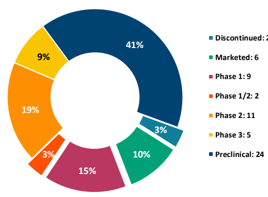
Aditi Bhardwaj, Amit Sonawane*, Lakshmi Subramanian*, Anupriya Verma, Preethiba Gopi, Sreekala Nair*
**Authors contributed equally*

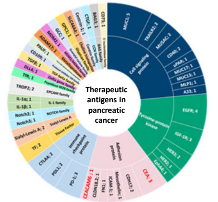
Syngene International Ltd, Biocon Park, SEZ, Bommasandra Industrial Area - Phase-IV Bommasandra-Jigani Link Road, Bangalore 560 099, India.

19th Annual Discovery on TARGET
September 27-30, 2021
BOSTON, MA | SHERATON BOSTON & VIRTUAL [EDT]

Motivation

- Pancreatic cancer is highly aggressive and 7th leading cause of cancer-related deaths accounting for 2.5% of all new cancer diagnoses and 4.5% of all cancer deaths.[1]
- Members of carcinoembryonic antigen cell adhesion molecule (CEACAM) family have been identified as markers in pancreatic cancer.[2]
- CEACAMs 1, 5 and 6 are upregulated in pancreatic cancer; molecules against CEACAM5 and 6 are in early developmental phase.
- Therapeutic antigens targeted with molecules in advanced developmental phases belong to immune checkpoint, cell signaling, tyrosine-protein kinase family.
- Competitive landscape of therapeutic molecules against pancreatic cancer indicates promising opportunities for best-in-class in LM space with only 10% (6/59 molecules) of the space covered by marketed entities.



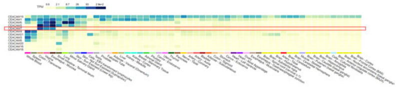


Using an informatics-driven target prioritization strategy, we probed to identify a hitherto untargeted member of CEACAM family as a potential target for pancreatic cancer.

[1]Bray F et al. CA Cancer J Clin. 2018;68(6):394-424. [2]Gebauer F et al. PLoS One. 2014;9(11):e113023. Sources: TABS and AdisInsight

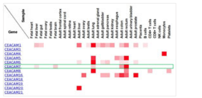
Expression profile of CEACAM7 shows potential for therapeutic targeting

Restricted mRNA expression of *CEACAM7* among 12 family members, in normal tissues e.g., colon, esophagus and vagina, minimizing chances of off-organ effects



Protein expression of CEACAM7 in normal adult tissues is limited to rectum (high), colon (moderate) and lung (low)

- CEACAM7 was reported to be observed only on the apical surface of differentiated epithelial cells in the colorectal mucosa and isolated ductal epithelial cells within the pancreas.[1]
- DICE and HPA databases report low immune cell expression of CEACAM7.



9 CEACAMs are observed to be differentially expressed in cancers, with *CEACAM7* showing significant expression in 5 cancers: ~2.9- and ~2.1-fold overexpression in pancreatic (PAAD) and gastric (STAD) cancers

Differential analysis by ANOVA with a log_2FC (fold change) threshold of 1.0 and q-value cut-off of 0.01 was used to determine significant fold changes for overexpression and downregulation.

[1]Schölzel S et al. Am J Pathol. 2000;156(2):595-605. Sources: GTEx, HPM, GEPIA2, HPA, DICE

Cell lines having high expression of CEACAM7 are useful for *in vitro* profiling

Pancreatic cancer cell lines

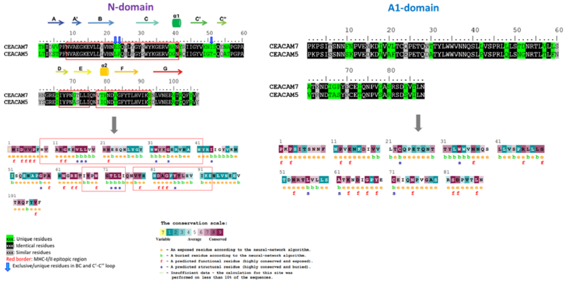
Cell lines	Tissue and Cancer type	Harmonizome: Data Source	Harmonizome: Standardized value	Harmonizome: Clean Value	DepMap: log 2(TPM+1)
BxPC-3	Pancreatic Cancer	GDSC	2.84	9.10	2.94
LS-1034	Cecum adenocarcinoma	GDSC	2.71	9.01	0.33
CL40	Colorectal Cancer	CCLE	2.41	8.25	5.14
HT29	Colorectal Cancer	BioGPS	2.28	2.60	0.61
NCIH508	Colorectal Cancer	CCLE	2.25	8.10	4.54
SNU-C1	Colorectal Cancer	GDSC	2.24	8.36	1.89
CAPAN-1	Pancreatic ductal adenocarcinoma	GDSC	2.18	8.23	2.83
MDAMB175VII	Breast Cancer	CCLE	2.14	7.45	4.09
LoVo	colon adenocarcinoma	GDSC	1.85	7.52	0.86
HT55	Colorectal Cancer	GDSC	1.69	6.83	0.77
NCIH2029	Lung Cancer	CCLE	1.68	5.50	2.78
NCIH1395	Lung Cancer	CCLE	1.61	5.22	2.27
HCC2935	Lung Cancer	CCLE	1.53	5.09	2.41
T3M4	Pancreatic Cancer	CCLE	1.49	4.99	2.37
SW1990	Pancreatic adenocarcinoma	GDSC	1.44	5.61	NA
SNUC1	Colon/Colorectal Cancer	CCLE	1.44	4.93	1.89
PANC-08-13	Pancreatic ductal adenocarcinoma	GDSC	1.42	5.56	0.11
KCIMOH1	Pancreatic Cancer	CCLE	1.38	4.84	1.79

Standard value ≥2nd quartile _CEACAM7 specific_entire dataset

Sources: Harmonizome, Depmap portal

Prediction of targetable unique residues in N and A1 domains of CEACAM7

Despite moderate conservation, several potentially targetable unique residues, in N & A1 domain, are identified by multiple sequence alignment and ConSurf server in case of CEACAM 7 over CEACAM5



Sources: Clustal omega, BioEdit tool, ConSurf server

- 24 unique residues are identified in N-domain of CEACAM7. 18/24 are surface exposed and 5 of these 18 are functionally relevant.
- 21 unique residues in A1-domain of CEACAM7. 16/21 are surface exposed, of which only 1 is functionally relevant.

Attribute	N domain	N domain (MHC-I/II epitopic region)	A1 domain
Unique and exposed residues	Tyr42, Asn50, Ser52, Asn55, Asn56, Thr66, Arg101, Tyr105	Ser24, His39, Ala40, His80, Glu93	Phe10, Asn11, Ile18, Leu42, Leu48, Thr51, Asp52, Leu59, Ile66, Pro68, Ile73, Gly78, Ser80, Pro84, Thr86
Exposed and functionally relevant residues	**Asn2, Val6, Val7**	**Thr79, Ala83**	**Gly67**
Literature reported unique residues[1]	• Glu23 and Ser24 in the BC loop* • Insertion of isoleucine between residues 51 and 52#		---

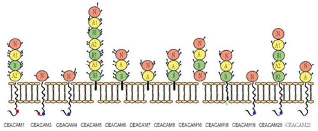
**Instead Leu23 and Pro24 are present in CEACAM5. Loss of Proline in CEACAM7 is reported to reduce the rigidity of the loop*

 # Causes C'' strand distortion leading to formation of tight dimers by CEACAM7

[1]Bonsor DA et al. Acta Crystallogr F Struct Biol Commun. 2015;71(Pt 9):1169-75.

The varied structural features of CEACAM family

CEACAMs are a large family of proteins comprising of 12 members belonging to immunoglobulin (Ig) supergene family, with diverse expression profiles and biological roles in cell adhesion, proliferation, signal transduction and tumorigenesis.[1]

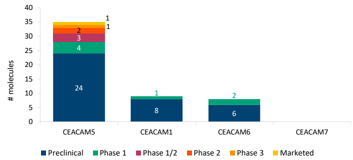


CEACAM proteins [2] comprise of

- Single extracellular **N domain** (orange) - variable (V)-like Ig domain (except CEACAM16 which has two N domains).
- Variable number of extracellular **A domains** (yellow) and **B domains** (green) that are constant C2-like Ig domains.
- **Membrane anchorage** - CEACAM5, CEACAM6, CEACAM7 and CEACAM8 are associated via GPI anchor, whereas CEACAM1, CEACAM3, CEACAM4, CEACAM19, CEACAM20, and CEACAM21 are anchored through transmembrane domains. CEACAM16 is a secreted protein.
- **Cytoplasmic domain** - CEACAM1 CTD has inhibitory ITIM motif, while CEACAM3, CEACAM4, CEACAM19, and CEACAM20 have activating ITAM motifs.

[1]Beauchemin N and Arabzadeh A. Cancer Metastasis Rev. 2013;32(3-4):643-71. [2]Han ZW et al. Invest New Drugs. 2020;38(6):1888-1898.

Clear IP space for CEACAM7 as an anti-cancer target



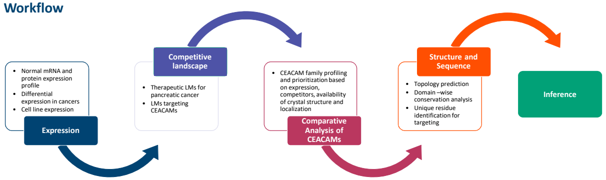
- LMs targeting CEACAMs 5,6 & 1 are reported in various early developmental phases.
- Modalities targeting CEACAM7 are not yet reported in any database providing an opportunity to develop first-in-class therapeutic molecule against CEACAM7.
- A recent study has reported that CAR-T cell therapy against CEACAM7 is likely to be efficacious in pancreatic cancer.[1]

[1]Raj D et al. Clin Cancer Res. 2021;27(5):1538-1552. Source: TABS

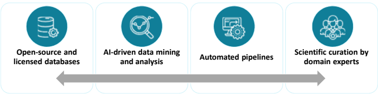
Comparative analysis of CEACAM7 and its paralogs exhibits therapeutic advantage for CEACAM7 and CEACAM5

Gene Symbol	Normal expression	Differential expression in cancer	Function	Therapeutic LMs	Crystal structure	Localization	Target Rating
CEACAM7	Restricted to 3 organs	Yes	Tumorigenesis	No	Yes	Apical	👍
CEACAM1	Ubiquitous	Yes	Tumorigenesis	Yes	**Yes**	Apical and secreted	👎
CEACAM6	Restricted but present in hematopoietic cells	Yes	Tumorigenesis, immuno-regulator, cell adhesion	Yes	**Yes**	Apical	👎
CEACAM8	Blood	Yes	Cell adhesion (in a calcium-independent manner)	No	**Yes**	Membranous	👎
CEACAM5	Restricted to 5 organs	Yes	Tumorigenesis	Yes	Yes	Apical	👍
CEACAM21	Blood, spleen, lymphocytes	Yes	Not clear	Yes	No	Membranous	👎
CEACAM16	Minimal- Nil	No	Maintaining the integrity of the tectorial membrane	No	No	Secreted	👎
CEACAM18	Minimal- Nil	No	Not clear	No	No	Membranous	👎

Workflow



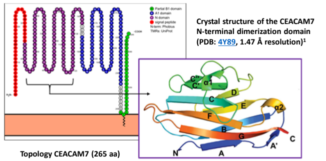
Multi-faceted informatics expertise



By integrating biology and bioinformatics our target diligence endeavor accelerates informed decision-making in the early discovery phase through providing scientific rationale

- Target biology driving the disease
- Protein structure & sequence
- Competitive landscape to aid in therapeutic design
- Unintended off-organ effects by expression studies
- Off-target effect prediction through homolog analysis

N-terminal and A1 domains are moderately conserved among CEACAM7 paralogs

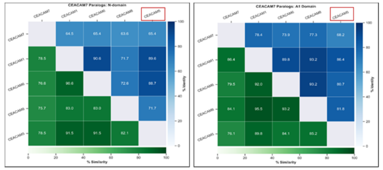


Crystal structure of the CEACAM7 N-terminal dimerization domain (PDB: 4Y89, 1.47 Å resolution)[1]

Topology CEACAM7 (265 aa)

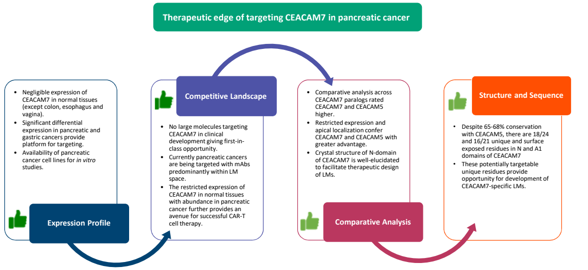
Source: Protter

- N-domain is essential for dimerization and cell adhesion in CEACAM7.[1]
- Known inhibitors of CEACAM 5, 6 (MN-3 and MN-15) bind to N and A1B1 domains.[2,3]



Sources: MatGAT tool, In-house script

- pBLAST analysis indicated CEACAM 1, 5, 6 & 8 are to be closest sequence paralogs of CEACAM 7 N- and A1 domains.
- Structurally, closest homologs of CEACAM7 are CEACAM1 and CEACAM5.[1]

[1]Bonsor DA et al. Acta Crystallogr F Struct Biol Commun. 2015;71(Pt 9):1169-75. [2]Blumenthal RD et al. BMC Cancer. 2007;7:2. [3]Blumenthal RD et al. Cancer Res. 2005;65(19):8809-17.

Inference