UNIONE EUROPEA
Fondo Sociale Europeo

Ministero dell'Università e della Ricerca

PON RICERCA E INNOVAZIONE

Oncoxx Biotech Srl

UPSTATE MEDICAL UNIVERSITY

UNIVERSITÀ DEGLI STUDI DI MESSINA

UNIVERSITY OF MESSINA

DOCTORATE THESIS IN MEDICAL AND SURGICAL BIOTECHNOLOGIES

CYCLE XXXIV

Genomic Risk for Cancer - Activation Profiles of Trop-1 and Trop-2

Author: Khouloud BOUJNAH

Coordinator: Prof. Giovanni *SQUADRITO*

Tutor: Prof. Saverio *ALBERTI*

Co-Tutor/Industrial Tutor: Dr. Emanuela *GUERRA*

US Tutor: Prof. William. G. *KERR*

2021-2022

UNIVERSITA DEGLI STUDI DI MESSINA

DIPARTIMENTO DI MEDICINA CLINICA E SPERIMENTALE

DOTTORATO DI RICERCA PON INDUSTRIALE IN BIOTECNOLOGIE MEDICHE E CHIRURGICHE

_XXXIV_CICLO

Rischio Genomico per Cancro – Profili di Attivazione di Trop-1 e Trop-2

Dottoranda

Khouloud BOUJNAH

Coordinatore del corso:

Prof. Giovanni *SQUADRITO*

Co-Tutor/ Tutor Industriale:

Dr. Emanuela *GUERRA*

Tutor:

Prof. Saverio *ALBERTI*

US Tutor:

Prof. William. G. *KERR*

2021-2022

DÉDICACE

Je dédie notamment ce projet à :

*Mes très chers parents ; A l'âme de mon père « **Ali** », puisse Allah avoir pitié de votre âme, qui m'a quitté sans voir le fruit de son éducation. Lui qui m'a transmis l'amour de vivre, l'amour de sacrifice et celui de continuer à donner sans limite, pour tous ses efforts et ses aides morales et matérielles. Ma mère « **Hassna** » un cœur tendre et plein d'affection qui m'a énormément soutenu et comblé avec son amour. La lumière de mes jours, la source de mes efforts, la flamme de mon cœur, ma vie et mon bonheur.*

Je ne peux pas vous remercier assez pour tout l'amour, les prières, les soins et les sacrifices pour m'avoir éduqué et préparé pour l'avenir.

J'espère avoir répondu aux espoirs que vous avez fondé en moi, que Dieu garde ma mère et la procure santé et longue vie et qu'il offre le paradis à mon père pour sa dernière demeure.

Je vous rends hommage par ce modeste travail en guise de ma reconnaissance éternelle et mon infini amour.

*A mon amour de vie « **Ahmed Ghali** » pour tous ses patiences, son soutien et ses encouragements à tous moments pour atteindre mon objectif.*

*Ma chère sœur « **Sonia** » qui m'a beaucoup soutenu et qui à travers ses encouragements tout au long de ma vie m'a donné la confiance en moi et en ma capacité à atteindre mes buts dans la vie.*

*Mon cher et meilleur frère « **Ahmed** » pour tous ses efforts et ses aides morales et matérielles, pour tous ses encouragements et son soutenu tout au long de ma vie, pour sa confiance en moi, je place entre vos mains le fruit de longues années d'études et d'inquiétude.*

*Mes deux frères « **Walid** » et « **Ramzi** » qui m'ont permis d'apprécier la vie en famille, vous résumez si bien le sens de mot frères.*

*A ma deuxième famille, mon beau-père « **Mahmoud** » ma belle-mère « **Halima** » pour leurs encouragements*

*Ma meilleur amie et sœur « **Khaoula** » qui m'a soutenu dans les moments critiques de ma vie, et qui n'a jamais cesser de fournir des sacrifices pour mon bonheur et bien-être.*

UNIONE EUROPEA
Fondo Sociale Europeo
Ministero dell'Università e della Ricerca
PON RICERCA E INNOVAZIONE
Oncoxx Biotech Srl
UPSTATE MEDICAL UNIVERSITY
UNIVERSITÀ DEGLI STUDI DI MESSINA

ACKNOWLEDGMENTS

Finally, three years of Doctorate finished, and my graduation day is in the horizon after so many struggles. I still can't believe it; it seems like yesterday when I started studying here. These years past in a blink of an eye, I will miss all people I met during this period!

First and foremost, praises and thanks to Allah, the Almighty, for His showers and blessings through my Doctorate to complete the research successfully.

This work was carried out at the ***University of Messina, Department of Clinical and Experimental Medicine*** *and* ***Department of Biomedical and Dentistry Sciences and Morphological, Functional Imaging,*** *in collaboration with* ***Oncoxx Biotech****, the* ***Laboratory of Cancer Pathology, Center for Advanced Studies and Technology (CAST) of the University of Chieti*** *and* ***SUNY University, Syracuse, New York.***

*I warmly thank my Professor "****Giovanni Squadrito****", **Coordinator of the Research Doctorate in Medical and Surgical Sciences** at the **University of Messina** for entrusting me with this work and giving all his confidence for its completion. I am very thankful to him for his guidance and advice that he devoted to me during my Doctorate.*

*I would like to express my gratitude and appreciation for my Tutor and supervisor Professor "****Alberti Saverio****", **Head of Medical Genetics Section, Department of Biomedical and Dentistry Sciences and Morphological, Functional Imaging, University of Messina** for his human and professional qualities, for his advice and instructive criticism, and his encouragement, that have been invaluable throughout this study. I am very grateful to him for his guidance and advice that he devoted to me throughout this project. Without his precious support it would have not been possible to conduct this research.*

*My sincere admiration and deep gratitude to my industrial Tutor Dr. "****Emanuela Guerra****" at **Oncoxx Biotech S.r.l**, for her supervision, trust, help, availability, encouragement, advice, and instructive criticisms which have been of great help to me for the smooth running of this work, and for conducting to realization the molecular and genetic research.*

*All my sincere gratitude for professor "**Marco Trerotola**", head of **Laboratory of Cancer Pathology (CAST)** at the **University of Chieti**. for all his support, encouragement, and availability. I would like to express all my deep gratitude for having welcomed me in his research laboratory and for having offered me the means and many technical facilities in my experiments during my study project*

*A very special gratitude goes out to my Co-Tutor Professor "**William Kerr**" at the **Department of Microbiology and Immunology**, **University of New York - Upstate Medical University, Syracuse** for having allowed me the opportunity to join his team and secondly to all his laboratory members.*

*I could not forget to thank my colleagues at the **Laboratory of Cancer Pathology (CAST)** **Martina Ceci** for her help, support and standing my endless questions and **Ludovica Pantalone** for her encouragement, together with all my colleagues at the **Medical Genetics Laboratory** at the **University of Messina***

*I would like to thank the members of the **Exam Committee** who agreed to judge my work, the **Coordinator**, the **Reviewers** and the **"Collegio dei Docenti"**.*

I place on record, my sense of gratitude to one and all, who directly or indirectly, have lent their hand in this venture.

I want to dedicate this thesis to everyone I met during my Doctorate Course, to everyone that supported me and encouraged me! Do not give up, if you fall seven times, stand up eight times. Don't let the fear and life circumstances bring you down, you all are great and special!

Finally, I would like to thank my family, in particular my love, my wonderful mother who constantly helped and encouraged me during my long years of study.

Thank you again!

AUTHOR CONTRIBUTIONS

***Article I*: «Trop-2 cleavage by ADAM10 is an activator switch for cancer growth and metastasis » April 2021 [1]**

Marco Trerotola*:* Conceptualization, Investigation, Formal analysis, Writing – Original draft, Visualization.

Emanuela Guerra*:* Conceptualization, Investigation, Formal analysis, Writing – Original draft, Visualization.

Zeeshan Ali*:* Investigation, Formal analysis.

Anna Laura Aloisi*:* Investigation, Formal analysis.

Pasquale Simeone*:* Investigation, Formal analysis, Data curation.

Angela Acciarito*:* Investigation, Formal analysis.

Paola Zanna*:* Investigation, Formal analysis.

Giovanna Vacca*:* Investigation, Formal analysis.

Martina Ceci*:* Investigation, Formal analysis, Data curation.

Antonella D'Amore*:* Investigation, Formal analysis.

Khouloud Boujnah*:* Investigation, Formal analysis, Data curation.

Valeria Garbo*:* Investigation, Formal analysis, Data curation.

Antonino Moschella*:* Investigation, Formal analysis, Data curation.

Rossano Lattanzio*:* Resources, Investigation, Formal analysis.

Saverio Alberti*:* Conceptualization, Formal analysis, Writing – Original draft, Visualization, Supervision, Project administration, Funding acquisition

Article II: «Trop-2 induces ADAM10-mediated cleavage of E-cadherin and drives EMT-less metastasis in colon cancer» September 2021 [2]

E.G., M.T., V.R. R.T., G.V., M.C., K.B., R.Z., P.S., and **S.A.** performed the cell biology tests, biochemical analyses and *in vivo* assays. R.d.L. and U.H.W. performed the DNA macroarray and microarray analyses for tumor cell line profiling. M.T. identified the gene interaction networks. **R.L., D.F.A., R.D., M.T.R., A.P., P.Q., M.P., E.B., D.A., G.P., C.D.L., M.P.** and **X.-F.S.** collected the human tumor samples, constructed the tissue microarrayS and performed the IHC analyses. **S.B., A.C., P.D.R., A.M., V.G.** and **S.A.** performed the quantitative meta-analyses of surgical pathology and *in silico* data. **L.A.** and **R.L.** performed the statistical analyses. **E.G., M.T.** and **S.A.** planned the studies, evaluated the data, and wrote the article.

Emanuela Guerra: Conceptualization, Investigation, Formal analysis, Data curation, Writing – Original draft, Writing – review & editing. Visualization.

Marco Trerotola: Conceptualization, Investigation, Formal analysis, Writing – Original draft, Writing – review & editing. Visualization.

Valeria Relli: Investigation, Formal analysis.

Rossano Lattanzio: Resources, Investigation, Formal analysis.

Romina Tripaldi: Investigation, Formal analysis.

Giovanna Vacca: Investigation, Formal analysis.

Martina Ceci: Investigation, Formal analysis.

Khouloud Boujnah: Investigation, Formal analysis, Data curation.

Valeria Garbo: Investigation, Formal analysis, Data curation.

Antonino Moschella: Investigation, Formal analysis, Data curation.

Romina Zappacosta: Investigation, Formal analysis.

Pasquale Simeone: Investigation, Formal analysis, Data curation.

Robertde Lange: Resources, Formal analysis.

Ulrich H. Weidle: Resources, Formal analysis.

Maria Teresa Rotelli: Resources, Formal analysis.

Arcangelo Picciariello: Resources, Formal analysis.

Raffaella Depalo: Resources, Formal analysis.

Patrizia Querzoli: Resources, Formal analysis.

Massimo Pedriali: Resources, Formal analysis.

Enzo Bianchini: Resources, Formal analysis.

Domenico Angelucci: Resources, Formal analysis.

Giuseppe Pizzicannella: Resources, Formal analysis.

Carla Di Loreto: Resources, Formal analysis.

Mauro Piantelli: Resources, Investigation, Formal analysis.

Laura Antolini: Investigation, Formal analysis.

Xiao-Feng Sun: Resources, Formal analysis.

Donato F. Altomare: Resources, Formal analysis.

Saverio Alberti: Conceptualization, Formal analysis, Writing – Original draft, Writing – review & editing. Visualization, Supervision, Project administration, Funding acquisition

STATEMENT OF ETHICAL CONDUCT

In this thesis, all of the procedures involving animals and their care were conducted in compliance with institutional guidelines, national laws, and international protocols (D.L.No.116, G.U., Suppl.40, Feb. 18, 1992; No.8, G.U., July 1994; UKCCCR Guidelines for the Welfare of Animals in Experimental Neoplasia; EEC Council Directive 86/609, OJL358. 1, Dec. 12, 1987; Guide for the Care and Use of Laboratory Animals, United States National Research Council, 1996). Preclinical protocols were approved by the Italian Ministry of Health (Prog. 19, 2006) and by the Interuniversity Animal Research Ethics Committee (CEISA) of Chieti-Pescara and Teramo Universities (Prot.26/2011/CEISA/PROG/16).

Studies on human tumor samples were approved by the Italian Ministry of Health (RicOncol RF-EMR-2006-361866, 2006) and by the ethical committee of the University of Regensburg (Germany).

CONTENT

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ADAM	A Disintegrin And Metalloproteinase
ADC	Average Difference Change
ADP	Adinosin Difosfate
AEBSF	4-(2-Aminoethyl)Benzenesulfonyl Fluoride Hydrochloride
AIRE	Autoimmune Regulator
ALK	Alk Receptor Tyrosine Kinase
AMV	Avian Myeloblastosis Virus
ANNOVAR	Annotate Variation
ANOVA	Analysis Of Variance
APC	Adenomatous Polyposis Coli
APP	Amyloid Beta Precursor Protein
ARE	Annotate Variation
ASB4	Ankyrin Repeat And SOCS Box-Containing 4
BACE	B-Site App Cleaving Enzyme 1
BAD	Browser Extensible Data
BAM	Binary Alignment Map
Biotin-11-CTP	Biotin-11-Cytidine-5′-Triphosphate, Tetralithium Salt
BLT	Bead-Linked Transposomes
BRAF	B-Raf Proto-Oncogene, Serine/Threonine Kinase
BRCA2	Breast Cancer 2
BSA	Bovine Serum Albumin
BWA	Burrows-Wheeler Aligner
CADD	Combined Annotation-Dependent Depletion
CDK	Cyclin Dependent Kinase
CDKN2A	Cyclin-Dependent Kinase Inhibitor 2a
CIP	Calf-Intestinal Alkaline Phosphatase
CNV	Copy Number Variation
COSMIC	Catalogue Of Somatic Mutations In Cancer
CPD	Cysteine-Poor Domain
CRC	Colorectal Cancer
CRD	Cysteine-Rich Domain
CREB	Camp Responsive Element Binding Protein 1
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CSV	Comma-Separated Value
CTE	Congenital Tufting Enteropathy
CUSHAW2	Long Read Alignment Based On Maximal Exact Match Seeds
DAG	Deacylglycerol
DAPI	4′,6-Diamidino-2-Phenylindole
DAVID	Database For Annotation, Visualization, And Integrated Discovery
dbGaP	Database Of Genotypes And Phenotypes
dbSNP	Single Nucleotide Polymorphism Database
dbVar	NCBI Database Of Human Genomic Structural Variation
DGV	Database Of Genomic Variants
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
EBSS	Earle's Balanced Salt Solution
ECD	Extracellular Domain
ECM	Extracellular Matrix
EDTA	Ethylenediamine Tetraacetic Acid
EGF	Epidermal Growth Factor
EGFP	Enhanced Green Fluorescent Protein
EGFR	Epidermal Growth Factor Receptor
EGP-1	Tumor Associated Calcium Signal Transducer 2
EGTA	Ethylene Glycol-Bis(B-Aminoethyl Ether)-N,N,N′,N′-Tetraacetic Acid
EMBOSS	European Molecular Biology Open Software Suite
EMT	Epithelial-Mesenchymal Transition
ENCODE	Encyclopedia Of DNA Elements
ENIGMA	Enhancing Neuroimaging Genetics Through Meta-Analysis
EPCAM	Epithelial Cell Adhesion Molecule
ERG	Ets Transcription Factor Erg
ESLI	Ethylene Glycol-Bis(B-Aminoethyl Ether)-N,N,N′,N′-Tetraacetic Acid
ESP	Exome Sequencing Project
FACS	Fluorescence-Activated Cell Sorting
FAK	Protein Tyrosine Kinase 2
FANTOM	Functional Annotation Of The Mammalian Genome
FAP	Familial Adenomatous Polyposis
FASTA	Fast Adaptive Shrinkage Threshold Algorithm
FAVR	Filtering And Annotation Of Variants That Are Rare
FBS	Fetal Bovine Serum
FCS	Fetal Calf Serum
FreeBayes	Bayesian Haplotype-Based Genetic Polymorphism Discovery And Genotyping
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GATK	Genome Analysis Toolkit
GDLD	Gelatinous Drop-Like Corneal Dystrophy

GERP	Glyceraldehyde-3-Phosphate Dehydrogenase
GFP	Green Fluorescent Protein
gnomAD	Genome Aggregation Database
GRC	Genome Reference Consortium
GSC	Goosecoid Homeobox
GSK	Goosecoid Homeobox
GTP	Gtp-Binding Proteins
HDR	Homology Directed Repair
HEPES	4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic Acid
HET	Heterozygote
HGMD	Human Gene Lesion Database
HLA-DR	Human Leukocyte Antigen – Dr Isotype
HUGO	Hugo Gene Nomenclature Committee
ICD	Intracellular Domain
IFG	Insulin-Like Growth Factor 1
IGF-1	Nsulin Like Growth Factor 1
IGF-1R	Insulin Like Growth Factor 1 Receptor
IGV	Viewer For Integrative Genomics
IHC	Immunohistochemistry
IPMN	Intraductal Papillary Mucinous Neoplasms
IRAK-1	Interleukin 1 Receptor Associated Kinase 1
IRAK-3	Interleukin 1 Receptor Associated Kinase 3
IRIS	Immuno-Reactivity Identification System
ITF-2	Immunoglobulin Transcription Factor-2
JAK	Janus Kinase 1
JAK3/2	Janus Kinase 3/
JNK	Mitogen-Activated Protein Kinase 8
KEGG	Kyoto Encyclopedia Of Genes And Genomes
KLF8	Kruppel Like Factor 8
LEF	Lymphoid Enhancer Binding Factor
LOVD	Leiden Open Variation Database
LRT	Likelihood Ratio Test
LSAB	Labelled Streptavidin Biotin
LSM	Laser Scanning Microscopy
MALDI MS	Matrix-Assisted Laser Desorption/Ionization Mass Spectometry
MANTA	Microbiota And Phenotype Correlation Analysis Platform
MAPK/ ERK	Mitogen-Activated Protein Kinases
MCN	Mucinous Cystic Neoplasm
MDK	Midkine
MEK1	Mitogen-Activated Protein Kinase Kinase 1
MHC class II	Class Ii Major Histocompatibility Complex Transactivator
MIA	Melanoma Inhibitory Activity
MINT	Molecular Interaction Database
MMP (2/7/9)	Matrix Metallopeptidase
MNPs	Multinucleotide Polymorphisms
MODELLER	Program For Comparative Protein Structure Modelling By Satisfaction Of Spatial Restraints
MOLMOL	Molecular Analysis And Display Program
MOPS	3-(N-Morpholino)Propanesulfonic Acid
MOWSE	Molecular Weight Search
MS	Mass Spectometry
MSI-H	Microsatellite Instability (High Level)
MTE-4-14	Mouse Thymic Epithelium-4-14
MYC	Myc Proto-Oncogene, Bhlh Transcription Factor
NCBI	National Center For Biotechnology Information
NF-rB	Nuclear Factor Kappa-Light-Chain-Enhancer Of Activated B Cells
NGS	Next Generation Sequencing
NHEJ	Non-Homologous End Joining
NHLBI	National Heart Lung And Blood Institute
NK/T	Natural Killer T Cells
NRG	Neuregulin 1
NSCLC	Non-Small-Cell Lung Carcinoma
OMIM	Online Mendelian Inheritance In Man
Opti-MEM	Reduced-Serum Medium- Minimal Essential Medium
P53	Tumor Protein P53
PAM	Protospacer Adjacent Motif
PB	Binding Buffer
PBS	Phosphate-Buffered Saline
PCDNA	Plasmid Cloning Dna
PCR	Polymerase Chain Reaction
PDAC	Pancreatic Ductal Adenocarcinoma
PDB	Protein Data Bank
PDGF	Platelet-Derived Growth Factor
PD-L1	Programmed Death-Ligand 1
PFAM	Protein Families (Database)
PFAM	Protein Families (Database)
Phylop score	Is A Measure Evolutionary Conservation At Individual Alignment Sites
PIN	Prostatic Intraepithelial Neoplasia
PKC	Protein Kinase C
PLC	Phospholipase C
PMF	Probability Mass Function
PMSF	Phenylmethylsulfonyl Fluoride
PolyPhen2	Polymorphism Phenotyping V2

PROCHECK	Is To Checks The Stereochemical Quality Of A Protein Structure
PRRX1	Paired Related Homeobox 1
PS-2	Presenilin 2
PTEN	Phosphatase And Tensin Homolog
PU.1	Spi-1 Proto-Oncogene
PVDF	Polyvinylidene Difluoride
RACE	Rapid Amplification Of Cdna Ends
RAS	Rat Sarcoma Viral
RIP	Regulated Intramembrane Proteolysis
RISC	RNA-Induced Silencing Complex
RNA	Ribonucleic Acid
RPMI	Roswell Park Memorial Institute Medium
RTK	Receptor Tyrosine Kinase
SAGE	Serial Analysis Of Gene Expression
SAM	Sequence Alignment Map
SDS-PAGE	Sodium Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis
SEM	Standard Error Of Mean
SIFT	Sorting Intolerant From Tolerant
SIX1	Six Homeobox 1
SMAD4	Smad Family Member 4
SMART	Simple Modular Architecture Research Tool
SMART	Simple Modular Architecture Tool
SNAI2	Snail Family Transcriptional Repressor 2
SNP	Single-Nucleotide Polymorphisms
snpEff	Snp Effect
SNV	Single-Nucleotide Variant
SOAP	Short Oligonucleotide Analysis Package
SPIDEX	Is Annotation Database
SRC	Src Proto-Oncogene, Non-Receptor Tyrosine Kinase
STAT1	Signal Transducer And Activator Of Transcription 1
SWISS	Wisconsin Hierarchical Analysis Tool For Incidental Findings
TACE	Tumour Necrosis Factor-Alpha Converting Enzyme
TACSTD1/2	Tumour-Associated Calcium Signal Transducer1/2
TAE	Tris-Acetate-EDTA Buffer
TBS	Tris-Buffered Saline
TCF-1	Hnf1 Homeobox A
TCF-4	Transcription Factor 4
TCGA	The Cancer Genome Atlas
TEV	Tobacco Etch Virus
TGFβ	Transforming Growth Factor Beta 1
TIDDIT	
TMA	Tissue Microarrays
TNFα	Tumor Necrosis Factor Alpha
TRADD	Tnfrsf1a Associated Via Death Domain
TROP2	Trophoblast Cell-Surface Antigen2
TSV	Tab-Separated Values
TWIST1	Twist Family Bhlh Transcription Factor 1
TWIST2	Twist Family Bhlh Transcription Factor 2
VarDict	A Novel And Versatile Variant Caller For Both DNA- And RNA-Sequencing Data
VCF	Variant Call Format
VEGF	Vascular Endothelial Growth Factor
VSVG	Vesicular Stomatitis Virus G Glycoprotein
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing
WNT	Proto-Oncogene Int-1 Homolog
XHMM	Exome Hidden Markov Model
YOABS	Yet Other Aligner Of Biological Sequences
ZEB1	Zinc Finger E-Box Binding Homeobox 1
ZEB2	Zinc Finger E-Box Binding Homeobox 2
Δcyto	Cytoplasmic-Tail Deletion Mutant
ΔHIKE	Hike Deletion Mutant

ABSTRACT

To understand cancer, it is important to interpret, analyze and understand the patient's genomic background, and its profoundly influence on treatment outcomes. However, all the necessary informations are hidden inside the 3 billion bases of the human genome. Hence, new technologies were designed to analyze genomic variation and to develop bioinformatic tools for analyzing sequence data, to identify new biomarkers capable to predict the individual patient's response to specific cancer treatments and correlated toxicity profiles.

As a result of the introduction of next-generation sequencing (NGS), it has become much easier to establish correlations between sequence variants in the human genome and illness risk. It is the goal of the Moli-sani study to uncover genetic and environmental risk factors for cancer and chronic degenerative illnesses in a Mediterranean-derived population. In this study, a large-scale sample (26,000 participants) was recruited and followed up for an average of 10 years after the start of the study [3-5]. The final goal of this project, which involves 1.500 subjects, is the definition of cancer disease risk factors through Next Generation Sequencing (NGS) approaches performed on Mediterranean populations, taking into consideration information such as information on lifestyle and life environment. The mutational landscape, in specific the presence and frequency of sequence mutations (coding and non-coding sections) and changes in the number of copies of the gene, obtained by NGS provide the basis to study multi-factorial interactions. The genetic variants that are identified as associated with cancer disease risk are validated by molecular and cellular analyses. Among them, cancer-related changes in the activation profile of the Trop-1 and Trop-2 genes were identified as cancer risk factors in humans, and as benchmarks for genetic risk factor genomic, transcriptomic, and functional analysis.

Studies in this thesis take advantage of these technologies and explore the possibility to identify broad genomic factors behind cancer development. The studies here presented have taken advantage of genomic-transcriptomic technologies to explore the functional role of Trop-2 in cancer progression. This was investigated in particular through the expression of designer mutants of the Trop-2 molecule, somatic knock-out and/or chemical inhibition of survival mechanisms in cancer cell systems (**Article I**, **Article II**).

In the work presented in **Article I** [1] we revealed that Trop-2 is cleaved at the first thyroglobulin domain loop of its extracellular region, between R87 and T88, as shown by antibody targeting and N-terminal Edman degradation. Molecular modeling revealed that this cleavage alters the Trop-2 structure and biological function. We have demonstrated that Trop-2 recruits ADAM10 and is activated by ADAM10-mediated cleavage to trigger molecular pathways for tumor growth and metastasis. We have then demonstrated that Trop-2 activation by post-translational processing promotes tumor growth and metastasis. Our findings that the ADAM10-mediated cleavage of Trop-2 at the R87-T88 site is necessary to activate the transformed growth stimulatory activity of Trop-2 and promote its metastasis driver function may pave the way for the development of new therapeutic precision strategies.

In the work presented in **Article II** [2] we identified, through transcriptomic analysis by NGS of colon cancer cells (metastasis models), that Trop-2 is a unique upregulated gene. Transcriptomic analysis of distinct cases of primary tumors and metastases, showed no down-regulation of CDH1 by transcription factors for epithelial-to-mesenchymal transition, thus suggesting that the pro-metastatic activity of Trop-2 operates through alternative mechanisms. In addition, and for the first time, we showed that the binding of Trop-2 to E-cadherin stimulated the proteolytic cleavage of the intracellular domain of E-cadherin by ADAM10. Detachment of E-cadherin from β-actin, loss of cell-cell adhesion, acquisition of invasive capability, and membrane-driven activation of β-catenin signaling, were further enhanced by the tail-less ΔcytoTrop-2 mutant. The Trop-2/E-cadherin/ β-catenin program led to anti-apoptotic signaling, increased cell migration, and enhanced cancer-cell survival.

I. GENERAL INTRODUCTION

The advent of Next Generation Sequencing (NGS), which can reveal genetic variations across whole genomes, has greatly increased the potential for the identification of genetic "risk profiles" for cancer. This also poses challenges in the handling and interpretation of such "big data", which calls for robust, validated laboratory and bioinformatic approaches that can reliably detect meaningful associations. In this context, NGS datasets from small, well defined experimental groups allow to benchmark models and strategies of analysis that can then be translated in "big data" settings. Diseases with a suspected genetic base can provide such experimental groups, where NGS approaches produce a wealth of data that overcome the limits of the traditional "one gene at a time" approach, but at the same time challenge our ability to discriminate between causative genetic variations and noise.

Cancer is characterized by the loss of control of the growth, division, and spread of transformed cells, which then leads to the formation of malignant tumors, that invade and destroy adjacent tissues or other regions of the body through metastasis [6]. Tumorigenesis is a multistep process that involves genetic and epigenetic changes, resulting in the reorganization of the cellular regulatory networks and interactions with the microenvironment. A key feature of this process is the accumulation of mutations in oncogenes and tumor suppressor genes.

The advent of genotyping and "high-throughput" sequencing has allowed the identification of high-frequency variants associated with cancer. The final goal of this thesis has been the definition of cancer disease risk factors through Next Generation Sequencing (NGS) approaches performed on Mediterranean populations. Association of genomic and transcriptomic variants that have been identified as associated with cancer risk was validated by molecular and cellular analyses. Genetic, bioinformatic and functional analyses that are performed on experimental sets with reduced complexity allowed to develop and optimize laboratory and bioinformatic approaches.

NGS approaches were also applied to whole-transcriptome analyses of the networks that orchestrate cancer cell signaling. Previous studies showed that the overexpression of the transmembrane signal transducer Trop-2 [7, 8] induces aggressive metastasization of cancer cells

and disanchors E-cadherin from the β-actin cytoskeleton for β-catenin activation. Metastasis has been associated with epithelial-mesenchymal transition (EMT) by metastasizing cells. Hence, we investigated the transcriptome of paired spleen cancer-liver metastases from Trop-2 expressing cells in preclinical models for the presence of a canonical EMT signature.

In vitro and in vivo approaches were implemented to validate and dissect the functional role and the mechanism of action of key molecules that drive the oncogenic process, such as the cancer driver Trop-2 [7, 8]. The functional role of Trop-2 in cancer progression was further investigated through expression of designer mutants, somatic knock-out and/or chemical inhibition of survival mechanisms in cancer cell systems (Published articles on Neoplasia 1 and 2) [1, 2].

II. BIBLIOGRAPHIC SYNTHESIS

This chapter presents an overall description of the thesis and introduces the main components and technology used in this research.

Cancer is the second largest cause of mortality in the United States and Italy (N.d.C), but it is also a significant cause of emotional and physical anguish. Patients and most of the public frequently inquire on whether a single cure for cancer will be discovered. The real answer to this fundamental question is that cancer is a collection of distinct mutations, each having a unique natural history and therapeutic response.

A. FREQUENT CANCER TYPES

Cancer can affect most organs in the body. The most frequent cancer types are:

1. Breast Cancer

Breast cancer remains the most frequent cancer in the world. As its molecular hallmarks have been extensively characterized, the way breast cancer is viewed has changed dramatically, and now includes immunohistochemical markers (e.g., ER, PR, HER2, ERBB2, and proliferation marker protein Ki-67, MKI67), genomic markers (e.g., BRCA1, BRCA2, and PIK3CA), and immunomarkers (eg, tumour-infiltrating lymphocytes and PD-L1) [9]. This complex disease occurs when there is dysregulation in signalling networks in the breast cells due to genetic and epigenetic changes. Many signaling pathways have been linked to the development of breast cancer. The HER-2 tyrosine kinase route [10], the hedgehog signaling pathway [11], the p53 pathway, and the PTEN pathway are the most important.

2. Colorectal Cancer

Colorectal cancer (CRC) is the third most common diagnosis and the second most lethal malignancy in both genders combined. CRC is associated with both strong environmental and genetic risk factors. Normal colonic epithelial cells are transformed into precancerous lesions (adenomas) and ultimately into invasive carcinoma by the accumulation of somatic and/or germline genetic mutations.

According to the clonal mutation evolution theory, colonic carcinogenesis is a process in which cells gain survival advantages and allows to develop more mutations providing other cancer hallmarks, such as proliferation, invasion, metastasis [12, 13].

Chromosomal instability, mismatch repair, and hypermethylation are three major molecular pathways linked to CRC. The chromosomal instability pathway is characterized by a gain of mutations that disrupts the balance of oncogenes and tumor suppressors, as seen with mutations in the adenomatous polyposis coli (APC), a hallmark of familial adenomatous polyposis (FAP)[14] [12]. Cells with a lack of DNA mismatch repair (dMMR), commonly MLH1 or MSH2, accumulate errors within the genome, which are then transmitted over cell generatioins, resulting in high levels of microsatellite instability (MSI-H), a defining feature of Lynch syndrome. BRAF and MLH1 gene expression can be activated or silenced by CpG hypermethylation of DNA. Sporadic oncogene somatic mutations (RAS, SRC, MYC) have been linked to CRC, with Ki-RAS having the most clinical significance [13, 15]

3. Prostate Cancer

Prostate cancer is the second most diagnosed cancer in men and the fifth most frequent cause of death globally. Advanced age, race, genetic variables, and family history are prostate cancer risk factors [16].

The prostate undergoes a multistep malignant transformation that begins with prostatic intraepithelial neoplasia (PIN), progresses to localized prostate cancer, advanced prostate adenocarcinoma with local invasion, and finally metastatic prostate cancer. As a result, metastatic illness is the major cause of death from prostate cancer. Lymph nodes adjacent to the primary tumors are frequently the initial sites of metastases [17], that then spread to the liver, lungs, and bones. Once prostate cancer cells have colonized the bone marrow, they interact with the bone microenvironment, causing a "vicious cycle" of bone creation and destruction, which promotes cancer cell survival and tumor growth. Based on multiple published studies by Logothetis [18], Guise [19], Body [20] and their colleagues, growth factors secreted by prostate cancer cells, such as endothelin 1 (ET-1) and adrenomedullin, as well as fibroblast growth factors (FGFs), platelet-derived growth factor (PDGF), and bone morphogenetic proteins (BMPs), can stimulate osteoblast

activation and new bone formation via paracrine signaling. Several large-scale genomic studies in primary prostate tumors and Metastatic Castration-Resistant Prostate Cancer (MCRPC) have found recurrent DNA copy number changes, mutations, rearrangements, and gene fusions.

Furthermore, multiple genomic and epigenomic studies have demostrate that signature genetic alterations target different pathways that are important in prostate development and function, regulation of cell survival and energy metabolism, cell proliferation, and DNA repair. Such signature alterations encompass the Androgen signaling pathway (AR), PI3K–PTEN, TGF-/SMAD4 pathway, and WNT, in most metastatic prostate cancers and in a high proportion of primary prostate cancers [21].

4. Pancreatic Cancer

Pancreatic cancer is one of the most lethal diseases worldwide. Recent research has discovered that pancreatic tumors include a high percentage of cancer stem cells (CSCs), which are resistant to chemotherapy and hence escape chemotherapy, promoting tumor recurrence. In pancreatic cancer, the epithelial to mesenchymal transition (EMT) is strongly linked to metastasis, the formation of CSCs, and therapy resistance.

Pancreatic ductal adenocarcinoma (PDAC), acinar cell carcinoma, pancreatoblastoma, solid pseudopapillary neoplasm, serous cystadenoma, and pancreatic endocrine tumors are examples of distinct pancreatic malignancies that resemble the normal cellular counterparts in the pancreas. Pancreatic intra epithelial neoplasia (PanIN), mucinous cystic neoplasm (MCN), and intraductal papillary mucinous neoplasms (IPMN) are three different forms of precursor lesions that lead to the most frequent pancreatic neoplasm (PDAC).

The mutational landscape of pancreatic cancer was investigated using whole genome sequencing analysis. This revealed an average of 119 somatic chromosomal structural variants per patient. Intra-chromosomal deletions, duplications, tandem duplications, inversions, fold back inversions, amplified inversions, and intra chromosomal rearrangements made up most of these variants. At least 12 separate fundamental signaling pathways were altered in 67-100 percent of the cancers.

The Kirsten rat sarcoma viral oncogene homolog (Ki-Ras), P16/ cyclin-dependent kinase inhibitor 2A (CDKN2A), tumor protein P53 (P53), breast cancer 2 early onset (BRCA2), and SMAD family member 4 (SMAD4)/ deleted in pancreatic carcinoma 4 are the most observed signature genetic lesions in pancreatic cancer (DPC4). In pancreatic cancer patients, Ki-Ras mutations are frequently thought to be the first event to occur in adult cells, followed by P16 mutations and then P53 and SMAD4 loss [22].

B. ONCOGENES AND TUMOR SUPRESSOR GENES MUTATIONS

Tumorigenesis is a multistep process that involves genetic and epigenetic changes, resulting in the reorganization of the cellular regulatory networks and altered interactions with the microenvironment. A key factor in this process is the accumulation of mutations in oncogenes and suppressor genes.

Suppressor genes are critical in the development of malignancies. Tumor suppressor genes encode proteins that help control cell development, and frequently inhibit cell proliferation. Among proto-oncogenes are:

- ✓ Membrane receptors for growth factors, for example erbB, the EGF (epidermal growth factor) receptor.
- ✓ G proteins (GTP binding proteins) such as proto-oncogenes of the ras family.
- ✓ Proteins that control genes in the nucleus: this is the case for fos, jun, c-myc, erbA ...

An oncogene is a gene that stimulates the growth of cancer. Oncogenes largely derive from alteration or overexpression of a normal gene involved in cell division control, known as a proto-oncogene [23]. In human neoplasms, three genetic processes are involved in the conversion of a proto-oncogene to an active oncogene.

Mutations activate proto-oncogenes by causing structural changes in their encoded protein (by substitutions, deletions, and insertions). A second mechanism is gene amplification, i.e., an increase in the number of copies of a gene within a cell's genome. A third mechanism is represented by chromosome rearrangements, which primarily involve translocation or inversions, which lead to rearrangement of parental genes within the chromosomes, which are typically located at the

borders of the structural rearrangement. This change places the proto-oncogene in a new genomic context, where it will be regulated in a different manner.

The following sections will briefly present the main hallmarks and signaling pathways involved in the most common cancer types, breast, colorectal, prostate, and pancreas. Years of research led to identify two key proteins, Trop-1 and Trop-2, that play important roles in various types of human cancers and are among the most overexpressed proteins in the major types of cancer cells. Trop-1 and Trop-2 are the main research area of my laboratory group's work, and have been central to my thesis work, as presented in the successive sections.

1. EpCAM: gene, protein, structure, and function in normal and cancerous (malignant) cells

In the 1970s, functional screens for novel tumor-specific cell surface antigens were performed, and monoclonal antibodies were generated by immunization of mice with colorectal cancer cells and the production of antibody-secreting hybridomas [24].

EpCAM, or Epithelial Cell Adhesion Molecule (known also on the name TACSTD1 [25]), is a type-1 transmembrane glycoprotein that was discovered as a new family of transmembrane cell adhesion molecules (CAMs) to previous families that included cadherins, selectins, integrins, and immunoglobulin (Ig)-like CAMs [25].

Since then, EpCAM has been investigated in a wide range of human carcinomas and normal epithelial tissues. EpCAM expression is dramatically increased in adenocarcinomas and squamous cell carcinomas in humans [26].

EpCAM is being developed as a marker for prognosis, diagnosis, and therapeutic intervention for epithelial malignancies due to its overexpression in cancer and accessibility on the cell surface. EpCAM serves a morpho-regulatory role in normal epithelia and stem/progenitor cells, and it may actively drive tumor growth in cancer cells. This has led EpCAM to being targeted as a molecular biomarker. Several EpCAM-targeting treatment techniques have been developed [27].

a) EpCAM gene

The human EpCAM gene is found on chromosome 2 (location 2p21) and is estimated to be~42 kb in size [25]. The EpCAM encoding gene contains 9 coding exons, according to a comparison of genomic and cDNA sequences: Exons 1 to 6 encode EpCAM's extracellular domain, which includes the signal peptide, whereas exon 7 encodes the transmembrane area. Exons 8 and 9 code for the intracellular domain . The EpCAM gene is found in many different species, including mice, rats, and zebrafish [28].

b) EpCAM protein, structure, and functional domain

As we mentioned previously, the EpCAM gene encodes a polypeptide of 314 amino acids (aa) which constitute ~ 40kDa in molecular weight. Structurally, is divided into three essential domains: an extracellular domain including the signal peptide (242 aa), a single transmembrane domain (23aa), and a short intracellular domain (26aa) [25].

(1) The extracellular domain of human EpCAM (EpEX)

The extracellular domain of EpCAM is constituted firstly of a stretch of signal peptide (Met1-Ala23) at the N-terminus of EpCAM that is cleaved during synthesis. As a result, the mature EpCAM protein's amino acid sequence begins only at Gln24. A shorter signal peptide is cleaved by signal peptidase at Ala21 may exist [29]. EpCAM's second segment runs from Gln24 to Lys265. This region makes up the EpCAM ectodomain, also known as the EpCAM cleaved extra-cellular domain (EpEX).

Chong and Speicher identified three domains within the extracellular part in 2001, addressed by disulfide linkage and glycosylation analysis: a completely novel cysteine-rich N-terminal domain (ND), a thyroglobulin type-1A (TY) domain, and a unique C-terminal domain with no cysteine residues (CD) [30]. The crystal structure of the extracellular part of human EpCAM confirmed these assignments in 2014, providing insight into EpCAM structure and function on multiple levels. First, it was discovered that the domains are not linearly organized, but rather form a triangular shape, with the main protease-sensitive site Gly79–Arg80–Arg81 located on a loop protruding from the plane of the ND–TY–CD triangle, which Thyroglobulin type 1 domains are found in a variety of proteins and are capable of binding to and blocking certain

cathepsins (cysteine proteases) that are involved in the growth of cancer [28]. Following that, the structure revealed that the ND domain is structurally distinct, whereas the CD domain shares a fold with some otherwise unrelated proteins.

Three N-glycosylation sites have been identified in EpCAM [29, 31]. Asn198 is not glycosylated in insect cells (Cysteine-poor area), but Asn111 is fully glycosylated and Asn74 is moderately glycosylated. Point mutations of the possible N-glycosylation sites in human and murine cell lines, however, revealed that all three sites are glycosylated. EpCAM's cell surface expression and protein stability appear to be highly dependent on Asn198 glycosylation [32]. EpCAM expression was reduced when the Asn198 glycosylation site was mutated. Furthermore, when compared to wild type EpCAM, removing all three glycosylation sites lowered the protein's half-life period from 21 to 7 hours. Because several cell surface molecules, including Notch, E-Cadherin, integrins, and CD44, are glycosylated differently in carcinoma versus normal epithelia, differential glycosylation of EpCAM could be a key factor causing differences in EpCAM function between healthy and malignant tissue [28].

(2) The transmembrane domain (EpTM)

The transmembrane domain is involved in the association with tight junction protein Claudin-7 [28].

(3) The intracellular domain (EpICD)

A small cytoplasmic domain, consisting of only 26 amino acids, runs from Ser289 to Ala314 in the protein. EpCAM Cleaved IntraCellular Domain is the name given to this cytosolic area (EpICD). EpCAM's cytoplasmic domain (IC) is made up of 26 amino acids, 14 of which are charged. Balzar and colleagues used EpCAM mutants with intracellular domain deletions to identify two potential β-actinin binding sites at positions 289 to 296 and 304 to 314. Which are reportedly important for EpCAM's localization at cell-cell contacts and thus for the molecule's cell adhesion properties. A putative PDZ binding site can be found at the C-terminus of amino acids Leu312, Asn313 and Ala314. In the major case of cell-cell protein contact, the hydrophobic C-terminal aa can interact with multi-PDZ domain proteins that are important in the building of complexes with signaling or structural proteins [33].

c) EpCAM function and signaling

(1) Proteolytic cleavage

EpCAM's functions and activity are heavily reliant on proteolytic processing. As previously stated, a signal peptidase cleaves off the N-terminal signal peptide of EpCAM after either Ala23 or Ala21 (in 1% of cases) to target the endoplasmic reticulum (ER) during protein maturation. Serine and cysteine proteases, on the other hand, break EpCAM at the TY-repeat between Arg80 and Arg81 [29]. The cleaved peptide has a molecular weight of 6 kDa, and the disulphide link between Cys6 and Cys7 keeps it attached to the 32 kDa component [29].

Protease inhibitor EpCAM has been demonstrated to be able to suppress cathepsin activity either via its TY domain or by serving as a "decoy" substrate for cathepsin to break the Arg80/Arg81 link, which is a key step in the progression from tumor cell to tumor cell metastases. Because cathepsins are frequently expressed by metastatic cancer cells, overexpression of EpCAM can result in increased inhibitory effects, potentially protecting these cancer cells from cathepsins as the tumor progresses.

(2) EpCAM and adhesion

EpCAM or epithelial cell adhesion molecule was firstly described as a Ca2+-independent cell–cell adhesion molecule by Litvinov and colleagues [34], in which they showed that EpCAM capable of mediating cell-cell adhesion in cells that normally have lack or absent of cell-cell interaction through direct homophilic interaction between EpCAM molecules on adjacent cells [34, 35].

EpCAM can increase homophilic cell–cell contacts, according to Balzar et al [35], but its functional antagonism on E-Cadherin-mediated adhesions shows that it functions as a modulator of cell adhesion strength rather than a promoter of epithelial cell aggregation and junctional complex formation [36]. EpCAM's anti-adhesive function, in contrast to Cadherin-mediated adhesions, may be strictly regulated and coordinated throughout morphogenesis and tissue regeneration, but disturbed during neoplasm growth [36].

(3) Regulated Intramembrane Proteolysis (RIP)

Regulated intramembrane proteolysis (RIP) is a conserved signal-transducing system that uses two-step cleavage process to convey information across cellular compartments [37]. With RIP, the metalloprotease Tumour Necrosis factor-Alpha Converting Enzyme (TACE/ADAM17) or β-secretase 1 (β-site APP Cleaving Enzyme 1) (BACE) [38] cleaves first EpCAM near to the extracellular side of the plasma membrane, resulting in the shedding of EpCAM's extracellular domain, known as EpEX [39].

TACE and BACE cleavages, often known as α and β-cleavages, take place in different parts of the cell. Practically, α-cleavage occurs at the plasma membrane, while β-cleavage occurs following EpCAM internalization, since BACE is predominately located to the trans-Golgi network [40].

Presenilin 2 (PS-2), a protease component of the γ-secretase complex, cleaves the C-terminal intracellular domain (EpICD) in the second step, yielding a 6 kDa peptide [29, 39]. Schnell et al. discovered a new RIP pathway for EpCAM in 2013, in which the EpICD is released by cleavages at two places inside the cysteine-poor region in EpCAM's ectodomain, followed by intramembrane proteolysis [29, 41].

Following RIP, the EpICD fragment is translocated from the cytoplasm to the nucleus, where it forms a complex with FHL2 (Four and A Half LIM Domains 2), beta-catenin, and Lef-1 (Lymphoid Enhancer Binding Factor 1). The resulting EpIC–FHL2–β-catenin–Lef1 signaling complex promote transcription cell proliferation-related genes like cyclin A2-D1-E (known as CCNA2-CCND1-CCNE1) and of c-Myc [42]. In immunodeficient mice, this has been demonstrated to stimulate cell proliferation, enhance hyperplastic development, and be carcinogenic [39, 43, 44].

(4) Tetraspanin-Enriched Microdomains (TEMs)

The proteolytic process occurred with EpCAM-mediated proliferation through RIP in which requires cell-to-cell contacts, that represent the first trigger events [43]. In addition, EpCAM's presence is also localized in specific subdomains of plasma membrane that are important for signalling, such as the Tetraspanin Enriched Microdomains (TEMs) [45]. Tetraspanins, integral transmembrane proteins that may interact with other tetraspanins, specific

lipids, and a range of transmembrane and cytosolic proteins create TEMs [46, 47], which are structured macromolecular complexes [48, 49]. EpCAM has been identified as a new molecule in TEMs [50] by forming a complex with tetraspanins CD9 and CO-029 (Tetraspanin8), as well as CD44 variant isoforms (CD44v4-7) [51] which have a role in tumor progression [52]. It has been discovered that Claudin-7, a tight junction protein, is another protein involved in the EpCAM-CD44v-tetraspanin complex [53]. EpCAM specific function to promote cell proliferation, apoptosis resistance and tumorigenicity occurs with the direct binding with Claudin-7 [54].

In addition, metalloproteases such as ADAM10 [48] have been shown associated with CD9 and other tetraspanins. Maetzel et al [39], has shown that ADAM10 and ADAM17 are involved in EpCAM as we said previously which are keys players in ectodomain shedding or RIP[48, 55].

d) EpCAM in health and disease

(1) EpCAM expression in normal tissues

In healthy adult tissues, EpCAM is found in most epithelial cells and is expressed during embryogenesis[56], on the basolateral membrane of pseudo-stratified and transitional epithelia [57, 58]. Normal squamous stratified epithelia do not show EpCAM expression. EpCAM is thought to facilitate cell adhesion in a Ca2+-independent way by engaging with other EpCAM molecules on neighboring cells (homophilic contact) in early investigations [34, 59].

Through oligomerization of EpEX in cis and trans interactions in the intercellular space, it was later shown that this homophilic interaction was weak. EpICD is hypothesized to be anchored to the α-actinin cytoskeleton when the two molecules engage [35]. When EpCAM and Claudin-7 work together [54], they keep the intestinal lining intact because EpCAM plays a role in creating functional tight junctions and modulating claudin dynamics [60] by directly binding to claudin-7 [36, 53, 61] as we explained in previous section.

Furthermore, according to several studies, it interacts with a variety of critical CAMs and modulates the adhesive characteristics between cells and their surrounding cell-matrix [61], including tight junctions [60], adherents junctions, desmosomes, and hemidesmosomes, among others [62]. As example, myosin contractility is increased and cadherin-mediated cell-cell

adhesion is impaired, as a result of EpCAM's EpICD tail, which inhibits protein kinase C (PKC) [63].

However, whether EpCAM really works as a cell adhesion mediator or not is still a matter of debate among researchers. The overexpression of the EpCAM, for example, has a negative impact on the strength of the classical CAM's E-Cadherin-mediated cell adhesion, resulting in a weakening of adherens junctions' properties [36, 64].

A further study using CRISPR-Cas9 to knock out EpCAM in a FaDu hypopharynx carcinoma cell line found that the absence of EpCAM had no significant influence on cell-matrix or cell-cell adhesion properties [40].

It has also been shown that inhibiting cleavage activity has no effect on EpCAM cell adhesion properties, indicating that EpCAM cell surface expression (EpEX) is unaffected by RIP signaling. Although, oligomerization is a characteristic of adhesion complexes, the resolved EpCAM crystal structure does not reveal that EpCAM forms a higher-order arrangement, notably more than dimerization [65], as has been suggested before. According to a recent research [66], EpCAM does not homo-oligomerize with another EpCAM on the other cell, contrary to what was previously thought. It is difficult to conclude that EpCAM has a direct role in homophilic adhesion molecule since oligomerization is a sign of homophilic contact that is capable of creating cell-cell adhesion activity [40, 66].

It should be mentioned that the closest homologue to TROP-1 (EpCAM alias) is TROP2, with which it has 67 percent amino acid sequence similarity[67]. It has been demonstrated by Fornaro et al. [67] that neither TROP1 nor TROP2 play any function as homophilic adhesion molecules, but rather act as signal transducers.

Even though EpCAM works as an adhesion molecule in the early stages of vertebrate embryonic development but later in adulthood it is not required [68], this shows that EpCAM may play distinct roles at different times. Recent studies, using single-cell RNA-seq datasets, have demonstrated that EpCAM is spatiotemporal patterning in mice embryonic development, with EpCAM seeming to be responsible for endodermal differentiation (e.g., lungs, colon epithelium) but negatively regulating mesodermal differentiation (e.g., heart) [69].

(2) EpCAM in cancer

EpCAM levels are high in most epithelium-derived cancers. Epidermal cell adhesion molecule (EpCAM) expression levels beyond a specific threshold have been shown to be related with a poor prognosis in breast and ovarian malignancies, as well as in pancreatic, urothelial, and gallbladder cancers [70-75].

Exceptional cases of cancer include renal and thyroid carcinomas, in which elevated EpCAM levels have been associated with increased survival [76, 77]. EpCAM seems to have a dual function in the progression of various cancer types, either encouraging or preventing tumor development [78]. In terms of tumor prognosis, Van der Gun and colleagues [78] and Patriarca and colleagues [58] present overviews of EpCAM.

Additionally, the predictive utility of EpCAM has been examined in conjunction with E-Cadherin and the tight junction protein Claudin-7, both of which are structurally and functionally related to EpCAM [79]. Co-expression analysis did not increase EpCAM's prognostic value for the incidence of nodal metastases in oral and oropharyngeal squamous cell carcinoma [79]. EpCAM might be cleaved using a well-regulated proteolytic procedure.

According to studies using an antibody against EpCAM's intracellular domain, nuclear/cytoplasmic EpCAM staining correlates with the aggressiveness of thyroid cancer and overall patient survival, and that this nuclear/cytoplasmic EpCAM staining can be found specifically in epithelial cancers as opposed to normal tissue [77, 80, 81].

Even though the staining findings were equivalent in oral and oropharyngeal squamous cell carcinoma material, is unable to rule out artifacts caused by non-specific binding of the main antibody. If the necessary controls are not in place, immunostaining artifacts that appear because of routine activities may be readily overlooked [82].

e) EpCAM Loss of Function studies: Animal knock-down models

It has been shown that inactivating germ-line mutations of the human EPCAM/TROP1/TACSTD1 [25] are related with congenital tufting enteropathy (CTE) [83], a life-threatening intestinal abnormality that manifest from birth. CTE is characterized by

macroscopic abnormalities in the intestinal epithelium, including villous atrophy, crypt hyperplasia, and focal crowding of enterocytes (tufts) [84, 85].

Because of a scarcity of genetic loss-of-function research, the in vivo function of EpCAM remains a poorly known phenomenon. EpCAM knockout animal models have just lately begun to appear in greater numbers. The absence of EpCAM in the developing zebrafish resulted in reduced epithelial plasticity and adhesiveness, with hyper-proliferation as a secondary consequence, which may have occurred because of the lack of contact inhibition. Increased susceptibility to bacterial infections and increased inflammation were also seen in patients with compromised skin integrity [86].

EpCAM-/- mice that were homozygously lacking in EpCAM were previously suggested to die in utero, possibly because of placental abnormalities. EpCAM -/+ mice were found to be alive and to be free of abnormalities [87]. Using a conditional knock-out of murine EpCAM in epidermal Langerhans cells, which represent skin-resident dendritic cells, Gaiser et al found that the cells motility and migration were reduced, and that their ability to regulate the skin inflammatory response was hampered [62].

For a further development Dr. Guerra and her colleagues of Cancer Pathology Laboratory at the University of Chieti, succeeded in developing two live lines of EpCAM knock-out mice, both of which had intestinal abnormalities, which are consistent with the findings in CTE, as it shown in the figure (Fig. 1 , [85]).

Those gene-trapped EpCAM knock-out animal that is alive until fourth day of postnatal and has been developed [85], and that it exhibits the same histological characteristics as reported in people suffering from CTE [83].

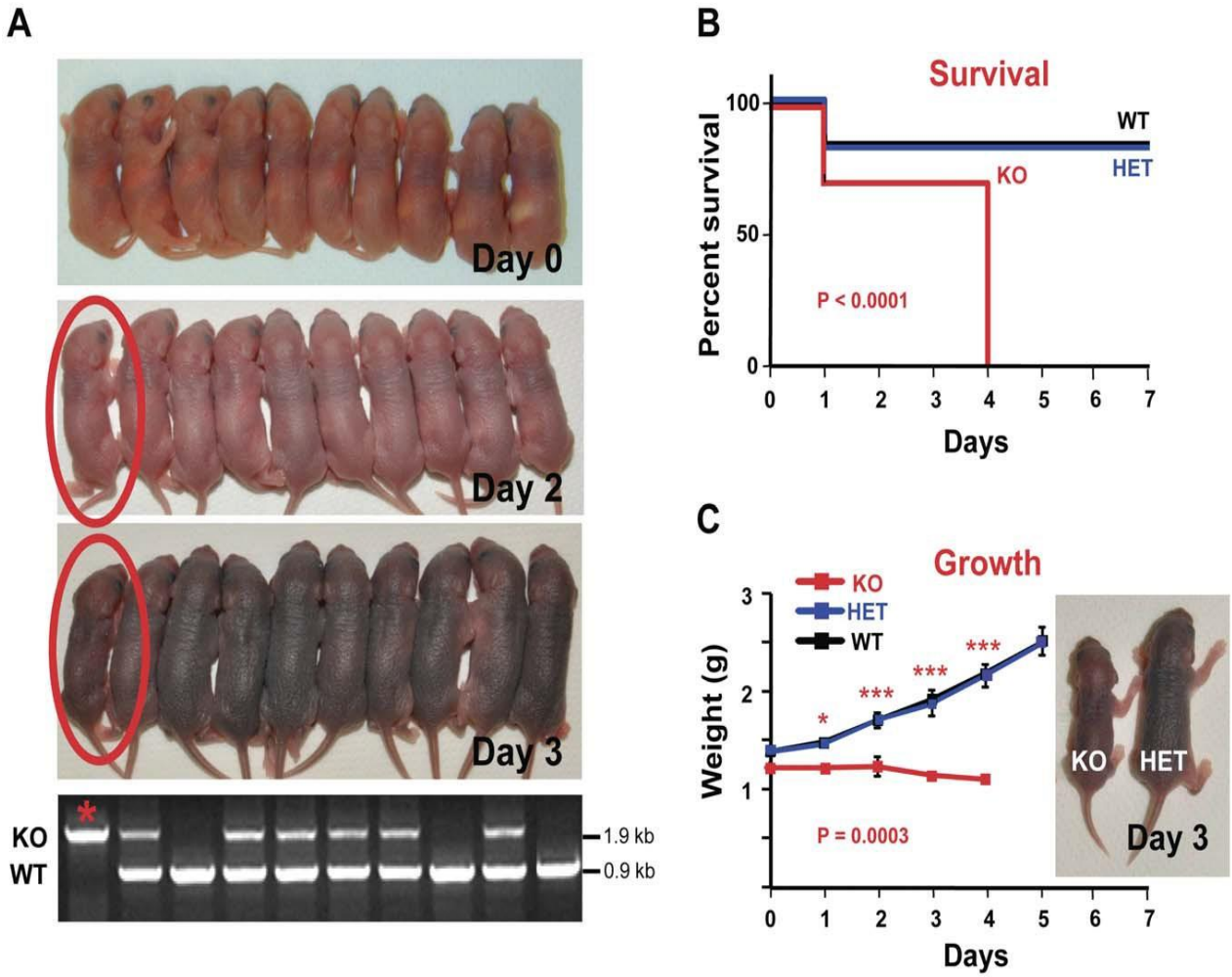


Fig. 1 : Growth arrest and early death of mTrop1-null mice. (A) Whole litter from a representative Heterozygous HET crossing. (B) Kaplan-Meier survival curves of WT, HET and KO newborns. (C) Growth curves of WT, HET and KO pups

To achieve this results, Guerra et al identified two gene-trapped ES clones, i.e., RST412 and RST413, where mTrop-1 was demonstrated to be inactivated by insertion of a promoter-less βGEO cassette (Fig. 2A, [85]) with 5′ rapid amplification of cDNA ends (RACE) used for sequence validation. Blastocyst injection was performed with both clones. Clones RST412 and RST413 produced three and seven chimeric mice. The mTrop-1 status of all the littermates was evaluated using gene-specific genotyping throughout the breeding process.

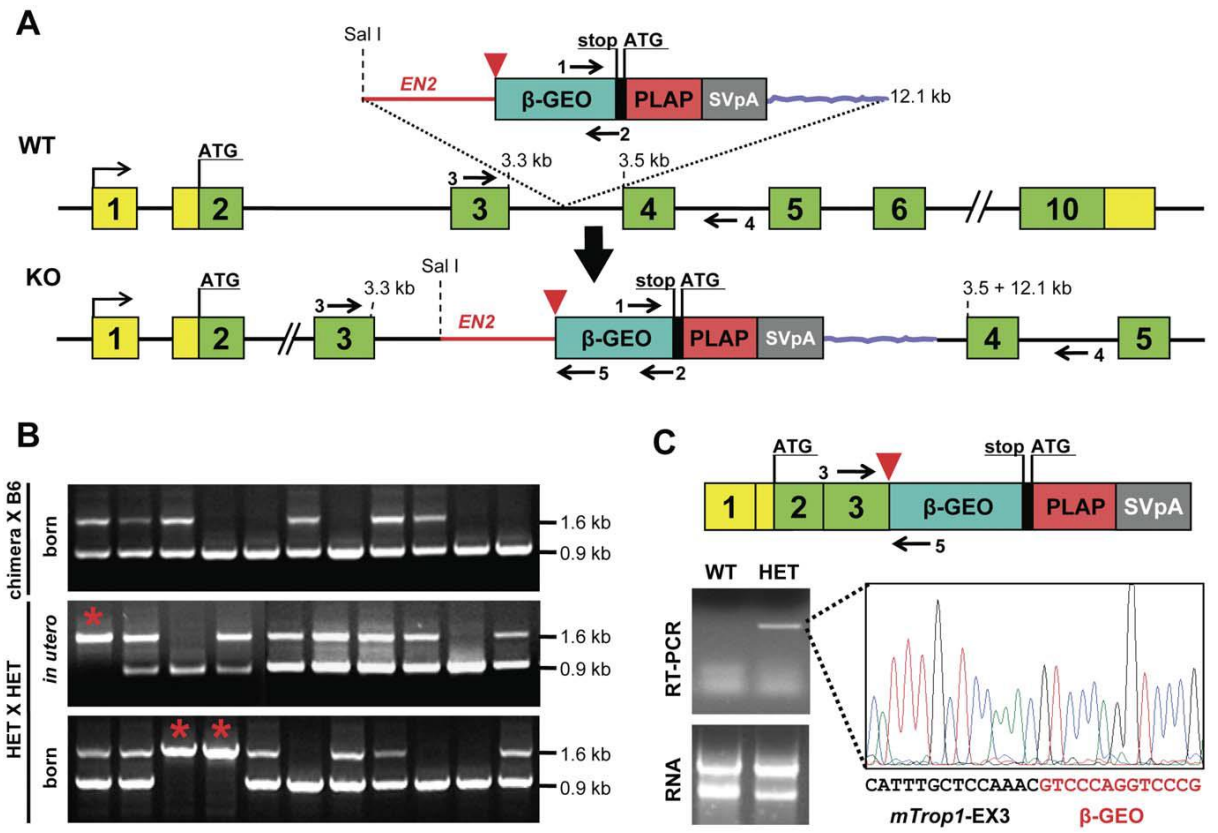


Fig. 2 : mTrop1 gene trapping

To confirm the success of used approach, genomic and transcriptomic characterization of RST412 F1 mice are proceeded showing (Fig. 2B, C, [85]). HET mice were able to reproduce and were viable. There was no additional study of RST413 F1 mice, who were bGEO-positive but had lost their imprisoned mTrop-1 due to in-vivo genomic recombination [85].

To investigate embryonic development defects brought about by mTrop1 ablation, they used the timed matings between HET mice from the RST412 colony, they analyzed the litters in utero at embryonic day (E: Mother uterine tissue) and as result it was hard to differentiate between Homozygous gene-trapped KO embryos, and wild-type (WT), and their HET siblings in size, developmental stage, body symmetry and somite architecture size. Immunofluorescence analyses with the anti-mTrop-1 G8.8 monoclonal antibody (mAb) demonstrated that mTrop-1 was absent in KO embryos which confirm the ablation of the protein (Fig. 3, [85]), while it is expressed in the intestine, pharingeal cleft, nose placode, limb buds and other body-lining epithelia lining epithelia of WT and HET mice [85].

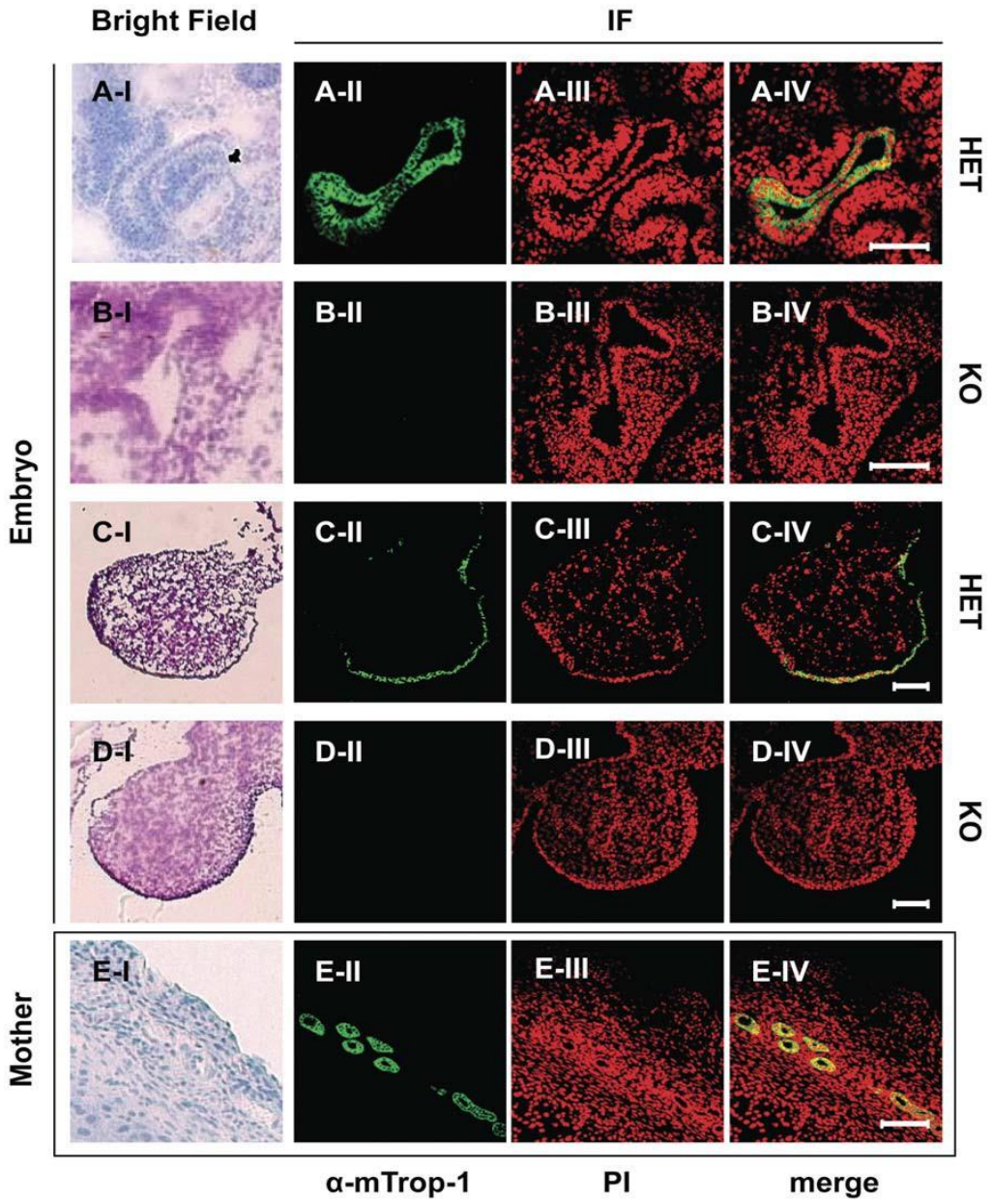


Fig. 3 : mTrop1 protein expression in the developing mouse embryo. (A, B) Embryo intestine. (C, D) Embryo forelimb bud. (E) Mother uterine tissue

In addition the absence of expression of mTrop-1 in the intestines of newborn KO mice was confirmed by immunofluorescence (Fig. 4, [85]) indicating that the gene had been completely silenced,. The gene-trapping b-gal marker, on the other hand, was only seen in KO and HET mice as a result of the method of gene inactivation [85].

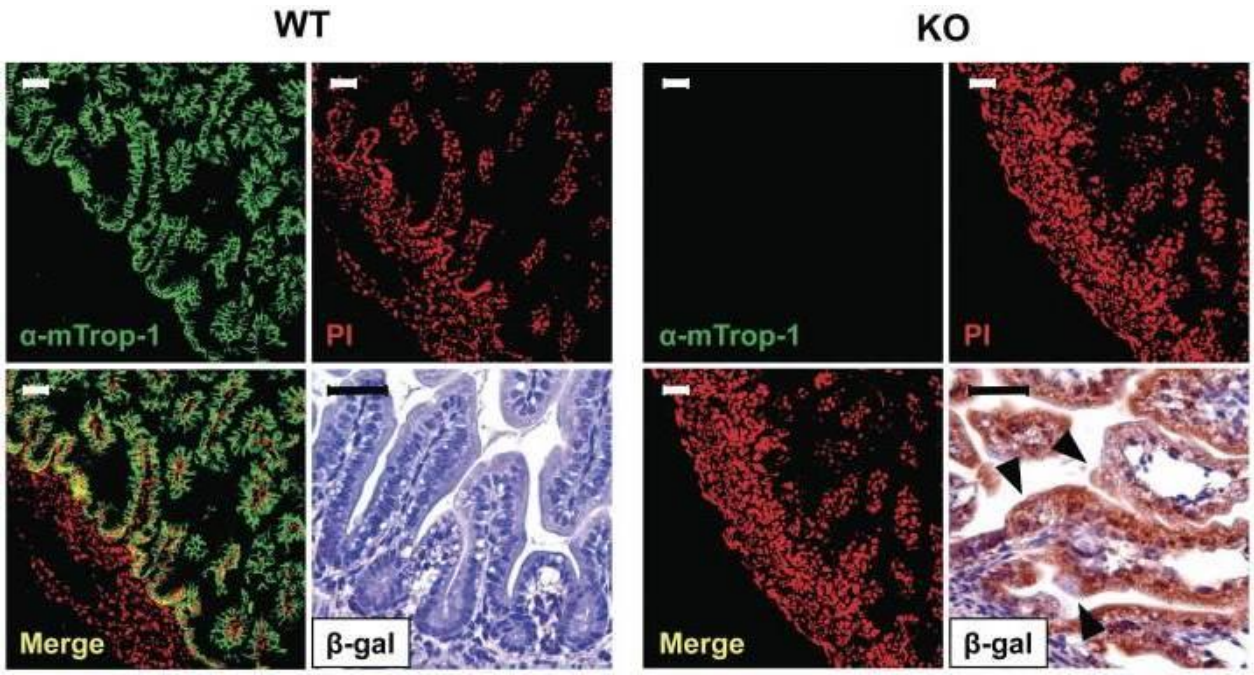


Fig. 4 : mTrop-1 expression in the intestine

These results indicated a different pathogenetic course from that described previously by Nagao et al [87]. They demostrate the absence of negative selection in utero against the KO. The difference between KO from their WT and HET appeared only after postnatal as we described previously (Fig. 1A, C , [85]), which suggested that mTrop1 inactivation had a harmful impact on the fetus shortly after delivery.

Furthermore, gene expression analysis (SAGE) and microarray hybridization profiles of embryonic tissues revealed that mTrop1 expression is highest in the intestinal epithelium. In line with this, they discovered significant levels of mTrop-1 expression in the intestinal mucosa epithelial layer in newborns.

These data suggested that the intestine would be a primary target for defects linked to mTrop1 ablation which is confirmed also through systematic macroscopic analyses that reveal the intestine of 3-day old KO pups (Fig. 5,[85]) was smaller than the WT in contrast with other organ that they didn't show any macroscopic morphological defects [85].

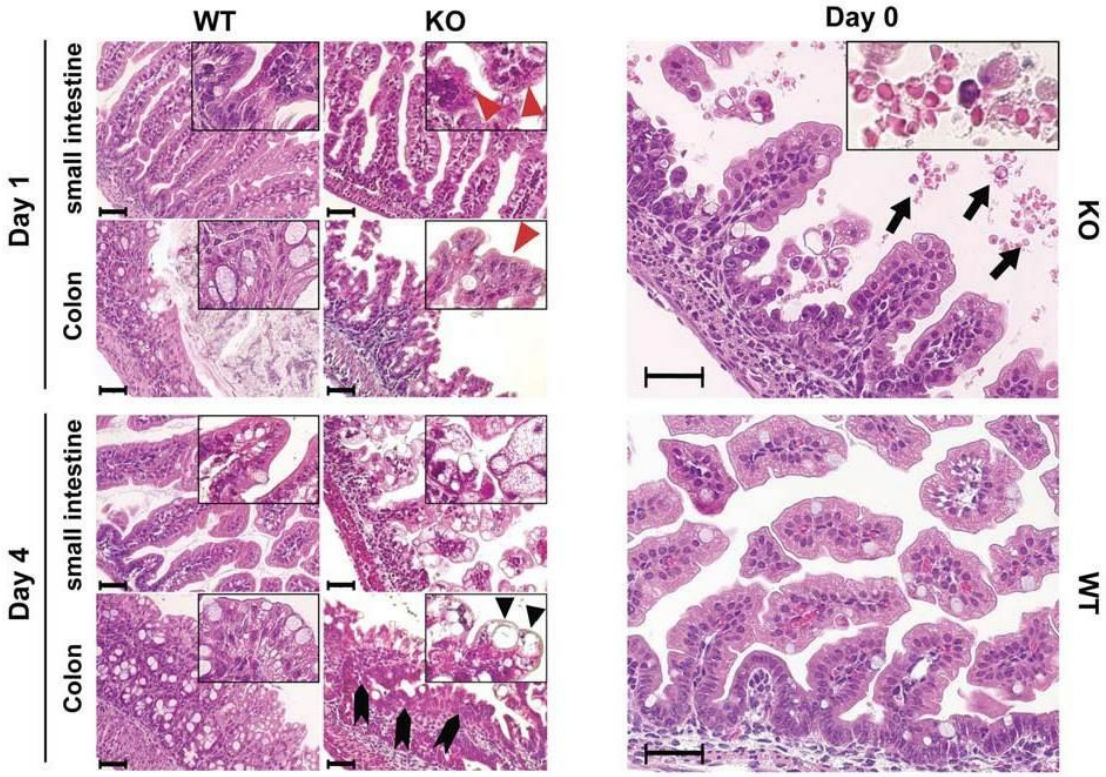


Fig. 5 : Tufting enterophathy in mTrop1-null mice

In addition, they confirmed the macroscopic result with Histopathology analyses that demostrate that the villous atrophy of the intestine was shown to be growing in severity, from minor abnormalities at birth to the complete loss of normal mucosal architecture by day 4. (Fig. 5,[85]). Furthermore, two viable EpCAM knock-out mice have been generated, by Lei and her colleges, both of which exhibit intestinal abnormalities [61], which are consistent with what has been discovered in CTE and has been described by Guerra et al [85].

With this information, they concluded that CTE is caused by mTrop1 deficiency, and that the intestinal mucosa structure is severely altered, with the loss of epithelial architecture and barrier function [88]. CTE in humans was originally linked to Trop-1 deficiency, whereas total loss of Trop-1 expression was found in other instances of CTE [83]. However, a less severe form of CTE is linked to TROP1 gene mutation, such as c.498insC [89], which may indicate residual Trop-1 function or expression, as in TROP1 exon 4 skipping instances [85].

Immunohistochemistry studies showed that the loss of mTrop-1 affected E-cadherin and β-catenin expression [85]. From birth to day 4, they revealed that both E-cadherin and β-catenin were both detected in the basolateral membrane compartment of WT epithelial cells. Also, result shown that E-cadherin present the highest expression level in the intervillar epithelium and developing crypts, which are the sites of the most active cell proliferation. E-cadherin expression along the villi was weak at birth, and markedly increased at cell-cell junctions in the following days. In the other hand also the expression of β-catenin was the strongest in the intervillar region, to day 4, where it reached the highest levels in the villous epithelium. As a conclusion, immunohistochemistry studies revealed essential role for Trop-1 in the maintenance of intestinal architecture and functionality, through regulation of E-cadherin/β-catenin expression and cellular localization [85].

Besides the important founding provided in this animal model for understanding the pathogenesis of intestinal alteration [85] and besides the important of cell surface of EpCAM for normal development [90], it still unclear whether the molecular mechanism that lead to the intestinal abnormalities are due to lower levels of tight junction proteins in CTE mouse models [85].

f) EpCAM's homologue TROP-2

The TACSTD2 (Tumour-associated calcium signal transducer 2) gene that encodes TROP-2 (Trophoblast cell-surface antigen-2) is a member of the TACSTD gene family [25]. It is the only known molecule that is homologous to EpCAM known also as TACSTD1 (Tumour-associated calcium signal transducer 1) which encodes TROP-1 (Trophoblast cell-surface antigen-1), with a sequence identity of around 49% and a similarity of 67% when conserved substitutions are considered [25, 91, 92] second to the alignment of their protein sequences. Sequence analysis indicates that exon shuffling happened throughout the development of the TACSTD gene family, since certain exons of TACSTD genes are similar to exons of genes encoding thyroglobulin, HLA-DR-associated invariant chain, and maybe nidogen [93].

Unlike TACSTD1[25, 91], TACSTD2 is intronless [67] suggesting that it was produced by retroposition via an mRNA intermediate. Since this retroposition occurred prior to the split of

the avian and mammalian lineages, TACSTD genes are over 300 million years old. The TACSTD1 and TACSTD2 genes are very well conserved across species [93, 94].

A region in the extracellular domain that is the thyroglobulin repeat domain, and the single transmembrane region are the regions with the greatest degree of similarity. Different domains are present in the intracellular tails of TROP2 and EpCAM, indicating that their intracellular signaling mechanisms are distinct.

TROP-2 and EpCAM, preserve the locations of the 12 cysteines residues, as well as the overall distribution of hydrophilic and hydrophobic residues [25, 92].

Each protein has a varied number of glycosylation sites. TROP-2 (36 kDa) has four N-glycosylation sites instead of three which are representing in EpCAM, two of which are conserved between both proteins [92]. TROP-2 contains a HIKE domain that contains a putative PIP2 binding site as well as a serine (S303) that has been observed phosphorylated within the cytoplasmic domain at Ser303 by PKC [95, 96].

Trop-2 serin residues S303 and S322 are not conserved in EpCAM [25, 97], suggesting that the cytoplasmic domains of the two proteins have distinct signaling properties. The common tyrosine residue present in both proteins has never been known to be phosphorylated. Tumor growth-promoting signaling may need protein kinase C phosphorylation of TROP-2 [96], whereas EpCAM-mediated signaling does not seem to require this. EpCAM's intracellular tail has two potential α-actinin binding sites, involved in EpCAMs cell adhesion properties [33], as well as a putative PDZ-binding site in which PDZ domains are required for transmembrane protein anchoring to the cell cytoskeleton [28].

TROP-2 is mostly expressed in epithelial tissue, much as EpCAM. Tissues with high levels of EpCAM expression, such as the colon and lungs, and EpCAM-negative epithelium may have low levels of TROP-2 expression, although the two do not seem to be linked [91]. Low patient survival and increased tumor aggressiveness and metastasis have been associated to high levels of TROP-2 expression in several late-stage epithelial carcinomas (pancreatic, colorectal, gastric, and oral) [96]. EpCAM participates in cell-cell adhesion and cell signaling through c-myc and cyclin A and E to enhance proliferation, migration, invasion, and differentiation in many cell types [96].

Overexpression of EpCAM is associated with a poorer prognosis in breast and ovarian cancers, but not in renal cell carcinoma [98]. In tiny adenocarcinomas, Trop-2 and EpCAM exhibit opposite biological effects: Trop-2 has a bad result, whilst EpCAM has a good one [99].

2. TROP-2: gene, protein, structure, and function in normal and cancerous (magliante) cells

In some tumors, but not all, the transmembrane glycoprotein Trop-2 is strongly expressed, and its expression varies in normal tissues [100]. Trop-2 crosses the cellular membrane, consisting of three domains: an extracellular domain, a transmembrane domain, and an intracellular domain, as well as a cytoplasmic tail that is required for signaling [25, 96].

a) The TROP-2 gene

TACSTD2, tumor-associated calcium signal transducer 2 gene, which is located on chromosome 1p32 and encode Trop-2 glycoprotein [25]. It was first described in 1981. It is around 36 kDa in size [25, 101]. Tacstd2 could be influenced by both genetic and epigenetic factors. It has a CpG island, which is a DNA sequence of at least 200 base pairs that encompasses a DNA fragment greater than 4 kb that is located upstream of the transcriptional start site and has a GC content of more than 50% and a CpG dinucleotide ratio of more than 0.6. TACSTD2 has 78 CpGs and four SP1 sites (GGGCGG) [102].

b) Trop-2 protein, structure, and domains

Trop-2 is composed of 323 amino acids, as showing in the figure (Fig. 6, [103]) below. Trop-2 is a member of the TACSTD family of proteins (Alias GA733 family)[25]. Trop-2 might have a role in EpCAM signaling as a modulator and/or enhancer [98].

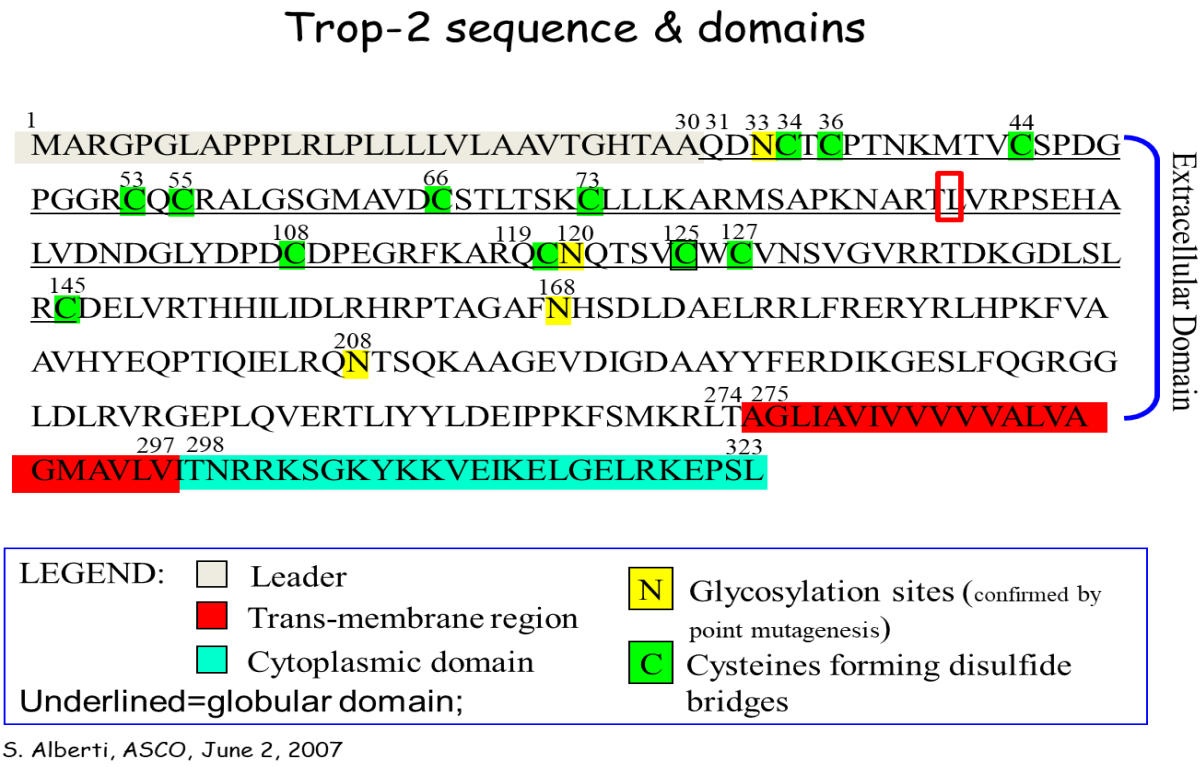


Fig. 6 : TROP2 Sequence and domains

Trop-2 is large transmembrane (TM) glycoprotein that contain four N-linked glycosylation (Fig. 6, [103]). It begins with a hydrophobic leader peptide of 26 amino acids (highlighted in gray) [25, 96]. The ectodomauyhin is the N-terminal extracellular region of the molecule (Trop-2-EC). It has three domains connected by a single TM helix (highlighted in red) and a 26 amino acid intracellular tail (highlighted in water green) (Trop-2-IC). Small N-terminal cysteine-rich domain (CRD) and cysteine-poor domain (CPD) (in green fluorescent) [25, 104]. The 33, 120, 168, and 208 position [25, 96] are putative N-linked glycosylation sites in Trop-2. The extracellular EGF-like repeat domain contains amino acids 1-274 [102]. The extracellular domain (EC) contains a thyroglobulin type-1 (TY) repeat domain between amino acids 70-145 .

The extracellular domain of Trop-2, globular section, contains a cysteine-rich region (Fig. 6, [103]) that contain a GA733 type 1 motif [30], followed by a thyroglobulin domain [96]. The globular region is followed by a stem segment, devoid of cysteines [67], that connect the globular moiety to hydrophobic TM region that composed of 22 aa (Fig. 6, [103]). The intracellular domain

(cytoplasmic tail) of Trop-2 is 26 amino-acid long and contains a HIKE domain [101], and two protein kinase C (PKC) phosphorylation sites, at Ser303 [95] and Ser322 [105] and a phosphatidylinositol 4,5-bisphosphate (PIP2) binding site [94].

Regulated intramembrane proteolytic cleavage separates the Trop-2EC and Trop-2IC from the TM. The Trop-2IC produced indirectly regulates cell proliferation and self-renewal via the β-catenin signaling pathway [104].

c) TROP-2 function and expression in Healthy Tissue

(1) Embryogenesis

Trop-2 was expressed in different healthy epithelial cells of many organs including respiratory tract, cervix, endometrium, fallopian tubes, placenta, seminal vesicles, thymus, vagina, esophagus, skin, tonsils, cornea, breast, kidney, pancreas, prostate, salivary glands, uterus, lung, stomach, colorectum, and bile duct epithelium of the liver [107, 108] but the highest expression was in trophoblast cells during the process of placental implantation which explain the origin of Trop-2' name [96]. During embryonal and fetal development, Trop-2 found expressed in divers cells such as ureteric bud, urogenital sinus, renal tubules, lung, tooth, hair follicles, skin, intestine, stomach, bladder, kidneys and brain [8, 109-113].

Trop-2 is expressed at E15.5 in the intervillus zone and freshly formed villi. Trop-2+ cells progressively decrease between E16.5 and birth, owing to villi loss [114] . As we explained previously (section II.B.1.e) were the mice born alive but died soon after owing to a lack of weight gain and hemorrhagic diarrhea, illustrating the significance of Trop-2 in early development and survival [85].

(2) Fetal Growth

Trop-2 levels rise during fetal lung development and expansion. McDougall et al, demonstrate in their study how Trop-2 expression was reduced by 75%, resulting in a 50% drop in proliferating fibroblasts. Trop-2 may therefore have a role in embryonic lung development. Trop-2 may enhance cell proliferation in the developing lung since organogenesis and

tumorigenesis share regulatory pathways [111]. In a setting with low Wnt and high Bmp stimulatory tones, Trop-2/Cnx43+ cells form fetal intestine [114].

(3) Expression in Normal Tissues

Many normal tissues express Trop-2, which is vital to keep in mind while targeting cancerous tissues that express Trop-2. It is found in type II alveolar epithelial cells (AECs), interstitial fibroblasts, smooth muscle cells, myofibroblasts, and airway epithelial cells in fetal rats [111]. Stratified, cuboidal, and columnar epithelial cells have membrane localized Trop-2 expression. Other tissues, expressing Trop-2 include the stratum basale epidermis, breast, cervix, cornea, the epithelial secretory tissue of the endocrine and exocrine glands, the esophagus, heart, kidneys (distal convoluted tubules and collecting duct), larynx, lungs, pancreas, prostate, salivary gland, skin, thymus, tonsils, trachea and uterus [108].

d) TROP-2-mediated Signaling and Interaction Network

More than two decades ago, published research under the direction of Professor Alberti Saverio were the first to reveal that Trop-2 operates as a calcium signal transducer by raising the concentration of calcium (Ca2+) in the cytosol [115].

Because the signal is transmitted even when there are no external calcium ions present, the researchers hypothesized that Trop-2 causes calcium to be released from internal stores in the cell. This calcium signal transduction function has been attributed to the intracellular domain that contains the PIP2-binding sequence, which is located inside the cell [94].

Trop-2 functions as a calcium signal transducer, and so a more thorough and mechanistic understanding of its activity is urgently required. On the other hand, multiple investigations have documented activation of Akt kinase by Trop-2, mostly in cancer cell lines [116], but also in murine mesenchymal stem cells [117]. Recent findings from a comprehensive proteome investigation revealed that Trop-2 regulates over 100 signaling molecules, with the PTEN/PIK3CA/Akt/GSK3ß pathway being the most often activated pathway in cancer cells [118].

In addition to the signaling pathways identified in this large screen and confirmed in various studies, other signaling pathways identified in this large screen include MAPK/ERK [119],

ErbB [120], TGFβ [121], Wnt/ β-catenin [122], JAK/STAT [123], integrin signaling [124], adherent and tight junction signaling [124]. It should also be mentioned that the downregulation of the Akt and MAPK/ERK pathways by Trop-2 has also been reported , indicating that the action of Trop-2 may be cell context-specific [102, 125, 126].

Being a transmembrane protein, Trop-2 is synthesized in the cytoplasmic reticulum and can accumulate in cytoplasmic vesicles. Hence, "cytoplasmic-like" localization is revealed by immunohistochemistry [109, 120, 127]. In breast cancer specimens, the expression of Trop-2 retention in the cytoplasm is positively correlated with phospho-Akt [118].

Trop-2 is subjected to regulated intramembrane proteolysis (RIP), according to a study by Stoyanova et al. [122] While the extracellular domain of the protein is cleaved, the intracellular domain is liberated and enters the nucleus, where it accumulates in conjunction with catenin, which subsequently triggers transcription of downstream targets.

While the extracellular domain (ECD) of Trop-2 is cleaved primarily by the metalloproteinase tumor necrosis factor-α converting enzyme (TACE), the intracellular domain (ICD) is cleaved mostly by the presenilin-1 and presenilin-2 proteins. Prostate cancer cells have been shown to have Trop-2 ICD, which indicates that it is involved in the progression of prostate cancer [122]. Trop-2 and β-catenin have also been reported to interact physically in gastric cancer cell lines, where Trop-2 is necessary for catenin nuclear accumulation and is related with the development of EMT [121].

Further research is required to determine the precise processes that govern the processing, cellular localization, and posttranslational changes of Trop-2, as well as to determine the consequences of these mechanisms on the biological function of Trop-2 itself. Based on several report that have shown that with the exception of β-catenin [121] [122] and claudins 1/7 [128], only a few numbers of Trop-2-interacting partners have been found so far such as PKC [101], Occludin [129], α5β1 integrin/Talin complex [130, 131], IGF-1 [102, 132], MDK [132] and NRG-1, [120] most of which function via their signaling capabilities, and some are briefly detailed below in more detail.

(i) IGF-1

Triggering Trop-2's downstream mediators PIP2 and Ca2+ may be possible using IFG-1 (insulin-like growth factor 1), which is a ligand of Trop-2. There is an extracellular EGF and thyroglobulin type-1 repeat in Trop-2's phosphatidylinositol (PIP2)-binding domain that includes a phosphorylation site (S303) [132]. Trop-2's PIP2-binding domain may compete with the IGF-1 binding protein by binding to IGF-1. IGF-1R signaling might also be prevented if Trop-2 and IGF-1 create an interfering complex [102] (**Erreur ! Source du renvoi introuvable.**, [106]).

(ii) Claudin-1 and Claudin-7

Trop-2 ectodomain binds to claudin-1 and claudin-7 transmembrane proteins (EC). Tight connections in the epithelial barrier rely on these proteins. As an anchor or transporter during the rearrangement of claudin, or as a stabilizer to prevent its destruction through the ubiquitin-proteasome system, Trop-2 has been postulated [104].

Trop-2 is also critical to the integrity of the tight junctions. Using the P1 integrin/RACK1 (receptor for activated protein kinase C) complex formation, Trop-2 may be able to influence cell adhesion to fibronectin [104]. Trop-2 loss results in a reduction in protein expression and a reorganization of protein subcellular localization, which affects the barrier function of the epithelium. In the GDLD, this is explicitly stated. In theory, the claudin-4 protein is found within a few micrometers of Trop-2 but does not bind it, since it is assumed to interact with claudin-1 or claudin-7. Because EpCAM interacts with claudin-7 through the AxxxG motif, it is believed that Trop-2 and claudins-1 and -7 bind to one another. When it comes to the actual mechanism, we don't know yet [128].

(iii) PKC

HIKE of Trop-2 interacts to and is phosphorylated by protein kinase C (PKC). Alberti et al, demonstrate that the HIKE sequences are potential candidates for binding to pleckstrin-homology domains, which can be present in the β subunit of G proteins, kinases, ankyrin, and kinesin [133]. Indeed, a calcium signaling protein called calmodulin (CaM) may attach to G protein's HIKE domain and control the binding of other proteins. HIKE-like CaM binding sequences are found in both neurogranin and neuromodulin, which are phosphorylated by PKC

and bind phosphatidylinositol (PI) [133]. TROP2 may have a role in calcium signaling because of the presence of HIKE domain, PIP2 binding site, and serine phosphorylated by PKC togherther.

In order to control downstream effector binding to phosphorylation at S/T residues, PKC modulates the binding of downstream effectors. PKCa is an isoform of PKC that binds to the tail of Trop-2 and regulates its activity [101]. DAG and Ca2+ bind to it and activate it, which then moves to the plasma membrane. PKCa activation is required for the production of podosomes.

Trop-2's cytoplasmic tail is phosphorylated at S303 by PKC. There is some evidence that PIP2 (phosphatidylinositol 4,5-bisphosphate) may influence PKC's phosphorylation of S303. C-phospholipase C may be concentrating the Trop-2 tail for hydrolysis by binding it to PIP2 (PLC). It is possible that the cytoplasmic tail's binding to PIP2 may be reversed when position S303 is activated.

IP3 (inositol 1,4,5-triphosphate), DAG (deacylglycerol), and Ca2+ release from the endoplasmic reticulum is increased as a consequence of PIP2 being exposed to PLC for cleavage. Free Ca2+ and deacylglycerol might activate additional PKC in a positive feedback process, which could lead to more Trop-2 and Raf/NF-B pathways being activated by PKC. If the phosphorylation of S303 arises from Ca2+ before or after Trop-2 signaling or if the phosphorylation of S303 releases PIP2 [96] was unclear till the novel research of Guerra et al , demonstrate that Trop-2 binds the cell membrane Na+/K+-ATPase, and that clustering of Trop-2 induces an intracellular Ca2+ rise followed by membrane translocation of PKCα, which in turn phosphorylates the Trop-2 cytoplasmic tail. This feed-forward signaling is promoted by the binding of Trop-2 to the PKCα membrane-anchor CD9 [134].

(iv) Regulated intramembrane proteolysis

Several intracellular signaling pathways have been linked to Trop-2. Intrinsic calcium signals are sent via Trop-2. Internal Ca2+ mobilization may explain why Trop-2-induced signal transmission is possible without external Ca2+. Trop-2 is cross-linked using specialized antibodies. Ca2+ levels in the cytoplasm grow significantly as a result of this cross-linking [96]. Trop-2 is a key signaling molecule for cells that need to proliferate, survive, replenish themselves, and invade [135]. In addition to Claudin-1, claudin-7, Trop-2 may also bind to IGF-1. Having stem

cell properties, Trop-2 governs cell development, transformation, regeneration and proliferation, which is why its overexpression may lead to cancer progression. Stem/progenitor cells express it on the surface, and it has a function in the maintenance of tight junction integrity [104].

Cell signaling generated by EpCAM may be modulated and/or enhanced by Trop-2. EpCAM proliferates and migrates into liver parenchyma when Trop-2 is modulated [96]. Trop-2 can promote cell migration even when growth factors are not present. The inability to sustain cell growth and mobility is reflected in the creation of induced foci [135].

Trop-2's accelerated cell proliferation and self-renewal activity in prostate cancer is dependent on the regulation of intramembrane proteolysis (RIP). Trop-2 is cleaved by the TNF-converting enzyme (TACE) and then cleaved by -secretase in the transmembrane domain via RIP. Presenilin 1 (PS-1), the dominant enzyme, and presenilin 2 (PS-2) are involved in cleavage (PS-2). The extracellular and intracellular domains of this cleavage have been identified [122].

Only the plasma membrane and the cytoplasm contain the ECD. An increase in sphere size but not number of cells shows that ECD promotes the proliferation of progenitor cells, especially prostate stem cells, through increasing cytoplasmic activity. Using ECD to treat prostate cells resulted in the development of tiny 6 kD fragments, which suggests cleavage of Trop-2. Trop-2 cleavage may be induced by the ECD via interactions with specific binding partners or directly through hydrophilic contacts [122].

ICD is mostly liberated from the membrane and accumulated in the nucleus. Only cancer specimens can detect nuclear ICD. To perform its transformation function, it requires cleavage and activation [122]. However, it might possibly be linked to other malignancies. Trop-2's ICD is its most important component. Involved in a β-catenin-dependent signaling cascade, it promotes self-renewal, begins prostatic intraepithelial neoplasia (PIN). The process of RIP activity and ICD interacts with β-catenin [135][136].

Trop-2 may lose its ability to function and enter the cytoplasm, where it plays a role in cancer growth, if it is produced in an incomplete or aberrant manner [102].

Trop-2 accumulates in the Golgi apparatus owing to a faulty transport in GDLD [101]. phospholipase C (PLC) cleavage is critical for the release of Ca2+ from the cytoplasmic tail of

Trop-2 as well as Trop-2-mediated cell signaling and cell cycle progression [135]. We still don't know exactly how Trop-2 regulates cell signaling.

e) Expression and Multiple Functions of Trop-2 in cancer

More than 40 years of study on the connection between cancer and Trop-2 and EpCAM have been conducted. As we reported in previous section, that recent studies have shown the importance role of EpCAM [137] in carcinogenesis and tumor cell dissemination [138], and its ability to be used as a prognostic marker [139] and therapeutic target in cancer treatment [140].

(1) The overexpression of Trop-2 in several cancer types: Up and Downregulation

Like EpCAM, Trop-2 was found to be overexpressed in a wide range of human carcinomas including for example gastric [141, 142], Colorectal [143-145], Lung adenocarcinoma [146], Breast [147], Oral squamous cell [148], ovarian [100], pancreatic [149], prostate [130, 131, 150], stomach [108], Papillary thyroid [151], urinary bladder [108], and uterine [152]. At the same time TROP2 can be downregulated in other specific cancer type such as esophageal [153], head [153], neck [153], hepatocellular [154], live fluke-associated cholangiocarcinoma [126] and lung tumors [102]. Referred to several published articles [121, 149, 155, 156], unfavorable prognosis and an increased risk of metastasis generally go hand in hand when Trop-2 is overexpressed.

In certain cancers, downregulation is associated with a worse prognosis, even though this tendency is becoming more and more common [154]. Even more interesting, Trop-2 may have a predictive value depending on its cellular location in the tumor. A study by Ambrogi et al. [147] found that breast cancer patients with membranous Trop-2 were less likely to survive, but those with intracellular Trop-2 were more likely to survive and have reduced relapses frequency.

Cancers of non-epithelial origin, such as melanomas [157], extranodal nasal NK/T cell lymphoma [158], glioblastomas and gliomas [123] have also been shown to have elevated Trop-2 expression. Another study found that Trop-2 was overexpressed in pituitary tumors [159].

Even though Trop-2 is frequently overexpressed during carcinogenesis, the TACSTD2 gene exhibits minimal point mutations and copy number alterations. There is no structural

alteration in Trop-2 in cancer, but rather a transcriptional and post-transcriptional dysregulation, according to COSMIC and Cancer Genome Atlas data [8, 160]. For identifying TROP2 regulatory pathways in tumors, Guerra et al. [7] uncovered a network of transcription factors that govern Trop-2 expression in cancers. It includes TP63/TP53, ERG, GRHL1/Get-1, HNF1A/TCF-1, SPI1/PU.1, WT1, GLIS2, AIRE, FOXM1, and FOXP3 in this structure. Further studies have shown that CREB controls the breast cancer-related transcription factor Trop [161].

In cancer cells, Trop-2 activates the ERK/MAPK pathway, which alters cell proliferation, migration, and invasion [162]. In which wild-type Trop-2 expression is proven to be both essential and sufficient for tumor development [8]. For the development of human cancer cells in vitro and in pre-clinical models, Trop-2-driven signaling is required [7]. It plays a role in the acceleration of the cancer cell cycle as well as the stimulation of cell growth [127]. The overexpression of Trop-2 in several malignancies is linked to poor patient survival as well as enhanced tumor aggressiveness and metastasis [127].

Trop-2 has been linked to angiogenesis in several studies. On the other hand, downregulation of Trop-2 reduces VEGF expression in glioblastoma cells, whilst overexpression of Trop-2 enhances it [123]. Vascular endothelial growth factor (VEGF) is a powerful angiogenic agent that stimulates the formation of blood vessels, allowing tumor growth and dispersion to take place [123]. Trop-2 expression is associated with increased microvessel density in gliomas and hilar cholangiocarcinomas [163, 164].

According to a recent research, Trop-2 is responsible for the neuroendocrine phenotype of prostate cancer, and it's overexpression results in a large increase in Poly (ADP-ribose) polymerase 1 (PARP1) [156, 165], is a critical enzyme for DNA repair regulation, replication, transcription, and chromatin remodeling [165]. Although DNA replication and damage accumulation are enhanced by Trop-2 overexpression even if PARP1 and the other DNA repair proteins are highly expressed in these cells [156]. Trop-2 may regulate PARP1 via upregulating c-MYC, however this is yet to be confirmed [106]. Neuroendocrine phenotype is reversed, and tumor growth and metastasis are reduced by PARP1 inhibition in vivo [106].

Signaling pathways that modulate Trop-2 expression include cyclooxygenase-2 [166] and tumor necrosis factor-α in colon cancer cells [167], PTEN and lipoxygenase in prostate cells [168]

[169], TGF--β in Langerhans cells [170] and fibroblast growth factor in embryonic lungs [110]. These pathways are often disturbed during carcinogenesis, which may explain in part why Trop-2 expression is changed in cancerous cells. Intriguingly, mechanical forces have been shown to activate Trop-2 expression [111]. During carcinogenesis, mechanical stress is a common occurrence, and Mechano-transduction may have a role in Trop-2 regulation [171]. Nonetheless, this has yet to be shown scientifically.

Epigenetic factors may also affect the expression of Trop-2 in cancer cells. It was recently revealed by Hao et al. that a circRNA–miR-488-3p–Trop-2 regulatory loop was released in head-and-neck squamous cell carcinoma cell lines [172]. Tumors of the head and neck and the urothelium of the bladder have also been reported to have miRNAs that target Trop-2 [173, 174]. A number of studies have revealed that the methylation of the TACSTD2 gene promoter in cancer cells has a direct effect on Trop-2 expression [102, 126, 154, 175, 176]. The TACSTD2 gene promoter is hypermethylated in lung adenocarcinoma tissues and cell lines, which explains the low Trop-2 expression in this malignancy [102]. All these characteristics may be relevant in different malignancies as the source of Trop-2 heterogeneity, as single-cell studies may indicate.

(2) Relation between Trop-2 and EMT

Adhesion and motility are critical for the successful colonization of secondary organs by metastatic cancer cells. Epithelial-Mesenchymal Transition (EMT), an evolutionarily conserved transcriptional pathway, allows polarized epithelial cells to acquire the phenotypic of mesenchymal cells with improved migration and invasiveness. [177].

A decrease in epithelial markers like E-cadherin and the upregulation in mesenchymal markers like Vimentin, N-cadherin and Fibronectin have been found to be linked to EMT. Epithelial-Mesenchymal Transition (MET) and its inversed process, mesenchymal-to-epithelial transition (MET) [178], are necessary for metastasis development. On the other hand, EMT may be a route that leads to the creation of cancer cells that have stem cell properties [178].

According to recent research, an extensive panel of human and murine breast and prostate cancer cell lines, as well as human cancers, showed that Trop-2 membrane expression is favorably

associated with E-cadherin expression and negatively associated with the mesenchymal gene signature [176].

In addition, Remsik et al., suggested that Trop-2 expression on the surface of breast and prostate tumors is linked to the epithelial phenotype [176]. Trop-2 expression is also inhibited by DNA hypermethylation or EMT transcription factors, such as ZEB1, during EMT [176]. E-cadherin protein levels are dramatically reduced in immortalized murine keratinocytes when mTrop-2 (murine Trop-2) is absent [179].

The mTrop-2 expression is lost during mesenchymal trans-differentiation in cells that are mTrop-2-positive. In addition, mRNA expression of TACSTD2 was shown to be lower in a subset of primary head and neck squamous cell carcinomas with EMT features [179].

EMT may be influenced by cancer type or cell environment, as shown by the fact that Trop-2 has been linked to mesenchymal phenotype in a number of distinct cancer types. While overexpression of Trop-2 results in an increase in Vimentin and decrease in E-cadherin expression in gallbladder cancer both in vitro and in vivo [180], When Trop-2 levels are decreased, the reverse occurs. Trop-2 has been demonstrated to be linked to the PI3K/Akt pathway in this model, since the downregulation of Trop-2 decreases Akt phosphorylation and increases the expression of PTEN, a negative regulator of Akt activity [180].

EMT may be caused through the PI3K/Akt pathway, which is well-known [181]. Li et al findings are in accordance with previous studies that have shown a correlation between increased Trop-2 expression and decreased E-cadherin expression in cervical, gallbladder, and gastric cancer cell lines. Nuclear factor-B (NF-B) pathway activation by Trop-2 causes EMT in nasopharyngeal cancer cells [182].

Trop-2-induced invasion and EMT are reduced when NF-B is blocked in these cells. To reverse this effect, it is necessary to downregulate Trop-2, which causes an increase in N-cadherin, Vimentin and Twist, but a decrease in E-cadherin. In gastric cancer cell lines, Trop-2 binds to -catenin and promotes its nuclear translocation and accumulation, increasing its transcription activity and contributing to EMT [121]. Recent studies have revealed that Trop-2 knockdown inhibits EMT in endometrial cancer cells [183].

Researchers' results on Trop-2 and EMT are mixed, and it's not apparent what its exact role or significance in this process is. That's why we think it's crucial that we understand how EMT-related intra-tumoral variation in Trop-2 expression affects the response to therapies targeting Trop-2-expressing cells. Cell Proliferation and Trop-2.

(3) Trop-2 influencing Cell Proliferation

Trop-2 promotes cell proliferation and growth in a wide range of tumors. The ERK signaling pathway is regulated by Trop-2 in cervical cancer cells [119]. cyclin D1, cyclin E, CDK2, and CDK4 expression are all lowered when Trop-2 is knocked out, whereas p27, a CDK inhibitor is elevated. Similarly, Trop-2 overexpression reduces p27 protein levels and increases cyclin E1 expression in human bladder cancer cell lines [184].

Curcumin seems to exert some of its anti-tumor action through modulating Trop-2 expression, and overexpression of Trop-2 in combination with curcumin therapy partly reverses these effects. mTrop-2 also raises ERK1/2 phosphorylation levels in murine pancreatic cancer cell lines, leading to an increase in proliferative activity in low serum circumstances [135].

Ectopic Trop-2 has been shown to promote cell proliferation, with a larger percentage of cells in S phase of the cell cycle in a wide panel of immortalized and transformed cell lines [8]. When Trop-2 is expressed at high levels in the body, it enhances tumor development in vivo. The anchorage-independent growth ability of colon cancer cells is similarly improved by Trop-2 [185]. This signaling pathway is activated in glioblastoma cells, and Trop-2 acts as a positive regulator of the production of downstream molecules such as cyclin D1, surviving, matrix metalloproteinase 2, and vascular endothelial factor [123]. (VEGF). Additionally, a loss of Trop-2 in glioblastoma cells results in a decrease in the phosphorylation of JAK/STAT.

There was also a report of the capacity of Trop-2 to inhibit cell growth. Trop-2 expression is often downregulated in lung cancer [102]. The silencing of Trop-2 in one lung cancer cell line stimulates cell proliferation, whereas the enhanced expression of Trop-2 in another cell line inhibits cell growth. Cholangiocarcinoma cell lines with Trop-2 knockdown show enhanced cell proliferation and motility, as well as an increase in ERK phosphorylation [126].

This also boosts cervical cancer cell proliferation, motility, and invasion capabilities in vitro [132] and promotes tumor formation in vivo [132]. Knockdown of Trop-2 modestly boosts proliferation in MCF7 breast cancer cells relative to normal cells [186]. Given the above-mentioned data, it is likely that Trop-2's involvement in controlling proliferation is a multifaceted phenomenon that depends on the cell type and the organ.

(4) Trop-2 and Cell Adhesion and Migration

On the other hand, Trerotola et al., have postulated that Trop-2 may have a role in the migration and invasion of cancer cells in a variety of cancers. A decrease in cell adhesion to the prostate protein Fibronectin, caused by Trop-2, has been shown to increase prostate cancer cell motility [150].

Trop-2's direct interaction with the 51 Integrin complex and Talin, which results in their relocalization from focal adhesions to the leading edges, RACK1 translocation to the cell membrane, Src and FAK activation, and the destabilization of the 51 Integrin complex's bond to Fibronectin, is a complex mechanism that is responsible for this [150].

Exosomes positive for Trop-2 protein have been shown to promote T2-negative prostate cancer cells to migrate on Fibronectin through a pathway that is presently unclear [130]. By activating the MAPK ERK/JNK signaling cascade and increasing the production of MMP2 in thyroid cancer cells, Trop-2 promotes tumor invasion [187].

For colorectal cancer cells to be able to migrate, Trop-2 must be phosphorylated at serine 322 and Claudin 7 must be re-localized. In cells that produce phospho-mimetic Trop-2, the ability of the cells to migrate dramatically rises when this phosphorylation site is blocked [105].

(5) Trop-2 as Tumor Suppressor

In a wide range of malignancies, increased Trop-2 expression has been shown to enhance tumor development [8], and this has been shown to be associated with poor prognosis and metastasis indicating that Trop-2 operates as a potential oncogene. Some research, on the other hand, has shown that Trop-2 may work as a tumor suppressor in some cases [188]. A well-established function of proteins is their capacity to perform multiple roles (either oncogenic or tumor suppressor) [189, 190]. The discovery that both the acquisition and loss of Trop-2 function

lead to significant phenotypes suggests that this protein is a signaling protein with robust signaling activity. As with EpCAM, which modulates both adhesion [34] and intracellular signaling [39] via the generation of cleavage products, Trop-2 may have a pleiotropic mechanism of action as well [179].

Trop-2's reduced expression in lung adenocarcinomas may be due to TACSTD2 promoter hypermethylation or loss of heterozygosity [102].

Insulin-like growth factor 1 (IGF-1) receptor binding is disrupted by Trop-2 in lung cancer cell lines. Also suppressed are Akt and ERK kinase activity, -catenin and Slug expression. Trop-2 expression is lowered, which results in an increase in cell proliferation and tumor formation in the body. Trop-2 overexpression was observed to improve the prognosis of patients with NSCLC, according to a study by Pak et al. [125].

Other studies have shown that increased Trop-2 expression is associated with poor prognosis in NSCLC patients. Lung cancer cell lines that overexpress Trop-2 show increased cell proliferation, migration, and invasion whereas those that knockout Trop-2 show an increase in cell death [146]. For example, Trop-2 has tumor-suppressive properties in cervical cancer cell lines, where it similarly suppresses the activation of IGF-1R and ALK, potentially via the binding of their ligands IGF-1 and MDK [132]. Similarly, Trop-2 expression in the mesenchymal subtype of squamous head and neck cancer has been diminished or totally eliminated [179].

ErbB3, commonly known as HER3, is more phosphorylated when Trop-2 expression is reduced in this cancer type. For ErbB3, NRG-1 is an important ligand, and inactivating Trop-2 increases the quantity of NRG-1 on the cell surface, which is subsequently cleaved by TACE [120].

Activation of ErbB3 by the NRG-1 extracellular domain leads to cell proliferation and tumor development. In squamous cell carcinomas, Trop-2 loss is inversely proportional to ErbB3 loss during advancement [153]. Trop-2 protein expression is related with resistance to anti-ErbB3 therapy in head and neck squamous cell carcinoma [191]. It is more effective to use anti-ErbB3 and anti-Trop-2 antibodies in combination than to use either antibody alone to combat the tumor [191].

Furthermore, Trop-2 expression has been observed to be diminished in cholangiocarcinomas and hepatocellular carcinomas linked with liver flukes [126, 129, 154]. Both Trop-2 downregulation and TACSTD2 promoter hypermethylation are epigenetically regulated. In hepatocellular carcinomas, low Trop-2 expression is associated with poor overall survival, invasion, and poor differentiation [154], but not in cholangiocarcinomas [126]. Remšík, J. et al. found that reduced TACSTD2 mRNA expression in individuals with lymph node-positive breast cancer and prostate cancer was associated with a worse prognosis [176].

Depending on the cancer cell's overall genetic background, Trop2's specific involvement may differ. Non-oncogene addiction, a phenomena that has been well studied, may help to explain these inconsistencies [192]. Non-oncogenic genes have a critical role in the survival tumor cells in a physiologically challenging microenvironment. According to reviewed study, Trop2's involvement in carcinogenesis varies over time, from early tumorigenesis when it functions as a tumor suppressor to later stages where its reactivation increases metastasis and is associated with a worse prognosis [106].

This is consistent with the findings of the reviewed research [192] [106]. In contrast to the "genetic context theory," this progressive rise in Trop-2 explains its ubiquitous expression in cancer, independent of its dual involvement in other linked pathways. This hypothesis, on the other hand, necessitates additional research in related systems. Furthermore, how different Trop2 types and subcellular locations affect cancer cell activity remain a mystery. The signaling capabilities, oncogenic or tumor suppressive roles, and prognostic value of these many forms in malignancies require additional extensive studies, regarding the few numbers of studies addressed this [106, 111] .

For further investigation on the functional role of Trop-2 in cancer, several robust techniques are required, and they have been used to explore new features of Trop-2 pathways and interaction. One of the most important is NGS (WGS/WES/Transcriptomic/Micro-Arrays) approach and siRNA/shRNA/CRISPR technology that will be detailed in next section and applied in our studies.

C. High Throughput Sequencing Analysis (NGS)

Referring to the definition of Nature.com "Next-generation sequencing refers to non-Sanger-based high-throughput DNA sequencing technologies. Millions or billions of DNA strands can be sequenced in parallel, yielding substantially more throughput and minimizing the need for the fragment-cloning methods that are often used in Sanger sequencing of genomes".

Whole Genome Sequencing (WGS), and Whole Exome Sequencing (WES), and targeted resequencing of specific region are examples of genomic assays that are used to uncover variations that are involved in cell function or illness. NGS-based transcriptome analysis (RNA-seq) [193] includes quantitative gene expression profiling, finding of new transcribed sequences [194], and non-coding RNA species such as microRNA and long non-coding RNA (lncRNA) [195] [196]. DNase-seq [197], ATAC-seq [198], DNA methylation [199], and histone modification ChIP-seq [197] are examples of epigenome approaches that are focused on chromatin structure [200]. In the Epigenome section, we have also included TF ChIP-seq data [201-204].

In this section I will briefly introduce some used technology in our study and their required step during analysis on both side from library preparation to bioinformatic analysis and tools

1. WGS/WES Sequencing

With whole genome sequencing (WGS), the complete genome is sequenced instead of just the protein coding regions, which make up less than one percent of the genome. The sensitivity of SNP identification using the WGS technique is also improved [205, 206].

The apparent benefit WGS over targeted approaches is the volume of data generated: the complete genome is sequenced as opposed to simply protein coding areas, which account for around 1 percent of the genome. Furthermore, the WGS approach has been shown to increase the sensitivity of SNP discovery [205]. [206]. Using WGS, researchers can explore genetic or somatic variants, opening up new paths for the investigation of normal and disease phenotypes. WES, on the other hand, concentrates on collecting and sequencing protein-coding regions (exomes), restricting the data to a more functionally relevant portion of the genome [207, 208].

Library preparation, instrument sequencing, quality control, alignment & variant identification, and data analysis are all part of a typical procedure. For Illumina NOVASEQ 6000 sequencer, Nextra DNA protocol steps for sample preparation are:

a) Extraction of DNA/RNA and Library preparation

Generally, the stages involved in library preparation include shearing the DNA sequence into small fragments, with insert sizes ranging from 100 base pairs to several thousand base pairs, depending on the quantity of input DNA, technology, and the preparation procedure used. The fragments are then ligated together using adaptors at the 5' and 3' ends.

(1) Whole Genome Sequencing (WGS)

The DNA were collected from blood lysis or saliva lysis sample, specific purification beads were used to extract DNA followed by a bead-linked transposomes (BLT) for tagmenting DNA in small fragment and tag each sequence with adapter that are necessary barcode legation. The second step consist of post tagmentation cleanup before proceeding PCR amplification. A double purification with magnetic bead at high and low concentration to collect the desired fragment length for sequencing.

(2) Whole Exome Sequencing (WES)

For WES sequencing step are bit longer and differ from WGS. An enzymatic fragmentation of the Input DNA. Once we obtained our fragment library preparation is proceeded by repair abd dA-tail the DNA ends (A-T overhang), ligation the molecular barcode and adaptor, purification of samples using AMPure XP beads, amplification the adaptor-ligated library and purification of the amplified library with AMPure beads. The next step is hybridization DNA samples with the probe and capturing the hybridized DNA using streptavidin coated beads. The final step is purification the amplified captured libraries using magnetic bead and preparation the pool for sequencing. Once the fragments have been ligated, they are ready for cluster creation and sequencing.

b) Quality Control

For optimal results, researchers would want to obtain sequencing data that has precisely the same content as human DNA, with reads mapping uniformly and having 50/50

paternal/maternal alleles, among other things. Although there are various aspects that influence the quality of the data, some are more significant than others. These issues include contamination and DNA purity as well as technological limitations and technical mistakes. Performing quality control on sequencing reads is essential for determining the accuracy and dependability of the data collected.

(1) Average depth of water

Average depth, also known as depth of coverage, is calculated by adding up the depth of all targets sequencing bases and dividing the total number of bases by the total number of bases. Higher depth at a single base location translates into greater statistical relevance of base calling.

(2) The percentage of maps having at least X depth.

This evaluation is frequently done in conjunction with the calculation of average depth. As the term "average depth" implies, not all the target DNA is covered in at same amount of time. One explanation is that there are low complexity or non-unique sections in the DNA that are difficult to map or unmappable, which makes it difficult to detect mutations. Using the percentage of mappings with at least X depth, where X represents the number of depths with sufficient statistical significance, it is possible to determine whether not the sequencing data has strong enough base calling quality.

The Fig. 7 explain that if there are two identical sections in the DNA (D) and the size of F is much lower than the size of D, reads that cannot be uniquely mapped to the reference genome will not be guaranteed to be mapped evenly during the read mapping stage, as described above. Alternatively, it is conceivable that 100 percent of these reads will be mapped to just one D, with the other reads having no mapping at all. As a result, one D has twice the average depth while the other has zero depth.

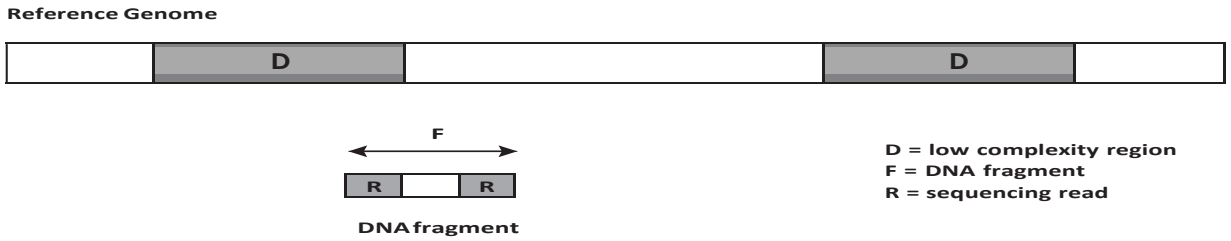


Fig. 7 : Illustration of unmappability

(3) The percentage of GC content

Base calling error is not evenly spread across all base replacements, according to research. Furthermore, as the readings go, the mistake rate rises [209]. The errors are typically preceded by base G statistically. A > C substitution is the most frequent error, whereas C > G substitution is the least common. Sequencing bias during library preparation and base calling might be indicated by an unusual distribution of GC content.

(4) Reads that are duplicated

PCR amplification has always been necessary as part of the library preparation process. However, this procedure might result in PCR duplicates, which are sequence readings that include the same DNA fragment. The PCR duplicates might lead to erroneous copy number variation detection. Furthermore, if there are mistakes in the readings, they may spread throughout the variant calling process, resulting in an artefact.

c) Alignement and identification of variants

The purpose of this phase is to produce a list of variations from sequencing reads. This stage includes read alignment and variant calling in general. Additional procedures, such as Mark Duplicate, Indel recalibration, Variant recalibration, and variant quality score recalibration, may be added to increase the quality of variant identification [210, 211]. Base substitution, tiny insertion, small deletion, copy number variation, and structural rearrangement are examples of variations that may be discovered.

(1) Alignment of the read

The process of appropriately aligning reads to the human reference genome is known as read alignment. The Genome Reference Consortium (GRC) maintains the human reference genome, and various institutions, such as the University of California at Santa Cruz (UCSC), offer interfaces and extra resources [212, 213]. Read mapping, indel realignment, and base recalibration are among the alignment techniques involved. The process of mapping raw read data to the reference is known as read mapping. Many bioinformatics software are available to do this kind of task such as Bowtie2 [214], BWA- BWA-MEM [215, 216], YOABS [217], CUSHAW2 [218], SOAP [219] [220], and Stampy [221] are just (Lunter and Goodson, 2011).

Processing time, memory use, read size, sequencing equipment, licensing type, and multi-threading are all factors that influence their strengths. After mapping readings, the next step is to realign indels. Indels and SNPs may be incorrectly identified based on independently mapped reads, particularly if indels are at the end of the reads. To identify indels, the indel alignment procedure uses all previously mapped reads. Genome Analysis Toolkit (GATK) is one of the most extensively used tools available [210, 211, 222]. Base recalibration is the last stage in read alignment. The method for calling variants is significantly influenced by the quality scores of individual bases in reads, and the estimated scores supplied by sequencing equipment are prone to a variety of systematic technical flaws. Base recalibration's purpose is to change quality ratings depending on data and recognized variances.

(2) Single nucleotide variant and small insertion-deletion calling

The most difficult part of this stage is reducing the number of false positives and negatives. Base quality scores and base ratios are the most common pieces of information used by germline variant callers to detect zygosities. GATK [222], BCFtools [223], and FreeBayes [224] are some of the software applications that have been created. ExScalibur [225], Fermikit [226], BAYSIC [227], FAVR [228], and VarDict Koboldt [229]. The amount of the indels, computational time, memory use, multi-threading, accuracy, platform specificity, and read depth are among the variations.

(3) Copy Number Variant calling

To identify copy number occurrences, there are a few computational methods. As proof, one method is to employ paired or split readings. If a paired read is mapped to a point that is distant from the predicted position given the insertion size, for example, it might be used as evidence for structural deletion. TIDDIT [230] and MANTA are two tools that may be used with this strategy [231]. Another option is to employ genome-wide sequencing coverage. This implies that the depth of the genome is about equal. Any locations with average depths that are significantly lower or higher than predicted indicate structural deletion or duplication events. CNVnator is a tool that uses this approach [232] . CNV identification may be done utilizing another sample, or a group of samples, as a reference for exome sequencing data with unequal depth. Exome Depth [233] and XHMM are two tools that employ this strategy [234].

(4) Structural rearrangement

Paired readings are used as evidence in the procedure for identifying structural rearrangement. As a result, CNV detection algorithms like TIDDIT [230] and MANTA [231] may also identify structural rearrangement.

Variations must be annotated to offer biological importance, which aids in filtering and prioritizing disease-causing variants, in order to turn a list of variants into actionable information. Frequency-, structural-, prediction-, and evidence-based data are all examples of information that may be tagged to variations. Aside from annotation, visualization is critical to the understanding of the results since it converts information from a computer-readable format into a human-readable visual representation.

d) Analysis data process

In most cases, exome sequencing may uncover up to 30,000 variations in a single sample [235]. To reduce the number of candidate variations to a manageable number, variant annotation is used to combine data from many sources and forecast variant importance. VariantStudio (Illumina), IonReporter (Life Technologies), Geneticist Assistant (Softgenetics), and Expressionist (Softgenetics) are commercial software packages that come coupled with sequencing devices (GeneData). Basic information, such as gene names, transcripts, and regulatory areas, may be added to the variations using open-source tools such as ANNOVAR [236], the

Ensembl Variant Effect Predictor [237], and snpEff [238]. With regard to known protein domains, PFAM [239] and SMART [240] can predict functional relevance. It's vital to remember that silent variations, which don't affect the translated protein downstream, may sometimes cause problems in mRNA splicing [241]. SPIDEX [242], ENCODE [243], and FANTOM can annotate non-coding variations, regulatory areas, and splice sites [244, 245]. The chance of a sequence change affecting protein function is predicted using in silico pathogenicity prediction. Three tactics, evolutionary conservation, structural or biophysical features, and machine learning approaches, were used to construct the predictors. SIFT [246, 247], PolyPhen2 [248], LRT [249], MutationTaster [250], Phylop [251], GERP++ ++ [252], and CADD [253].

Previous discoveries and clinical evidence may be found in a number of databases. SNPs, indels, microsatellites, multinucleotide polymorphisms (MNPs), heterozygous sequences, and named variations are all included in the Single Nucleotide Polymorphism Database (dbSNP) [254]. DGV [255] and dbVar [255] both have genomic structural variants. The database of Genotypes and Phenotypes (dbGaP) [256, 257] collects and disseminates data and findings from research that look at how genotype and phenotype interact in humans. OMIM [258] is a database that contains knowledge-based information on human genes and genetic diseases.

Exome sequencing data and associated phenotypic data from populations with heart, lung, and blood problems may be found in the NHLBI Exome Sequencing Project (ESP). The Human Gene Lesion Database (HGMD) [259] is a database of known gene lesions that cause human hereditary illness. The Leiden Open Variation Database (LOVD) [260] is a platform for collecting and displaying DNA variants that is concentrated on genes. ClinVar is a database that compiles data on genetic variation and its impact on human health [261]. The most comprehensive library of DNA variations re-sequenced in the genes that contribute to gastrointestinal cancer is housed and maintained by InSiGHT [262]. ENIGMA is a multinational group of researchers dedicated to curing sequence abnormalities in BRCA1, BRCA2, and other breast cancer genes that are known or suspected. ClinVar now incorporates curated genomics data from InSiGHT and ENIGMA on a regular basis. COSMIC is a system that saves and displays information on somatic variants [263]. The Cancer Genome Atlas (TCGA) project sequenced human cancers and healthy tissues at the DNA, RNA, protein, and epigenetic levels using high throughput methods. Exome sequencing

data from a number of large-scale sequencing efforts is aggregated and harmonized by the Genome Aggregation Database (gnomAD) [264].

Visualization is a critical stage in high-throughput data processing because it converts data from a computer format that requires technical expertise to grasp into a user-friendly visual display. Most annotator software can output comma-separated values (CSV) and tab-separated values (TSV) files, which may subsequently be loaded into any spreadsheet reader. Web-based browsers that incorporate diverse sources of annotations include the Ensembl Genome Browser [265] and the University of Santa Cruz (UCSC) Human Genome Browser [266] [267]. Viewer for Integrative Genomics (IGV) [268] is an independent browser that supports array-based and massively parallel sequencing data as well as genomic annotations.

e) File formats

The scientific community has standardized file formats for communication between each operation in high throughput sequencing research.

(1) FASTA

A FASTA file is a text-based file used to contain nucleotide or protein sequences. FASTA files are often used in Human DNA analysis to store human reference sequencing, such as GRCh37 and GRCh38. Almost all techniques in DNA sequencing analysis employ reference FASTA files.

(2) FASTQ

A file in the FASTQ format is a text-based file used to record nucleotide sequences and their associated quality ratings. A sequencing read is represented by each sequence. A FASTQ file usually has four lines for each read. The ID of the read appears on the first line. The nucleotide sequence is on the second line. The third line begins with "+" and is followed by sequence information that is optional. The quality scores are on the fourth line. In paired-end sequencing, one sample is frequently split into two FASTQ files (Fig. 8, From my analysis file), each with an equal number of reads. A pair of reads is represented by sequences with the same ID. In general, sequencer files in the FASTQ format are considered full products. The files are then aligned with a reference genome using an alignment program.

Fig. 8 : FASTQ file format

f) SAM

A SAM file is a text-based file that stores nucleotide sequences from related FASTA files that have been aligned to a reference genome. The files are the output of an alignment tools (BWA-mem / Samtools). A header and an alignment section are included in a SAM file. The header of the file or sample comprises general information, such as the genome reference utilized during the alignment process. The alignment information may be found in the alignment section. There are 11 obligatory tab-limited fields on each line. One alignment is represented by a single line (Fig. 9) .

g) BAM

The binary/compressed version of a SAM file is a BAM file. A SAM file is compressed into a BAM file once it has been aligned (Fig. 9). The most frequent intermediate files used in sequencing analysis are BAM files. IGV may be used to see the file's alignments [268]. All the recalibration techniques for better alignment quality use BAM as an input and output format [210]. Most SV calls need BAM input [230-232]. Input in BAM format is required for all SNV callers [210, 223].

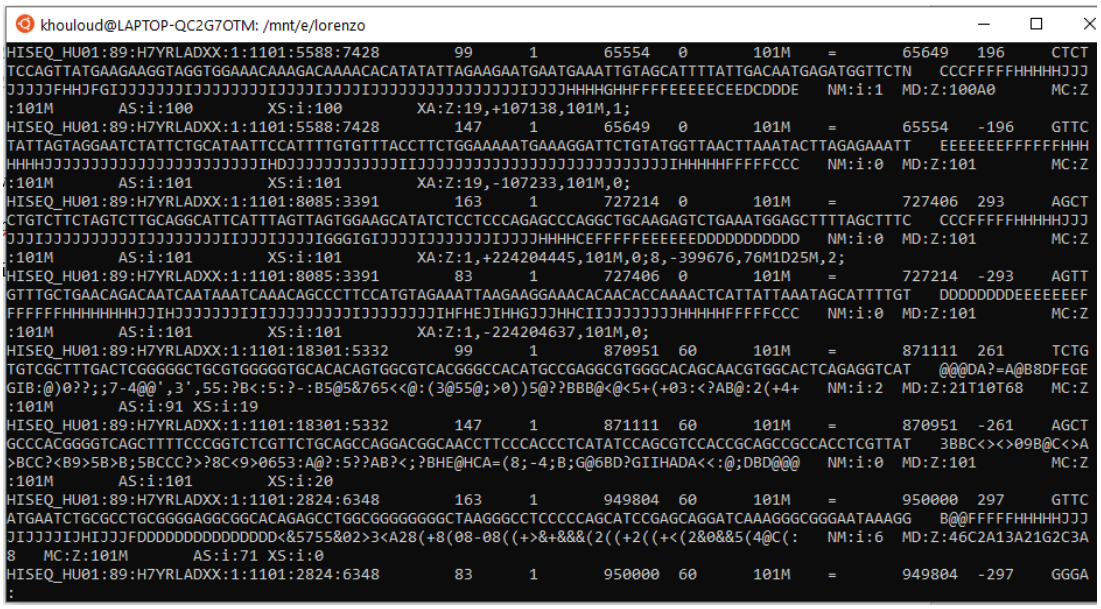


Fig. 9 : Bam File Format

h) VCF

A VCF file is a text-based file used to store genetic variants. The file contains the output of all SNV and SV callers [230, 231]. A header section and a variation section make up the file. The file's general information is included in the header section, which includes the sample name, genotyping data format, and annotation data format. The name of the columns in the variation section appears on the final line of the header section. The variation section uses a tab-limited syntax for each line. One option is represented by a single line. Depending on the number of samples in the file, the number of columns in the variation section varies. Many annotation tools [236, 238] [237] have a VCF export option.

i) TSV

A tab-limited text-based file containing "tab" as an intent for tabular organization that can subsequently be imported into Excel is known as a TSV file. TSV format is supported by a few variant callers [232, 236]. Even though TSV files are more versatile and easier to read than VCF files, they are not widely utilized for storing genetic variations due to their size and inefficiency in storing complicated information such as multiallelic variants or variants with several transcripts.

j) CSV

A CSV file is a text-based, comma-limited file having the same purpose as a TSV file: to be used as a tabular structure. However, because of its non-standardized format, it is not very common. If there are commas or new lines in the text, errors might occur. The file is sometimes used as an intermediary export to be loaded into Excel.

2. Transcriptomic analysis

Whole-transcriptome sequencing using RNA sequencing (RNA-Seq), combined with suitable bioinformatics, is a powerful technique for identifying new mutations that may be used to aid in the diagnosis of genetic disorders [208]. High-throughput RNA sequencing allows a huge amount of comprehensive qualitative and quantitative mapping of all transcripts [269].

Total RNA is extracted, and ribosomal RNA is either removed or polyA-tailed RNA is isolated using poly(T) oligomer magnetic beads (Truseq Illumina Library prep). RNA sample is used to select mRNA for sequencing by its polyadenylation tail, which is then fragmented using sonication in preparation for Illumina sequencing, which provides "short reads" consisting of hundreds of bases.

The mRNA is reverse transcribed into a library of complementary DNA (cDNA) followed end repair, adaptor ligation, and indexing and will be sequenced later on. After being amplified and quantified in a library, RNA readings are mapped to known transcripts and the whole genome in order to enable the identification and quantification of new and existing transcripts. Reads from the sequence are tallied and mapped to a reference transcriptome using a variety of commun alignment software such as (SOAPdenovo trans [270]) [208].

3. Microarrays

In contrast to RNA-seq, Microarrays Expression based on hybridizing cDNAs labeled on DNA fixed on a non-solid surface porous (glass slide). Microarrays are composed of small nucleotide oligomers, referred to as "probes," that are arranged on a solid substrate. The amount of transcripts in a sample is determined by the hybridization of fluorescently tagged transcripts to these probes. The amount of fluorescence at each probe location on the array shows how many

transcripts there are for that probe sequence. Microarrays need some previous information about the organism of interest, for example, in the form of an annotated genome sequence or in the form of an EST library that may be used to produce the probes for the array.

D. Gene Silencing Methods:

1. Gene KNOCKOUT

Gene KNOCKOUT may occur when a double stranded break is created in the DNA of a gene's coding region, causing the gene to stop functioning [271]. One of two repair mechanisms is activated as a result of the event: either non-homologous end joining (NHEJ) or homology directed repair (HDR) (if a repair template is provided). Because NHEJ is a prone to error process, it will often result in the insertion or deletion of nucleotides throughout the process (called InDel mutations) [272]. Because each cell will go through a different editing process, it is essential to check for cell lines that have frameshift mutations that result in a significant change in the translated protein before using them (often truncation). It is possible for a protein to be totally eliminated from a cell line that has had a frameshift mutation in all alleles of the gene. This signifies that no functional protein will be expressed in the cell as a result of this.

CRISPR stands for Clustered Regularly Interspaced Short Palindromic Repeats, is a gene editing technology. The engineered CRISPR system is composed of two essentials parts, which are a single guide RNA (sgRNA) and Cas9 endonuclase. This complex is directed by the sgRNA of 20 pb to cleave double DNA strands in specific sequence called Protospacer Adjacent Motif (PAM) to the 3' of the target sequence and by consequence Error-Prone NHEJ repair of this site.

2. Gene KNOCKDOWN

When the expression of a gene is decreased but not entirely silenced, this is referred to as gene knockdown. This is performed most typically by targeting the mRNA transcript of a gene rather than the DNA of the gene. Either the mRNA is degraded, or the translation process is interrupted. This indicates that certain gene expression is frequently escapes the regulatory mechanisms. RNA interference (iRNA) takes advantage of the endogenous system for miRNA-induced gene silencing to an artificial inhibition of expressed gene via transcriptional regulation.

Short RNAs may be delivered into the cell in two forms: shRNAs (short hairpin RNAs) and siRNAs (short interfering RNAs) (small interfering RNA). Both shRNAs and siRNAs are around 21 bp in length and are engineered to have complementarity to the target mRNA they are intended to target. shRNAs are dsRNAs that have a loop structure and are processed into siRNA by the host enzyme DICER, which is a DICER-dependent enzyme. siRNA are dsRNA with 2 nt 3' end overhangs [273]. A single strand of siRNA will be loaded into the RISC once it has been processed (RNA-induced silencing complex). Complementarity determines whether or not the siRNA will bind to its target. As long as the bonding between the siRNA and the mRNA is perfect, the RISC will be capable of cutting the mRNA. Because of the imperfection in the binding between the siRNA and the mRNA, the siRNA will produce translational inhibition but will not cause mRNA cleavage [274].

III. MATERIALS AND METHODS

A. CANCER GENE FUNCTIONAL STUDIES: SURVIVAL AND ROLE OF TROP2 IN THE REGULATION OF THE AUTOPHAGIC PROCESS

1. Cells

Human pancreas BxPC-3, prostate DU-145 and colorectal KM12-SM cancer cells lines were grown in RPMI 1640 medium supplemented with 10% fetal calf serum (FCS) and 100 IU/ml penicillin and 100 µg/ml streptomycin (Euroclone, Milano, Italy).

For starvation experiment BxPC3 and KM12-SM cells were grown in EBSS, Earle's Balanced Salt Solution, starvation medium that provide essential element for cell survival and after reaching maximum confluency, they were treated by a unique dose of Chloroquine which is autophagy inhibitor at concentration [12.5 µM]. BxPC-3 cancer cells devoid of Trop-2 expression were obtained by somatic CRISPR-Cas9-dependent knock-out (KO) of the *TACSTD2* gene. By using the same technology, Trop-2 expression in KM12-SM cell lines was generated by transfection with a CRISPR-Cas9 resistant version of the gene (KM12-SM/T2). Vector-alone transfected cells were used as negative control (KM12-SM/V). Expression levels of Trop-2 were

assessed by flow cytometry as describe [275]. Fluorescence analyses were carried out on fluorescence-activated cell sorters (FACS Calibur and Canto II, Becton Dickinson).

Absence of mycoplasma contamination in cell cultures was routinely confirmed by PCR analyses [276].

2. In-vitro cell growth assays

BxPC-3 and KM12-SM cancer cells were seeded at 4.0 $\times 10^3$ cells and 25.10^3/well in 96-well plates (three replica wells per data point). Cells were incubated for 11 days (264 hours), mimicking growth in fresh medium followed by growth in spent/exhausted medium, where significant reduction of nutrients is expected. Cell numbers at time zero. Cell numbers were quantified every 12h / 24h and 48h by staining with crystal violet [277]. Cell numbers were normalized against standard reference curves of crystal violet absorbance of two-fold serially diluted cell samples.

B. CANCER GENE FUNCTIONAL STUDIES: TROP2 DESIGNER MUTANTS FOR FUNCTIONAL INVESTIGATION

All steps of **subcloning** that are proceeded to clone Trop-2 mutants are the following:

1. PCR amplification of inserts

The human TROP2 DNA mutants (Trop-2; Proteolytic mutant, Glycolytic mutant, TEV, T88) were amplified by PCR from about 6 ng of DNA plasmid per 50 µl reaction volume (Table I). PCRs for amplification of cloning inserts were carried out using Platinum SuperFi PCR Master Mix DNA Polymerase (Invitrogen Thermo-fisher, Monza, MB, Italy) following the manufacturer's instructions (Table I).

Optimal PCR primer annealing temperatures were chosen according to the oligonucleotide's specific Tm. In most cases forward and reverse primers were designed to anneal the specific sequence of each mutant. Sequences of all oligonucleotides used in this study are listed in the supplementary **(**Table II**)**. We invariably used 35 cycles, by which all amplification reaction program is summarized in (Table III).

Table I : Composition of the reaction mixture for PCR reactions

Reagents	PCR (50 ul)	Templet's origin plasmids
2x Platinum SuperFi PCR Master Mix	1x	
10 uM Forward Primer F	0.5 uM	
10 uM Reverse Primer F	0.5 uM	
Template DNA	6 ng	- DNA template
5x SuperFi GC Enhancer	1x	
Water, nuclease-Free	7 ul	

Table II : Sequences and positions of used primers

Primers Eurofins	Nucleotide sequence	Target	Purpose
hT2-pBabe-BamHI.For	GCG GGA TCC ATG GCT CGG GGC CCC GGC (27)	Trop-2 : Mut prot/ gly/Tev	PCR
hT2-pBabe-EcoRI.Rev	GCG GAA TTC CTA CAA GCT CGG TTC CTT TCT CAA (33)	Trop-2: Mut prot/ gly/Tev/Stem/T-88	PCR
hT2stem-pBabe-BamHI.For	GCG GGA TCC ATG CCC ATG GGG TCT CTG CAA C (31)	Trop-2 Stem	PCR
hT2-T88-pBabe-BamHI.For	GCGGGATCCgccaccATGGCGCCCCCG CAGGTCCTCGCGTTCGGGCTTCTGCT TGCCGCGGCGACGGCGACTTTTGCC GCAGCTACGCTGGTGCGGCCG	Trop-2 threonine 88	PCR
pBabe-seq-3'	ACC CTA ACT GAC ACA CAT TCC (21)	Insert in pBabe	Sequencing
pBabe-seq-5'	CTT TAT CCA GCC CTC AC (17)	Insert in pBabe	Sequencing

Table III : Programs and cycles of enzymatic amplifications used

Amplification parameters	Number of cycles
30s at 98°C	1
20s at 98°C; 20s at 60°C; 1 min at 72°C	35
5 min at 72°C	1
Hold at 4°C	1

To check whether there is DNA amplification, electrophoresis on agarose gel of 1% concentration was carried out, 1g of agarose dissolved in TAE buffer (Tris-Acetate-EDTA, 1x) with ethidium bromide staining, the DNA amplificants are stained with a purple loading buffer (Invitrogen) and are migrated in the same TAE buffer under a constant voltage of 85 V. Fluorescence of DNA fragments are visualized instantaneously by UV transilluminator (280 - 320 nm). The size of DNA fragments is revealed using a size marker 1 Kb Plus DNA Ladder (Invitrogen, Monza, MB, Italy)

Ethanol precipitation (1:10) V of 3M sodium Acetate, 2 V of Ethanol 100% and 1 ul of glycogen) of PCR products followed by enzymatic digestion of both Trop-2 mutant's fragments and pBabe Puromycin Vector with 100 U of BamHI and 60 U of EcoRI restriction enzymes in CutSmart buffer 10x (New England BioLabs, Italy) to creat sticky-end. The double-digestion reactions were incubated at 37 °C for 3 hours. Restriction endonucleases were purchased from New England Biolabs (Italy) and used according to manufacturer's instructions.

In parallel, digestion of both plasmid pEX-k168-Hu2G10-ScVf and pBabe Puro separately with BamHI/SalI in NEBuffer 1.1 (New England BioLabs, Italy) and Buffer O (Thermofisher) for 2h/2h.

2. GenClean Gel purification of digested product

The success of the digestions was verified by a control gel electrophoresis of the samples (insert/vector) that were digested. Then were fractionated by electrophoresis on a long agarose gel (1%) to mainly separate interested fragment. After electrophoresis, the gel was placed onto the transilluminator and longitudinally sliced in between the parallel lanes. A quick excise of DNA band using low light intensity of UV were recommended to avoid DNA damaging. DNA was recovered from the agarose gel using GenClean gel extraction protocol; An equivalent volume to the band of TE buffer (10 mM Tris pH 8, 1 mM EDTA pH 8) was added and mixed with 3 volume of 3 M Sodium iodide.

After 5-10 min incubation in 50°C, 30 ul of prepared suspension glass milk solution was added to isolate and purify DNA followed by ice incubation for 10 min with periodic resuspension. 3 washes of the pellet with 1 ml of cold wash buffer (10 mM Tris-HCl pH 7.5-8, 1 mM EDTA, 50 mM NaCl

and 50% Ethanol) [278]. To separate the DNA from glass milk binding an equivalent volume to the pellet of TE was added followed by incubation for 5 min at 55°C. The amount of purified DNA (Elution s) was quantified using nanodrop (Nanophotometer® N120) and verified by a control gel electrophoresis 1%.

3. Quick CIP vector dephosphorylation and ligation

To prevents re-ligation of linearized plasmid DNA a heat-labile recombinant version of calf intestinal alkaline phosphatase (CIP) (New England BioLabs #M0510S) was used on pBabe puromycin digested with BamHI/EcorI and BamHI/SalI. A 1 pmoli of vector was dephosphorylated with 1 ul (10U/ul) of enzyme in 10 min in 1x of Cutsmart buffer and irreversibly inactivated by heating it at 80°C for 2 minutes.

Ligation is the process in which an insert is annealed into a vector by a covalent bond called phosphodiester bond. DNA ligases catalyze the end-to-end joining of the DNA by forming a phosphodiester bond between the 3'hydroxyl and the 5'phosphate ends of nucleic acid molecules [279]. To obtain a recombinant vector, 500 fmole of fragment (Trop-2 mutants) were annealed to 50 fmole of vector using 40 units of T4 DNA ligase (New England Biolabs M0202S) followed by incubation for 10 min at RT°.

4. Transformations of pBabe-Puro-Trop-2 mutants

Competent cells were prepared according to the rubidium chloride protocol. E. coli Stbl3 Chemically Competent cells (Invitrogen™) were transformed; 10 µl of each ligation mixture of plasmid pBabe-puro-Trop-2 mutants was added to 100 µl StBl3, in Falcon™ Round-Bottom Polypropylene Tubes, under sterilizing conditions using Bunsen burner.

Samples (tubes are semi closed aero-anaerobic conditions) were incubated on ice for 20 minutes followed by a 45 second heat shock in a 42 °C water bath. Immediately following, the samples were placed on ice for two minutes. 900 ul of sterile LB media was added to each sample and incubated in a 37 °C shaking incubator for 1 hour at 200 rpm. After centrifugation at 3000 rpm for 5 min, cells were plated on LB plates containing 0.1 mg/mL ampicillin and incubated at 37 °C O.N (~16 hours) for colonies growth.

5. Mini-prep Qiagen plasmid extraction, digestion, and sequencing

Fives single isolated colonies from each transformation were inoculated in selective LB-Agar array plate for maintaining and separately in 5 ml of LB media containing 0.1 mg/mL ampicillin for plasmid mini-prep extraction (Cat No./ID: 27106, QIAprep Spin Miniprep Kit (250)) following the manufacturer's instruction. Incubation at 37 °C Overnight under agitation for LB media falcons (200 rpm).

Cell cultures were harvested by centrifugation at 3000 rpm for 15 min at 15°C. Harvested cells were resuspended with 250 ul of solution P1 containing RNase enzyme, lysed with 250 ul of solution P2 for 5min followed by precipitation of debris and E. coli chromosomal DNA with 350 ul of solution N3. Plasmid DNAs were recuperated in spin column resin. A wash buffer PB was added to eliminate debris flowed by another wash using PE buffer (10 mM Tris-HCl pH 7.5, 80% ethanol). Finally, elution of DNA in 30 ul of Elution buffer (10mM Tris-HCl, pH8.5).

In order to verify that the gene was successfully cloned (creating Trop-2-(Gly/MutProt/TEV/T88)-pBabe B/E), the miniprepped samples were digested with restriction enzymes. 5 µl of purified plasmid was digested with 6 units of each enzyme (either BamHI-HF and EcoRI-HF for pBabe-Trop-2). The purified plasmids were analyzed by 1% agarose gel electrophoresis. To avoid the presence of mutation of our insert during the first step of amplification, the final recombinant plasmid pBabe should be sequenced. According to BMR Genomic company protocol, 600 ng of DNA plasmid mixed with 6.4 pmole of Prime Forward or Reverse were lyophilized at 60 °C and sent back to the company for Sanger sequencing. The universal primer of pBabe plasmid were regrouped in (Table I). Electropherogram was analyzed using the program Finch-TV. Once confirmed, plasmids are ready to transfection process.

6. Lentiviral production and Cell transfection

To generate lentiviral particles, HEK293FT cells in 10 cm culture plates (3 million cells per plate) were co-transfected with 4 µg of pBabe-Trop-2 vector, 3.6 µg of pUMVC (gag/pol fusion protein), 0.4 µg of pCMV-VSVG-retro (VSV-G envelope glycoprotein) plasmids using Lipofectamine 3000 (Invitrogen L3000-008) in Opti-MEM® medium followed by incubation for

7h a 37°C + 5% CO2. Lentivirus-containing supernatants were harvested 48 h after transfection in DMEM medium, followed by filtration through a 0.45 µm pore diameter filter, and used to infect cultures of Mouse Thymic Epithelium-4-14 (MTE-4-14) and KM12-SM cell line in presence of polybrene (4ug/ml for Mouse cells,8 ug/ml for human cells) to ameliorate the efficiency of retroviral infection. Polybrene medium were changed after 7h with fresh supplemented medium (DMEM/RPMI). After 48h, stably infected cells were cultured in the presence of puromycin selector (2ug/ml MTE-4-14 and 3ug/ml KM12-SM) for at least one week, and expression of Trop-2 was analyzed by Flow Cytometry [275] where cell were subjected to single staining with Conjugated Antibody T16-488 [8].

C. CANCER GENE FUNCTIONAL STUDIES : TROP-2 CLEAVAGE BY ADAM10 IS AN ACTIVATOR SWITCH FOR CANCER GROWTH AND METASTASIS (Article I)

1. Cells

The human ovarian OVCA-432, mammary MCF-7, pancreatic BxPC3, and colorectal HT-29, HCT-116 and KM12SM [280] cancer cell lines, and the murine myeloma NS-0 cells, were grown in RPMI 1640 medium supplemented with 10% fetal calf serum. The human 293 kidney and the murine MTE 4-14 immortalized thymic epithelium [281] and L fibrosarcoma [282] cell lines were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum. The cells were transfected with purified DNA [283] using Lipofectamine 2000 or LTX (Invitrogen). Stable transfectants were propagated in complete medium supplemented with 100 µg/mL G418. MTE4-14/Trop-2 transfectants acquire specific features of cancer cells, as they become tumorigenic in vivo, at variance with MTE4- 14/vector transfectants [118]. Growth in suspension was obtained by culturing cells in 6-well ultra-low attachment plates. Cells were seeded at 106 per well, and analyzed after 5 d.

2. Tumor patient case series

A human breast cancer case series [147] was analyzed, where eligible patients showed N0, T1/T2 tumors. Snap-frozen material was available for 55 cases. The frozen tissues were ground in

liquid nitrogen and processed as for Western blotting. Studies on human tumour samples were approved by the Italian Ministry of Health (RicOncol RF-EMR-2006-361866, 2006).

3. Plasmids

The pBJI-neo vector was provided by Dr. M. Davis. The Bluescript vector was obtained from Stratagene (La Jolla, CA). The pcDNA3 expression vector was obtained from Invitrogen (Gröningen, The Netherlands). The pEYFP-N1 was obtained by Clontech (Palo Alto, CA).

TROP2 constructs. A 970 bp full-length *E1/TROP2* cDNA from cells from the ovarian cancer fragment FE was isolated by RT-PCR. The amplified band was digested with BamHI and EcoRI and subcloned in the corresponding sites of the pcDNA3 expression vector using the following primers:

F: 5'-CTCGGATCCATGGCTCGGGGCCCCGGCCTC-3'

R: 5'-CTCGAATTCCAAGCTCGGTTCCTTTCTCAA-3'

A pEYFP corresponding vector devoid of the coding sequence of EYFP was used to express Trop-2 in mammalian cells. All pΔEGFP-expressing cells are indicated as "control", or "vector-alone" cells. Wild-type *TROP2* was obtained by PCR from the original full length *TROP2* clone [67], and inserted in the vector at the XhoI/KpnI sites, using the following primers:

F: 5'-GCGATTctcgagTCCGGTCCGCGTTCC-3'

R: 5'-GCGCCggtaccAAGCTCGGTTCCTTTC-3'

The A87-A88 Trop-2 mutant was generated by site-specific mutagenesis of Trop-2 proteolytic site with the Quikchange® Site-directed mutagenesis kit (Stratagene) following the instruction of the manufacturer, using the following primers:

F: 5'-AGCGCCCCCAAGAACGCCGCCGCGCTGGTGCGGCCG-3'

R: 5'-CGGCCGCACCAGCGCGGCGGCGTTCTTGGGGGCGCT-3'

The mutagenized coding region was entirely sequenced, to verify the absence of any Taq polymerase-induced mutations.

4. siRNA and shRNA

Complementary strategies were utilized for siRNA design, including those following Tuschl' criteria [284] and the Sonnhammer procedure [285]. siRNA were synthesized and the corresponding annealed oligos were subcloned into the pSUPER vector, kindly provided by Dr. R. Bernard [286]. siRNA expression constructs were transiently transfected in MTE4-14 transfectants, and cell growth was measured. siRNA-targeted transcript levels were quantified by real-time PCR. Negative control siRNAs directed towards irrelevant targets were used; these were chosen after extensive testing for lack of off-target influence on cell growth.

siRNA targeting murine *ADAM10* (NM_007399.3, nt 292-309)

F: 5'-gatccccAGACATTATGAAGGATTATttcaagagaATAATCCTTCATAATGTCTtttttggaaa-3'

R: 5'-agcttttccaaaaaAGACATTATGAAGGATTATtctcttgaaATAATCCTTCATAATGTCTgggg-3'

siRNA targeting murine *MMP9* (NM_013599, nt 1780-1798)

F: 5'-gatccccCCGCAGACCAAGAGGGTTTttcaagagaAAACCCTCTTGGTCTGCGGttttttggaaa-3'

R:5'-agcttttccaaaaaCCGCAGACCAAGAGGGTTTtctcttgaaAAACCCTCTTGGTCTGCGGgggg-3'

Scramble and ADAM10 shRNA cloned in the pLVTHM-GFP vector as well as the lentivirus helper plasmids psPAX and pMD2g were kindly provided by Dr. A. Ludwig [287]. Corresponding lentiviruses were generated by co-transfection of HEK293T cells with the three plasmids, and viral particles were collected from the culture medium after 48 h.

Target cells were then infected with the lentiviruses pre-mixed with Polybrene (8 µg/ml). GFP^+ cells were sorted after one week by flow cytometry to obtain enriched populations of cells expressing the corresponding shRNAs.

5. Immunoprecipitation

Cells were washed with cold wash buffer (10 mM HEPES, 150mM NaCl, 1mM $CaCl_2$,1 mM $MgCl_2$, pH 8) and lysed with 3 ml of lysis buffer (1% Triton, 0,1% SDS, 20 mM NaF, 1 mM $NaVO_4$, 2 mM PMSF, 5 µM Pepstatin, 5 µM Leupeptin). Four mg of lysate were incubated with rotation at 4 °C for 3 h with T16-NHS Sepharose. The resin was washed 3 times with PBS and T16-bound Trop-2/protein complexes were eluted 3 times with 150 µl of glycine buffer pH 2.5. Eluted solutions were immediately neutralized with TRIS 1M pH 11.5.

6. Antigen purification and protein sequencing

Trop-2 was purified by affinity chromatography over a Sepharose-E1 mAb column. Briefly, monolayers of cells from the ovarian cancer sample FE were extensively washed in PBS and lysed for 20 min at 4 °C in 20 mM TRIS pH 7.4, 150 mM NaCl, 0.5 % Triton X-100, 50 U/ml aprotinin, 1 mM AEBSF. The cell lysate was centrifuged at 1,500 g for 10 min; the supernatant was cleared by centrifugation at 12,000 g for 20 min and passed through a hydroxyapatite column (DNA-grade Bio-Gel HTP, Richmond, CA). The unbound fraction was loaded onto a Sepharose-E1 mAb column. Bound material was eluted with 0.1 M glycine pH 2.7, 0.1 % Triton X-100. The eluate was immediately neutralized with 1 M TRIS pH 9, and eluted fractions were analyzed by SDS-PAGE on 4-18 % gradient gels and Western blotting. Fractions containing the E1-immunoreactive material were pooled, concentrated and transferred to a PVDF membrane. N-terminal sequences of the blotted protein samples were obtained by Edman degradation. The N-terminus of uncleaved Trop-2 appeared blocked/resistant to sequencing procedures.

7. Proteomic chip analysis

Transfectants for vector-alone or *TROP2* [283] were homogenized in lysis buffer containing Ser/Thr phosphatase inhibitors (30 mM NaF, 60 mM β-glycerophosphate and 20 mM sodium pyrophosphate) and a Tyr phosphatase inhibitor (1 mM Na_3VO_4). Fifty µg of protein lysates were labeled with 2 mg/ml Cy-3 or Cy-5. Labeled vector- and Trop-2-transfectant lysates were incubated side-by-side on a KAM-1.1 antibody microarray (www.kinexus.ca). Two samples were analyzed per chip, in duplicate for each sample. Each microarray contained 609 antibodies directed against phosphorylation sites or framework regions of 240 protein kinases or proteins

regulating cell proliferation, stress and apoptosis (N=111). Array images were acquired with a ScanArray scanner (Perkin Elmer, Wellesley). Spots with uneven morphology were excluded from the analysis. Signal median, background median, signal-to-noise ratios and percent error ranges were quantified. Background-subtracted net signals were normalized against global signals in the slide. Thresholds of percent change from control ≥20%, absolute net signal counts ≥1,000 and a t-test P value ≤0.05 were imposed to identify relevant changes.

8. Immunohistochemistry (IHC) analysis

Tumor samples were fixed in phosphate-buffered formalin and embedded in paraffin, following standard procedures. Five-µm sections were mounted on silanised slides, deparaffinized and rehydrated. Endogenous peroxidase was inactivated with 3% hydrogen peroxide. After blocking with 1% BSA, 2% rabbit serum, sections were incubated with the 162-46.2 monoclonal anti-Trop-2 antibody [282], at room temperature for 30 min. Bound mAbs were detected with anti-mouse horseradish peroxidase-labelled polymerized secondary antibodies (DAKO EnVision™ System), incubated with 3,3'-diaminobenzidine (DAKO). Counterstaining was performed with Mayer hematoxylin. Slides were mounted with Immunomount (Shandon, Pittsburgh, PA) and scored for percentage of expressing cells and staining intensity (1: low/nil; 2: moderate; 3: strong).

9. MS analysis

Differentially expressed spots were excised and incubated overnight at 37 °C with sequencing-grade porcine trypsin (Promega). A 1 ml aliquot of each peptide mixture was analyzed by MALDI MS (www.york.ac.uk/depts/biol/tf/proteomics/). Positive-ion mass spectra were obtained using a Bruker Ultraflex III in reflectron mode, equipped with a Nd:YAG smart-beam laser. Identifications were obtained by MS-Fit web-application of Protein Prospector v5.7.3 (prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msfitstandard; UCSF).

The peaks list by the mass spectrometer was used as input. We imposed as costant modification the carbamidomethylation, by reaction with iodoacetamide. Methionine oxidation, N-terminus acetylation, and transformation of peptide N-terminal Gln to pyroGlu was imposed like recognizable modifications. The mass tolerance was set at 100 ppm. Contaminant masses of the matrix α-cyano-4-hydroxycinnamic acid were excluded from analysis. Protein identifications

were ranked on MOWSE score. The MOWSE score reported by MS-Fit is based on an updated scoring system [288].

10. Proteolytic cleavage sites MS identification

To identify proteolytic cleavage sites of Trop-2, PMF-mass spectrometer-peak lists was analyzed by MS-Fit. To identify activatory cleavage *in vivo*, spectra were scanned for peptide sequences that did not fit trypsin cleavage consensus sites. The mass tolerance was set at 20 or 50 ppm (50 ppm data are shown). Contaminant masses of matrix α-cyano-4-hydroxycinnamic acid were excluded from analysis.

11. Western blotting

Western blotting was performed as described [94]. Tumors from experimental animals were rapidly frozen upon surgical removal and stored at –80 °C. Frozen samples were quickly thawed and homogenized with a Dounce homogenizer. Cleared homogenates of tumors and cultured cell lysates were analyzed by SDS-PAGE and transferred to nitrocellulose membrane for subsequent processing by Western blotting. Ponceau red staining was routinely used as control of protein loading [289]. Quantification of the protein band signals was performed using Fiji/ImageJ.

12. Quantitative RT-PCR

One µg of total RNA was reverse transcribed with the ImProm-II Reverse Transcriptase (Promega) according to standard protocols. Quantitative RT-PCR was performed using an ABI-PRISM 7900HT Sequence Detection System and Power SYBR® Green PCR Master Mix (PE Applied Biosystems, Foster City, CA) according to the manufacturer instructions [290], using the listed primers:

Murine *ADAM10* F: 5'-AGCAACATCTGGGGACAAAC-3'

Murine *ADAM10* R: 5'-TTGCACTGGTCACTGTAGCC-3'

Murine *MMP9* F: 5'-GTGCGACCACATCGAACTT-3'

Murine *MMP9* R: 5'-GGATGCCGTCTATGTCGTCT-3'

Murine *B2M* F: 5'-GAGCCCAAGACCGTCTACTG-3'

Murine *B2M* R: 5'-AAAGAAGGTGATGTGTACATTGCT-3'

Each sample was assayed in triplicate. The $2^{-\Delta\Delta CT}$ method was used to calculate relative changes in gene expression [291]. The ß2-microglobulin (*B2M*) housekeeping gene was used as an internal control. For set-up curves, ΔC_T ($C_{T,\text{ target gene}}$ - $C_{T,\text{ GAPDH}}$) were calculated for each cDNA dilution. The data were fit using least-squares linear regression analysis. As relative amplification efficiency was invariant over the range of RNA amounts used [292-294], amplification curves were used to calculate cross-over point values for siRNA-treated samples. Each sample was routinely assessed for genomic DNA contamination by using non-retrotranscribed RNA isolates as templates for PCR reactions.

13. Sequence analysis

Subcloned PCR bands were sequenced using the method of Sanger [295], to confirm the correctness of the construct and the absence of PCR-induced mutations. DNA and protein sequences were analyzed using GCG and EMBOSS (www.uk.embnet.org/Software/EMBOSS/) programs. Analysis of ADAM10 substrates was performed through the Merops database (merops.sanger.ac.uk/cgi-bin/protsearch.pl). The interrogation output was cross-checked with additional, updated published data on native-protein cleavage-sites.

14. 3D structure analysis and modelling

The structure of the thyroglobulin domain of the p41 isoform of the invariant chain of MHC [296] class II (PDB code 1ICF, chain I) was used as a template for the homology-modelling of the Trop-2 thyroglobulin region. The Trop-2 and p41 thyroglobulin domains were aligned using GAP and NEEDLE software. Alignments were manually refined in regions with lower conservation, using conserved cysteines and secondary structures as anchor sites. A short α-helix around the first cysteine of the Trop-2 thyroglobulin domain was concordantly predicted by secondary-structure prediction programs (PHDsec at cubic.bioc.columbia.edu/predictprotein/; jpred^2 at jura.ebi.ac.uk:8888/; 3D-pssm at www.bmm.icnet.uk/~3dpssm/), in good correspondence to that of p41. A model of the tertiary structure of the Trop-2 thyroglobulin domain was built using the programs MODELLER-4 [297] and WHATIF (bioslave.uio.no/Programs/MVL/index.php3). Similar results were obtained with the two softwares and only the model generated by

MODELLER was further analyzed. The stereochemical properties of the resulting models were assessed with the program PROCHECK.

The graphic representations of the MODELLER outputs were prepared with the MOLMOL, Swiss-PdbViewer (www.expasy.ch/swissmod/SWISS-MODEL.html), RasMol (www.umass.edu/microbio/rasmol/) and Raster3D (asdp.bnl.gov/asda/LSD/Modeling/Raster3D.html) [298] programs.

Modelling and model evaluation were performed on a Silicon Graphics Octane r12000 workstation. PDB files of ADAM10-cleaved proteins were retrieved from the PDB databank (www.rcsb.org/pdb/home/home.do) and analyzed as described above. Structure images were generated with PyMol (www.pymol.org/).

15. Antibodies

Balb/c mice were immunized with cells from an ovarian cancer fragment (sample FE). Cell fusion and hybridoma cloning were carried out as previously described [299]. Screening for cell-surface-reactive hybridomas was performed by immunohistochemistry of ovarian cancer cells [300], and by flow cytometry of Trop-2 transfectants. The E1 (IgG2ak by immuno-isotyping), 162-46.2 [282], and T16 [301] anti-Trop-2 [302] hybridomas were grown in serum-free medium. The RS7 anti-Trop-2/EGP-1 ascites was kindly provided by Rhona Stein. Rabbit polyclonal anti-Trop-2 antisera were generated by subcutaneous immunization with the recombinant extracellular domain of human Trop-2 synthesized in bacteria [94], or with KLH-conjugated, N-ter biotinylated peptides corresponding to the cytoplasmic tail of human Trop-2.

Anti Trop-2 polyclonal antibodies were purified by affinity chromatography on recombinant Trop-2 conjugated to NHS Sepharose (GE Healthcare) or biotinylated Trop-2 cytoplasmatic tails conjugated to Streptavidin Agarose (Sigma-Aldrich). The AF650 polyclonal goat anti-Trop-2 antibody was from R&D Systems; the rabbit polyclonal anti-ADAM10 antibody was from Merck-Millipore (AB19026); and the goat polyclonal anti-ADAM10 antibody (sc-31853) was from Santa Cruz Biotechnology.

16. ADAM10 inhibitors

Cells were treated with ADAM10 inhibitors for 24 h at the minimum concentration found to be effective for inhibition of Trop-2 cleavage, and then they were assayed as indicated. The GM6001 metalloprotease family inhibitor (Calbiochem) was dissolved in dimethyl sulfoxide and used at 6.5 µM to 13 µM for 24 h. The GI254023X ADAM10 and MMP-9 inhibitor was kindly provided by A. Ludwig and was dissolved in dimethyl sulfoxide. The cells received two doses of 10 µM GI254023X every 24 h. Control cells received vehicle alone.

17. Flow cytometry

Cell staining for flow cytometry was performed as described previously [115, 275, 282, 303, 304]. Reconstituted mixtures of L cells and Trop-2 transfectants [303] were used for E1 mAb binding and competition studies of E1 with other anti-Trop-2 mAbs, where the cell mixtures were preincubated with 100-fold excess of the indicated ascites.

18. Immunofluorescence microscopy

Cells grown on glass coverslips were fixed with 4% paraformaldehyde in phosphate-buffered saline for 20 min. The cells were permeabilized and blocked in 10% fetal bovine serum, 0.1% saponin [305], and then stained with the 162-46.2 [282], T16 [301], E1 anti-Trop-2 [302], and anti-ADAM10 antibodies. The cells were analyzed by confocal microscopy (LSM-510 META and LSM-800; Zeiss).

19. Confocal time-lapse microscopy

Live cells cultured on glass slides were analyzed in Leibovitz's F15 culture medium without phenol red and bicarbonate, supplemented with 10% fetal bovine serum, 100 IU/mL penicillin, 100 µg/mL streptomycin (Euroclone), and 2 mM N-acetylcysteine (Sigma), to reduce free-radical damage. The cells were analyzed using a confocal microscope (LSM-510 META; Zeiss), with images captured at 1 min intervals.

20. In vitro cell growth assays

MTE 4-14 and KM12SM transfectants were seeded at 1.5-3.0 × 103 cells/well in 96-well plates (5 replica wells per data point). Cell numbers were quantified by staining with Crystal violet [277].

21. Experimental tumors and metastases

Cancer cell lines and TROP2-transfectants were injected subcutaneously into 8-wk-old female athymic Crl:CD1-Foxn1nu mice (Charles River Laboratories). Subcutaneous tumor growth curves were obtained by weekly measurements of tumor volumes (d2 × D/2; shortest diameter2 × longest diameter/2), and computed as described previously [306].

To assess for impact on metastatic dissemination, KM12SM colon cancer cell [280] transfectants were injected into the spleen of 8-wk-old female athymic Crl:CD1-Foxn1nu mice. After 4 weeks, the mice were euthanized, and tumor growth and dissemination to the liver and other organs were determined by comparisons of the tumor volumes between the experimental groups. All the autoptic samples underwent microscopy histopathology analysis to detect minimal tumors and the metastatic burden.

22. Statistical analysis

Student t tests were used for comparisons between mean protein levels in the control and TROP2 transfectants in the antibody microarrays. The normality of the distributions of the assay values was verified (www.graphpad. com). Spearman nonparametric correlation coefficients were computed for protein expression levels in human cancer samples.

Two-way ANOVA and post-hoc Bonferroni's t tests were used for growth curve comparisons. The data were analyzed using SigmaStat (SPSS Science Software UK Ltd.) (www.spss. com/software/science/sigmastat/) and GraphPad Prism (GraphPad Software Inc., La Jolla, CA, USA) (www.graphpad.com).

D. GENOMIC ANALISIS OF TROP2 (TRANSCRIPTOM/ NGS): TROP-2 INDUCES ADAM10-MEDIATED CLEAVAGE OF E-CADHERIN AND DRIVES EMT-LESS METASTASIS IN COLON CANCER (Article II)

1. Cells

The KM12C, KM12SM and KM12L4A human colon carcinoma cell lines were obtained from I. J. Fidler, M. D. Anderson Cancer Center. The KM12SM and KM12L4A were derived from the non-metastatic KM12C, by in vivo selection for acquisition of metastatic ability [307]. The HCT116 and HCT116U5.5 human colon carcinoma cell lines [308] were obtained from Roche Diagnostics GmbH (Penzberg, Germany). The HCT116U5.5 clone was derived from the non-metastatic HCT116, by in vivo selection for acquisition of metastatic ability [308]. Trop-2 low and Trop-2high subpopulations were derived from the parental HT29 colon cancer cell line by fluorescence-activated cell sorting. Frozen master stocks were prepared for each cell line within the first two months from receipt. Experiments were performed on cells that had been resuscitated from master stocks and passaged for no more than 6 months. Cell cultures were periodically checked for absence of mycoplasma contamination by PCR analyses and DAPI staining. The IGROV-1, HT29, HCT116 and KM12 cells were maintained in RPMI 1640 medium with 10% or 20% FBS respectively, sodium pyruvate, non- essential amino acids, L-glutamine and vitamins (Invitrogen). Murine MTE 4-14 cells [1] were grown in DMEM with 10% FBS. Lipofectamine 2000 or LTX (Invitrogen) were used for DNA transfection. Stable transfectants were selected with G-418 and enriched for expression as described [8].

2. Coimmunoprecipitation and Western blotting

Coimmunoprecipitation was performed as described [131]. Cell lysates and coimmunoprecipitates were analyzed as previously described [94]. Blots were incubated with relevant primary antibodies (as detailed in Supplementary Methods) in Tris-buffered saline and 5% non-fat dry milk. Hybridized filters were washed in TBS, 0.1% Tween-20. Antibody binding was revealed by chemiluminescence, using horseradish-peroxidase-conjugated anti-mouse

(#401215; Calbiochem), anti-rabbit (sc-2004; Santa Cruz Biotechnology) or anti-goat (sc-2020; Santa Cruz Biotechnology) pAbs. Cleavage of E-cadherin was assessed by comparing the signals from the antibodies recognizing the extracellular and intracellular domains of human E-cadherin. Ponceau red staining was routinely used as control of protein loading [289]. Signal intensity was quantified with ImageJ 1.47, using as reference a Kodak gray-scale standards power curve.

3. Cell aggregation

Cells at 70-90% confluence were detached from tissue culture plates and seeded in low-attachment plates to minimize interference by substrate binding [150]. Cell aggregation was assessed after 1h and 24h incubation at 37°C. Aggregates were quantified by image analysis.

4. Invasion assay

Inverse matrigel invasion assay was carried out as described in [309]. Briefly, 100 µl of a matrigel/serum-free medium mixture (1:1) was pipetted into Transwell inserts in a 24-well tissue culture plate and incubated for 30 min at 37°C for settling. Then, the Transwell inserts were inverted and 100 µl of cell suspension (containing 10,000 cells) was pipetted onto the filter. Inverted plates were incubated for 4h to allow for cell attachment. The plate was then turned right-side-up and each Transwell insert placed in a well containing serum-free medium. Complete medium was gently pipetted on top of the set matrigel and incubated for 5 d at 37°C. Transwell inserts were placed in fresh 24-well dishes containing 1 ml of 4 µM Calcein AM stain solution. Additional 300 µl of Calcein AM solution was pipetted on top of each filter. After 1h of incubation at 37°C, cells were imaged by confocal microscopy with a 20x objective with optical Z-sections scanned at 15-µm intervals moving up from the underside of the filter into the matrigel.

5. β-catenin reporter assays

KM12C and KM12SM cells were co-transfected with the β-catenin reporter pTOPFLASH-EGFP (a kind gift from Prof. Rolf Kemler) and plasmids encoding for wtTrop-2, ΔcytoTrop-2 or with control empty vectors. At 48 hours after transfection cells were detached and stained with the T16 anti-Trop-2 mAb conjugated with Alexa-633 for 20 min at 4°C. GFP and Trop-2 expression were analyzed by flow cytometry.

6. Flow cytometry

Cell analyses and sorts were performed as described [302] (FACSCalibur, FACSCanto II, FACSAria III flow cytometers/cell sorters BD Biosciences). Subtraction of cell autofluorescence and displacement in the red channel were performed for Alexa488-stained cells [8].

7. Pre-clinical models

KM12 cell transfectants were injected in the spleen of 8-week-old female athymic Crl:CD1-Foxn1^{nu} mice (Charles River Laboratories). After 4 weeks mice were euthanized; tumor growth and diffusion to the liver or other organs were determined. All autoptic samples underwent microscopy histopathology analysis to detect minimal primary tumor or metastatic burdens.

Pre-clinical protocols were approved by the Italian Ministry of Health (Prog. 19, 2006) and by the Interuniversity Animal Research Ethics Committee (CEISA) of Chieti-Pescara and Teramo Universities (Prot.26/2011/CEISA/PROG/16).

8. Patients

IHC analysis for Trop-2 expression was performed on normal colon mucosa, colon cancer and colon cancer metastases. Trios of normal tissue, primary tumor, and metastasis from individual patients bearing breast, uterus, ovary, or colon cancers were analyzed by cDNA RT-PCR/dot blot assay. Kaplan-Meier curves and Cox proportional hazard models were utilized to quantify impact on relapse-free and overall survival of cancer patients.

Studies on human tumor samples were approved by the Italian Ministry of Health (RicOncol RF-EMR-2006-361866, 2006) and by the ethical committee of the University of Regensburg (Germany).

9. Plasmids supps

The pEYFP-N1 expression vector, upon removal of the coding sequence of enhanced yellow fluorescent protein (EYFP), was used to express wtTrop-2 and ΔcytoTrop-2 in mammalian cells. The wtTrop-2 was obtained as described [1]. The tail-less ΔcytoTrop-2 mutant was obtained

by PCR from the original full length TROP2 clone [67] and inserted into the vector at the BamHI/EcoRI sites, using the following primers:

F: 5'- CTCGGATCCACCATGGCTCGGGGCCCC -3'

R: 5'- CTCGAATTCCCGGTTGGTGATCACGTC -3'.

The amplified region was sequenced to verify the absence of mutations induced by Taq polymerase. shRNA for ADAM10 and scramble control cloned in the pLVTHM-GFP vector were kindly provided by Dr. A. Ludwig [287]. Lentiviruses were generated in the packaging HEK293T cell line and used to infect target cells as described [1]. GFP+ cells expressing the corresponding shRNAs were sorted by flow cytometry.

10. 2D Cell growth assays (monolayer)

MTE4-14 and IGROV-1 transfectants were seeded at 1.5-3.0 ×103 cells/well in 96-well plates (five replica wells per data point) [293]. Cell growth was quantified by staining with crystal violet [7], upon normalization against standards of cell-reference curves of two-fold serially diluted cell samples. For starvation experiments, KM12SM transfectants were seeded at 25 ×103 cells/well in 96-well plates (5 replica wells per data point) [293], after 72 hours the growth medium was substituted with Earle's Balanced Salt Solution and surviving cells were quantified as above for the following 48 hours (3 replica wells per data point).

11. Wound healing

Cell monolayers of MTE4-14 transfectants were scratched using µm-volume pipet tips, washed twice with PBS, and incubated in low-growth factor (2.5% FBS) medium. Images were taken at fixed coordinates, at pre-determined time points, by phase-contrast microscopy.

12. Antibodies

The T16 anti-Trop-2 monoclonal antibody (mAb) [301] was used for flow cytometry and immunofluorescence analyses. IHC was performed with the 162-46.2 (ATCC clone HB187) anti-Trop-2 mAb or the AF650 anti-Trop-2 goat polyclonal antibody (pAb; R&D Systems). Target/antibody pairs were: activated β-catenin, as unphosphorylated at S37/T41 [310] (clone 8E7,

05-655; Millipore), β-actin (JLA20, Oncogene Research Products), E-Cadherin N-terminal 5 extracellular domain (Santa Cruz 67A4) mouse mAbs; E-Cadherin C-terminal cytoplasmic domain (24E10; Cell Signaling #3195), β-catenin (D10A8; #8480) rabbit mAbs (Cell Signaling Technology); ADAM10 (AB19026 Merck-Millipore), Rb (C-15; sc-50), phospho-Akt (Thr 308; sc- 16646-R) rabbit pAbs; ADAM10 (sc-31853), phospho-GSK-3β (Ser 9; sc-11757); Akt (C-20; sc-1618) goat pAbs (Santa Cruz Biotechnology).

13. Multiplex Western blotting

Cleared cell homogenates (300 μg each) were loaded on a single, gel-wide lane for electrophoresis and transferred to a nitrocellulose membrane. A manifold with 20 separate channels was overlaid over the membrane, and mixtures of antibodies against framework or phosphorylation-dependent epitopes were applied to each slot (KinetworksTM analyses; Kinexus) (www.kinexus.ca/index.html). Antibody binding was revealed by chemiluminescence (Amersham). Chemiluminescence images were acquired using a 16-bit camera (Bio-Rad Fluor-S Max Multi- Imager). Images were analyzed with the Bio-Rad Quantity One quantitation software. Each immunoblotted sample was scanned for its maximum scan time, to ensure that the signal for the strongest immunoreactive protein on the immunoblot was just below saturation. This ensured the detection of minor immunoreactive proteins and linear quantitation over a 2,000-fold concentration range [311]. The raw data from the scans were normalized to counts per minute (cpm) versus the average total signal of all samples. Detected bands were identified with the IRIS programme (immuno-reactivity identification system, Kinexus). Band identification and quantification were performed under operator supervision.

14. Protein microarray analyses

KM12SM cells stably transfected with vector alone, wtTrop-2 or ΔcytoTrop-2 were homogenised in lysis buffer (20 mM MOPS, 2 mM EGTA, 5 mM EDTA, 1 mM PMSF, 3 mM benzamidine, 5 μM pepstatin, 10 μM leupeptin, 0.5% Triton X-100, 1 mM dTT), Ser/Thr phosphatase inhibitors (30 mM NaF, 60 mM β-glycerophosphate and 20 mM sodium pyrophosphate), Tyr phosphatase inhibitor (1 mM Na3VO4), pH 7.2. Fifty μg of protein lysate were labelled with 2 mg/ml Cy-3 or Cy-5. Labelled sample lysates were incubated on a KAM-1.1

antibody microarray (Kinexus). In each chip two samples were analysed in duplicate. Each microarray contained 609 antibodies (phospho-sites or framework regions of 240 protein kinases and 111 others signaling proteins) (www.kinexus.ca/kinex.htm). After probing the array was washed and the images were acquired by ScanArray scanner (Perkin Elmer, Wellesley). Two images (Cy-3 and Cy-5) were obtained (ImaGene, BioDiscovery, El Segundo, CA). Spots with irregular morphology and/or high local background were excluded from the analysis. The parameters that were quantified were signal median, background median, signal-to-noise ratios, and percent error ranges. Net signals 6 (background-subtracted) were calculated and normalized versus the total signals of the slide. % Change measures the average change in normalized signal intensity between Trop-2 transfectants and control. Cut off used were: % changes >10%, absolute net signal counts >500 and t-test P value <0.07.

15. Immunofluorescence

Cells were cultured on glass coverslips, fixed in 4% paraformaldehyde for 10 min and incubated with relevant antibodies for 30 minutes at 4°C. Cell nuclei were counterstained with 0.5 µg/ml propidium iodide, 0.3 µM DAPI (D3571; Thermo Fisher Scientific) or 0.1 µM DRAQ5 (#62251; Thermo Fisher Scientific). Secondary Alexa Fluor-conjugated antibodies (488- and 546-goat antirabbit IgG, 488-goat anti-mouse IgG) were from Invitrogen. Confocal images were obtained with the LSM-510 META or the LSM-800 AiryScan confocal microscope and analyzed for subcellular signal localization using the ImageJ software.

16. Tissue microarrays (TMA)

Tumor specimens were fixed in phosphate-buffered formalin, pH 7.2, and embedded in paraffin following standard procedures. Representative tumor areas were selected in haematoxylin & eosinstained slides. High-density TMA (2-mm diameter, 0.8 mm spacing) were assembled using a tissue puncher/arrayer system (Beecher Instruments).

17. IHC analysis

Five µm sections from tumor samples, tissue microarray (TMA), or “artificial tissues” from cell cultures were mounted on silanised slides. Tissue peroxidase activity was blocked with 3%

H2O2 for 5 minutes. Slides were quenched with 0.3% BSA for 30 min. Antigen retrieval was performed by microwave treatment at 750 W for 10 min in 10 mM sodium citrate buffer pH 6.0 (β-catenin and E-cadherin staining) or 1 M urea buffer pH 8.0 (Trop-2 staining). Slides were then incubated at room temperature for 30 min with the relevant antibodies. Anti-mouse (K4001, EnVision kit; DAKO) and anti-goat (K0679, **LSD** kit; DAKO) secondary pAbs were used for signal amplification, as appropriate. Control sections were treated with isotype-matched immunoglobulins or non-immune serum. Slides were washed in Tris-buffered saline-Tween 20 and incubated for 10 min in 3,3'-diaminobenzidine (DAKO). Counterstaining was performed with haematoxylin. Slides were mounted with Immunomount (Shandon). Stained slides were scored by consensus of the pathologists (MP, RL) for percentage of expressing cells and intensity of staining. Evaluation was performed blind to the clinicopathologic information from patients. β-catenin positivity threshold was ≥1% of tumor cells; that for E-cadherin was an IHC score ≥1 [312]. Expression levels of membrane/activated Trop-2 were analyzed with the goat anti-Trop-2 pAb AF650 (R&D Systems).

Target expression was quantified as percentage of stained cells and as intensity of the staining. An IHC score (H-score) was obtained, with a range from 0 to 12 [70]. For this purpose, a positivity score was first determined according to 5 categories: 0 (0% of positive cells), 1 (<10 % of positive cells), 2 (10–50% of positive cells), 3 (50–80% of positive cells), 4 (>80% of positive cells). An intensity score classified the average intensity of the positive cells: 1 (weak staining), 2 (moderate staining) and 3 (strong staining). The positivity and intensity scores were then multiplied to obtain the H-score. To perform the crosstab analysis (chi-square test) between Trop-2 expression and clinic-pathological features of patients, the protein H-score was dichotomized using a cut-off at H=4 [70].

18. Cancer cell spheroids

One thousand or 5,000 HCT116/Vector and HCT116/Trop-2 cells spheroids were obtained as 3D hanging drop spheroids [313]. After 4 days of growth, spheroids were embedded in 0.5% or 2.5% Matrigel and examined after 14 hours of incubation.

19. Transcriptomics and Next-Generation Sequencing

a) RNA extraction

Total RNA was isolated from formalin-fixed, paraffin-embedded tissue sections using the High Pure RNA Paraffin Kit (Roche Diagnostics). Briefly, sections were de-paraffinized in a xylene bath and incubated with SDS/Proteinase K at 55°C overnight. RNA was recovered by adsorption to glass fibres in 5 M guanidine-thiocyanate. After elution from the column, residual DNA was digested with RNase-free DNase I. After incubation with Proteinase K, the RNA was recovered over glass fibres and eluted in low salt buffer. RNA viability for RT-PCR analysis was verified by amplification of a 123 bp G6PDH band using LightCycler-h-G6PDH amplification and LightCycler-DNA master hybridization probes kits (Roche Diagnostics). Total RNA was extracted from cell lines using the TRI Reagent (Sigma-Aldrich). Spleen tumors and liver metastases from KM12SM cell transfectants were macroscopically dissected, flash frozen and stored at -80 °C. Tissues were ground in liquid N2 and total RNA was extracted by lysis in guanidinium thiocyanate (4 M guanidinium thiocyanate, 25 mM sodium citrate, pH 7; 0.5% sarcosyl, 0.1 M 2-mercaptoethanol), followed by addition of 0.1 volume 2 M sodium acetate pH 4, 1 volume acidic phenol, 0.2 volume chloroform-isoamyl alcohol mixture (49:1). One volume of isopropanol was added to the aqueous phase, before centrifugation to recover the RNA pellet.

b) Transcriptome profiling by DNA array hybridization

Total RNA from non-metastatic and metastatic cell lines was isolated using the RNeasy mini kit (Qiagen) and converted into double stranded cDNA using a cDNA Synthesis System kit from Roche Diagnostics. The VN-anchored oligo(dT24) primer 5'-GGCCAGTGAATTGTAATACGACTCACTATGGGAGGCGG-(T)24-VN – 3' containing a T7 RNA polymerase promoter site was utilized to prime retro-transcription.

Double-stranded cDNA was purified with the High Pure RNA Tissue Kit from Roche Diagnostics. Biotinylated complementary RNA (cRNA) was synthesized by in vitro transcription in the presence of biotin-11-CTP and biotin-16-UTP (MegaScript T7 kit; Ambion). Labelled cRNA was purified using RNeasy mini-spin columns (Qiagen) and randomly fragmented to an average size of 50 bases by a mild alkaline treatment (35 min at 94°C in 200 mM Tris-acetate, pH

8.1, 500 Mm potassium acetate, 150 mM magnesium acetate). Biotinylated cRNA probes (50 µg/ml) were hybridized to Affimetrix Human Genome U95Av2 GeneChip® micro-array slides in 50 mM MES (2 (NMorpholino) ethanesulfonic acid), 0.5 M NaCl, 10 mM EDTA, 0.005% Tween 20, 100 µg/ml herring sperm DNA and 500 µg/ml acetylated BSA. Control cRNAs corresponding to bacterial orphage genes (bioB, bioC, bioD and cre, 1.5, 5, 25 and 100 pM, respectively) were added to the mixture and were used to normalize for hybridization efficiency across arrays. Hybridization was carried out at 45°C for 16h. Arrays were washed and stained in an Affymetrix fluidics station. Hybridization was revealed with streptavidin (Roche Diagnostics), followed by biotinylated goat anti-streptavidin (Vector Laboratories) and by streptavidin-phycoerythrin (Molecular Probes). Hybridized arrays were scanned using an argon-ion laser confocal microscope (Hewlett-Packard G2500A Gene Array Scanner, Affymetrix). Each hybridization experiment was carried out in triplicate.

c) Affymetrix data analysis

Digitized DNA array images were processed using the Affymetrix® Microarray Suite 4.0 software. The quality of the cRNA was verified by hybridization with oligonucleotide probes for ß-ACTIN (ACTB) or GAPDH. These oligonucleotide sets included sequences corresponding to the 5', middle and 3' regions of the genes. cRNA samples with a 3'/5' sequence ratio of less than four were considered to be of sufficient quality for further analysis. Overall fluorescence intensity was normalized (Roche Affymetrix Chip Analysis, RACE-A). Average difference (AD) of hybridization signals measured the mean difference of fluorescence intensity between perfect-match and central-mismatch oligonucleotides in each probe set. The average difference change (ADC) is the difference between normalized AD values. Change factor (CF) quality was the ratio (ADC between two cRNA samples)/ (sum of the two AD standard deviations) for each probe set. ADC ≥2 and change factor quality >1 were used as thresholds to identify truly differentially expressed genes. Functional annotation was performed using the DAVID Bioinformatics Resources 6.7 [314]. Complete-linkage hierarchical clustering was applied using the Gene Cluster program; results were displayed using TreeView (rana.lbl.gov/EisenSoftware.htm).

d) Northern blot analysis

Total RNA was electrophoresed on a denaturing 1% agarose formaldehyde gel and blotted by capillary transfer (NorthernMax^{TM} blotting kit, Ambion) to a nylon membrane (BrightStar-Plus^{TM}, Ambion). After UV cross-linking, blots were hybridized with (α- ^{32}P) dATP-labelled cDNA. Equal loading and transfer of total RNA to the membrane was assessed by hybridization with a [95]-labelled β-ACTIN cDNA. A 204 bp-long TROP2 probe was obtained from the full-length clone (X77753) [67] by RT-PCR using the following primers:

F: 5'- GCGCAAAGGAGACGTTTATC - 3'

R: 5' - ACATAACGCTGTGCCATCC - 3'.

The probe falls in the 3´-untranslated region of TROP2.

e) Tumor cDNA targeted gene expression analysis

One μg of total RNA was reverse transcribed with avian myeloblastosis virus (AMV) reverse transcriptase, primed with random hexamers (Roche Diagnostics). cDNA amounts were quantified by ethidium bromide fluorescence in solution [315]. PCR reactions were performed in a Mastercycler EP, with HotMaster Taq DNA Polymerase (Eppendorf) (95 °C, 2 min; 95 °C, 30 sec; 60 °C, 30 sec; 72 °C, 30 sec; 35 cycles), using the following primer pairs:

CDH1

F: 5' - CACCAAAGTCACGCTGAATACAG - 3'

R: 5' - TCTTCTGAGGCCAGGAGAGG - 3'

CTNNB1

F: 5' - CGACTAGTTCAGTTGCTTGTTCGT - 3'

R: 5' - TACCCGAGCTAGGATGTGAAGG - 3'

GAPDH

F: 5' - AGGGCTGCTTTTAACTCTGGT- 3'

R: 5’ - CCCCACTTGATTTTGGAGGGA - 3’.

Amplification of GAPDH was used as an internal standard. Quantitative real-time PCR reactions were performed in a LightCycler® apparatus using the LightCycler-DNA Master SYBR Green I kit (Roche Diagnostics) and the following forward and reverse primer pairs:

TROP2

F: 5’ - GCAGGACAACTGCACGTGTC - 3’ (frozen samples)

F: 5’ - GTATCCCCTTTCGGTCCAA - 3’ (formalin-fixed samples)

R: 5’ - ATCGTTGTCCACGAGCGCGT - 3’

CALNEXIN

FOR: 5’ - ATTGTCAGATCGTTCATTGC - 3’ (frozen samples)

REV: 5’ - ATGGAACAGGTAACCAGCAT - 3’ (frozen samples)

FOR: 5’ - GTCTGCAGGTTTCTCCTTGA - 3’ (formalin-fixed samples)

REV: 5’ - ATACTGTCAACGGTGGGTGA - 3’ (formalin-fixed samples)

CALNEXIN was found equally expressed in colon cancers and corresponding metastases and was used as internal standard. Results were expressed as ratios between TROP2 and CALNEXIN and normalized to the average expression of TROP2 in primary tumors. Standard curves were obtained from serial dilutions of known amounts of TROP2 and CALNEXIN. To ensure correct target amplification, melting curves were obtained after PCR and paired to electrophoresed samples (2% agarose-gels). To minimize quantitative differences caused by reverse transcriptions with different efficiencies, three independent cDNA syntheses were performed and analyzed from each patient sample. Each tissue sample was routinely assessed for genomic DNA contamination by using non-retrotranscribed RNA isolates as templates for PCR reactions.

f) Dot blot analysis

The expression of TROP2 in human tumors was assessed by dot blot analysis of a human tumor profiling array (Clontech). A TROP2 PCR fragment was labelled with (α-^{32}P) dATP and hybridized, following the manufacturer instructions

g) RNA sequencing

One µg of total RNA was reverse transcribed with AMV reverse transcriptase primed with random hexamers (Roche Diagnostics). cDNA amounts were quantified by ethidium bromide fluorescence in solution [315]. The Ribozero kit was used to deplete samples of ribosomal RNA. One-end transcriptome sequencing was carried-out over 4 lanes of Illumina Hiseq 1000, as indicated by the manufacturer. Genes were differentially identified as expressed by tumor cells (human) or by stromal or infiltrating cells (murine). The analysis pipeline included TopHat 2.0 (for spliced alignment of the reads), HT-Seq 0.5.4 (to count reads upon annotation) and edgeR 1.0 (a statistical analysis software suite in R), as indicated by the manufacturer. Sequence alignment over reference genomes of Mus musculus and Homo sapiens identified reads that uniquely mapped to each species. Cancer-associated metastasis-module components were sought after subtraction of murine stroma sequences. Statistical analysis of Trop-2-driven versus native expression levels was performed as indicated below.

20. Xenotransplant transcriptome analysis

EMT paradigm signatures were recapitulated in luminal MCF-7 versus mesenchymal MDA-MB 231 breast cancer cells [316]. Additional transcriptional EMT benchmarks encompassed transcription factors, vimentin, myosins, TGFβ signaling [317-320]. Epithelial differentiation benchmarks [294, 317, 321-323] are listed in Table S3H. β-catenin transcriptional target genes (web.stanford.edu/group/nusselab/cgi-bin/wnt/target_genes) were identified by DNA microarrays analysis of cells treated with β-catenin siRNA or CRISPR/CAS9 [324, 325]. β-catenin targets are listed in Table S4B.

21. Image analysis

Signal intensity in Western Blotting and confocal microscopy was quantified with Fiji-Image J (https://imagej.net/). Fiji gray scale lookup table (256 channels) was normalized versus Kodak (www.kodak.com) gray scale optical density standards (power scale). Signal volume was obtained by multiplying signal area (pixel2) by average gray intensity channels. Signal volume was measured in voxels (SV).

22. Network analysis

Parental cell or Trop-2 transfectant transcriptomes were examined with the Ingenuity Pathways Analysis software (IPA 8.8, Ingenuity Systems, www.ingenuity.com). mRNA/genes were mapped onto the Ingenuity knowledge base as focus points (HUGO identifiers). Networks of focus molecules were generated by maximizing interconnectedness of input identifiers. Networks were then ranked by a score based on a hypergeometric distribution, which considers network size, number of network-eligible molecules, and the ratio between the latter and the total number of molecules in the knowledge base. The String software (string-db.org) was used to build protein-protein networks. Input proteins were those differentially expressed in KM12SM cells transfected with wt or ΔcytoTrop-2 versus vector (Table S5B-D). String uses interaction information from databases (Mint, KEGG), experimental data and text-mining; gene fusion, co-occurrence and co-expression data were excluded (confidence score ≥0.7). Subnetworks were defined with Kmeans, score ≥5.

23. Statisticals analyses

Normality of distribution of assay values was verified (www.graphpad. com). Student t-test was utilized for comparison between protein level means between experimental groups. Two-tailed Fisher exact test was utilized to compare tumor take rates and metastasis frequency. Spleen primary tumors and liver metastasis parameter counts (pseudo-glandular differentiation, mitotic figures, and apoptotic bodies) were displayed graphically through boxplot representations. Equality of distributions was tested by the Wilcoxon rank sum test, adjusting the significance level for multiple comparisons by the Bonferroni correction (0.05/2).

IV. RESULTS AND DISCUSSION

A. CANCER GENE FUNCTIONAL STUDIES: SURVIVAL AND ROLE OF TROP2 IN THE REGULATION OF THE AUTOPHAGIC PROCESS

Trop-2 is a transmembrane glycoprotein that induces tumor growth [8] by triggering converging signaling pathways, such as those of Akt [118], ERK, NFkB, and cyclin D1 [7]. In this section of functional study, we investigated whether Trop-2 could support cell suvival mechanisms through induction of autophagy.

The cellular models used in the research project were pancreatic tumor cells, BxPC3, endogenously expressing Trop-2 (line) and on which a functional Trop-2 knockout (KO) was performed using CRISPR / Cas9. Furthermore, we employed colorectal cancer cells, KM12SM which were transfected with the wild-type Trop-2 gene (KM12SM / T2) and with empty vector (KM12SM / V).

At the beginning, we evaluated the effect of Trop-2 on the growth of BxPC3 by comparing early and late incubation times in vitro for more hours regardless of standard growth curve. We performed growth assays for 264 hours without changing the culture medium, in order to measure the effect under conditions of consumption of serum and growth factors. The growth advantage is evident in Trop-2 expressing cells over KO cells, particularly at incubation times corresponding to the early stages of tumor development (<150 hours) (Fig. 10). This growth advantage is also maintained over time, indicating a possible intervention of this molecule also in the protection of cells from the induction of senescence and cell death programs.

Autophagy, autolysis, or auto-phagocytosis is a physiological, intracellular mechanism of protection and recycling of cellular elements: which plays a role in cellular homeostasis by facilitating the degradation and lysosomal recycling of macromolecules and intracellular organelles.

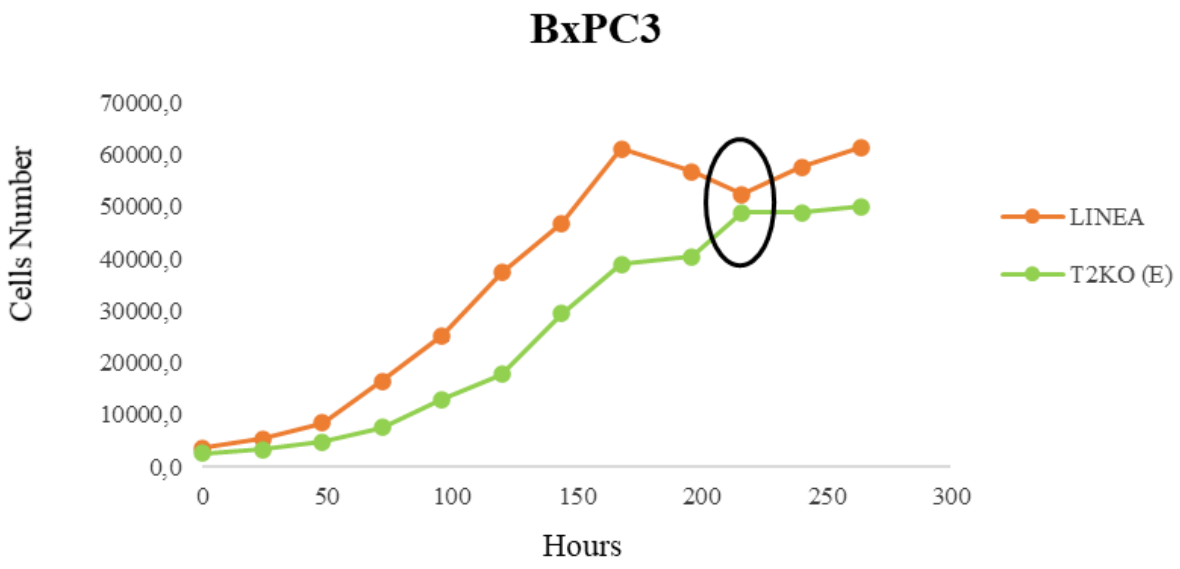


Fig. 10 : Growth curves of the BxPC3 line and BxPC3 / T2 KO

In cancer cells it has been observed that autophagy can act both by suppressing the tumor in its early stages of growth, and by supporting the survival of cancer cells by improving their resistance to necrosis and apoptosis in the later stages of progression [23, 326].

Considering the numerous studies and the previous experiment results, we began to explore the effect of Trop-2 on the autophagic process in cancer cells. BxPC and KM12-SM (Trop-2 + and Trop-2–) cells were subjected to different treatments to induce or block the autophagic cascade. Cells were grown in spent/exhausted RPMI medium for 3 days until we got maximum confluence and then treated for 48 hours by a unique dose of chloroquine with 12.5uM concentration (an inhibitor of autophagosome-lysosome fusion) in EBSS buffer solution to induce autophagy.

In support of the above, we observed that the two cell lines treated for 48 hours with EBSS individually (EBSS–) and in combination with chloroquine, (EBSS +) show a differential response dependent on the presence or absence of Trop-2 expression. In both cell models, the induction of autophagy by EBSS leads to a growth advantage in Trop-2 + cells while the combined EBSS-chloroquine treatment, by preventing autophagy, cancels the effect of Trop-2 on tumor growth (Fig. 11). Collectively, these data indicate that Trop-2 could induce autophagosome formation and degradation.

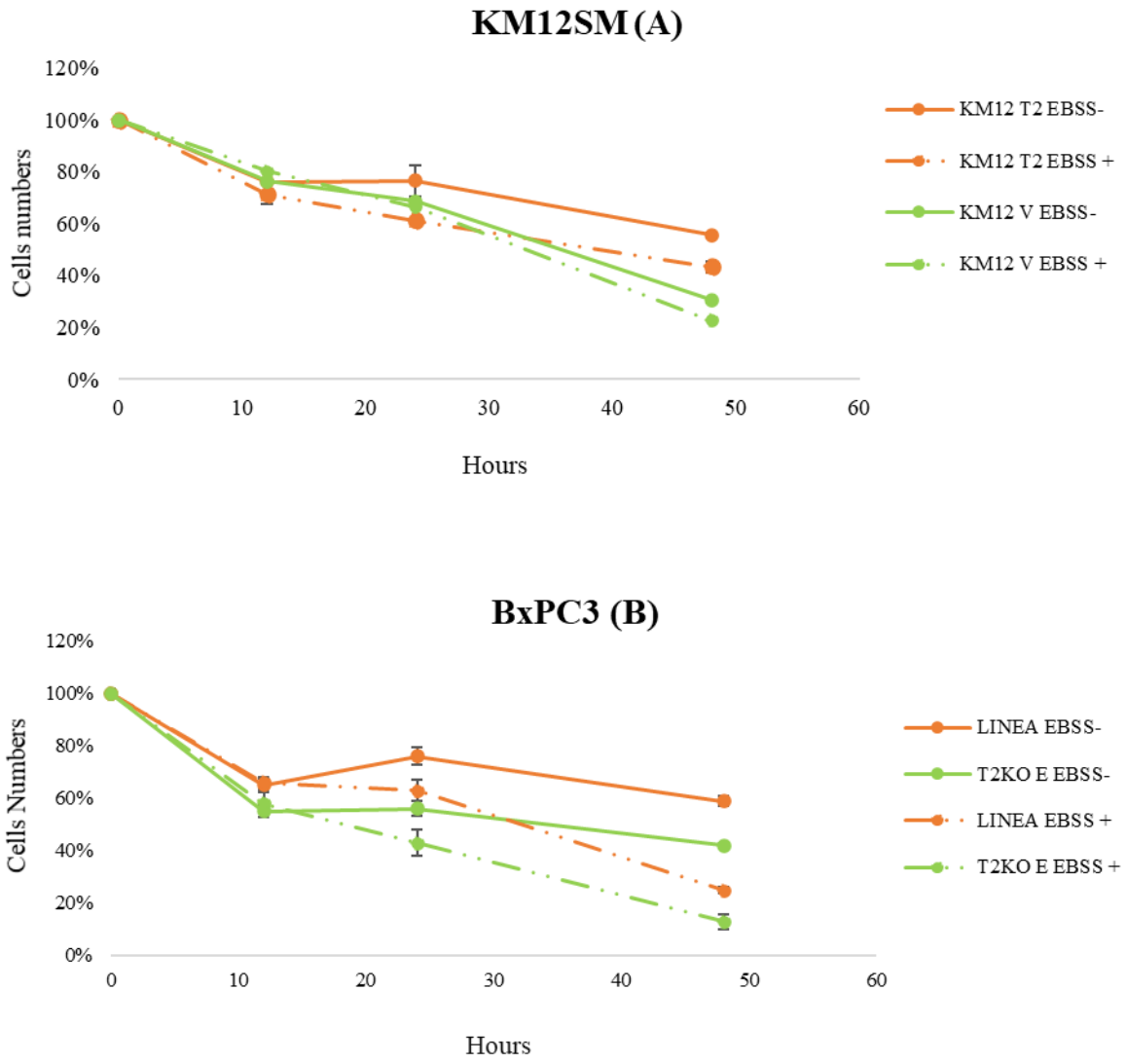


Fig. 11 : Growth curves of KM12SM / Trop-2 and KM12SM / Vector cells (A) and BxPC3 cell line and BxPC3 / T2 KO (B) treated with EBSS (EBSS-) and with EBSS + chloroquine (EBSS+)

B. CANCER GENE FUNCTIONAL STUDIES: TROP2 DESIGNER MUTANTS FOR FUNCTIONAL INVESTIGATION

As we described previously, Trop-2, is a glycoprotein that composed from 323 amino acids, comprised of a large extracellular domain, a single hydrophobic transmembrane domain, and a short intracellular, or cytoplasmic tail [131].

As a tool for functional studies of the Trop-2 cancer signaling network a series of Trop-2 designer mutants had been previously obtained in the University of Chieti laboratory. These included:

- ***Trop-2 glycolytic mutant***, where all N-glycosylation sites were substituted with the amino acid alanine. N-Glycosylation is an important post-translational modification which plays a critical role in determining protein structure, function and stability [327].
- ***Trop-2 proteolytic mutant***, where the two amino acid arginine and threonine at position 87/88 were substitute with 2 alanines, for a less efficient proteolytic processing. Regulated proteolysis is a mechanism that leads to cleavage and activation of Trop-2 for cancer growth stimulation [122].
- ***Trop-2 TEV mutant***, where a sequence of 7 amino-acids (Glu-Asn-Leu-Tyr-Phe-Gln-Gly) was inserted at position 87/88 to replace the native arginine and threonine. This 7-aminoacid sequence inhibits native proteolytic cleavage by creating a specific recognition site for the TEV protease enzyme. The TEV protease is a highly specific cysteine protease that is not present in mammalian cells and can be instead used for in vitro cleavage of fusion proteins.
- ***Trop-2-T88 mutant***, where the Trop-2 sequence starts from the threonine at position 88, which could in principle mimic the fully cleaved Trop-2.

These Trop-2 mutants were subcloned in suitable vectors for mammalian cell expression. Subcloning is a basic procedure in molecular biology for transfer of DNA inserts from one vector

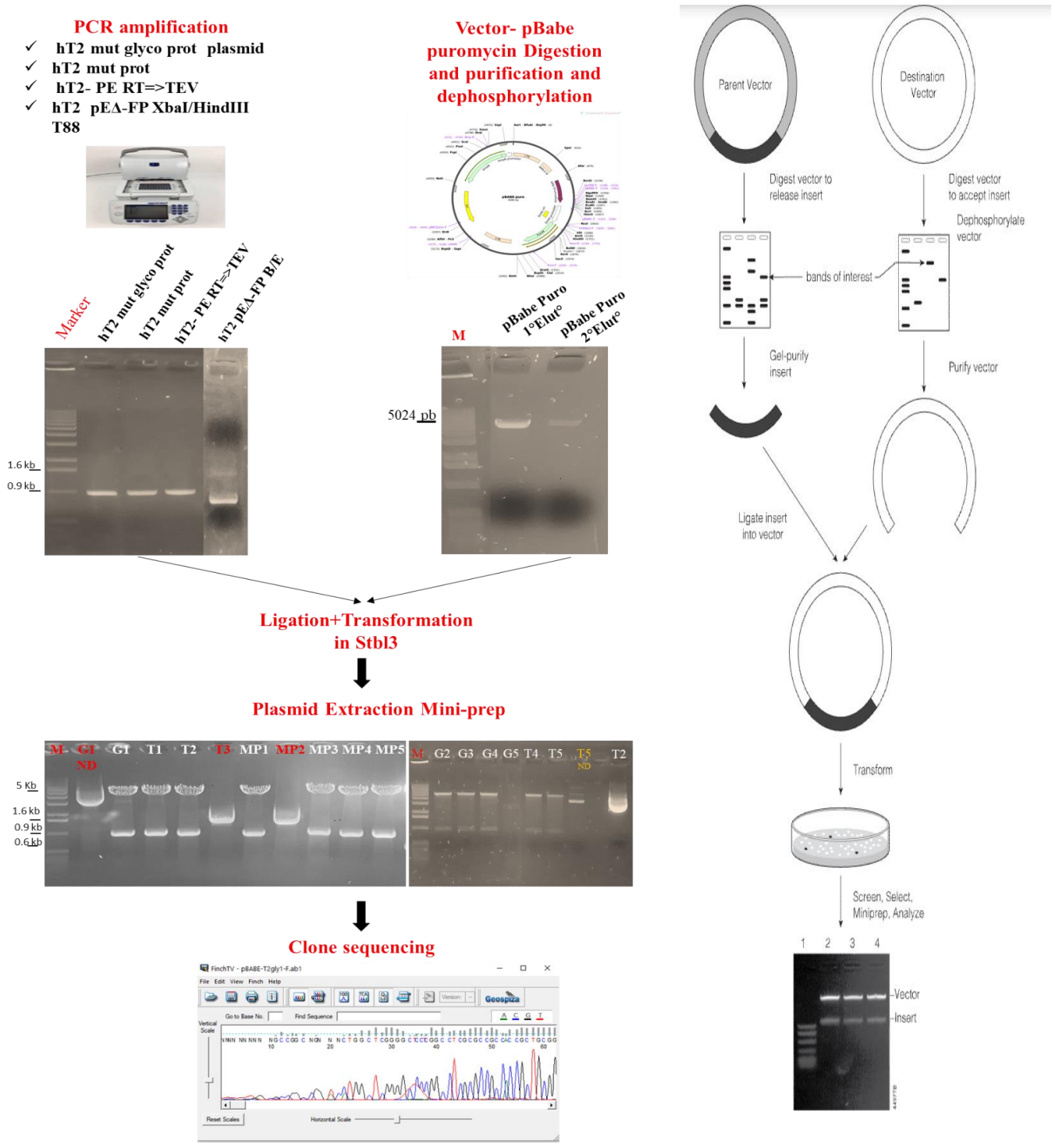


Fig. 12 : Subcloning process of pBabe-puro-Trop-2 Mutants

to another to gain functionality to study the sequence of interest. All subcloning reactions that we performed are illustrated in the figure below (Fig. 12).

We released and purified the insert from the parent vector through the amplification of the inserts using specific primers (Table II) or using an intermediate plasmid. The inserts were then ligated into viral puromycin-resistant pBabe vector followed by transformation of ligation reaction into competent bacterial cells.

Finally, positive clones were identified and sequenced to verify absence of mutation using bioinformatics tools such as Multiple sequence alignment (Clustal Omega and EMBOSS Needles), sequence translator (EMBOSS Transeq) and a restriction enzyme analyzer (Watcut).

All the recombinant pBabe plasmids were packaged into retroviral particles in the HEK293T cell lines. The virus containing HEK293T supernatants were then used to infect the MTE and KM12-SM cell lines. Infected cells efficiently expressed the Trop-2 transgene as shown by flow cytometry.

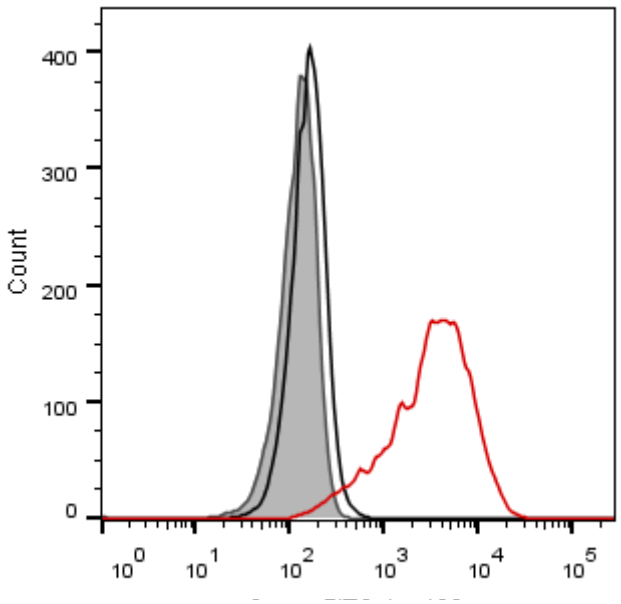


Fig. 13 : Flow Cytometry analysis profiles of Trop-2 expression in MTE-4-14/ Vector and MTE4-14/ Trop-2

We also subcloned these same Trop-2 mutants in the PCDNA3 mammalian expression vector (Fig. 13), to be directly transfected into recipient cell lines, as an alternative to retroviral infection.

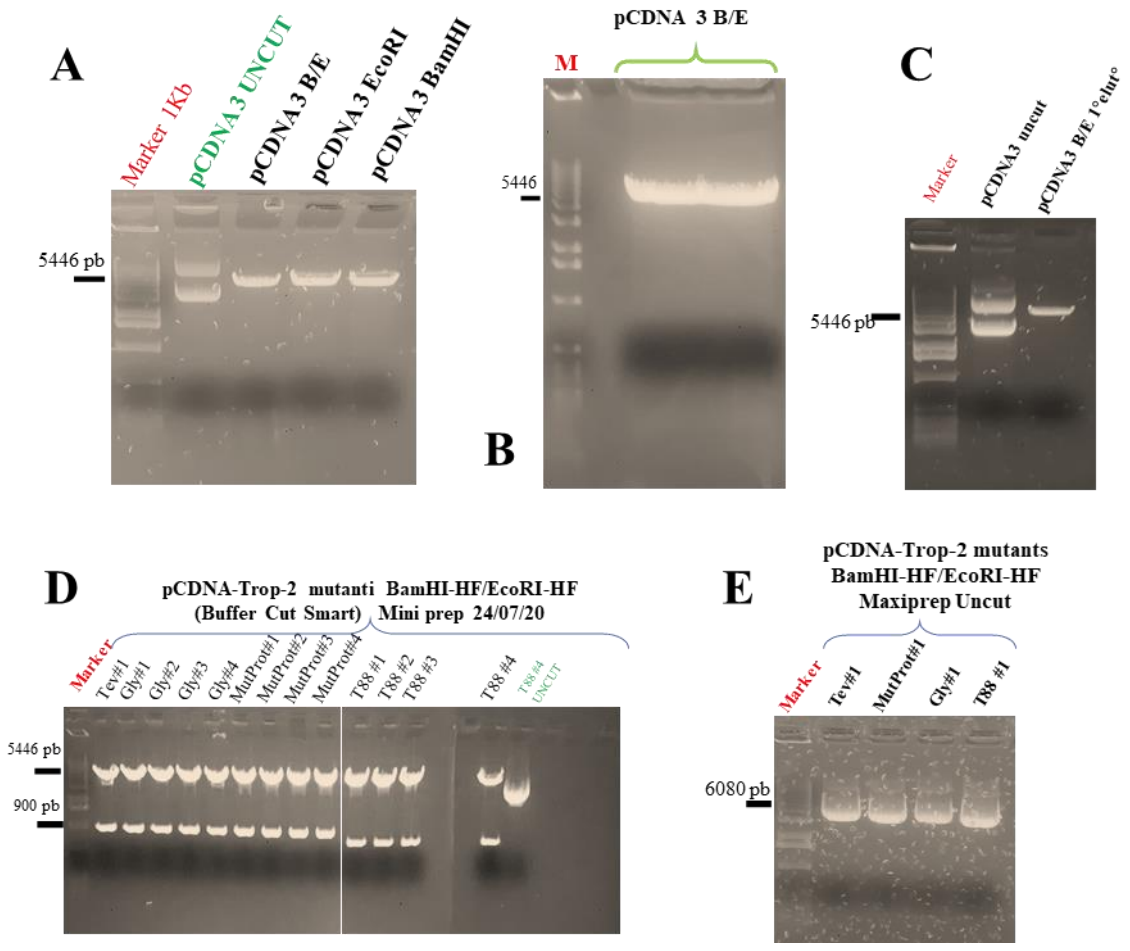


Fig. 14 : Subcloning steps of Trop-2 mutants in PCDNA.3 vector. A: PCDNA.3 Digestion with BamHI/EcorI restriction enzymes. B: Agarose gel 1% band separation. C: GenClean band purification. D: Mini prep plasmid extraction and clone verification with BamHI/EcorI enzymes. E: Maxi prep of recombinant positive plasmid extraction at high amount.

C. CANCER GENE FUNCTIONAL STUDIES: TROP-2 CLEAVAGE BY ADAM10 IS AN ACTIVATOR SWITCH FOR CANCER GROWTH AND METASTASIS (Article I)

1. Introduction

Trop-2 is a monomeric transmembrane signal transducer [95, 115] expressed by epithelial cells at various stages of differentiation [301, 328]. The TROP2 gene (TACSTD2; formerly M1S1) [25] is an intronless derivative of the TROP1 gene (EPCAM, TACSTD1; formerly M4S1) [25, 67, 93, 94], which suggests strong evolutionary pressure for conserved functional roles of the Trop molecules.

We and others have shown that Trop-2 is a driver of cancer development and progression [131, 156]. Pro-oncogenic Trop-2 signaling occurs through regulation of the expression and activity of cyclin D1, ERK, NFκB, Akt [7, 8, 118, 293], FAK, Rac1, and integrins [131, 150]. No Trop-2–activating mutations have been detected in cancers as yet, which indicates that alternative molecular mechanisms are required to account for the functional derangement of Trop- 2 [131].

Trop-2 and Trop-1 each contain a GA733 type 1 motif and a thyroglobulin repeat [30, 67, 93, 94, 98, 329]. These structural motifs share extended similarities with the corresponding regions of nidogen and IGF-binding proteins [67, 93, 94, 98], and have roles in the homophylic binding of Trop-1/Ep- CAM [35, 330]. We have shown that the proteolytic cleavage of the first loop of the Trop-1 thyroglobulin domain [29] activates the cell-growth stimulatory properties of Trop-1. We then speculated that a similar molecular switch may operate on Trop-2.

Proteolytic enzymes have critical roles in cancer development and progression and are central to multidirectional signaling networks that regulate processes such as tumor–microenvironment interactions, chemokine/cytokine crosstalk, and angiogenesis. Stoyanova et al. showed that regulated intramembrane proteolysis (RIP) is responsible for the cleavage of Trop-2 that is mediated by ADAM17/TACE at the A187-V188 site, followed by γ -secretase processing at G285-V286 [122]. This induces release of the Trop-2 intracellular domain, which then co-translocates with ß-catenin to the nucleus [122]. Recently, matriptase was reported to cleave Trop-

2 at the R87-T88 site [331], to induce decreased levels of claudins at the epithelial cell surface. Here, for the first time, we identify ADAM10 as a key interactor of Trop-2 at the cell membrane. ADAM10 is shown to cleave Trop-2 at R87-T88 in cancer cells, to activate Trop-2, and to induce cancer cell growth. Using tumor xenografts of human colon cancer cells that express the R87A-T88A Trop-2 mutant (designated as A87-A88 Trop-2), we have here demonstrated that Trop-2 proteolytic processing by ADAM10 correspondingly drives metastatic dissemination.

2. Results

a) Anti-Trop-2 reactivity of the E1 monoclonal antibody

Post-translational processing is a key activation step for several tumor growth inducers and adhesion molecules [332-335]. As Trop-2 activating mutations have not been detected as yet in cancers, we hypothesized that post-translational processing of the molecule might be specifically induced, to provide cancer cells with growth advantage over their normal counterparts. We have previously shown that Trop-1 undergoes activator cleavage at R80- R81 [29]. We explored here the existence and functionality of a corresponding activator mechanism for Trop-2.

To investigate Trop-2 post-translational processing, we generated anti- Trop-2 monoclonal antibodies (mAbs) through immunization and screening procedures aimed at obtaining efficient recognition of Trop-2 under native conditions in living cells. To this end, mAb-producing hybridomas [299] were screened for immunohistochemistry reactivity [300] of cell culture supernatants on ovarian cancer cells [336]. Flow cytometry analysis of live Trop-2 transfectants was then carried out, and anti-Trop-2 mAbs were selected for recognition of native Trop-2 in tumors and cell lines.

The E1 mAb was shown to mostly bind Trop-2–expressing cancer cells of epithelial origin (Table S1), such as mammary carcinomas, ovarian cancers, lung adenocarcinomas, colorectal cancers, pancreatic adenocarcinomas, prostate cancers, and choriocarcinomas. Specific binding of a Trop-2– negative L-cell murine fibrosarcoma transfected with genomic TROP2 or TROP2 cDNA (including a single-residue polymorphic cDNA clone obtained from the FE ovarian cancer) (Fig. 15A, B, D), provided formal proof for specific recognition of Trop-2. E1 recognized Trop-2

in Western blotting assays under non-denaturing conditions (Fig. 15D), but it did not react with reduced Trop-2.

Cross-blocking in competition binding assays showed that the E1 antibody binding site coincides with, or is in close proximity to, that of the main anti-Trop-2 mAbs that have been generated to date (Fig. 15A and [337]), including T16 [301], 162-46.2 (ATCC HB-187), AR47A6.4.2 [338], 77220 (R&D Systems,MN, USA), and the RS7 murine mAb, that has been developed into the anti-Trop-2 antibody–SN-38 drug conjugate [339] Sacituzumab govitecan-hziy (Trodelvy) [340]. Mapping using Trop-2 deletion mutants identified this immunodominant epitope in the extracellular region spanning D146-R178 [337].

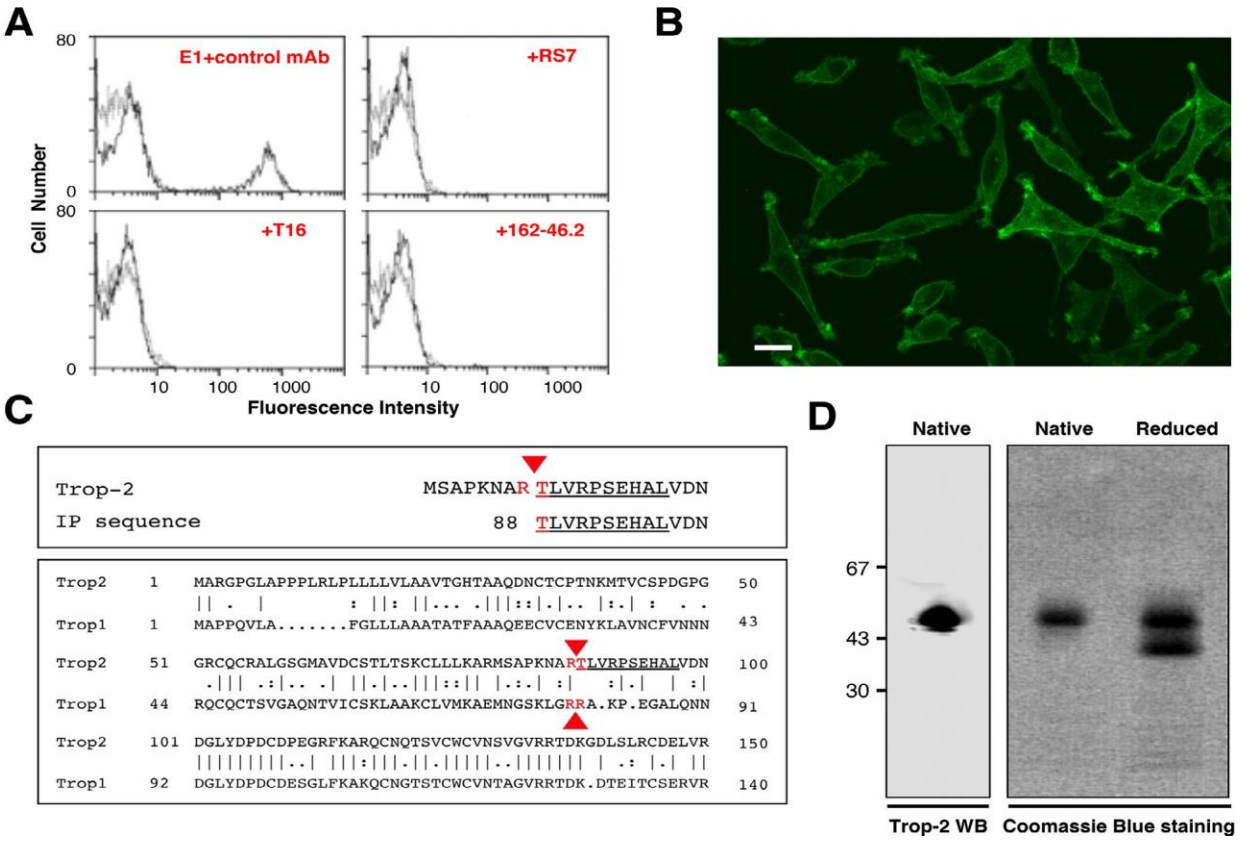


Fig. 15 : Purification, sequencing and analysis of Trop-2

b) Trop-2 purification and sequencing

We investigated the post-translational processing of Trop-2 using the E1 mAb to immunoprecipitate Trop-2 from ovarian cancer cells. Purification of Trop-2 was then performed

by affinity chromatography over Sepharose–E1 mAb columns. Western blotting analysis indicated an essential homogeneity of the isolated Trop-2 (>95% purity) (Fig. 15D).

N-terminal sequences were obtained from the E1-immunoreactive material by Edman degradation. The N-terminal Trop-2 sequence started at T88 (Fig. 15C, top), which first indicated a potential, endogenous cleavage of Trop-2 between R87 and T88 in the thyroglobulin domain. The unprocessed Trop-2 N-terminus was blocked and did not generate any amino-acid sequence. Post-translational cleavage at R87-T88 was expected to result in a large conformational rearrangement, leading to the generation of a two-chain molecule with a ~10-kDa fragment bound to the ~40-kDa membrane- bound segment by a single disulfide bridge between Cys73 and Cys108 (Fig. 16). Consistent with this, reduction of the disulfide bridges led to the loss of a ~10-kDa fragment from full-length Trop-2 (Fig. 15D).

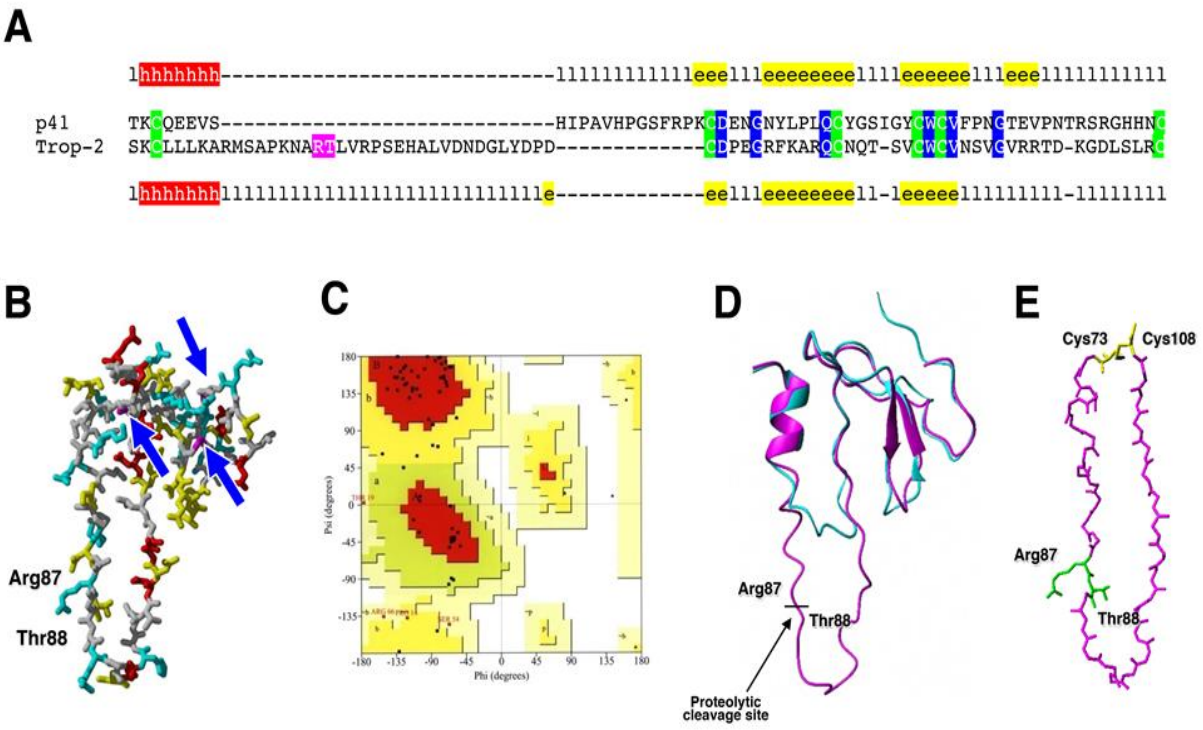


Fig. 16 : Structure of the Trop-2 thyroglobulin domain

c) Trop-2 immunoprecipitation and mass spectrometry analysis

The Trop-2 cleavage site was confirmed by immunoprecipitation/ PAGE/mass spectrometry (MS) peptide mass fingerprinting. The list of the peak fragments of Trop-2 is given in Table S2. The MS spectra were scanned for peptide sequences that did not fit trypsin-cleavage consensus sites, to determine the endogenous proteolytic cleavage sites of Trop-2. Four independent peptides were identified that showed cleavage at R87-T88, which confirmed the N-terminal Edman sequencing data (Fig. 15C, top; Table S2). Of note, the R87-T88 Trop-2 cleavage site shows exact correspondence to the activator cleavage site of the paralogous Trop-1 molecule [29] (Fig. 15C, bottom), suggesting a similar impact on an evolutionarily conserved function.

d) Three-dimensional modeling of the Trop-2 thyroglobulin domain

We generated a three-dimensional model of the thyroglobulin domain of Trop-2 using the crystal structure of the thyroglobulin domain of the p41 splice variant of the major histocompatibility complex class II-associated invariant chain [296] as a template (Fig. 16). Independent prediction of the secondary structure using the PHD program (≥80% accuracy; www. predictprotein.org) provided the corresponding subsegment identification, which supported its overall prediction accuracy. Ramachandran plot analysis allocated 79% of the residues of the Trop-2 model into the core region, and 17% into the allowed region of the plot. A corresponding analysis of the p41 crystal structure positioned 88% of the residues in the core region and 12% in the allowed region of the plot [296], in good correspondence with the Trop-2 model. The Trop-2 thyroglobulin domain showed comparable main-chain parameters versus the p41 crystal.

The root mean square deviation between the Trop-2 thyroglobulin domain and the p41 template was 0.399 A˚, supporting the reliability of the Trop-2 model [341]. Superimposing the Trop-2 model on the p41 template showed that the Trop-2 and p41 thyroglobulin domains have a relatively flat, wedge-shaped structure, with three loops that are largely aligned on a plane that extends from the three conserved disulfide bridges. The region with the highest structural divergence was between the first and second cysteine, where Trop-2 showed a much longer loop than p41. This region contains the proteolytic cleavage site. PROCHECK and MOLMOL

(PaintSurface) residue accessibility routines confirmed high solvent accessibility of the proteolytic cleavage region.

Cleavage of the first loop of Trop-2 generated free amino-acid chains of 14 and 20 residues and removed a rigid constraint to a planar configuration of the thyroglobulin domain (Fig. 16, Movie S1). Favorable main-chain parameters (i.e., Ramachandran plot quality, peptide bond planarity, bad nonbonded interactions, c-a tetrahedron distortion, main chain hydrogen bond energy, overall G-factor) and root mean square deviation were reached, thus allowing reliable analytical insight on the Trop-2 structure.

The first p41 thyroglobulin repeat loop binds the active site of cathepsin L [296]. On the other hand, Trop-2 showed a much longer loop at this position (Fig. 16D, E). Hence, the cathepsin L enzymatic pocket appeared too small to fit the first Trop-2 loop, thus suggesting cleavage by a different protease.

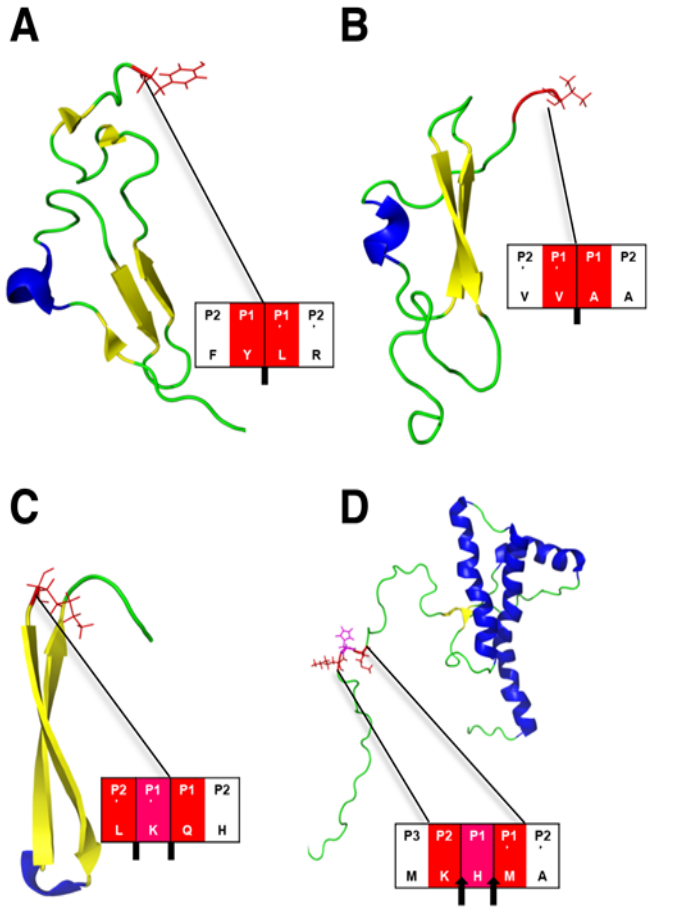
e) ADAM10 is a candidate Trop-2 protease at the R87-T88 site

We carried out immunoprecipitation of Trop-2, followed by MS analysis of its binding partners. This revealed four independent ADAM10 peptides across multiple identification procedures (Table S3), thus indicating ADAM10 as a Trop-2 interactor. ADAM10 is a member of the 'a disintegrin and metalloproteinase' (ADAM) family of membrane-bound zinc-dependent metalloproteases [342, 343], and it is involved in processing of several adhesion molecules and transmembrane receptors [332, 333, 344-348]. Peptide library screening combined with crystal structure analysis of known ADAM10 substrates [349] revealed preference for Pro at P5 (Calsyntenin, proTGFα, TNFα, cleaved peptides [349]), Arg at P1 (EGF, N-Cadherin, CD23), and Thr at P1' (amyloid β/A4 protein, HB-EGF) in target site subsets (Fig. 17). ADAM10 was also shown to be prone to cleave within the loop segments of target molecules (Fig. 17).

All of these sequence/structure features are present in the Trop-2 cleavage site (Fig. 15C; Table S2), making ADAM10 a candidate Trop-2-cleaving protease at the Trop-2 R87-T88 site. ADAM10 is expressed at high levels in most human cancer types and is up-regulated/activated in cancers of the stomach [350], liver [351], colon [352], uterus, ovary [353], oral cavity [354],

pancreas [355], lung, and in triple-negative breast cancers [348], as well as in melanoma and glioblastoma [348, 356].

Next-generation sequencing profiling of tumor transcriptomes of almost 8000 cancer patients in the TCGA dataset (www.proteinatlas.org) identified significant association of ADAM10 overexpression with worse prognosis for pancreas (Fig. 18A) and lung (Fig. 18B) cancers. Subgroup analysis [357] showed that the negative prognostic impact of ADAM10 mainly occurred on lung adenocarcinomas (LUADs) (P= 0.0013), and only marginally on lung squamous



E

ADAM10 protein target sequences cleavage sites

Substrate	Uniprot Accession	P5	P4	P3	P2	P1	P1'	P2'	P3'	P4'
amyloid beta A4 protein	P05067	Glu	Val	His	His	Gln	Lys	Leu	Val	Phe
amyloid beta A4 protein	P05067	Gly	Val	Val	Ile	Ala	Thr	Val	Ile	Val
amyloid beta A4 protein	P05067	Val	Gly	Gly	Val	Val	Ile	Ala	Thr	Val
amyloid beta A4 protein	P05067	Val	His	His	Gln	Lys	Leu	Val	Phe	Phe
Calsyntenin-1	O94985-2	Ala	Asn	His	Met	Ala	Ala	Gln	Pro	Gln
Calsyntenin-1	O94985-2	Glu	His	Ala	Asn	His	Met	Ala	Ala	Gln
Calsyntenin-1	O94985-2	Ala	Thr	Val	Val	Ile	Val	Val	Cys	Val
Calsyntenin-1	O94985-2	Pro	Ser	Thr	Ala	Thr	Val	Val	Ile	Val
Calsyntenin-2	Q9H4D0	Glu	His	Val	Asn	His	Leu	Ile	Val	Gln
Calsyntenin-2	Q9H4D0	Ala	Thr	Val	Val	Ile	Ile	Ile	Ser	Val
Calsyntenin-2	Q9H4D0	Pro	Ser	Ile	Ala	Thr	Val	Val	Ile	Ile
Calsyntenin-3	Q9BQT9	Ala	His	Pro	Ser	His	Val	Leu	Ser	Ser
Calsyntenin-3	Q9BQT9	Ala	Thr	Leu	Ile	Ile	Val	Val	Cys	Val
Calsyntenin-3	Q9BQT9	Pro	Ser	Ala	Ala	Thr	Leu	Ile	Ile	Val
CD40 ligand	P29965	Asn	Ser	Phe	Glu	Met	Gln	Lys	Gly	Asp
delta-like 1	Q61483	Leu	Ser	Glu	Arg	His	Met	Glu	Ser	Gln
ephrin type-B receptor 2	P54763	Ala	Glu	Tyr	Gln	Thr	Ser	Ile	Lys	Glu
HB EGF	Q06186	Thr	Pro	Ser	Lys	Glu	Arg	Asn	Gly	Lys
HB EGF	Q06186	Lys	Lys	Lys	Gly	Lys	Gly	Leu	Gly	Lys
HB EGF	Q06186	Glu	Asn	Pro	Leu	Tyr	Thr	Tyr	Asp	His
major prion protein	P04156	Thr	Asn	Met	Lys	His	Met	Ala	Gly	Ala
major prion protein	P04925	Ala	Tyr	Tyr	Asp	Gly	Arg	Arg	Ser	-
myelin basic protein	P02687	Tyr	Leu	Ala	Ser	Ala	Ser	Thr	Met	Asp
myelin basic protein	P02687	Gly	His	His	Ala	Ala	Arg	Thr	Thr	His
myelin basic protein	P02687	Tyr	Gly	Ser	Leu	Pro	Gln	Lys	Ala	Gln
myelin basic protein	P04370	Tyr	Gly	Ser	Leu	Pro	Gln	Lys	Ser	Gln
Neurogenic locus notch homolog protein 1	Q01705-1	Tyr	Lys	Ile	Glu	Ala	Val	Lys	Ser	Glu
p75NTR	P07174	Gly	Ser	Ser	Gln	Pro	Val	Val	Thr	Arg
p75NTR	P07174	Gln	Pro	Val	Val	Thr	Arg	Gly	Thr	Thr
protransforming growth factor alpha	P01135	Pro	Val	Ala	Ala	Ala	Val	Val	Ser	His
tumor necrosis factor	P01375	Pro	Leu	Ala	Gln	Ala	Val	Arg	Ser	Ser
BTC	P35070	Ser	Pro	Glu	Thr	Asn	Leu	Leu	Cys	Gly
BTC	P35070	Val	Asp	Leu	Phe	Tyr	Leu	Arg	Gly	Asp
EGF	P01133	His	Tyr	Ser	Val	Arg	Asn	Ser	Asp	Ser
EGF	P01133	Trp	Trp	Glu	Leu	Arg	His	Ala	Gly	His
E-cadherin	P12830	Arg	Lys	Ala	Gln	Pro	Val	Glu	Ala	Gly
N-cadherin	P19022	Thr	Asp	Val	Asp	Arg	Ile	Val	Gly	Ala
CD23	P06734	Ala	Glu	Gln	Gln	Arg	Leu	Lys	Ser	Gln
CD23	P06734	Gly	Asp	Gln	Met	Ala	Gln	Lys	Ser	Gln
CD23	P06734	Leu	Arg	Ala	Glu	Gln	Gln	Arg	Leu	Lys
FasL	P48023	His	Thr	Ala	Ser	Ser	Leu	Glu	Lys	Gln
FasL	P48024	Ser	Ser	Leu	Glu	Lys	Gln	Ile	Gly	His

Fig. 17 : 3D structure and sequence of the ADAM10 cleavage sites

cell cancers (LUSCs) (P = 0.013). Similar data were obtained by meta-analysis of DNA microarray data from 2437 NSCLC patients through the KMPlot database (www.kmplot.com) (Fig. 18B).

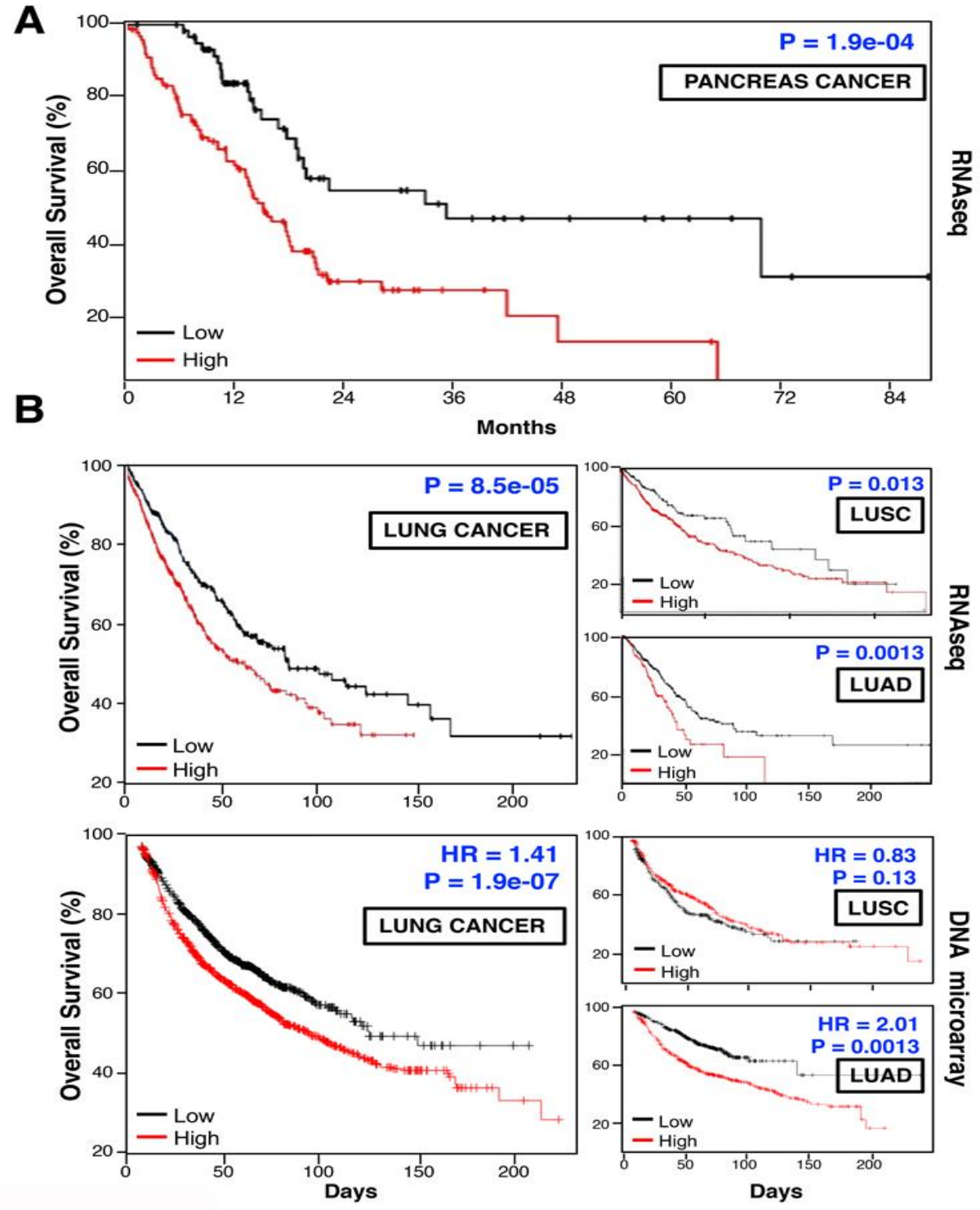


Fig. 18 : Prognostic value of ADAM10 expression in human cancer case series

This closely paralleled our earlier analysis of the prognostic impact of Trop-2 on LUADs versus LUSCs [323]. Taken together, these data prompted us to further investigate the interplay between ADAM10 and Trop-2.

f) ADAM10 interacts with Trop-2 at the cell membrane

Previous subcellular localization analysis had revealed that a minor fraction of mature ADAM10 is located in the endoplasmic-reticulum/plasma- membrane–enriched fractions [358], which suggested that most ADAM10 activity is located at the plasma membrane 68. Here, immunofluorescence (Fig. 19A) and time-lapse fluorescence microscopy (Movie S2) revealed full colocalization between ADAM10 and Trop-2 in 72.2% of the cells analyzed, with partial colocalization in 22.2%, and no colocalization in 5.5% (n = 54).

The specificity of the colocalization was verified by immunofluorescence analysis of cells infected with the pLVTHM-GFP lentivirus that co-expresses ADAM10 or scramble shRNA. The antibody used for recognition of ADAM10 did not stain the GFP+ cells co-expressing ADAM10 shRNA. On the other hand, uniform membrane staining for ADAM10 was observed in GFP+ cells that co-express the scramble shRNA (Fig. 19B). We further showed that expression and membrane localization of Trop-2 were not affected by ADAM10 inhibition (Fig. 19B).

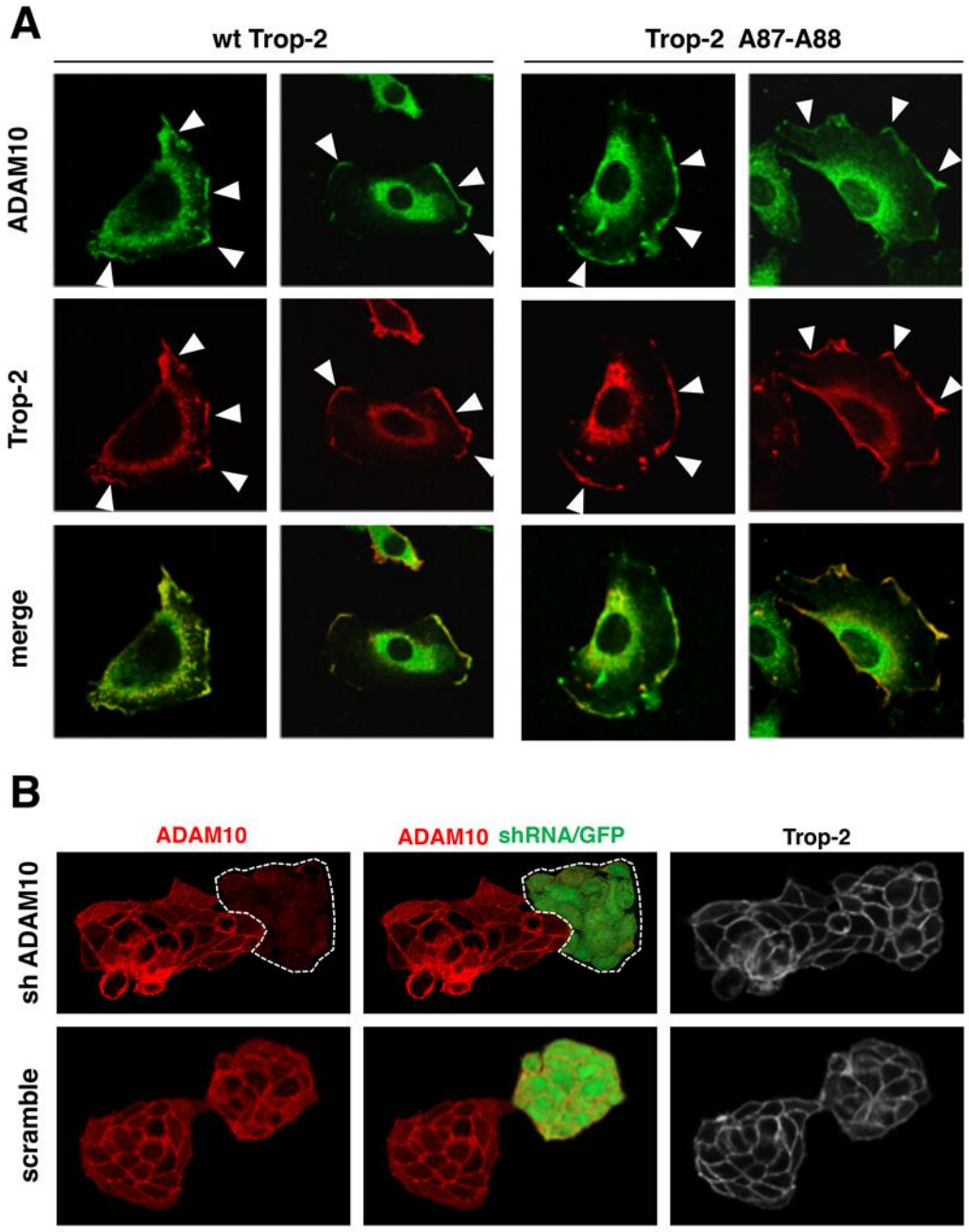


Fig. 19 : Colocalization of Trop-2 and ADAM10

g) Trop-2 proteolysis is impaired by ADAM10 inhibition

We questioned whether ADAM10 mediates the proteolytic processing of Trop-2 in vivo. We assessed Trop-2 cleavage upon inhibition of ADAM10 activity and expression. Treatment of Trop-2 expressing BxPC3 pancreatic cancer cells or MTE4-14/Trop-2 transfectants with the

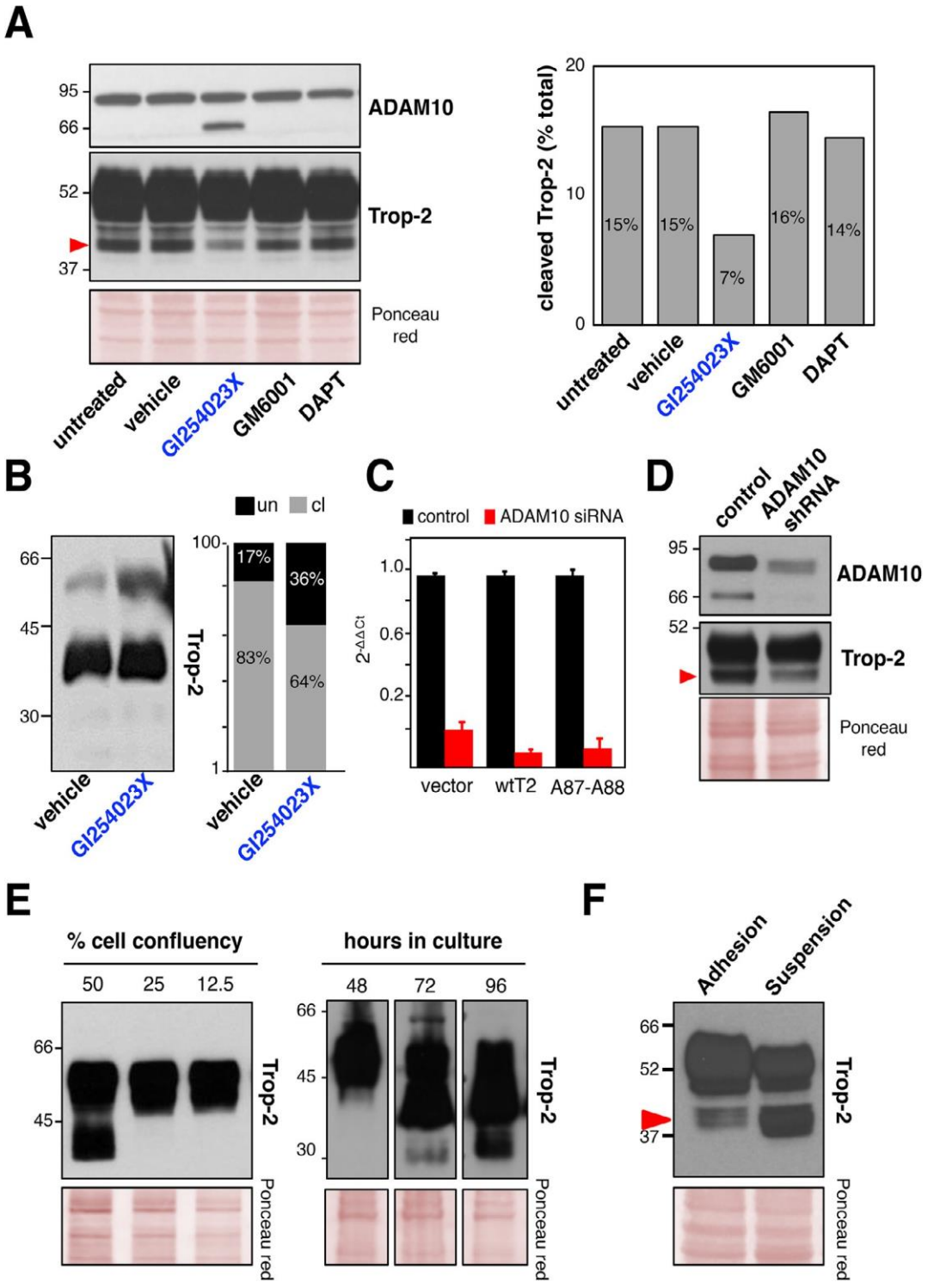


Fig. 20 : Trop-2 cleavage by ADAM10.

ADAM10 inhibitor GI254023X [359] led to a strong reduction in the 40-kD cleavage band, while parallel treatment with nonspecific protease inhibitors had no such effect (Fig. 20A, B). Accumulation of the ADAM10 mature form was seen in the GI254023X-treated cells (Fig. 20A), as a result of this inhibition of ADAM10 enzymatic activity [360]. Consistent with the Trop-2 cleavage reduction seen upon chemical inhibition of ADAM10 activity, RNAi-mediated knock-down of ADAM10 expression resulted in a dramatic reduction of the Trop-2 cleavage (Fig. 20C, D). Taken together, our findings indicate that ADAM10 is an effector protease of Trop-2 cleavage at R87-T88.

h) Proteolytic processing of Trop-2 is required for cell growth stimulation

MTE4-14/Trop-2 transfectants were shown to acquire traits of malignant transformation upon Trop-2 expression, such as the ability to form tumors in immunosuppressed animals [118] and to grow at high cell density. This was paralled by increasing levels of Trop-2 cleavage at increasing cell densities (Fig. 20E). Growth under reduced adherence to a substrate is a corresponding hallmark of malignancy [23]. Consistent with this, cells grown in suspension induced Trop-2 cleavage (Fig. 20F).

Cell-surface biotinylation/pull-down assays were used to reveal the status of Trop-2 at the cell membrane. This showed that wild type (wt)Trop-2 molecules were cleaved/activated at the cell surface (Fig. 21A). Further, we detected cleaved Trop-2 using antibodies that targeted either the extracellular domain or the intracytoplasmic domain of Trop-2 (Fig. 21A). This demonstrated that the R87-T88–cleaved Trop-2 still contained the cytoplasmic tail, thus indicating that R87-T88 cleavage precedes TACE cleavage, RIP, and release of the Trop-2 extra-cytoplasmic and intra-cytoplasmic segments [122] along the Trop-2 activation cascade.

The Trop-2 cleavage site was mutagenized by switching the R87-T88 residues to alanine. Flow cytometry analysis on live cells showed that A87-A88 Trop-2 was efficiently transported to the cell membrane, where it retained wt Trop-2–like expression pattern (Fig. 21B and Fig. 19). Under high cell density and prolonged culture conditions, the A87-A88 Trop-2 showed considerably reduced cleavage compared with wt Trop-2 (Fig. 21C).

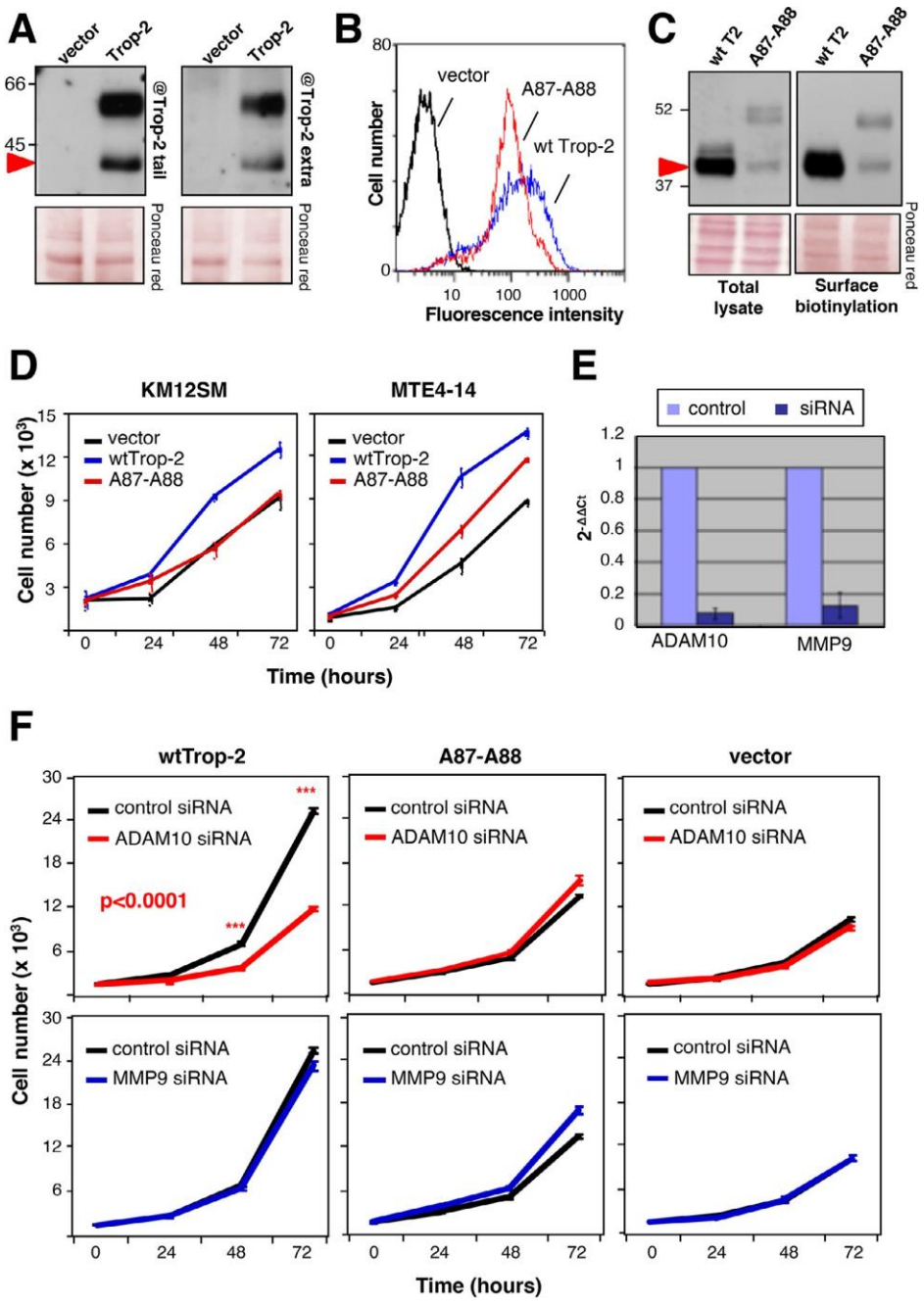


Fig. 21 : Trop-2 proteolytic processing-defective mutant versus wild-type

We then asked whether the cleavage at R87-T88 acted as an activator switch for the Trop-2 growth-stimulatory functions. In colon cancer KM12SM and murine MTE4-14 transfectants, the cleavage-

impaired A87-A88 Trop-2 mutant did not stimulate cell growth over baseline, at variance with the cell growth enhancement induced by wt Trop-2 (Fig. 21D).

We then performed in vitro proliferation assays for cells treated with control siRNAs or ADAM10 siRNAs (Fig. 21E). We here found that ADAM10 inhibition suppressed the growth of wt Trop-2-expressing cells (Fig. 21F). On the other hand, ADAM10 inhibition had no impact on the A87-A88 Trop-2 transfectants, as well as on the negative control transfectants. Of note, the growth rate of wt Trop-2 transfectants treated with ADAM10 siRNAs was reduced to the growth rate of the A87-A88 Trop-2 transfectants, which indicated a mandatory role for ADAM10 in Trop-2 cleavage/activation for induction of cell growth. The ADAM10 inhibitor GI254023X can also inhibit MMP9 [359]. Hence, we functionally tested the role of MMP9 in Trop-2 cleavage by using specific siRNA (Fig. 21E). MMP9 inhibition had no impact on the growth rate of either wt Trop-2 or the A87-A88 mutated Trop-2 cell transfectants (Fig. 21F), which supported the specific involvement of ADAM10 in Trop-2 processing and activity for tumor growth.

i) Trop-2 proteolytic processing at R87-T88 only occurs in cancer cells

As wt Trop-2 is strongly upregulated in cancer and no cancer-related mutations have been identified in the corresponding gene (TACSTD2) as yet [8], we hypothesized that this Trop-2 proteolytic processing acts as an activator switch and might be a key activating step for Trop-2 in transformed cells. Human epidermis expresses high levels of Trop-2 under native conditions [8, 301], which allows immediate comparisons of Trop-2 processing in normal epidermal cells versus skin cancers (Fig. 23, Fig. 22). Here, Trop-2 was not cleaved in normal human keratinocytes, whereas its cleavage occurred in the basal cell carcinomas. Minimal, if any, cleavage at R87-T88 was detected in squamous cell carcinomas (Fig. 23A).

We speculate that this was due to the prevalent expression of Trop-2 in well-differentiated, keratinizing regions of the tumors, thus suggesting retention of differentiation features of normal human epidermis (Fig. 22). This analysis was extended to include an indipendent set of normal human skin samples, where absence of Trop-2–activating cleavage was confirmed (Fig. 23B). On the other hand, Trop-2 cleavage was detected in most of the Trop-2-transformed and cancer cells

analyzed (i.e., ovarian, breast and colon carcinoma cells; transfectants of carcinoma, myeloma, fibrosarcoma cells) (Fig. 23C). The same cleavage patterns and extent were detected in vivo and in vitro, and across different ranges of cell surface expression (Fig. 23C), suggesting tight regulation of Trop-2 processing.

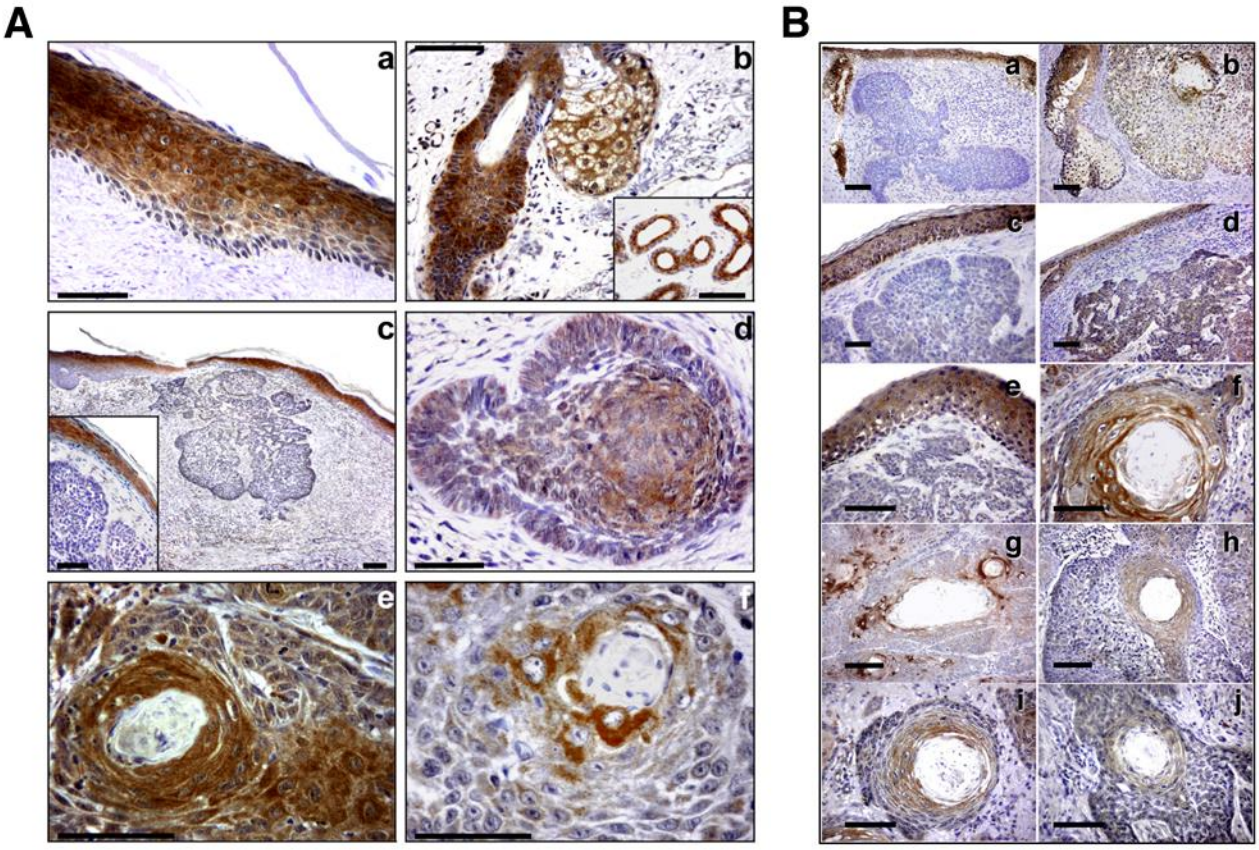


Fig. 22 : Trop-2 expression in normal and transformed keratinocytes

j) The signaling mechanism driven by Trop-2–activating cleavage

The signaling mechanisms that are triggered by activation cleavage of Trop-2 were explored using proteomic and phosphoproteomic chip analysis (Table SA, B). The wt Trop-2 versus vector-alone group and the A87- A88 Trop-2 versus vector-alone group were analyzed for concordant and discordant changes in the expression or activity of various signaling effectors (Table S4C, D). This allowed to reveal signaling steps that were specifically activated by wt Trop-

2. Protein modifications that operated as central nodes of cleaved/activated Trop-2 were shown to encompass Src, RSK1/2, glycogen synthase, MEK1, CDK9, p21, and BAD as main downstream effectors (Table S4D).

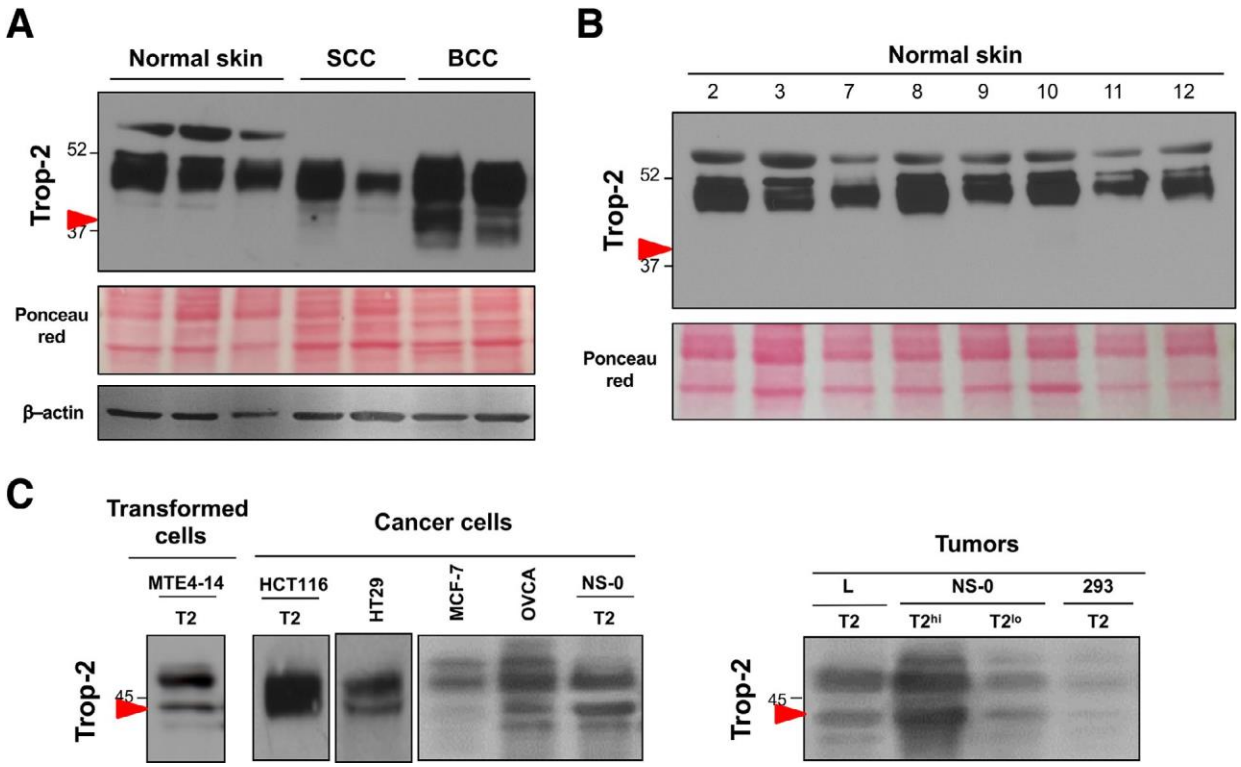


Fig. 23 : Proteolytic processing of Trop-2 in tumor cells

k) Trop-2 cleavage induces tumor growth

Our findings had indicated that Trop-2 proteolytic activation stimulates cell growth *in vitro*. We thus assessed the impact of this Trop-2 proteolytic activation on the growth of tumors xenotransplanted in murine models. Tumorigenic L fibrosarcoma and transformed HEK293 cells were transfected with the A87-A88 Trop-2 mutant and injected subcutaneously into immunosuppressed mice. The cells transfected with wt Trop-2 and with the corresponding empty vector were used as controls. Comparison of the tumor volumes between these groups indicated that while transfection of wt Trop-2 resulted in increased growth of both L fibrosarcoma and transformed HEK293 cells as tumors in vivo, the R87A-T88A mutagenesis led to complete loss of this growth-inducing activity (Fig. 24A). Hence, proteolytic processing at R87-T88 is mandatory for triggering the Trop-2 growth-inducing activity.

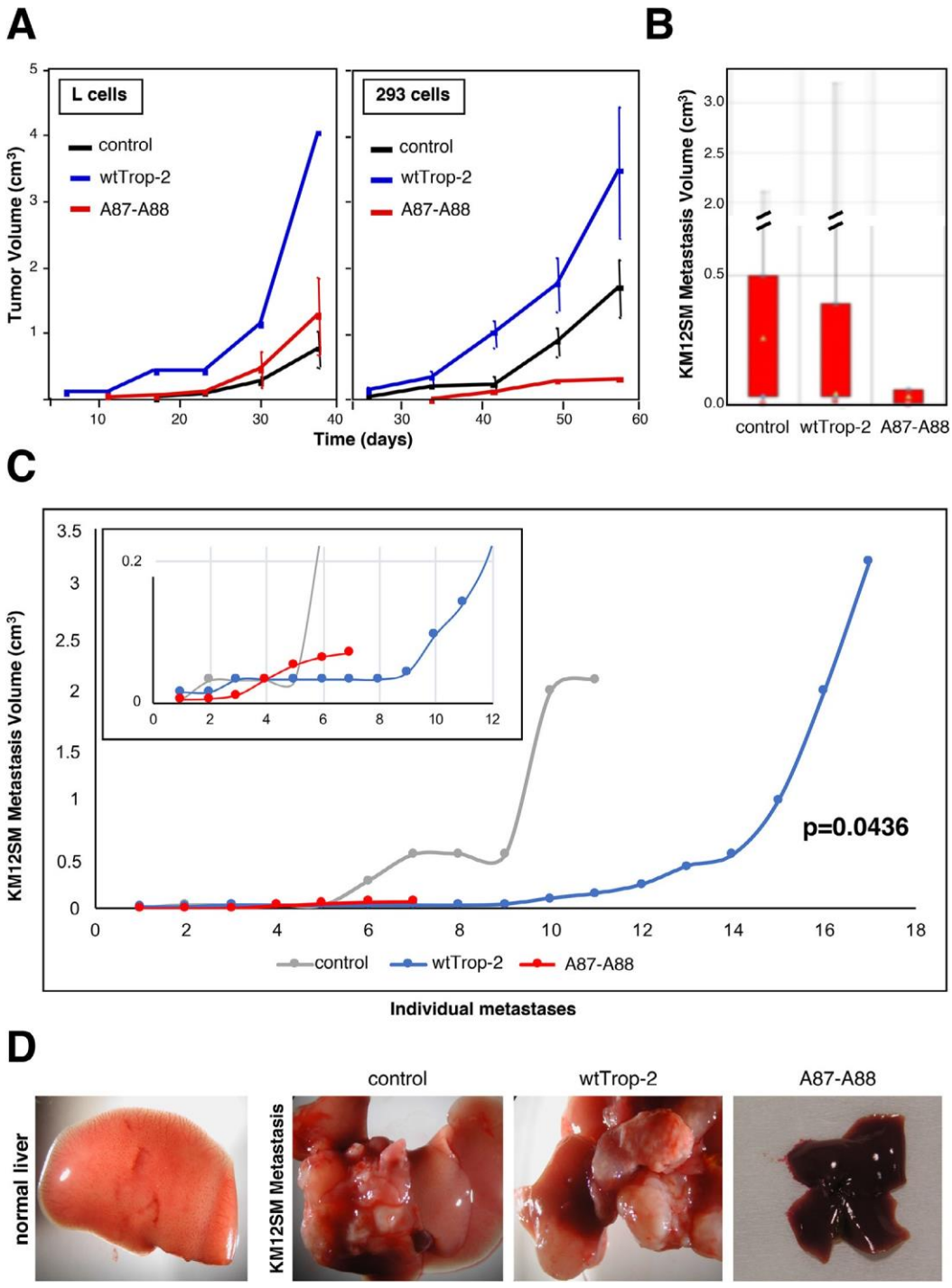


Fig. 24 : Tumor growth and metastasis dissemination require Trop-2 cleavage

l) Trop-2 cleavage induces metastasis

Cancer progression involves tumor growth in the primary site followed by metastatic dissemination to close and distant sites. Hence, we assessed whether Trop-2 cleavage is an activation step for induction of a pro-metastatic activity. To this end, KM12SM colon cancer cells were transfected with the empty vector, wt Trop-2, or A87-A88 Trop-2 and injected into the spleen of immunosuppressed mice. Dissemination to the liver was then determined by comparison of the tumor and metastasis volumes between these groups. All tissue samples from the three groups of xenografts underwent systematic histopathology analysis, to include detection of minimal tumor and metastasis burden and volumes. The individual metastasis volume distribution curves showed that the A87-A88 Trop-2 mutant invariably reduced, or outright abolished, metastasis growth (Fig. 24B-D). Boxplot distribution analysis showed that the dataset of A87-A88 mutant Trop-2 had the lowest median, together with the lowest maximum and minimum values of metastasis volume. The dimensions and volume distributions of mutant liver metastases versus vector-alone control cells and wt Trop-2 transfectants were assessed. These showed that the mutant transfectant metastases volumes were significantly reduced versus wt Trop-2 (nonparametric Mann-Whitin P = 0.0436), due to this loss of the required metastasis activation pathway.

m) Trop-2 activation in primary breast cancers

We then assessed whether this Trop-2–activating mechanism operates in primary tumors in cancer patients. Cleavage of Trop-2 was assessed by western blotting in 55 breast cancer samples from a N0, T1/T2 case series [147]. Trop- 2 expression was detected in all of the tumor samples, with high intensity (i = 3) in 44 samples (80.0%), low intensity (i = 1) in 3 samples (5.5%), and intermediate intensity (i = 2) in 8 samples (14.5%). Image analysis revealed broadly different levels of Trop-2 cleavage, which ranged from 85% to 1%–2%. However, no primary tumor showed an absence of Trop-2 cleavage, at variance with normal breast samples from control individuals, where we did not observe any cleaved Trop-2 (Fig. 25A). In 20% of the breast cancer samples, there were low levels of cleavage (≤15% of Trop-2 molecules). The majority of cases (67.3%) showed between 18% and 66% cleaved Trop-2 molecules, with high levels of proteolytic processing (67%–85% of Trop-2 molecules) in 12.7% of samples (Fig. 25A, B). These results extended and confirmed similar data obtained in samples of normal skin (Fig. 23A, B). We assessed the expression of ADAM10 in normal breast and in breast cancer samples (Fig. 25C). For best sensitivity, we pooled the cancer samples based on the cleavage

pattern of Trop-2 observed in the individual samples. We then quantified global levels of ADAM10 versus levels of activated/secreted form of ADAM10 [361]. This strategy allowed observing a tight direct correlation between the levels of the mature/active form of a secreted ADAM10 in breast cancer and the extent of Trop-2 cleavage.

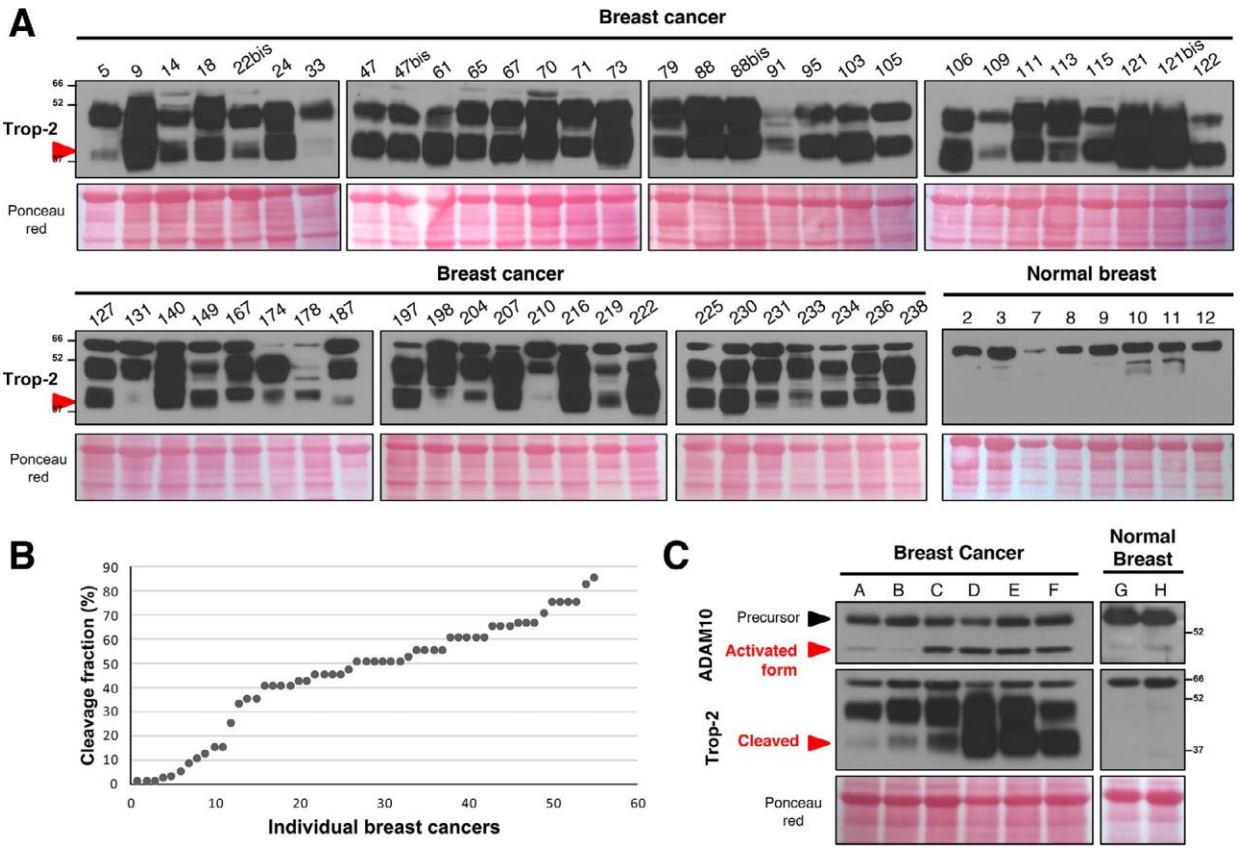


Fig. 25 : Breast cancer case series

3. Discussion

Better knowledge of the mechanisms underlying cancer cell growth and metastasis is an urgent need for cancer care. In this study, we set out to investigate the activation mechanisms of Trop-2, a signal transducer that is involved in tumor growth and metastasis.

The TROP2/TACSTD2 gene is an intronless derivative of TROP1/EPCAM [67, 94, 294]. The extracellular domain of the Trop molecules contains a GA733 type 1 motif and a thyroglobulin

repeat, which together have 12 conserved cysteine residues [25, 67, 94]. Both the GA733 type 1 and thyroglobulin domains have been proposed to be involved in homophilic intra-membrane and inter-membrane interactions of the paralogous Trop- 1/EpCAM [35, 294], which suggested a corresponding function for Trop-2.

We previously reported that proteolytic cleavage occurs between R80 and R81 in the first loop of the Trop-1 thyroglobulin domain [29], and that this activates the cell-growth stimulatory properties of Trop-1. We hypothesized that Trop-1 and Trop-2 may share the same activation mechanism. Here we have shown that this is indeed the case, as Edman degradation sequencing and peptide mass fingerprinting of Trop-2 purified to homogeneity revealed that the cleavage of the molecule occurs in the thyroglobulin domain, between R87 and T88. This site precisely aligns to the corresponding cleavage site of Trop-1. Cleavage at R87-T88 implies a large conformational rearrangement and potential swiveling of a 10 kDa subunit over the Trop-2 backbone, suggesting a large impact on the interactions of Trop-2 with binding partners. Cleavage of the extracellular domain(s) of transmembrane signal transducers is frequently catalyzed by MMPs. ADAM10 is a member of this family, and it has been shown to induce proteolytic processing of many type I and II transmembrane proteins, including cadherins [345, 347], CD44 [344],L1 [346], Her2 [362], HB-EGF [48, 363, 364], TNF-α [48], and APP [55].

Peptide library screening, combined with crystal structure analysis of known ADAM10 substrates [349] indicated preference for a consensus target site that includes Arg at P1 and Thr at P1', which are both present in the Trop- 2 candidate cleavage site. Similarly to Trop-2, ADAM10 is expressed by most human cancer types, and is upregulated in many of them [350-354, 365]. We have shown here direct interaction and dynamic colocalization of ADAM10 and Trop-2, in agreement with previous reports that signaling activation enhances the association of ADAM10 with its substrates, such as for HB-EGF and CD9, thereby modulating tumor invasion [363]. Inhibition of ADAM10 expression or activity in Trop-2–expressing cells was then shown to markedly reduce Trop-2 processing. These findings indicated that ADAM10 is effector protease at Trop-2 R87-T88.

Recently, matriptase was reported to recognize the R87-T88 Trop-2 cleavage site [331, 366], which raises the possibility of a finely regulated, multipronged post-translational processing

at this position. Stoyanova et al showed that TACE mediates proteolysis and ectodomain shedding of Trop- 2, which is followed by RIP through γ -secretase/presenilins, and nuclear translocation of the intracellular domain of Trop-2 upon release from the cell membrane [122]. Our membrane biotinylation/pull-down assays revealed that R87-T88–cleaved Trop-2 molecules are still anchored to the cell membrane. Furthermore, this R87-T88–cleaved Trop-2 remains linked to the cytoplasmic tail, which demonstrates that the Trop-2 cleavage at R87- T88 precedes TACE cleavage and RIP in the Trop-2 activation cascade.

R87A-T88A mutagenesis was shown here to lead to complete loss of wt Trop-2 pro-growth activity in vitro. Striking correspondence was observed also in vivo, where the A87-A88 Trop-2 mutant failed to stimulate tumor growth in both murine xenografts of L and HEK-293 cells.

We have previously shown that Trop-2 is highly expressed in most metastatic cancer cells and represents a key driver of metastasis. The proteolytic cleavage of Trop-2 at the R87-T88 site was shown to be a requirement for activation of Trop-2 for induction of the complete cascade from growth to metastasis in vivo.

Analysis of the Trop-2 processing in human samples here reveals that Trop- 2 is not cleaved in normal epidermidis and in squamous cell carcinomas, where Trop-2 could associate with retention of differentiation, whereas proteolytic cleavage occurs in basal cell carcinomas.

Correspondingly, no Trop-2–activating cleavage was detected in normal human breast samples. On the other hand, the majority of the breast tumors analyzed showed the Trop-2 cleavage, and a tight correlation between the extent of Trop-2 cleavage and the expression levels of the mature/active form of a secreted ADAM10 was observed. Together with the prognostic impact of Trop-2/ADAM10 on cancers of the lung and pancreas, this indicates a driving role for this Trop-2 activatory cleavage on the progression of malignant human tumors.

4. Conclusions

We have demonstrated here that Trop-2 recruits ADAM10 and is activated by ADAM10-mediated cleavage to trigger molecular pathways for tumor growth and metastasis. Although wt Trop-2 is strongly up-regulated in cancers, no cancer-related mutations have been as yet identified in the corresponding human gene (TACSTD2) [8]. Our findings here show that post-translational

processing of Trop-2 acts as the Trop-2 activator, to drive tumor growth and metastatic dissemination.

Our findings that the ADAM10-mediated cleavage of Trop-2 at the R87-T88 site is necessary to activate the transformed growth stimulatory activity of Trop-2 and promote its metastasis driver function now pave the way for the development of new therapeutic precision strategies. Many antibody-based therapeutic approaches that target Trop-2 are currently being developed; these include use of the humanized anti-Trop-2 antibody–SN-38 drug conjugate Sacituzumab govitecan-hziy (Trodelvy) in metastatic triple- negative breast cancers [340]. Inhibition of Trop-2 proteolytic processing may correspondingly open new perspectives for such specific therapies to control tumor growth and metastatic dissemination through targeting the Trop-2 activation steps in cancer cells.

D. GENOMIC ANALISIS OF TROP2 (TRANSCRIPTOM/ NGS): TROP-2 INDUCES ADAM10-MEDIATED CLEAVAGE OF E-CADHERIN AND DRIVES EMT-LESS METASTASIS IN COLON CANCER (Article II)

1. Introduction

Malignant cells from a primary tumor spread to distant organs in a complicated process known as cancer metastasis. Understanding this cycle at a molecular level might bring vital new insights into the illness and perhaps identify new therapy possibilities. The dynamic alterations that occur when tumor cells spread may be studied using transcriptome profiling of spontaneous cancer models. Due to their variability, however, such investigations are difficult to conduct. Using the well-established mouse model, this work sought to analyze the transcriptomes of metastatic colon cancer cells.

Metastatic disease is the dominant cause of death in cancer patients and is the greatest hurdle for cancer cure. Hundreds of proteins/genes have been linked to the metastatic phenotype [23]. However, neither proteomic analysis nor large-scale genome sequencing [321] succeeded in identifying reproducible markers of tumor aggressiveness and metastasis in cancer patients.

Metastasis-associated genes include not only drivers of the metastatic phenotype, but also secondary events, together with adaptive, counterbalancing changes [367]. Thus, to identify candidates with a decisive role in metastatic diffusion, we looked for genes that were concordantly dysregulated across multiple cancer metastasis models. This led to the discovery that TROP2 is a unique upregulated gene in metastatic colon cancer prototypes. Trop-2 is a tumor and stem cell growth inducer [8, 122, 368]. We and others showed that Trop-2 drives liver colonization of prostate cancer cells [131, 156]. Additional findings showed that Trop-2 expression associates to increase in vitro migration and invasion of oral squamous cell and gallbladder carcinoma cell lines [116, 180]. Trop-2 was also reported to be associated to altered expression of markers of epithelial-mesenchymal transition (EMT) in stomach and breast cancer [121, 155].

We recently discovered that proteolytic activation of Trop-2 by ADAM10 underlays its capacity to drive colon cancer malignant progression [1]. However, Trop-2-driven mechanisms of action and activatory pathways in metastasis remained unknown. We previously showed that Trop-2 is upregulated upon tumor progression in colon cancer [8]. Here we find that impairment of the membrane-to-nucleus signaling by removal of the Trop-2 cytoplasmic tail boosted the liver metastasis rate and stimulated colorectal cancer growth. NGS and targeted gene-expression profiling showed no increase in EMT transcription factors nor downregulation of the E-cadherin-encoding CDH1 gene. We found instead that Trop-2 triggered ADAM10 to cleave E-cadherin, leading to release from the scaffolding cytoskeleton, with disruption of cell-cell junctions, acquisition of invasive capability and activation of pro-metastatic signaling by β-catenin.

This Trop-2-driven metastasis program was found to associate to a significantly worse relapse-free and overall survival of patients with colorectal cancer, suggesting a pivotal impact on colorectal cancer progression. Sacituzumab govitecan-hziy is an anti-Trop-2 antibody-SN-38 drug conjugate, which was shown to be effective in metastatic breast cancer [339, 340, 369, 370] and was granted accelerated FDA approval as Trodelvy, for use in the clinics. Our findings support an indication for Trop-2-targeting therapies in metastatic colon cancer.

2. Results

a) Transcriptomic profile of colon cancer metastatic cells - DNA microarray analysis

The top upregulated genes in metastatic HCT116U5.5 cells were TM4SF1, TNFSF4, MIA, and TACSTD2/TROP2 (Fig. 26, Table S1A). The tetraspanin-like TM4SF1 is involved in formation of nanopodia [371] and promotes prostate cancer cell migration [372]. TNFSF4 encodes the OX40 ligand, which stimulates cell survival via sustained activation of PI3K and Akt [373]. Mia encodes the melanoma-derived growth regulatory protein, which enhances extravasation and metastasis of melanoma cells [374]. This protein has been shown to interact with fibronectin, suggesting involvement in the detachment of cells from the extracellular matrix (ECM) through competition with β1 integrins for ECM binding [375].

Consistent with this, our findings showed that Trop-2 inhibits cancer cell adhesion to fibronectin, by modulating the β1 integrin signaling network [150]. The most downregulated genes in HCT116U5.5 were NRIP1 and HSPG2 (Fig. 26, Table S1A). NRIP1 encodes RIP140, a strong transcriptional repressor that inhibits the mitogenic effects of estrogen receptor α through interaction with FHL1, a suppressor of cancer cell growth and migration [376]. HSPG2 encodes perlecan, an extracellular proteoglycan that is upregulated in desmoplastic ECMs and limits the process of tumor invasion [377].

The top upregulated genes in metastatic KM12 cells were GSTP1, SN2/SLC38A5, PAG and again TACSTD2/TROP2 (Fig. 26, Table S1B). GSTP1 encodes the glutathione S-transferase P1 and it is found hypomethylated in sentinel lymph node breast cancer metastases [378], though not in all cases [379]. The high affinity glutamine transporter SN2/SLC38A5 is induced by Myc and contributes to promoting mitochondrial glutamine metabolism (and glutamine addiction) in cancer cells [380]. The transmembrane adaptor PAG was found disregulated in non–small cell lung cancer [357, 381].

The most downregulated genes in metastatic KM12 cells were SLC6A8, ASB4, SLC38A10 and CALB1 (Fig. 26, Table S1B). Inactivation of the creatine transporter SLC6A8 results in creatine depletion via impaired creatine uptake [382]. A low creatine content has been

reported in several tumor types, together with a progressive decrease of cellular creatine with malignant transformation of skeletal muscle into sarcoma [383].

The ankyrin repeat and SOCS box-containing 4 (ASB4) modulates insulin signaling by increasing the degradation of the insulin receptor substrate 4 [384]. SLC38A10 is a putative sodium-coupled neutral amino acid transporter [385], which is involved in the regulation of cell volume in the mesenchyme [386]. The calcium binding Calb1 has been suggested to play a role in the downregulation of proapoptotic genes [387].

b) Transcriptome profiling of metastatic colon cancer cells

To identify drivers of metastatic diffusion, we profiled transcriptomes of unrelated cancer-metastasis systems, in a quest for concordantly dysregulated genes. To date, candidate metastasis genes primarily originated from mouse tumor models [388]. Hence, we focused on human metastatic cancer systems. Two independent human colon cancer metastatic models were analyzed, i.e. the non-metastatic HCT116 versus metastatic HCT116U5.5 cells [308] and the non-metastatic KM12C versus metastatic KM12SM and KM12L4A cells [307]. The HCT116U5.5 metastatic clone was derived from lung metastases of subcutaneously (SC) injected HCT116 cells [308].

Metastatic KM12SM and KM12L4A cells were derived from liver metastases of KM12C cells injected in the mouse spleen [307]. DNA microarray analysis for the expression of 12652 genes in the HCT116 and KM12 metastatic models, identified 70 and 159 differentially expressed genes, respectively (Supplementary Results, Table S1). The intersection of these independent datasets showed predominant, mutually exclusive expression changes (Fig. 26A), indicating widely divergent transcriptomic scenarios in the two colon cancer models.

Only two genes, TROP2/TACSTD2 and VIMENTIN, were upregulated in both systems (Fig. 26A, Table S1), first suggesting that pro-metastatic programs may only require few decisive determinants. mRNA findings were validated by RT-PCR (Fig. 26B). Differential expression of the encoded proteins was validated by Western blotting (Fig. 26B), immunofluorescence confocal microscopy and immunohistochemistry (IHC) (Fig. 26C).

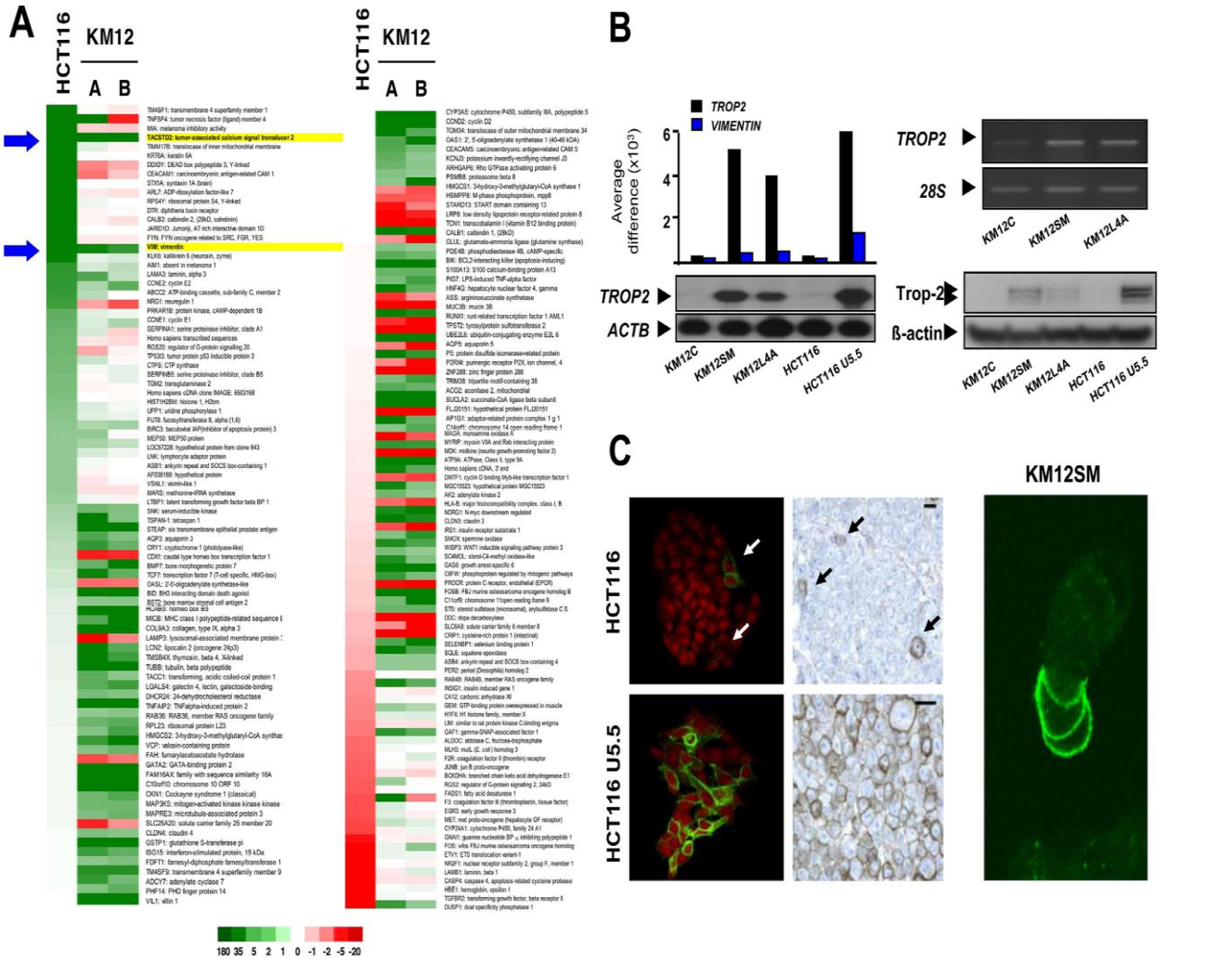


Fig. 26 Transcriptome profiling of metastatic colon cancer cells

c) The role of Trop-2 in metastatic diffusion

Trop-2 proteolytic cleavage by ADAM10 at R87-T88 is required for activation as a driver of primary tumor growth and malignant progression [1], while a two-step A187-V188 and G285-V286 intramembrane proteolysis releases the Trop-2 intracellular tail, which then co-translocates with ß-catenin to the nucleus [122]. To dissect the contribution of Trop- 2 nuclear signaling to colon cancer metastasis we designed a membrane- anchored mutant of Trop-2 devoid of the cytoplasmic tail (ΔcytoTrop-2). wtTrop-2 and ΔcytoTrop-2 KM12SM cells were selected by flow cytometry for membrane expression levels comparable to those found in human cancer (Fig. 27) [8]. KM12SM transfectants were then assayed for capacity to metastasize to the liver upon

injection in immunosuppressed mouse spleen (Fig. 27) [307]. wtTrop-2 strikingly increased the metastatic capacity of KM12SM cells, raising metastasis rates from 45% for control cells to 90% for wtTrop-2 transfectants (Fig. 28A, Table IV, Table S2A). The tail-less ΔcytoTrop-2 boosted metastatic growth, with metastatic livers reaching up to four times their normal size at four weeks after tumor injection (3.25± 0.64 cm3 versus 0.45 ± 0.18 cm3 for wtTrop-2, P<0.0001) (: Trop-2 drives metastasis in vivo.Fig. 28A, Table IV, Table S2A).

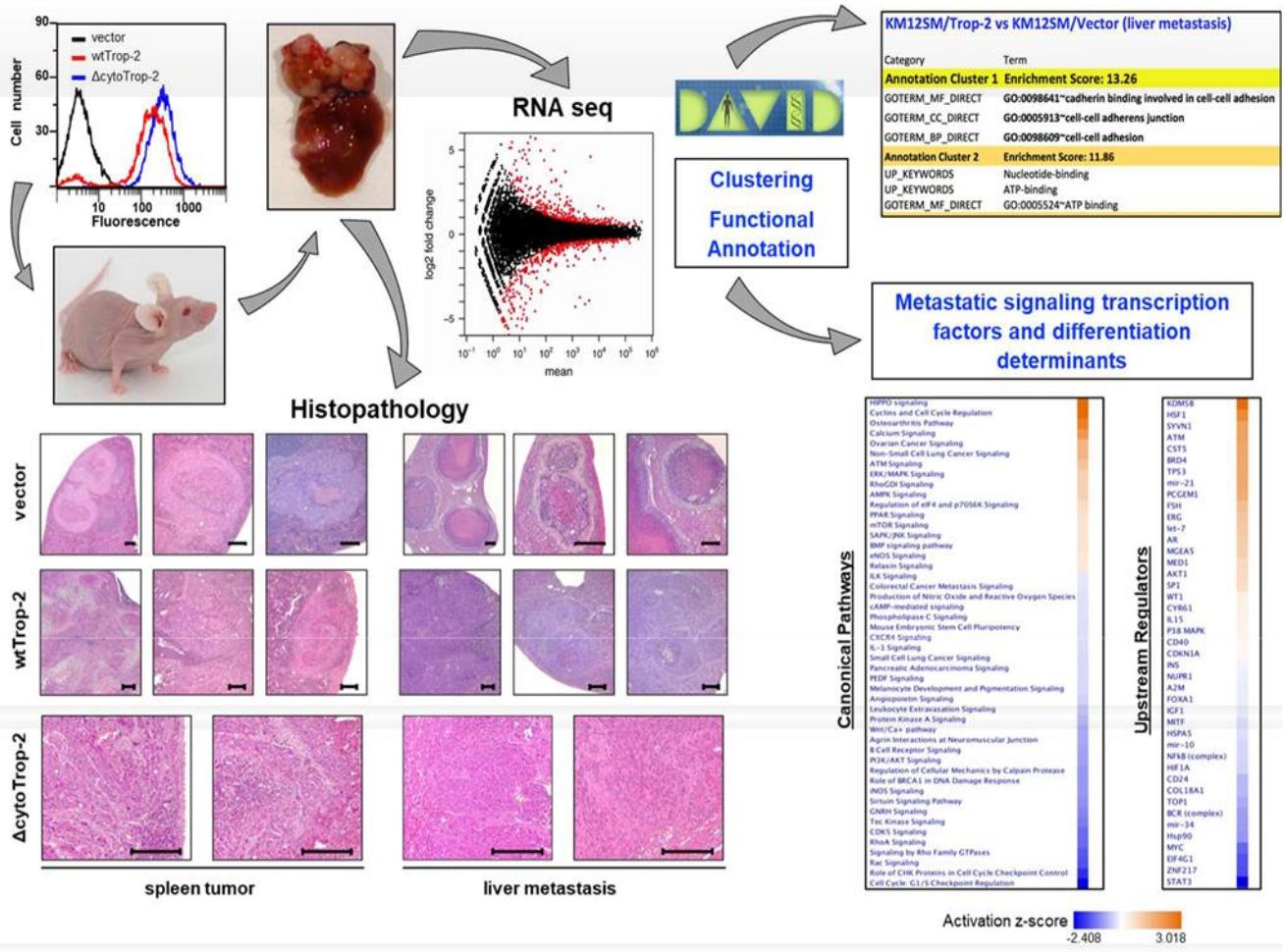


Fig. 27 : In vivo analysis of Trop-2 pro-metastatic ability

Histopathology analysis showed patterns of tumor growth of parental KM12SM cells as manifold nodules, with central necrosis and peripheral fibrous capsule. Residual tumor differentiation capacity was demonstrated by gland rudiments containing apoptotic bodies in their lumen. This closely mirrored human colon cancer, where such crypt-like architecture associates with low cancer grading and good prognosis [389] (Fig. 28B, Fig. 27, Table S2B). wtTrop-2-expressing metastases showed much reduced numbers of differentiated glands (mean±SD glands/20x field: 0.9 ± 0.97 versus 3.42 ± 0.86 in controls, P = 0.0003).

Poorly differentiated, smaller crypts were also reduced in frequency (2.54 ± 2.33 versus 10.58 ± 7.33 in controls, P = 0.0026). Thinner peri-metastatic capsule and pseudo-capsule and frequent mitotic figures (mean±SD mitotic figures/20× field 10.98 ± 9.18 versus 1.58 ± 2.11 in controls, P = 0.0004) were observed (Fig. 28B-C, Fig. 27, Table S2B). Augmented vascular invasion capacity was also observed (Fig. 28B). Loss of central-nodule apoptosis, and a shift toward smaller and peripheral necrotic areas, indicated a reduced apoptotic response to oxygen gradients/distance from blood vessels [7, 118] (Fig. 28B-C, Table S2B).

ΔcytoTrop-2 induced dramatic changes toward invasive growth patterns, with further loss of differentiated glands (0.12 ± 0.32, P <0.00001 versus control) and poorly differentiated glands (1.72 ± 1.44, P = 0.0007 versus control) (Fig. 28B, Fig. 27, Table S2B). Frequent mitotic figures were found (5.34 ± 4.39, P = 0.022 versus control) (Fig. 28C, Fig. 27, Table S2B).

Cancer cell invasion was detected across layers of hepatocytes or even within pre-existing metastatic areas (Fig. 28B), with patterns of invasive growth and little, if any, pseudo-capsule formation. Haemorrhagic necrosis indicated damage to vascular walls. Trop-2-dependent metastases showed diminished apoptotic features in activated ΔcytoTrop-2 versus parental KM12SM cells (mean±SD apoptotic bodies/20x field: 6.92 ± 5.36 versus 13.16 ± 12.75, respectively) (Fig. 28C, Fig. S1, Table S2B). Trop-2 prompted cancer cell rounding and diminished cell-cell adhesion in primary tumors and metastases (Fig. 2B, bottom panels; Fig. S1).

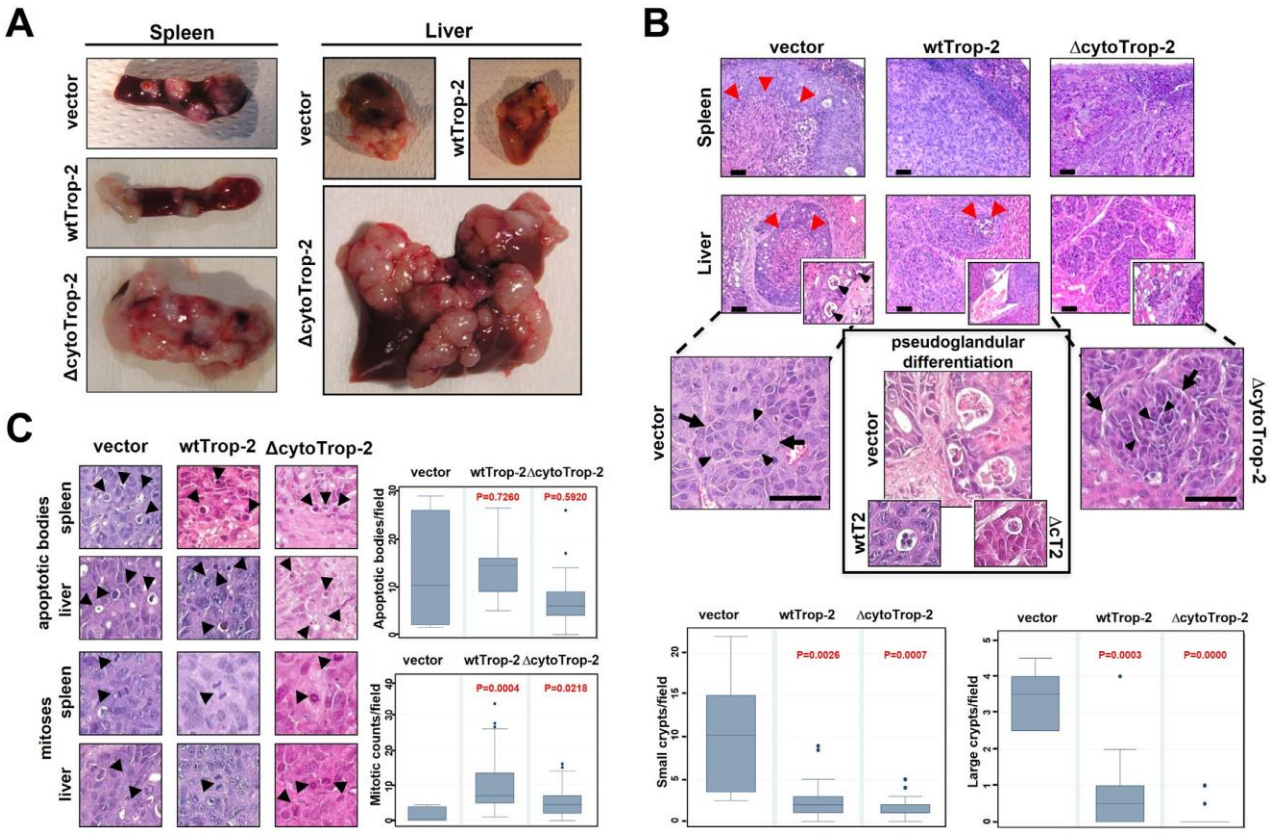


Fig. 28 : Trop-2 drives metastasis in vivo.

Table IV : In vivo metastatic capacity of colon cancer cells expressing Trop-2

	% spleen[a]	Size (cm^3)	% liver[b]	Size (cm^3)
Vector wtTrop-2	64.9 ± 8.7 69.0 ± 10.6	0.11 ± 0.04 0.09 ± 0.03	45.8 ± 14.8 **90.0 ± 12.0**[c]	0.50 ± 0.21 0.45 ± 0.18
ΔcytoTrop-2	**100.0 ± 0**[d]	**0.37 ± 0.23**[e]	76.9 ± 11.7	**3.25 ± 0.64**[f]
Vector wtTrop-2	64.9 ± 8.7 69.0 ± 10.6	0.11 ± 0.04 0.09 ± 0.03	45.8 ± 14.8 **90.0 ± 12.0**[c]	0.50 ± 0.21 0.45 ± 0.18

Vector : *n* = 37 ; wtTrop-2 : *n* = 31 ; *ΔcytoTrop-2* : *n* = 13.

Data were from 3 to 9 experiments per group. Values are expressed as mean±SEM. Bold: statistical significance or trend. primary tumors take rates (percentage of injected cases). b percent incidence of liver metastases. c Fisher's exact test: $P = 0.0114$ versus vector. d Fisher's exact test: $P = 0.0382$ versus wtTrop-2. e trend. f Student's t test: $P < 0.0001$ versus wtTrop-2.

d) Profiling of metastatic transcriptomes

NGS transcriptome analysis was conducted on independent KM12SM colon primary tumors and metastases (Fig. 27), to detect Trop-2- driven, differential expression of pro-metastatic determinants (Fig. 29). Compound EST analysis showed high intra-group R correlation coefficients, with robust segregation of parental cells, wtTrop-2 and ΔcytoTrop-2, primary cancer and metastasis expression profiles. Sequence alignment over Mus musculus or Homo sapiens reference genomes formally discriminated transcripts of human cancer origin versus murine stromal components (Table S3).

Human EST profiles of primary tumors showed invariant expression of 87.1% of 15782 transcripts between wtTrop-2-expressing versus parental cells. On the other hand, only 19.2% mRNA expression patterns overlapped in wtTrop-2 and control liver metastases, indicating that Trop-2 induces a distinct pro-metastatic program (Fig. 29A, Table S3). Wide modulation of the 1E-cadherin/ß-catenin/TCF-LEF and Wnt/Frizzled/DVL1/APC/ß- catenin pathway components was found, suggesting these to be drivers of Trop-2-induced metastasis. Modulated downstream targets included apoptosis regulators, such as PI3K/Akt [118], NFkB [7], p53 [390], cFOS, caspases 3 and 9, JAK/STAT1-3 and Cyclin D1 [7] (Fig. S2B).

Canonical EMT models predict progression to a metastatic state through loss of epithelial differentiation, via transcriptional downregulation of epithelial determinants, such as E-cadherin [23, 318, 319, 321, 391]. Hence, we explored EMT drivers versus epithelial differentiation markers in Trop-2-driven metastatic cells (Table S3H). Of 203 EMT biomarkers, 129 were transcribed in the colon cancer cells, 16 were differentially modulated by wtTrop- 2/ΔcytoTrop-2; 164

biomarkers were expressed in the tumor stroma, 76 of these were differentially modulated by wtTrop-2 and ΔcytoTrop-2.

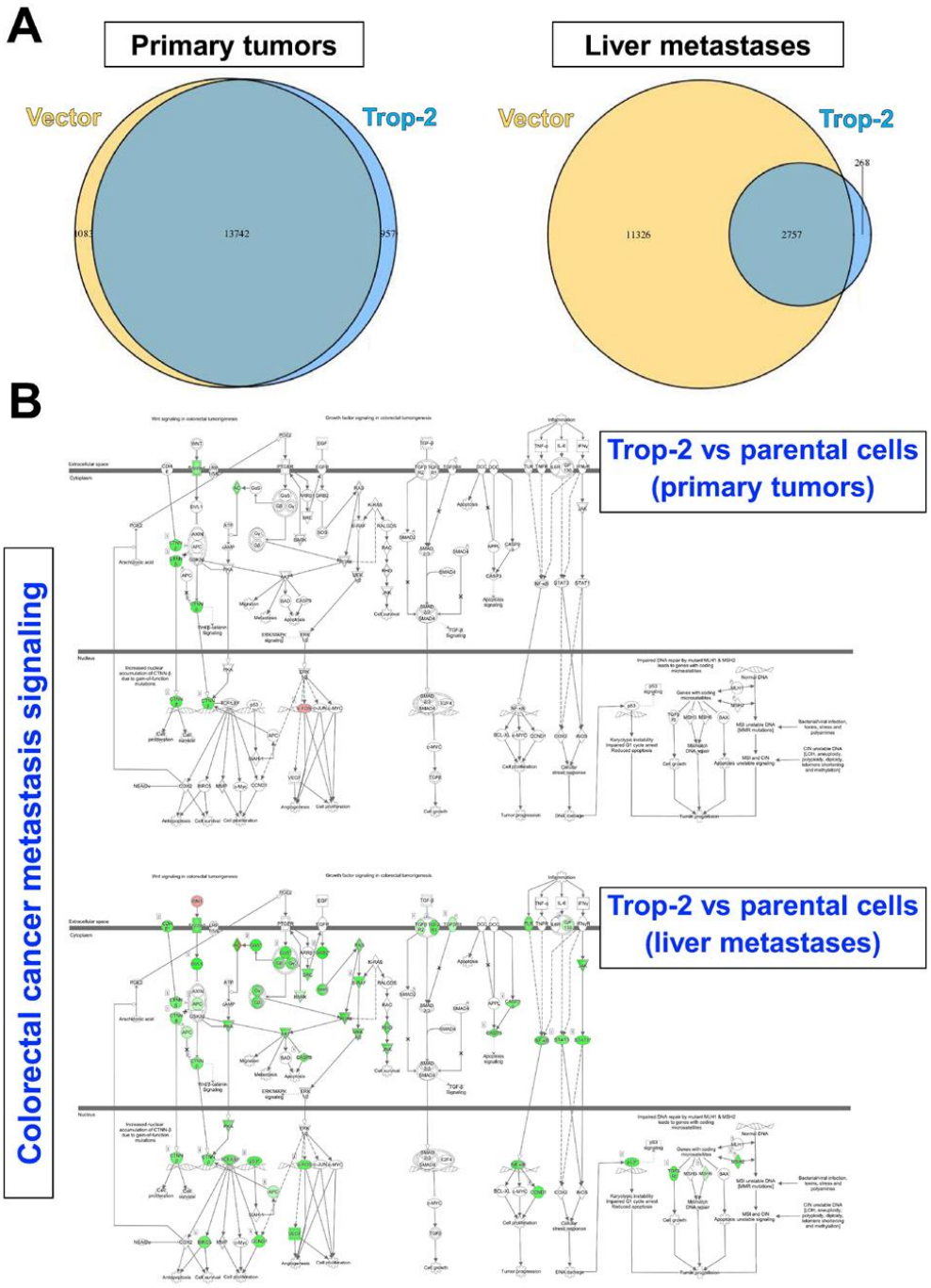


Fig. 29 : Trop-2-driven effector networks for metastasis

The EMT transcription factors TWIST1, TWIST2, ZEB1, ZEB2, SNAI2, KLF8, PRRX1, GSC, SIX1 and TCF4 were absent from Trop-2-driven colon cancer cells, SNAI2 was only expressed by the cancer stroma. TFAP2A and TCF3 were downregulated by wtTrop-2 in metastases. (Table S3H). Vimentin was not upregulated in metastatic colon cancer cells,but was induced by ΔcytoTrop-2 in stromal cells, consistent with a model whereby the metastasis stroma contributes to global EMT shifts [321].

Among epithelial differentiation determinants, MAPK6, EPCAM/TACSTD1/TROP1 [294] and most keratins were expressed at comparable levels in primary tumors and metastases. KRT80, which is associated with progression of colon cancer, was downregulated by wtTrop-2 in metastases (Table S3H). None of the α-integrins, including α6, nor the β-integrins, among them β1, β3, were significantly modulated in cancer cells or stroma. Myosins/β-actin are dynamically co-recruited at the cell membrane upon cross-linking with anti-Trop-2 antibodies. Consistent with this, numerous myosin isoforms were differentially modulated in metastases, whether in stroma (Myo 1a, 1b, 1c, 1f, 1g, H11, L9) or in cancer cells (Myo5a, 6, 7, 10, HC7B, HC10, HC15, myosin regulatory light chain - MRLC2). Conserved, high levels of CDH1 mRNA were revealed in metastases from wtTrop-2 and ΔcytoTrop-2 versus vector-transfected KM12SM cells as well as in primary tumors across all experimental groups (Table S3H). RT-PCR, FACS and Western blotting analyses also showed invariant levels of E-cadherin transcript and protein across transfectants and between control Trop-2-nil/non-metastatic KM12C and Trop-2-low/pro-metastatic KM12SM cells (Figs. S3A-B, D).

Taken together these data suggested limited impact of Trop-2 on epithelial differentiation, at variance with previous, more limited analyses [121, 155].

e) Trop-2 releases E-cadherin from the β-actin cytoskeleton and activates ß-catenin

Our findings showed that Trop-2 impacts on the E-cadherin/ß-catenin pathway in metastasis (Fig. 29B, Table S3H). Evidence of Trop-2-associated reduction in cell-cell adhesion had been obtained by morphometric analysis of metastases growing in vivo. Together, these

findings suggested a regulatory function of Trop-2 on intercellular junctions via alternative mechanisms to CDH1 downregulation.

Coimmunoprecipitation assays revealed that wtTrop-2 tightly interacts with E-cadherin (Fig. 30A). ΔcytoTrop-2 bound E-cadherin to a lesser extent, first suggesting a role of the Trop-2 cytoplasmic tail in macromolecular assembly with E-cadherin. We thus hypothesized an antagonistic function of Trop-2 versus E-cadherin/ß-catenin-mediated cell-cell adhesion that would critically contribute to metastatic diffusion [392]. This was explored in in vitro assays on low cell-attachment substrates. wtTrop-2 and ΔcytoTrop-2 markedly diminished KM12SM cell-cell adhesion (Fig. 30C). Distribution profiles of wtTrop-2 /ΔcytoTrop-2 cells indicated a highly significant reduction in aggregates with respect to control cells (χ 2 test P-value <2.2e−16). A shift of aggregate cluster size, toward smaller cell clusters and isolated cells, was correspondingly observed (Fig. 30C). Corresponding invasive capability of KM12SM transfectants was assesses by inverse matrigel invasion assays [393]. wtTrop-2 expression conferred KM12SM cells the ability to invade through the matrix, compared with the control vector-only cells. Invasive ability was maintained by the tail-less ΔcytoTrop-2 mutant (Fig. 30B), with no significant difference in the number of invading clusters with respect to the wt-Trop-2. ΔcytoTrop-2 clusters were smaller than the wt Trop-2 ones at 5 days from seeding, which could reflect differences in cell growth kinetics (see below).

We have previously shown that Trop-2 tightly interacts at the cell membrane with the ADAM10 metalloproteinase, which in turn leads to activation of Trop-2 via proteolytic cleavage at R87-T88 [1]. ADAM10 has also been shown to be responsible for E-cadherin processing in epithelial cell monolayers, which regulates cell adhesion and sorting [394].

Hence we investigated whether Trop-2 could drive E-cadherin functional inhibition via ADAM10 proteolysis. Western blot analyses of KM12SM tranfectants revealed that Trop-2 induces the release of the E-cadherin 30 kD intracellular domain (Fig. 32A). To validate this finding we assessed E-cadherin status in the HT29 colon cancer cell line, which displays a bimodal distribution of Trop-2 expression (Fig. 32B).

Western blot analyses of distinct HT29 subpopulations selected for low versus high levels of Trop-2 expression confirmed stimulation of E-cadherin cleavage by Trop-2 (Fig. 32B).

Downregulation of ADAM10 expression by a specific inhibitory shRNA abolished E-cadherin proteolysis in KM12SM/Trop-2 transfectants (Fig. 32C), thus indicating ADAM10 as the effector protease of Trop-2-induced E-cadherin shedding. Specificity was confirmed with a second, independent ADAM10 shRNA, which was equally effective in inhibiting Ecadherin processing (Fig. 31E).

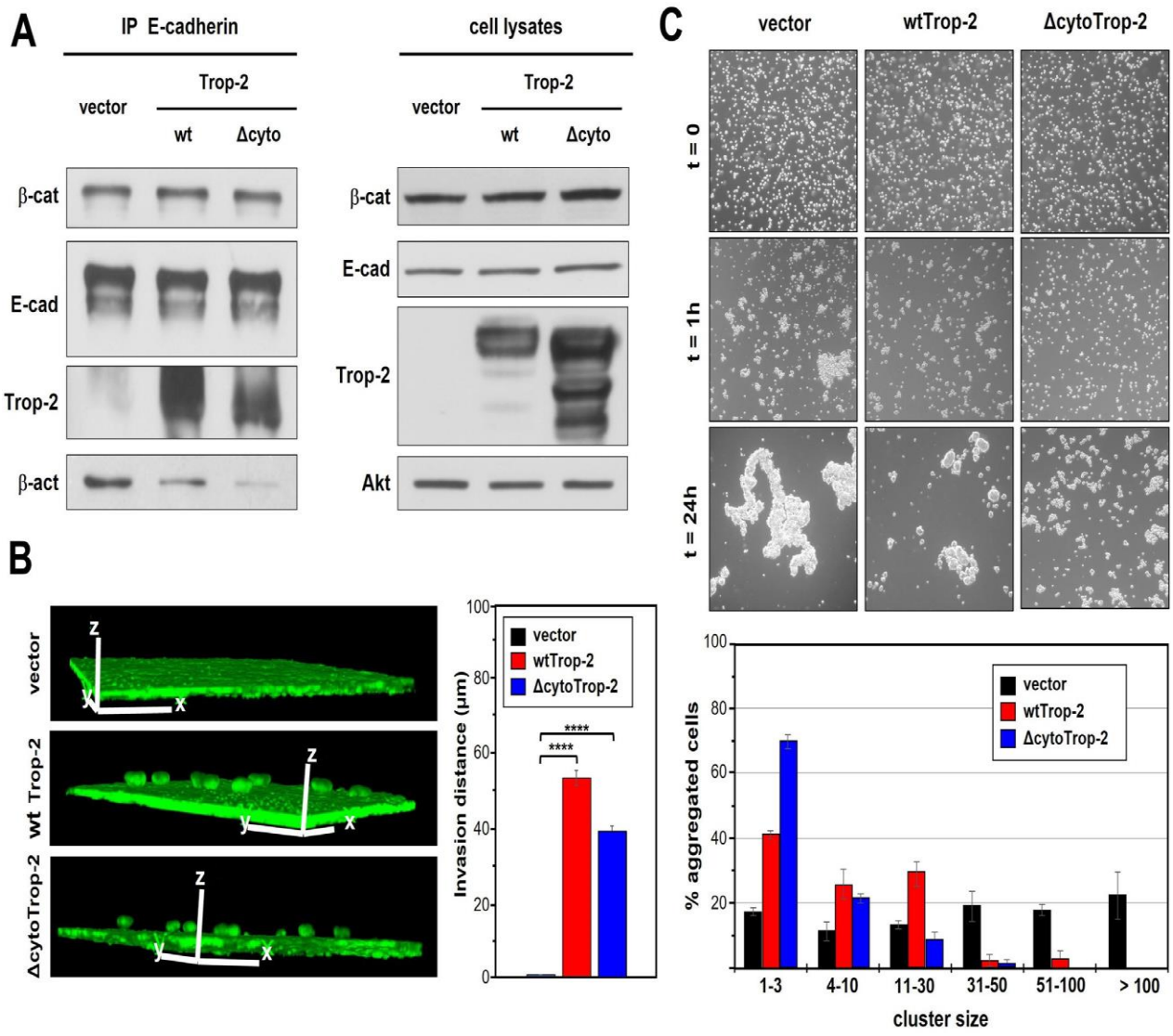


Fig. 30 : Trop-2 drives functional inactivation of E-cadherin.

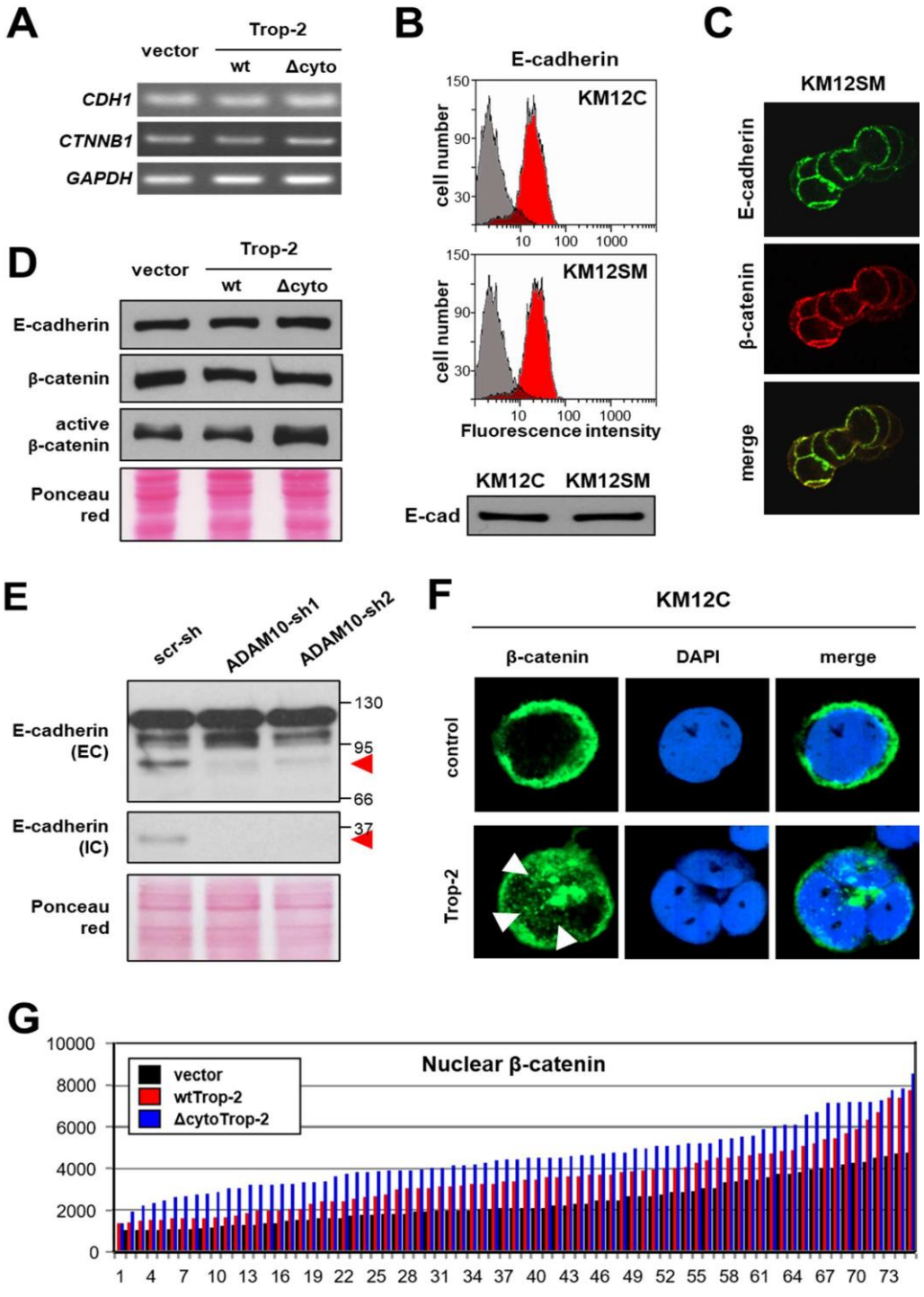


Fig. 31 : Trop-2-dependent E-cadherin and β-catenin activation status

The E-cadherin/β-catenin complex requires anchoring to the β- actin cytoskeleton for effective cell-cell adhesion [395]. wtTrop-2 severely diminished, and ΔcytoTrop-2 essentially abolished, E-cadherin binding to the actin cytoskeleton (Fig. 30A). Mechanistic analysis showed binding of Trop-2 to ezrin, indicating this as a requirement for bridging the Trop-2/E- cadherin/β-catenin complex to the cytoskeleton (Fig. 32D).

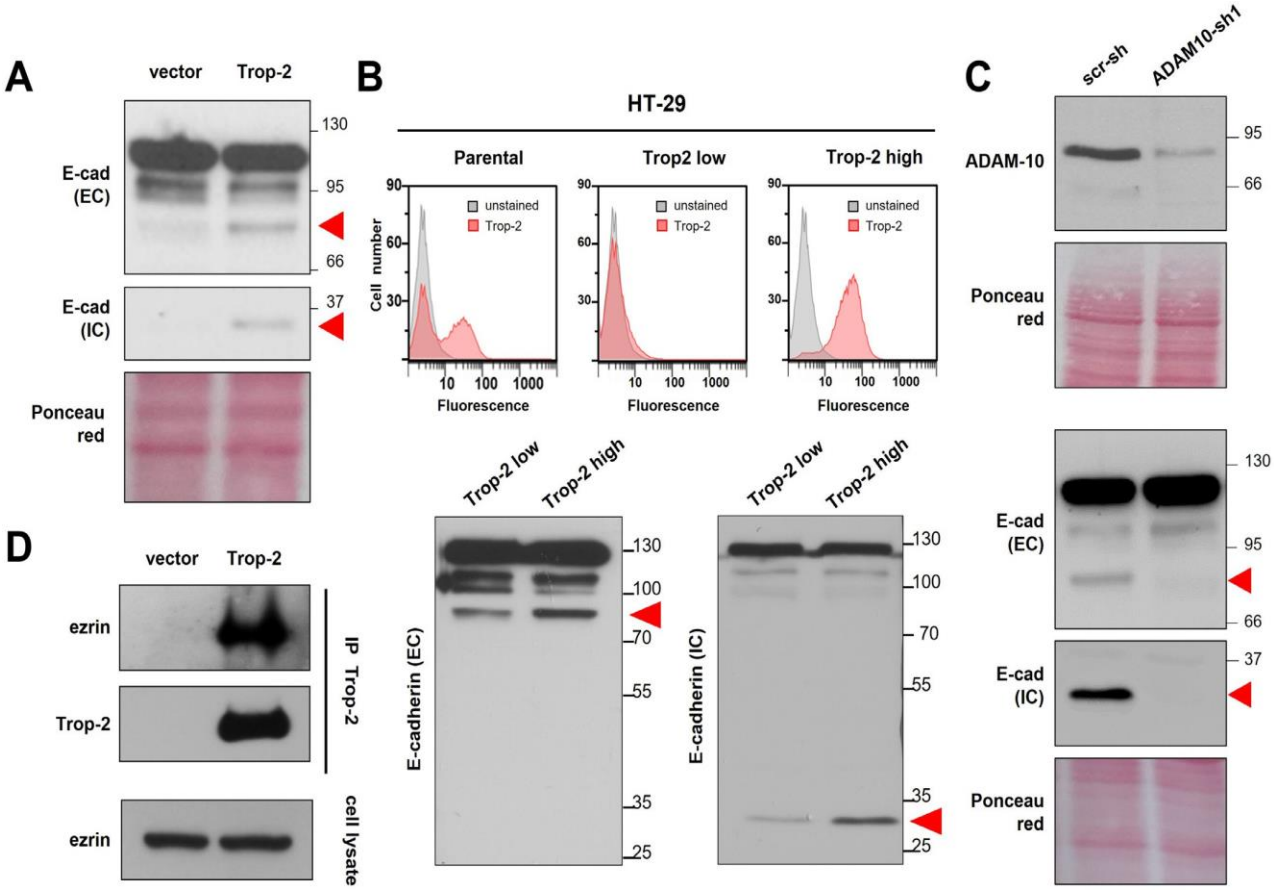


Fig. 32 : Trop-2 induces E-cadherin cleavage by ADAM10

E-cadherin inactivation leads to activation of the β-catenin transcriptional activity [395]. Such activation requires transit of β-catenin at the cell membrane, for dephosphorylation at S37/T41 [310]. KM12C, pro- metastatic parental KM12SM and KM12SM cells transfected with wtTrop- 2, ΔcytoTrop-2 or control vector were analyzed by confocal microscopy for β-catenin levels and localization (Fig. 31C, F-G, Fig. 33A-B, Table S4A). Consistent with a gradient of metastatic competence, the lowest levels of nuclear β-catenin were found in the non-metastatic KM12C < pro- metastatic parental KM12SM < wtTrop-2 < ΔcytoTrop-2. Quantitative image

analysis showed progressively higher mean values of nuclear/activated β-catenin in vector (mean±SEM 2604 ± 240 SV), wtTrop-2 (mean:3362±283 SV; P <0.0001 versus control vector) and ΔcytoTrop-2 cells (mean: 4431 ± 196 SV; P<0.0001 versus wtTrop-2) (Fig. 31G, Fig. 33A-B, Table S4A). This demonstrated that β-catenin signaling activation by Trop-2 in colon cancer does not require clustering with a membrane-released Trop-2 cytoplasmic tail [122].

We verified the transcriptional competence of Trop-2-activated β-catenin, using a β-catenin-RE-GFP reporter (Fig. 33B, C). GFP-expressing/β-catenin- responsive KM12C cells were 13%. wtTrop-2 increased the percentage of GFP-expressing/β-catenin-responsive KM12C cells to 24%; ΔcytoTrop- 2 led them to 28%. As predicted, Trop-2-positive metastatic KM12SM showed higher baseline transcriptional values (47%) than non-metastatic KM12C cells. Overexpression of wtTrop-2 in KM12SM cells led to a further increase of responsive cells to 54% (Fig. 33B, C). The highest β-catenin-driven GFP expression was induced by ΔcytoTrop-2 (67% of above-threshold cells).

β-catenin impact on transcription of target genes in vivo was assessed. β-catenin regulates genes associated with proliferation, differentiation, migration and angiogenesis. A set of 277 β-catenin target genes in colorectal cancer was retrieved (web.stanford.edu/group/nusselab/cgi-bin/wnt/target_ genes; Gene Expression Omnibus database). This included “direct” targets, as defined by the presence of TCF binding sites, and “indirect” targets, as obtained from large-scale β-catenin inactivation screening programs. β- catenin targets included the proto-oncogenes MYC and CCND1, as well as the genes encoding the basic helix-loop-helix proteins ASCL2 and ITF-2B [325]. This dataset was interrogated in primary tumors and metastases. Most remarkably, essentially no impact of Trop-2 on β-catenin target genes was revealed in primary tumors. Further, β-catenin was found to be downregulated in control, parental KM12SM metastases, with further loss of transcription of β-catenin targets. Massive changes were, on the other hand, revealed in β-catenin target genes in wtTrop-2 metastases (89 target transcripts, 61 genes overall) (Table S4B). Notably, control and wtTrop-2- modulated mRNAs showed no overlap, indicating that Trop-2 triggered a distinct control mechanism of metastatic diffusion.

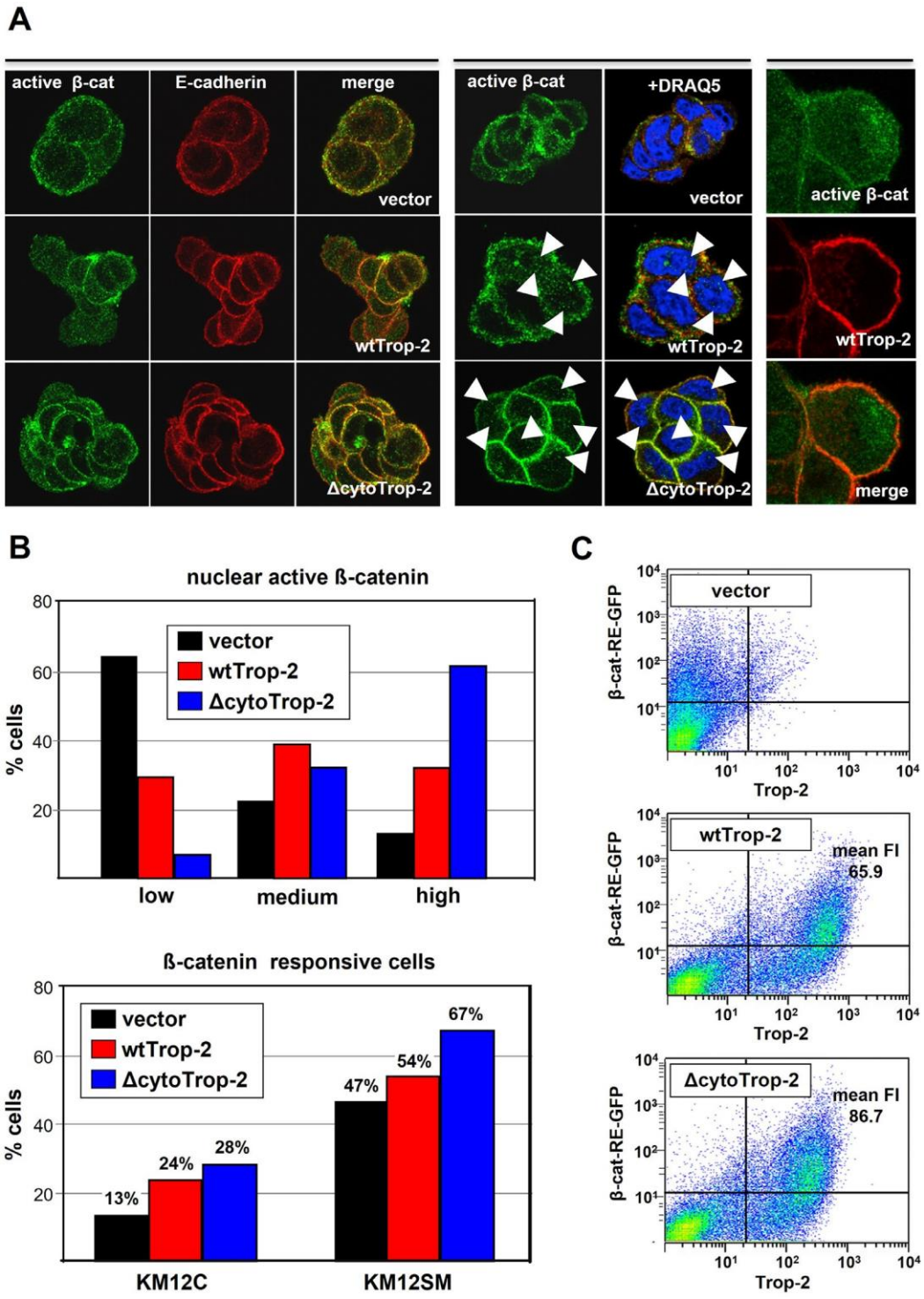


Fig. 33 : β-catenin release and activation by Trop-2. Activation of β-catenin transcriptional activity requires transit at the cell membrane and transport to the nucleus. Levels and localization of β-catenin were analyzed in KM12C, KM12SM cells transfected with wtTrop-2, cytoTrop-2 or control vector.

The vast majority (50) of modulated genes was shared between wtTrop-2 and ΔcytoTrop-2. Marked increased transcription was induced by ΔcytoTrop-2 versus wtTrop-2 in most cases (65/71 genes), consistent with the activated nature of ΔcytoTrop-2 and the aggressiveness of its induced metastases.

f) The Trop-2/E-cadherin/β-catenin pro-metastatic module in vivo

The Trop-2/E-cadherin/β-catenin pro-metastatic module was validated in colon cancer metastatic xenotransplants. E-cadherin and β-catenin levels were assessed by IHC analysis of matched primary tumor/metastasis pairs. wtTrop-2 KM12SM colon cancer cells showed higher levels of E-cadherin than control, parental cells. Further increase in levels of expression and recruitment to the cell membrane were detected in ΔcytoTrop-2 cells (Fig. 34A).

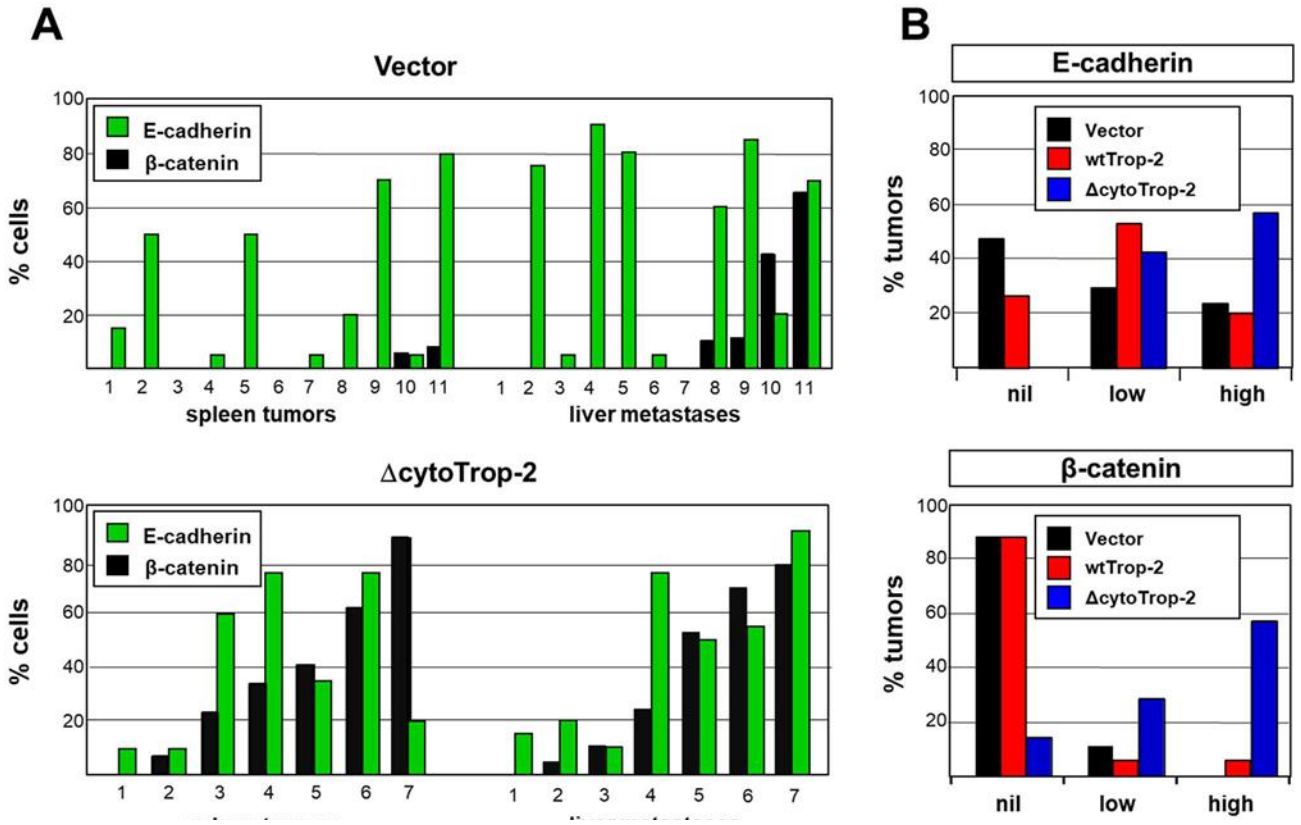


Fig. 34 : E-cadherin/β-catenin expression in Trop-2-driven tumors and metastases

Subgroup analyses of primary spleen tumors showed that 45% of controls were devoid of E-cadherin, and only 20% of cases reached high levels of expression. On the other hand, 70% of wtTrop-2 tumors expressed E- Cadherin, 20% of cases reaching high levels. No ΔcytoTrop-2 tumors were devoid of E-cadherin, and almost 60% of cases reached high expression levels (Fig. 34B).

In parallel assays, the lowest levels of β-catenin were found in parental KM12SM primary tumors. β-catenin expression increased in liver metastases (Fig. 34A). ΔcytoTrop-2 xenotransplants showed a vast increase of β-catenin expressing cells versus control tumors. Correspondingly high β-catenin expression was maintained in matched metastases (Fig. 34A). Subgroup analysis for expression of β-catenin showed that 90% of control tumors did not express the protein, and none reached high levels of expression. On the other hand, 10% of wtTrop-2 and 60% of ΔcytoTrop-2 tumors expressed high levels of β-catenin (Fig. 34B).

g) Trop-2 metastatic signalosome dissection

Metastatic diffusion requires resistance to apoptosis, enhanced migration and invasiveness. We modelled the contribution of these mechanisms to the Trop-2-driven metastatic process through global, multi-platform proteomic analyses (Fig. 35, Table S5) and in vitro functional assays (Fig. 36). Proteomic analysis of wtTrop-2 and ΔcytoTrop-2 metastatic cells showed modulation of RTK signaling, including Erb-B2, PDGFRα, FAK, the FAK- related kinase Pyk2 and PKCs, with Src/FAK-dependent JNK activation (Fig. 35, Table S5). wtTrop-2 induced activatory phosphorylation of Akt at T308 [118]. This was further increased (+80%) by the L'lcyto Trop-2 mutant, with Akt-mediated inactivating phosphorylation of Raf-1 (S259) and GSK3α/β (Ser21/Ser9) [118] (Fig. 35A,B, Table S5). Trop-2 induced Akt-mediated activation of c-Jun [396] (Table S5). This paralleled claudin-dependent, Akt-driven modulation of the FoxO1/TCF-4/β-catenin transcriptional repressor complex for cell-cell junction assembly [397]. Inactivation/downregulation of Rb contributes to metastatic spreading [398]. Consistent with this, we found reduced levels of Rb in KM12SM cells upon expression of wtTrop-2, and these paralleled activating phosphorylation of PKCε at S729 [399]; Rb levels were further reduced by the Δcyto mutant (Fig. 35A,B). Cyclin D1/cdk4 complexes trigger Rb phosphorylation at S807/811 and prevent binding to E2F-1 [400]. Consistent with a model of Trop-2-driven Rb

inactivation, these sites were found hyperphosphorylated in KM12SM/Trop-2 (+221.2%) (Table S5).

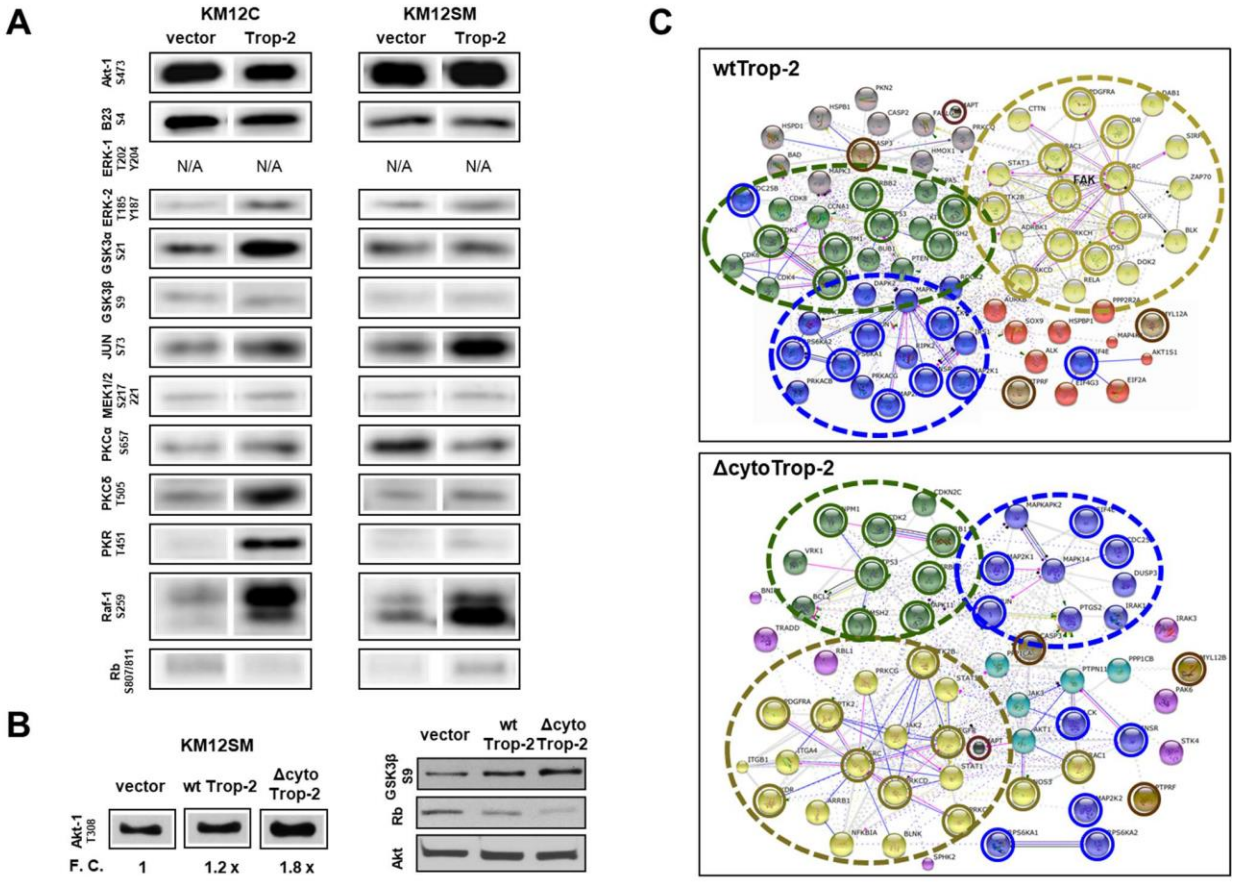


Fig. 35 : The Trop-2-driven signaling network

A large fraction of the 170 proteomic changes which were detected in the Trop-2 signalosome were found to be induced by both wt and ΔcytoTrop- 2 (Fig. 35C, Table S5D). In 23 instances, concordant variations of absolute amounts and/or extent of activatory phosphorylation of signaling proteins were recorded, suggesting these to be core Trop-2-dependent signaling events (Table S5D). ΔcytoTrop-2 downregulated absolute amounts and phosphorylation/activation of the p38/MAPK14 cell death inducers and downstream effectors IRAK-1, IRAK-3, TRADD, JAK3, STAT-1 (S727), JAK2 (Y1007-1008) (Table S5).

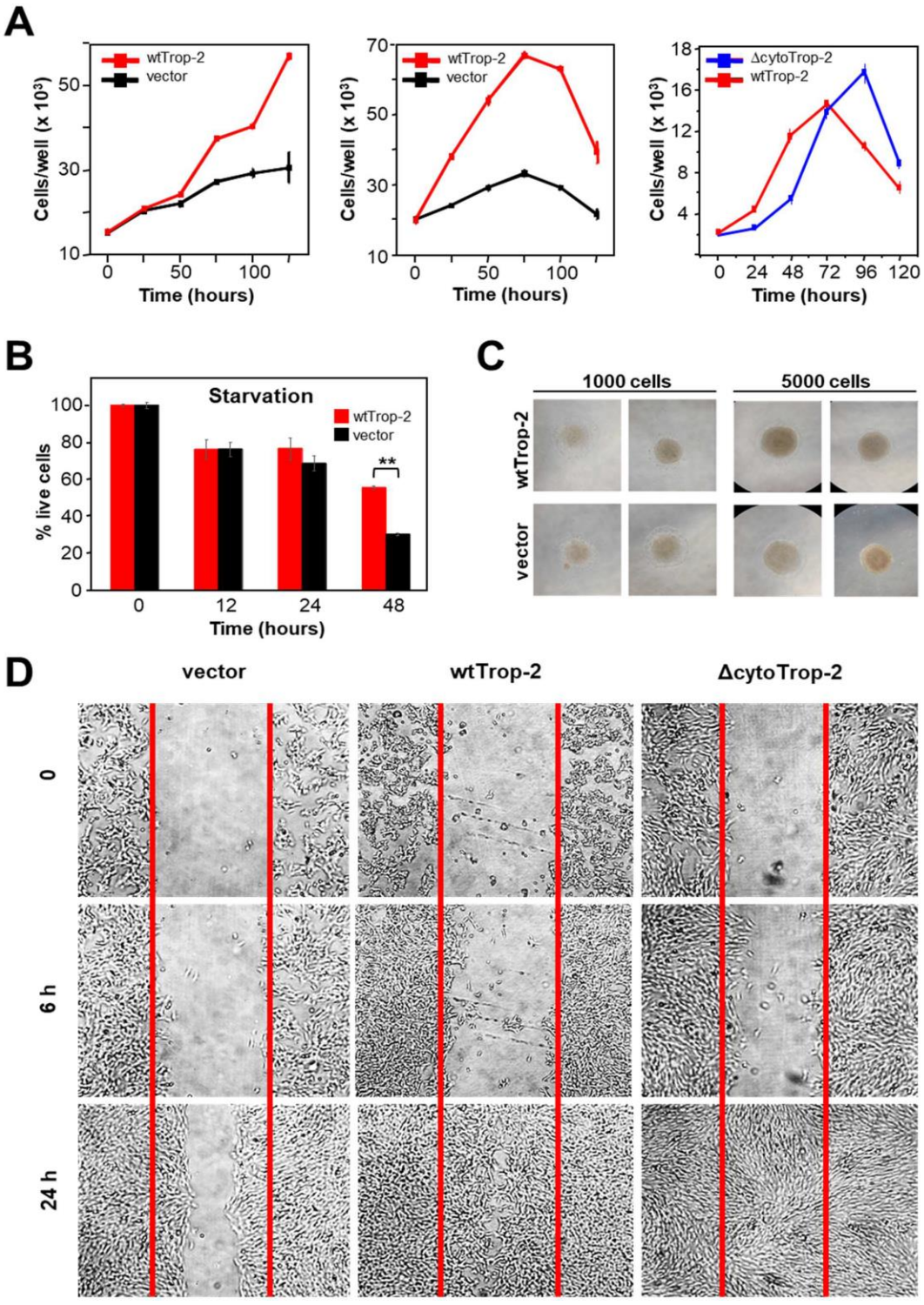


Fig. 36 : Trop-2 enhances cell survival and drives migration

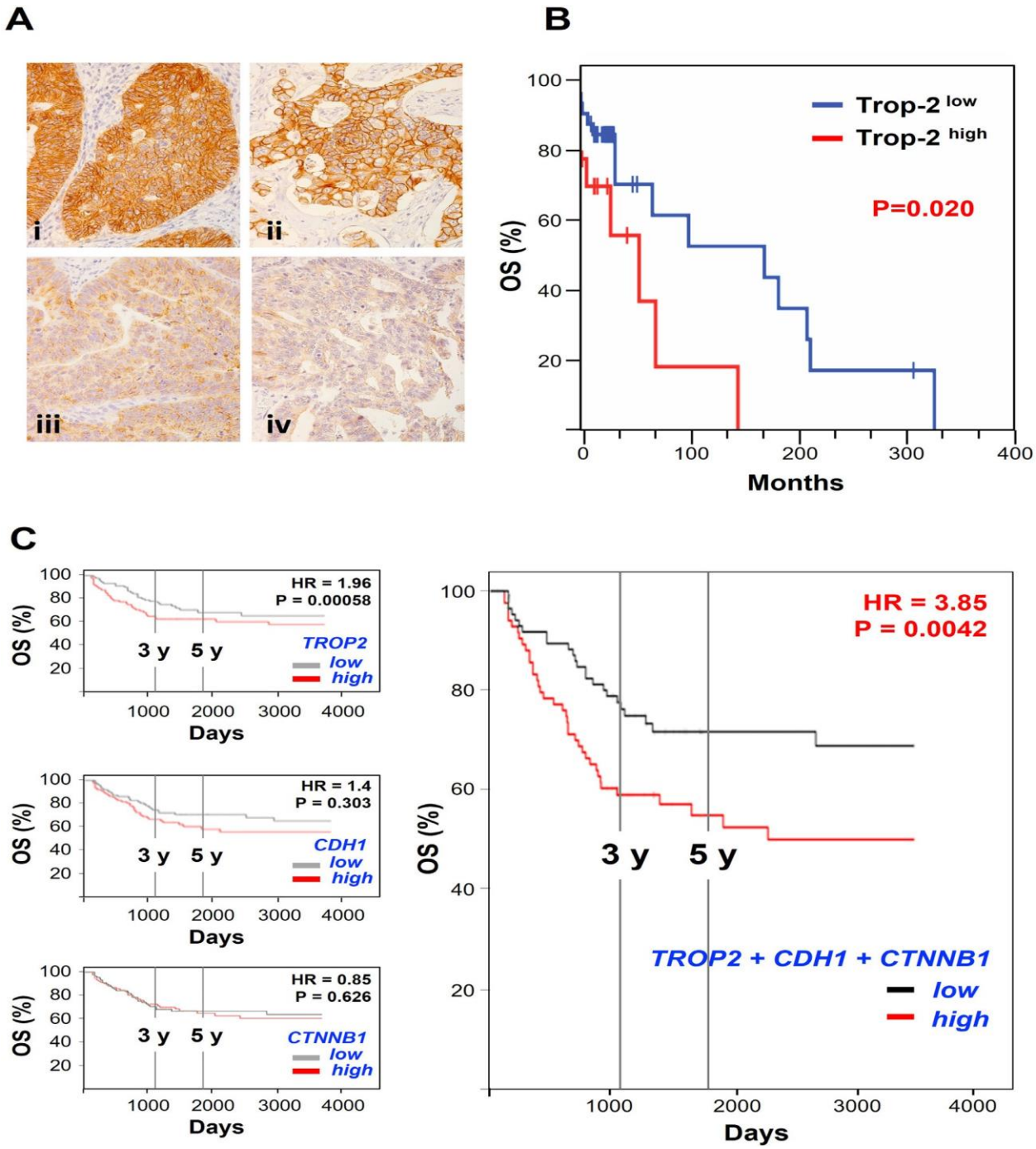


Fig. 37 : Impact of the Trop-2/E-cadherin/ β-catenin module on colon cancer patient survival

These findings, and the activation of the Ras/Raf/MEK, PTEN/Akt/GSK3, p38/MAPK14, and PKCε (S729) pathways, indicated impact of Trop-2 on pro-survival signaling [23]. A pro-survival role of Trop-2 was validated in in vitro functional assays in a range of Trop-2-responsive

cell systems (KM12SM colon cancer cells, IGROV-1 ovarian cancer cells, MTE4-14 murine thymic cells; Fig. 37A-B) [1, 8]. Serum-deprived IGROV-1 and KM12SM/Trop-2 transfectants reached much higher saturation values and survived markedly longer than vector- alone control cells (Fig. 36A-B). Activated ΔcytoTrop-2 had a larger impact than wtTrop-2 on cell culture saturation density and survival over time (Fig. 36A). Of note, ΔcytoTrop-2 induced slower growth of transfected cells than wtTrop-2 up to 48 hours from seeding, supporting a model of a distinct control of Trop-2 on cell growth versus survival [7], as well as versus invasion (Fig. 30B). This was validated in 3D models of metastatic colon cancer, as HCT116 cell spheroids [401] were not appreciably induced to grow by wtTrop-2 (Fig. 36C).

Increased cell migration underlies metastatic spreading [23, 318]. In wound- healing assays, wtTrop-2 induced effective wound closing (22% of wound area coverage at 6 h versus 14% in parental cells; 78% versus 17% at 24 h in parental cells). ΔcytoTrop-2 further increased cell migration ability, with complete wound closure at 24 h after seeding (Fig. 36D).

h) Selective pressure for Trop-2 expression in metastatic disease

Our findings predicted a critical impact of the Trop-2-driven pro- metastatic program in colon cancer. This was assessed in multiple case series of cancer patients. Proof of principle IHC analysis for protein expression showed progressive induction of Trop-2 from the Trop-2-negative normal colon mucosa, to colon cancer, to metastases, which showed strikingly high expression of Trop-2 (Fig. 38C). Trios of normal tissue, primary tumor, and metastasis from individual patients bearing colon, cancer was then analyzed by cDNA dot blot analysis. This showed overexpression of TROP2 in metastasis in 73% of cases (n = 11) (Fig. 38A, Table S6A). Such selective pressure for higher expression of TROP2 in metastasis, supported a driving role of TROP2 in metastatic diffusion. This applied also to organs and tissues where normal cells do not express TROP2, indicating reshaping of gene expression control networks according to a common malignant progression trajectory.

We verified the relevance of such findings in independent cancer patient case series. Analysis of 9 colon cancers and matching metastases indicated upregulation of TROP2 in

metastatic lesions (Fig. 38B). Such impact of a global TROP2 overexpression in metastases was challenged by IHC analysis of a validation case-series of colon normal tissues, primary cancers and matching metastases (n = 48) (Table S6A-B). This revealed that Trop-2 was overexpressed in metastases (Fig. 38D, Table S6A-B), with probability estimates (95% confidence intervals) for higher fractions of metastatic IHC- positive cells than in primary tumors of 88% (CI = 78-97%) of colon carcinomas.

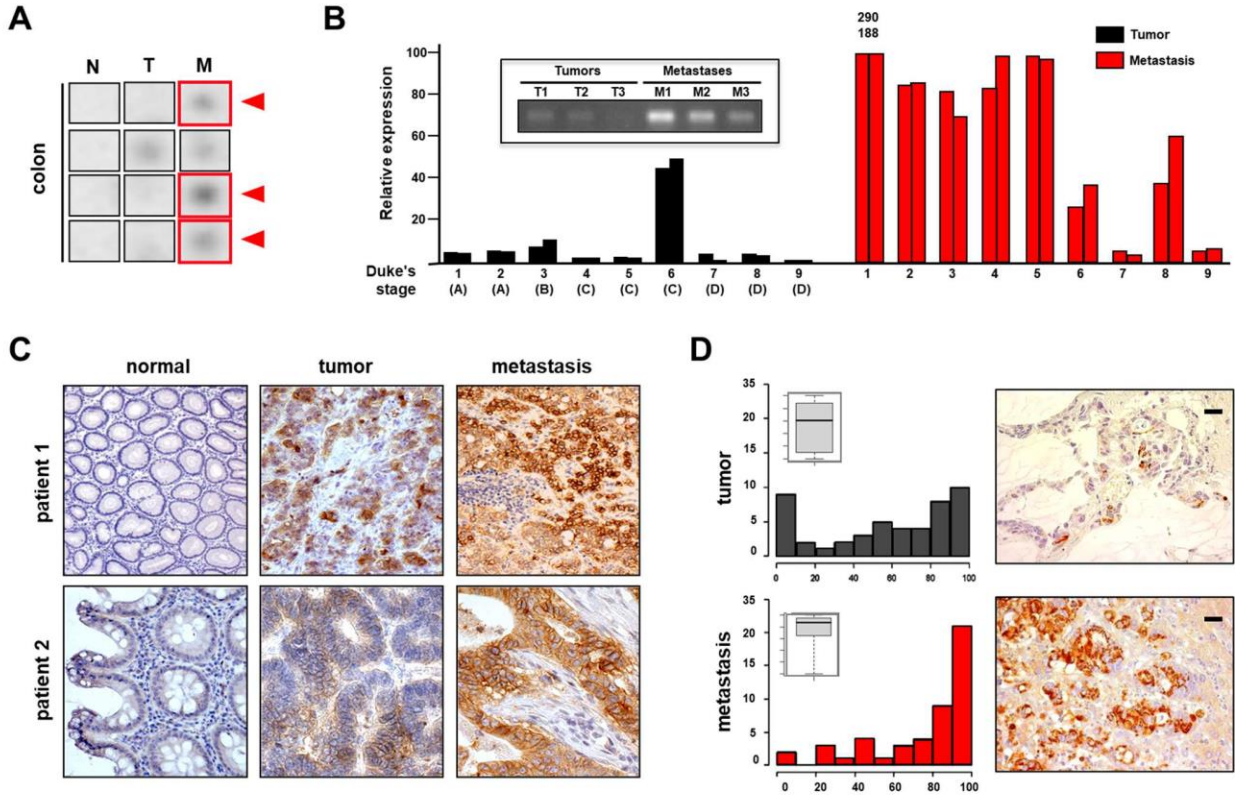


Fig. 38 : Trop-2 overexpression associates to metastasis in colon cancer patients

i) Trop-2 is a predictor of survival in colon cancer patients

We thus went on for assessing Trop-2 impact on colon cancer malignant progression and disease outcome. This was assessed in two independent colon cancer case series. A first patient

population (n=80) was examined, with median age of 72 years (range, 30-88) and heterogeneous tumor grading. Evaluable dissected samples indicated lymph-node metastasis at diagnosis in 32 patients, whereas 37 patients were metastasis-free. IHC analysis (Fig. 37A) showed expression of the Trop-2 protein in 65/80 (81.2%) cases. The fraction of Trop-2 expressing cells in the positive cases was up to 100%, with a mean ± S.E. of 41.6% ± 4.0. Among the 80 tumors, 40 (50.0%) exhibited Trop-2 overexpression (Trop-2high). Trop-2high was significantly associated with metastasis to lymph-nodes (P = 0.047) (Table S6C). No significant association was detected between Trop-2 expression and other clinicopathologic variables (age, sex, degree of tumor differentiation, tumor staging, tumor invasion).

To quantify the impact of Trop-2 expression on patient survival, a second series of 53 colon cancer patients was surveyed for an extended follow-up time (≈400 months). ROC analysis was applied to compute a Trop-2 positivity threshold according to the 0, 1-criterion. The optimal cut-off parameter for Trop-2 positive expression was 88%. Kaplan-Meier curves showed above-threshold expression as significantly associated with poor OS, with death of 57.1% of patients. Only 37.8% of Trop-2low cases encountered death (P = 0.020) (Fig. 37B, Table S6C).

We then went on to assess the impact of coordinate expression of Trop-2 (TROP2, TACSTD2), E-cadherin (CDH1), β-catenin (CTNNB1) on disease outcome and patient survival. Next-generation sequencing analysis of tumor transcriptomes of 597 cases of colon cancer in the TCGA dataset (www.proteinatlas.org) confirmed a negative prognostic impact of TROP2 mRNA overexpression on patient survival (P =3.67e-3). OS was then analyzed in colon cancer patients (genomics.jefferson.edu/proggene) for coordinate expression of TROP2, CDH1, CTNNB1 transcripts by DNA microarray analysis (160 cases, with 11 years of follow-up). Significant prognostic impact on patient OS for single-determinant analysis was revealed for TROP2 (HR = 1.96; P = 0.00058), with no significant impact of CDH1, CTNNB1 (Fig. 37C). Remarkably, though, coordinate assessment of TROP2, CDH1, CTNNB1 provided much higher estimates of relative risk on disease outcome and patient survival than any individual parameter alone (HR = 3.85; P=0.0042) (Fig. 37C), consistent with a coordinate impact of the Trop-2-driven module on cancer relapse.

3. Discussion and conclusion

In this work we identified Trop-2 as a unique upregulated gene across colon cancer metastasis models and showed for the first time that Trop-2 induces ADAM10 activation for E-cadherin cleavage and inactivation.

E-cadherin has a primary role in establishing cell-cell junctions and maintaining normal tissue architecture, whilst β-catenin takes part to adherent junction complexes and mediates the interactions of E-cadherin with the cytoskeleton by binding to α-catenin [402]. β-actin operates as a scaffold for the assembly of a Ras/Raf/ERK/ezrin complex, which is a key regulator of PI3K/Akt signaling through PI-phosphorylation. Consequently, E-cadherin requires anchoring to β-actin for function, and inhibition of tumor progression, through retention of β-catenin [402]. On the opposite side, release of β-catenin leads to translocation to the nucleus, and activation as a transcription factor on target genes, including those that promote cell proliferation, differentiation, migration and angiogenesis [402]. In colon cancer experimental models, we revealed a gradient of metastatic competence, with the lowest levels of nuclear β-catenin observed in the Trop- 2-nil non-metastatic KM12C colon cancer cells, higher levels measured in the Trop-2-prone pro-metastatic KM12SM cells, and the highest levels reached in the metastatic wtTrop-2 and ΔcytoTrop-2 KM12SM cells.

This paralleled the ability of wtTrop-2 and ΔcytoTrop-2 to induce ADAM10 cleavage of E-cadherin and disanchoring from the cytoskeleton, which then led to a corresponding loss of cell-cell aggregation. The transcriptional capacity of the released β-catenin was demonstrated in β-catenin-RE-GFP reporter assays in vitro. This was paralleled by massive induction of β-catenin transcriptional target genes in wtTrop-2 and ΔcytoTrop-2 -driven metastases in vivo. No trace of these changes was found in control metastasis, indicating that wtTrop-2 triggers a unique mechanism of metastatic dissemination.

E-cadherin loss has been widely considered to be required for EMT-dependent tissue invasion [403]. However, our findings showed no transcriptional downregulation of the CDH1 gene in Trop-2-activated cells, nor expression of EMT transcription factors that suppress the transcription of epithelial differentiation markers. Instances of incomplete/graded EMT have been recognized [318, 326, 388, 403]. Our findings add a novel turn to these scenarios, whereby Trop-2 induced E-cadherin cleavage and inactivation, as mediated by ADAM10. Unsurprisingly, this

associated to retention of epithelial differentiation biomarkers, high levels of β-catenin and of β-catenin transcriptional targets.

Our findings thus appear to recapitulate frequent surgical pathology assessments, that show inactivation/disarray of E-cadherin in aggressive tumors [322], for release of transcription-versed β-catenin into the cytoplasm of cancer cells and association to tumor progression [395], in the apparent absence of loss of epithelial differentiation. IHC analysis of distinct colon cancer case series showed that up to 88% of the tumors exhibited Trop-2 overexpression, and that overexpressed Trop-2 was significantly associated with metastasis to lymph nodes. An extended follow-up (≈400 months) of metastatic colon cancer patients, showed that high Trop-2 expression correlated with poor OS, and with fatal events in 57.1% of highly expressing cases.

Corresponding findings were obtained by transcriptomic analysis of independent colon cancer case series, where overexpression of Trop-2, β-catenin and E-cadherin showed heavy impact on patient survival. Notably, this impact was not matched by any of the risk factors taken individually, suggesting a pivotal pathological role of the Trop-2-driven module on colon cancer progression. Of note, the anti-Trop-2 mAb Sacituzumab govitecan hziy was shown to be effective in metastatic breast cancer patients [339, 340, 369, 370] and was granted accelerated FDA approval (www.fda.gov/drugs/drug-approvals-and-databases/drug-trial-snapshot-trodelvy). Our findings provide ground for potential major relevance of Trop-2 as a target in metastatic colon cancer.

V. LIST OF SUPPLEMENTARY TABLES

All TABLES ARE IN EXCEL FILES

ARTICLE I: TROP-2 CLEAVAGE BY ADAM10 IS AN ACTIVATOR SWITCH FOR CANCER GROWTH AND METASTASIS

Supplementary Table S1 E1 staining of Trop-2 in transformed cell lines.

Supplementary Table S2 Identification of Trop-2 cleavage site by mass-spectrometry analysis.

Supplementary Table S3 Identification of ADAM10 as a Trop-2 interactor by mass spectrometry analysis of Trop-2 immunoprecipitates.

Supplementary Table S4 Proteomic chip analysis of cleavage-less Trop-2 signaling.

Supplementary Movie S1 Structure of the Trop-2 thyroglobulin domain.

Supplementary Movie S2 Dynamic colocalization of Trop-2 and ADAM10.

ARTICLE II: TROP-2 INDUCES ADAM10-MEDIATED CLEAVAGE OF E-CADHERIN AND DRIVES EMT-LESS METASTASIS IN COLON CANCER

Supplementary Table S1 Transcriptome profiling of metastatic colon cancer cells.

Supplementary Table S2 In vivo pro-metastatic ability of Trop-2-expressing colon cancer cells.

Supplementary Table S3 NGS transcriptome analysis of Trop-2-expressing cancer cells in preclinical models.

Supplementary Table S4 Trop-2-dependent activation of β-catenin in colon cancer cells.

Supplementary Table S5 Proteomic analysis of Trop-2 signaling in colon cancer cells.

Supplementary Table S6 Impact of Trop-2 expression on relapse and overall survival of patients with colon cancer.

VI. REFERENCES

1. Trerotola, M., et al., *Trop-2 cleavage by ADAM10 is an activator switch for cancer growth and metastasis.* Neoplasia, 2021. **23**(4): p. 415-428.
2. Guerra, E., et al., *Trop-2 induces ADAM10-mediated cleavage of E-cadherin and drives EMT-less metastasis in colon cancer.* Neoplasia, 2021. **23**(9): p. 898-911.
3. Bonaccio, M., et al., *Adherence to the Mediterranean diet is associated with lower platelet and leukocyte counts: results from the Moli-sani study.* Blood, 2014. **123**(19): p. 3037-44.
4. den Hoed, M., et al., *Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders.* Nat Genet, 2013. **45**(6): p. 621-31.
5. Collaboration, N.C.D.R.F., *Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants.* Lancet, 2016. **387**(10026): p. 1377-1396.
6. Avendaño, C. and J.C. Menéndez, *Comprar Medicinal Chemistry of Anticancer Drugs*. 2008: Elsevier B.V.
7. Guerra, E., et al., *The Trop-2 signalling network in cancer growth.* Oncogene, 2013. **32**(12): p. 1594-600.
8. Trerotola, M., et al., *Upregulation of Trop-2 quantitatively stimulates human cancer growth.* Oncogene, 2013. **32**(2): p. 222-33.
9. Loibl, S., et al., *Breast cancer.* Lancet, 2021. **397**(10286): p. 1750-1769.
10. Pietras, R.J., et al., *HER-2 tyrosine kinase pathway targets estrogen receptor and promotes hormone-independent growth in human breast cancer cells.* Oncogene, 1995. **10**(12): p. 2435-46.
11. Kubo, M., et al., *Hedgehog signaling pathway is a new therapeutic target for patients with breast cancer.* Cancer Res, 2004. **64**(17): p. 6071-4.
12. Recio-Boiles, A. and B. Cagir, *Colon Cancer*, in *StatPearls*. 2022: Treasure Island (FL).
13. Malekpour, H., et al., *Gene expression analysis of colon high-grade dysplasia revealed new molecular mechanism of disease.* Gastroenterol Hepatol Bed Bench, 2018. **11**(Suppl 1): p. S111-S117.
14. Allen, J. and C.L. Sears, *Impact of the gut microbiome on the genome and epigenome of colon epithelial cells: contributions to colorectal cancer development.* Genome Med, 2019. **11**(1): p. 11.
15. Bien, S.A., et al., *Correction to: Genetic variant predictors of gene expression provide new insight into risk of colorectal cancer.* Hum Genet, 2019. **138**(7): p. 789-791.
16. Rawla, P., *Epidemiology of Prostate Cancer.* World J Oncol, 2019. **10**(2): p. 63-89.
17. Datta, K., et al., *Mechanism of lymph node metastasis in prostate cancer.* Future Oncol, 2010. **6**(5): p. 823-36.
18. Logothetis, C.J. and S.H. Lin, *Osteoblasts in prostate cancer metastasis to bone.* Nat Rev Cancer, 2005. **5**(1): p. 21-8.
19. Guise, T.A., et al., *Basic mechanisms responsible for osteolytic and osteoblastic bone metastases.* Clin Cancer Res, 2006. **12**(20 Pt 2): p. 6213s-6216s.
20. Body, J.J., S. Casimiro, and L. Costa, *Targeting bone metastases in prostate cancer: improving clinical outcome.* Nat Rev Urol, 2015. **12**(6): p. 340-56.

21. Wang, G., et al., *Genetics and biology of prostate cancer.* Genes Dev, 2018. **32**(17-18): p. 1105-1140.
22. Polireddy, K. and Q. Chen, *Cancer of the Pancreas: Molecular Pathways and Current Advancement in Treatment.* J Cancer, 2016. **7**(11): p. 1497-514.
23. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
24. Herlyn, M., et al., *Colorectal carcinoma-specific antigen: detection by means of monoclonal antibodies.* Proc Natl Acad Sci U S A, 1979. **76**(3): p. 1438-42.
25. Calabrese, G., et al., *Assignment of TACSTD1 (alias TROP1, M4S1) to human chromosome 2p21 and refinement of mapping of TACSTD2 (alias TROP2, M1S1) to human chromosome 1p32 by in situ hybridization.* Cytogenet Cell Genet, 2001. **92**(1-2): p. 164-5.
26. Went, P.T., et al., *Frequent EpCam protein expression in human carcinomas.* Hum Pathol, 2004. **35**(1): p. 122-8.
27. Simon, M., et al., *Epithelial cell adhesion molecule-targeted drug delivery for cancer therapy.* Expert Opin Drug Deliv, 2013. **10**(4): p. 451-68.
28. Schnell, U., V. Cirulli, and B.N. Giepmans, *EpCAM: structure and function in health and disease.* Biochim Biophys Acta, 2013. **1828**(8): p. 1989-2001.
29. Schon, M.P., et al., *Biochemical and immunological characterization of the human carcinoma-associated antigen MH 99/KS 1/4.* Int J Cancer, 1993. **55**(6): p. 988-95.
30. Chong, J.M. and D.W. Speicher, *Determination of disulfide bond assignments and N-glycosylation sites of the human gastrointestinal carcinoma antigen GA733-2 (CO17-1A, EGP, KS1-4, KSA, and Ep-CAM).* J Biol Chem, 2001. **276**(8): p. 5804-13.
31. Thampoe, I.J., J.S. Ng, and K.O. Lloyd, *Biochemical analysis of a human epithelial surface antigen: differential cell expression and processing.* Arch Biochem Biophys, 1988. **267**(1): p. 342-52.
32. Munz, M., et al., *Glycosylation is crucial for stability of tumour and cancer stem cell antigen EpCAM.* Front Biosci, 2008. **13**: p. 5195-201.
33. Balzar, M., et al., *Cytoplasmic tail regulates the intercellular adhesion function of the epithelial cell adhesion molecule.* Mol Cell Biol, 1998. **18**(8): p. 4833-43.
34. Litvinov, S.V., et al., *Ep-CAM: a human epithelial antigen is a homophilic cell-cell adhesion molecule.* J Cell Biol, 1994. **125**(2): p. 437-46.
35. Balzar, M., et al., *Epidermal growth factor-like repeats mediate lateral and reciprocal interactions of Ep-CAM molecules in homophilic adhesions.* Mol Cell Biol, 2001. **21**(7): p. 2570-80.
36. Litvinov, S.V., et al., *Epithelial cell adhesion molecule (Ep-CAM) modulates cell-cell interactions mediated by classic cadherins.* J Cell Biol, 1997. **139**(5): p. 1337-48.
37. Paschkowsky, S., et al., *The discovery of proteases and intramembrane proteolysis (1).* Biochem Cell Biol, 2019. **97**(3): p. 265-269.
38. Hachmeister, M., et al., *Regulated intramembrane proteolysis and degradation of murine epithelial cell adhesion molecule mEpCAM.* PLoS One, 2013. **8**(8): p. e71836.
39. Maetzel, D., et al., *Nuclear signalling by tumour-associated antigen EpCAM.* Nat Cell Biol, 2009. **11**(2): p. 162-71.
40. Tsaktanis, T., et al., *Cleavage and cell adhesion properties of human epithelial cell adhesion molecule (HEPCAM).* J Biol Chem, 2015. **290**(40): p. 24574-91.

41. Schnell, U., J. Kuipers, and B.N. Giepmans, *EpCAM proteolysis: new fragments with distinct functions?* Biosci Rep, 2013. **33**(2): p. e00030.
42. Sadeghi, S., Z. Hojati, and H. Tabatabaeian, *Cooverexpression of EpCAM and c-myc genes in malignant breast tumours.* J Genet, 2017. **96**(1): p. 109-118.
43. Denzel, S., et al., *Initial activation of EpCAM cleavage via cell-to-cell contact.* BMC Cancer, 2009. **9**: p. 402.
44. Martowicz, A., et al., *EpCAM overexpression prolongs proliferative capacity of primary human breast epithelial cells and supports hyperplastic growth.* Mol Cancer, 2013. **12**: p. 56.
45. Fagotto, F. and B.M. Gumbiner, *Cell contact-dependent signaling.* Dev Biol, 1996. **180**(2): p. 445-54.
46. Zoller, M., *Tetraspanins: push and pull in suppressing and promoting metastasis.* Nat Rev Cancer, 2009. **9**(1): p. 40-55.
47. Hemler, M.E., *Tetraspanin functions and associated microdomains.* Nat Rev Mol Cell Biol, 2005. **6**(10): p. 801-11.
48. Arduise, C., et al., *Tetraspanins regulate ADAM10-mediated cleavage of TNF-alpha and epidermal growth factor.* J Immunol, 2008. **181**(10): p. 7002-13.
49. Le Naour, F., et al., *Membrane microdomains and proteomics: lessons from tetraspanin microdomains and comparison with lipid rafts.* Proteomics, 2006. **6**(24): p. 6447-54.
50. Le Naour, F., et al., *Profiling of the tetraspanin web of human colon cancer cells.* Mol Cell Proteomics, 2006. **5**(5): p. 845-57.
51. Schmidt, D.S., et al., *CD44 variant isoforms associate with tetraspanins and EpCAM.* Exp Cell Res, 2004. **297**(2): p. 329-47.
52. Ponta, H., L. Sherman, and P.A. Herrlich, *CD44: from adhesion molecules to signalling regulators.* Nat Rev Mol Cell Biol, 2003. **4**(1): p. 33-45.
53. Ladwein, M., et al., *The cell-cell adhesion molecule EpCAM interacts directly with the tight junction protein claudin-7.* Exp Cell Res, 2005. **309**(2): p. 345-57.
54. Nubel, T., et al., *Claudin-7 regulates EpCAM-mediated functions in tumor progression.* Mol Cancer Res, 2009. **7**(3): p. 285-99.
55. Xu, D., C. Sharma, and M.E. Hemler, *Tetraspanin12 regulates ADAM10-dependent cleavage of amyloid precursor protein.* FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 2009. **23**(11): p. 3674-3681.
56. Gottlinger, H.G., et al., *The epithelial cell surface antigen 17-1A, a target for antibody-mediated tumor therapy: its biochemical nature, tissue distribution and recognition by different monoclonal antibodies.* Int J Cancer, 1986. **38**(1): p. 47-53.
57. Schmelzer, E. and L.M. Reid, *EpCAM expression in normal, non-pathological tissues.* Front Biosci, 2008. **13**: p. 3096-100.
58. Patriarca, C., et al., *Epithelial cell adhesion molecule expression (CD326) in cancer: a short review.* Cancer Treat Rev, 2012. **38**(1): p. 68-75.
59. Balzar, M., et al., *The structural analysis of adhesions mediated by Ep-CAM.* Exp Cell Res, 1999. **246**(1): p. 108-21.
60. Wu, C.J., et al., *Epithelial cell adhesion molecule (EpCAM) regulates claudin dynamics and tight junctions.* J Biol Chem, 2013. **288**(17): p. 12253-68.
61. Lei, Z., et al., *EpCAM contributes to formation of functional tight junction in the intestinal epithelium by recruiting claudin proteins.* Dev Biol, 2012. **371**(2): p. 136-45.

62. Gaiser, M.R., et al., *Cancer-associated epithelial cell adhesion molecule (EpCAM; CD326) enables epidermal Langerhans cell motility and migration in vivo.* Proc Natl Acad Sci U S A, 2012. **109**(15): p. E889-97.
63. Maghzal, N., et al., *EpCAM controls actomyosin contractility and cell adhesion by direct inhibition of PKC.* Dev Cell, 2013. **27**(3): p. 263-77.
64. Winter, M.J., et al., *Expression of Ep-CAM shifts the state of cadherin-mediated adhesions from strong to weak.* Exp Cell Res, 2003. **285**(1): p. 50-8.
65. Pavsic, M., et al., *Crystal structure and its bearing towards an understanding of key biological functions of EpCAM.* Nat Commun, 2014. **5**: p. 4764.
66. Gaber, A., et al., *EpCAM homo-oligomerization is not the basis for its role in cell-cell adhesion.* Sci Rep, 2018. **8**(1): p. 13269.
67. Fornaro, M., et al., *Cloning of the gene encoding Trop-2, a cell-surface glycoprotein expressed by human carcinomas.* Int J Cancer, 1995. **62**(5): p. 610-8.
68. Varki, N.M., R.A. Reisfeld, and L.E. Walker, *Antigens associated with a human lung adenocarcinoma defined by monoclonal antibodies.* Cancer Res, 1984. **44**(2): p. 681-7.
69. Sarrach, S., et al., *Spatiotemporal patterning of EpCAM is important for murine embryonic endo- and mesodermal differentiation.* Sci Rep, 2018. **8**(1): p. 1801.
70. Spizzo, G., et al., *Prognostic significance of Ep-CAM AND Her-2/neu overexpression in invasive breast cancer.* Int J Cancer, 2002. **98**(6): p. 883-8.
71. Spizzo, G., et al., *High Ep-CAM expression is associated with poor prognosis in node-positive breast cancer.* Breast Cancer Res Treat, 2004. **86**(3): p. 207-13.
72. Varga, M., et al., *Overexpression of epithelial cell adhesion molecule antigen in gallbladder carcinoma is an independent marker for poor survival.* Clin Cancer Res, 2004. **10**(9): p. 3131-6.
73. Brunner, A., et al., *EpCAM overexpression is associated with high-grade urothelial carcinoma in the renal pelvis.* Anticancer Res, 2008. **28**(1A): p. 125-8.
74. Scheunemann, P., et al., *Occult tumor cells in lymph nodes as a predictor for tumor relapse in pancreatic adenocarcinoma.* Langenbecks Arch Surg, 2008. **393**(3): p. 359-65.
75. Fong, D., et al., *Ep-CAM expression in pancreatic and ampullary carcinomas: frequency and prognostic relevance.* J Clin Pathol, 2008. **61**(1): p. 31-5.
76. Seligson, D.B., et al., *Epithelial cell adhesion molecule (KSA) expression: pathobiology and its role as an independent predictor of survival in renal cell carcinoma.* Clin Cancer Res, 2004. **10**(8): p. 2659-69.
77. Ralhan, R., et al., *EpCAM nuclear localization identifies aggressive thyroid cancer and is a marker for poor prognosis.* BMC Cancer, 2010. **10**: p. 331.
78. van der Gun, B.T., et al., *EpCAM in carcinogenesis: the good, the bad or the ugly.* Carcinogenesis, 2010. **31**(11): p. 1913-21.
79. Melchers, L.J., et al., *Lack of claudin-7 is a strong predictor of regional recurrence in oral and oropharyngeal squamous cell carcinoma.* Oral Oncol, 2013. **49**(10): p. 998-1005.
80. Ralhan, R., et al., *Nuclear and cytoplasmic accumulation of Ep-ICD is frequently detected in human epithelial cancers.* PLoS One, 2010. **5**(11): p. e14130.
81. Kunavisarut, T., et al., *Immunohistochemical analysis based Ep-ICD subcellular localization index (ESLI) is a novel marker for metastatic papillary thyroid microcarcinoma.* BMC Cancer, 2012. **12**: p. 523.

82. Schnell, U., et al., *Immunolabeling artifacts and the need for live-cell imaging.* Nat Methods, 2012. **9**(2): p. 152-8.
83. Sivagnanam, M., et al., *Identification of EpCAM as the gene for congenital tufting enteropathy.* Gastroenterology, 2008. **135**(2): p. 429-37.
84. Goulet, O., et al., *Intestinal epithelial dysplasia (tufting enteropathy).* Orphanet J Rare Dis, 2007. **2**: p. 20.
85. Guerra, E., et al., *mTrop1/Epcam knockout mice develop congenital tufting enteropathy through dysregulation of intestinal E-cadherin/beta-catenin.* PLoS One, 2012. **7**(11): p. e49302.
86. Slanchev, K., et al., *The epithelial cell adhesion molecule EpCAM is required for epithelial morphogenesis and integrity during zebrafish epiboly and skin development.* PLoS Genet, 2009. **5**(7): p. e1000563.
87. Nagao, K., et al., *Abnormal placental development and early embryonic lethality in EpCAM-null mice.* PLoS One, 2009. **4**(12): p. e8543.
88. Patey, N., et al., *Distribution of cell adhesion molecules in infants with intestinal epithelial dysplasia (tufting enteropathy).* Gastroenterology, 1997. **113**(3): p. 833-43.
89. Al-Mayouf, S.M., et al., *Tufting enteropathy and chronic arthritis: a newly recognized association with a novel EpCAM gene mutation.* J Pediatr Gastroenterol Nutr, 2009. **49**(5): p. 642-4.
90. Schnell, U., et al., *Absence of cell-surface EpCAM in congenital tufting enteropathy.* Hum Mol Genet, 2013. **22**(13): p. 2566-71.
91. Balzar, M., et al., *The biology of the 17-1A antigen (Ep-CAM).* J Mol Med (Berl), 1999. **77**(10): p. 699-712.
92. Szala, S., et al., *Molecular cloning of cDNA for the human tumor-associated antigen CO-029 and identification of related transmembrane antigens.* Proc Natl Acad Sci U S A, 1990. **87**(17): p. 6833-7.
93. Linnenbach, A.J., et al., *Retroposition in a family of carcinoma-associated antigen genes.* Mol Cell Biol, 1993. **13**(3): p. 1507-15.
94. El Sewedy, T., M. Fornaro, and S. Alberti, *Cloning of the murine TROP2 gene: conservation of a PIP2-binding sequence in the cytoplasmic domain of TROP-2.* Int J Cancer, 1998. **75**(2): p. 324-30.
95. Basu, A., D.M. Goldenberg, and R. Stein, *The epithelial/carcinoma antigen EGP-1, recognized by monoclonal antibody RS7-3G11, is phosphorylated on serine 303.* Int J Cancer, 1995. **62**(4): p. 472-9.
96. Cubas, R., et al., *Trop2: a possible therapeutic target for late stage epithelial carcinomas.* Biochim Biophys Acta, 2009. **1796**(2): p. 309-14.
97. Pavsic, M., et al., *The cytosolic tail of the tumor marker protein Trop2--a structural switch triggered by phosphorylation.* Sci Rep, 2015. **5**: p. 10324.
98. Linnenbach, A.J., et al., *Sequence investigation of the major gastrointestinal tumor-associated antigen gene family, GA733.* Proc Natl Acad Sci U S A, 1989. **86**(1): p. 27-31.
99. Olsen, J.V., et al., *Global, in vivo, and site-specific phosphorylation dynamics in signaling networks.* Cell, 2006. **127**(3): p. 635-48.
100. Bignotti, E., et al., *Trop-2 protein overexpression is an independent marker for predicting disease recurrence in endometrioid endometrial carcinoma.* BMC Clin Pathol, 2012. **12**: p. 22.

101. Ciccarelli, F.D., A. Acciarito, and S. Alberti, *Large and diverse numbers of human diseases with HIKE mutations.* Hum Mol Genet, 2000. **9**(6): p. 1001-7.
102. Lin, J.C., et al., *TROP2 is epigenetically inactivated and modulates IGF-1R signalling in lung adenocarcinoma.* EMBO Mol Med, 2012. **4**(6): p. 472-85.
103. Alberti, S., et al., *TROP2 is a major determinant of growth and metastatic spreading of human cancer.* Journal of Clinical Oncology, 2007. **25**(18_suppl): p. 10510-10510.
104. Vidmar, T., M. Pavsic, and B. Lenarcic, *Biochemical and preliminary X-ray characterization of the tumor-associated calcium signal transducer 2 (Trop2) ectodomain.* Protein Expr Purif, 2013. **91**(1): p. 69-76.
105. Mori, Y., et al., *Trophoblast cell surface antigen 2 (Trop-2) phosphorylation by protein kinase C alpha/delta (PKCalpha/delta) enhances cell motility.* J Biol Chem, 2019. **294**(30): p. 11513-11524.
106. Lenart, S., et al., *Trop2: Jack of All Trades, Master of None.* Cancers (Basel), 2020. **12**(11).
107. Lipinski, M., et al., *Human trophoblast cell-surface antigens defined by monoclonal antibodies.* Proc Natl Acad Sci U S A, 1981. **78**(8): p. 5147-50.
108. Stepan, L.P., et al., *Expression of Trop2 cell surface glycoprotein in normal and tumor tissues: potential implications as a cancer therapeutic target.* J Histochem Cytochem, 2011. **59**(7): p. 701-10.
109. Tsujikawa, M., et al., *Identification of the gene responsible for gelatinous drop-like corneal dystrophy.* Nature Genetics, 1999. **21**(4): p. 420-423.
110. Lu, J., et al., *Identification of FGF10 targets in the embryonic lung epithelium during bud morphogenesis.* J Biol Chem, 2005. **280**(6): p. 4834-41.
111. McDougall, A.R., et al., *The oncogene Trop2 regulates fetal lung cell proliferation.* Am J Physiol Lung Cell Mol Physiol, 2011. **301**(4): p. L478-89.
112. Tsukahara, Y., M. Tanaka, and A. Miyajima, *TROP2 expressed in the trunk of the ureteric duct regulates branching morphogenesis during kidney development.* PLoS One, 2011. **6**(12): p. e28607.
113. Fernandez Vallone, V., et al., *Trop2 marks transient gastric fetal epithelium and adult regenerating cells after epithelial damage.* Development, 2016. **143**(9): p. 1452-63.
114. Mustata, R.C., et al., *Identification of Lgr5-independent spheroid-generating progenitors of the mouse fetal intestinal epithelium.* Cell Rep, 2013. **5**(2): p. 421-32.
115. Ripani, E., et al., *Human Trop-2 is a tumor-associated calcium signal transducer.* Int J Cancer, 1998. **76**(5): p. 671-6.
116. Tang, G., et al., *TROP2 increases growth and metastasis of human oral squamous cell carcinoma through activation of the PI3K/Akt signaling pathway.* Int J Mol Med, 2019. **44**(6): p. 2161-2170.
117. Yang, J., et al., *Trop2 regulates the proliferation and differentiation of murine compact-bone derived MSCs.* Int J Oncol, 2013. **43**(3): p. 859-867.
118. Guerra, E., et al., *Trop-2 Induces Tumor Growth Through AKT and Determines Sensitivity to AKT Inhibitors.* Clin Cancer Res, 2016. **22**(16): p. 4197-205.
119. Liu, T., et al., *Overexpression of TROP2 predicts poor prognosis of patients with cervical cancer and promotes the proliferation and invasion of cervical cancer cells by regulating ERK signaling pathway.* PLoS One, 2013. **8**(9): p. e75864.

120. Zhang, K., et al., *Loss of Trop2 causes ErbB3 activation through a neuregulin-1-dependent mechanism in the mesenchymal subtype of HNSCC.* Oncotarget, 2014. **5**(19): p. 9281-94.
121. Zhao, W., et al., *The role and molecular mechanism of Trop2 induced epithelial-mesenchymal transition through mediated beta-catenin in gastric cancer.* Cancer Med, 2019. **8**(3): p. 1135-1147.
122. Stoyanova, T., et al., *Regulated proteolysis of Trop2 drives epithelial hyperplasia and stem cell self-renewal via beta-catenin signaling.* Genes Dev, 2012. **26**(20): p. 2271-85.
123. Hou, J., et al., *TROP2 promotes the proliferation and metastasis of glioblastoma cells by activating the JAK2/STAT3 signaling pathway.* Oncol Rep, 2019. **41**(2): p. 753-764.
124. Webb, D.J., et al., *FAK–Src signalling through paxillin, ERK and MLCK regulates adhesion disassembly.* Nature Cell Biology, 2004. **6**(2): p. 154-161.
125. Pak, M.G., et al., *Significance of EpCAM and TROP2 expression in non-small cell lung cancer.* World J Surg Oncol, 2012. **10**: p. 53.
126. Sawanyawisuth, K., et al., *Suppression of trophoblast cell surface antigen 2 enhances proliferation and migration in liver fluke-associated cholangiocarcinoma.* Ann Hepatol, 2016. **15**(1): p. 71-81.
127. Kobayashi, H., et al., *Expression of the GA733 gene family and its relationship to prognosis in pulmonary adenocarcinoma.* Virchows Arch, 2010. **457**(1): p. 69-76.
128. Nakatsukasa, M., et al., *Tumor-associated calcium signal transducer 2 is required for the proper subcellular localization of claudin 1 and 7: implications in the pathogenesis of gelatinous drop-like corneal dystrophy.* Am J Pathol, 2010. **177**(3): p. 1344-55.
129. Sekhar, V., et al., *Infection with hepatitis C virus depends on TACSTD2, a regulator of claudin-1 and occludin highly downregulated in hepatocellular carcinoma.* PLoS Pathog, 2018. **14**(3): p. e1006916.
130. Trerotola, M., et al., *Trop-2 is up-regulated in invasive prostate cancer and displaces FAK from focal contacts.* Oncotarget, 2015. **6**(16): p. 14318-28.
131. Trerotola, M., et al., *Trop-2 promotes prostate cancer metastasis by modulating β(1) integrin functions.* Cancer research, 2013. **73**(10): p. 3155-3167.
132. Sin, S.T.K., et al., *TROP-2 exhibits tumor suppressive functions in cervical cancer by dual inhibition of IGF-1R and ALK signaling.* Gynecol Oncol, 2019. **152**(1): p. 185-193.
133. Alberti, S., *HIKE, a candidate protein binding site for PH domains, is a major regulatory region of Gbeta proteins.* Proteins, 1999. **35**(3): p. 360-3.
134. Guerra, E., et al., *Trop-2, Na+/K+ ATPase, CD9, PKCα, cofilin assemble a membrane signaling super-complex that drives colorectal cancer growth and invasion.* Oncogene, 2022.
135. Cubas, R., et al., *Trop2 expression contributes to tumor pathogenesis by activating the ERK MAPK pathway.* Mol Cancer, 2010. **9**: p. 253.
136. Shvartsur, A. and B. Bonavida, *Trop2 and its overexpression in cancers: regulation and clinical/therapeutic implications.* Genes Cancer, 2015. **6**(3-4): p. 84-105.
137. Keller, L., S. Werner, and K. Pantel, *Biology and clinical relevance of EpCAM.* Cell Stress, 2019. **3**(6): p. 165-180.
138. Martowicz, A., A. Seeber, and G. Untergasser, *The role of EpCAM in physiology and pathology of the epithelium.* Histol Histopathol, 2016. **31**(4): p. 349-55.

139. Herreros-Pomares, A., et al., *EpCAM duality becomes this molecule in a new Dr. Jekyll and Mr. Hyde tale.* Critical Reviews in Oncology/Hematology, 2018. **126**: p. 52-63.
140. Eyvazi, S., et al., *Antibody Based EpCAM Targeted Therapy of Cancer, Review and Update.* Curr Cancer Drug Targets, 2018. **18**(9): p. 857-868.
141. Zhao, W., et al., *Trop2 is overexpressed in gastric cancer and predicts poor prognosis.* Oncotarget, 2016. **7**(5): p. 6136-45.
142. Riera, K.M., et al., *Trop2 is upregulated in the transition to dysplasia in the metaplastic gastric mucosa.* J Pathol, 2020. **251**(3): p. 336-347.
143. Fang, Y.J., et al., *Elevated expressions of MMP7, TROP2, and survivin are associated with survival, disease recurrence, and liver metastasis of colon cancer.* Int J Colorectal Dis, 2009. **24**(8): p. 875-84.
144. Ohmachi, T., et al., *Clinical significance of TROP2 expression in colorectal cancer.* Clin Cancer Res, 2006. **12**(10): p. 3057-63.
145. Sukhthankar, M., S. Alberti, and S.J. Baek, *(−)-Epigallocatechin-3-gallate (EGCG) post-transcriptionally and post-translationally suppresses the cell proliferative protein TROP2 in human colorectal cancer cells.* Anticancer research, 2010. **30**(7): p. 2497-2503.
146. Li, Z., X. Jiang, and W. Zhang, *TROP2 overexpression promotes proliferation and invasion of lung adenocarcinoma cells.* Biochem Biophys Res Commun, 2016. **470**(1): p. 197-204.
147. Ambrogi, F., et al., *Trop-2 is a determinant of breast cancer survival.* PloS one, 2014. **9**(5): p. e96993-e96993.
148. Tang, G., et al., *High expression of TROP2 is correlated with poor prognosis of oral squamous cell carcinoma.* Pathol Res Pract, 2018. **214**(10): p. 1606-1612.
149. Fong, D., et al., *High expression of TROP2 correlates with poor prognosis in pancreatic cancer.* Br J Cancer, 2008. **99**(8): p. 1290-5.
150. Trerotola, M., et al., *Trop-2 inhibits prostate cancer cell adhesion to fibronectin through the beta1 integrin-RACK1 axis.* J Cell Physiol, 2012. **227**(11): p. 3670-7.
151. Addati, T., et al., *TROP-2 expression in papillary thyroid cancer: a preliminary cyto-histological study.* Cytopathology, 2015. **26**(5): p. 303-11.
152. Lopez, S., et al., *Preclinical activity of sacituzumab govitecan (IMMU-132) in uterine and ovarian carcinosarcomas.* Oncotarget, 2020. **11**(5): p. 560-570.
153. Wang, F., et al., *Loss of TACSTD2 contributed to squamous cell carcinoma progression through attenuating TAp63-dependent apoptosis.* Cell Death Dis, 2014. **5**: p. e1133.
154. Sin, S.T.K., et al., *Down-regulation of TROP-2 Predicts Poor Prognosis of Hepatocellular Carcinoma Patients.* Hepatol Commun, 2018. **2**(11): p. 1408-1414.
155. Zhao, W., et al., *Trop2 is a potential biomarker for the promotion of EMT in human breast cancer.* Oncol Rep, 2018. **40**(2): p. 759-766.
156. Hsu, E.-C., et al., *Trop2 is a driver of metastatic prostate cancer with neuroendocrine phenotype via PARP1.* Proceedings of the National Academy of Sciences of the United States of America, 2020. **117**(4): p. 2032-2042.
157. Salerno, E.P., et al., *Human melanomas and ovarian cancers overexpressing mechanical barrier molecule genes lack immune signatures and have increased patient mortality risk.* Oncoimmunology, 2016. **5**(12): p. e1240857.
158. Chen, R., et al., *Increased expression of Trop2 correlates with poor survival in extranodal NK/T cell lymphoma, nasal type.* Virchows Arch, 2013. **463**(5): p. 713-9.

159. Chen, X., et al., *Overexpression of EpCAM and Trop2 in pituitary adenomas.* Int J Clin Exp Pathol, 2014. **7**(11): p. 7907-14.
160. Tate, J.G., et al., *COSMIC: the Catalogue Of Somatic Mutations In Cancer.* Nucleic Acids Res, 2019. **47**(D1): p. D941-D947.
161. Wu, M., et al., *A multi-layer inference approach to reconstruct condition-specific genes and their regulation.* Bioinformatics, 2013. **29**(12): p. 1541-52.
162. Leung, K., *(89)Zr-Desferrioxamine-anti-epithelial glycoprotein-1 hRS7 humanized monoclonal antibody*, in *Molecular Imaging and Contrast Agent Database (MICAD)*. 2004: Bethesda (MD).
163. Ning, S., et al., *TROP2 expression and its correlation with tumor proliferation and angiogenesis in human gliomas.* Neurol Sci, 2013. **34**(10): p. 1745-50.
164. Ning, S., et al., *TROP2 correlates with microvessel density and poor prognosis in hilar cholangiocarcinoma.* J Gastrointest Surg, 2013. **17**(2): p. 360-8.
165. Ray Chaudhuri, A. and A. Nussenzweig, *The multifaceted roles of PARP1 in DNA repair and chromatin remodelling.* Nat Rev Mol Cell Biol, 2017. **18**(10): p. 610-621.
166. Hidalgo-Estevez, A.M., et al., *Cyclooxygenase 2-Regulated Genes an Alternative Avenue to the Development of New Therapeutic Drugs for Colorectal Cancer.* Front Pharmacol, 2020. **11**: p. 533.
167. Zhao, P. and Z. Zhang, *TNF-alpha promotes colon cancer cell migration and invasion by upregulating TROP-2.* Oncol Lett, 2018. **15**(3): p. 3820-3827.
168. Lokody, I.B., et al., *Pten Regulates Epithelial Cytodifferentiation during Prostate Development.* PLoS One, 2015. **10**(6): p. e0129470.
169. Suraneni, M.V., et al., *Transgenic expression of 15-lipoxygenase 2 (15-LOX2) in mouse prostate leads to hyperplasia and cell senescence.* Oncogene, 2010. **29**(30): p. 4261-75.
170. Eisenwort, G., et al., *Identification of TROP2 (TACSTD2), an EpCAM-like molecule, as a specific marker for TGF-beta1-dependent human epidermal Langerhans cells.* J Invest Dermatol, 2011. **131**(10): p. 2049-57.
171. Liu, Q., et al., *Role of the mechanical microenvironment in cancer development and progression.* Cancer Biol Med, 2020. **17**(2): p. 282-292.
172. Hao, Y., et al., *miR-488-3p sponged by circ-0000495 and mediated upregulation of TROP2 in head and neck squamous cell carcinoma development.* J Cancer, 2020. **11**(11): p. 3375-3386.
173. Nakanishi, H., et al., *Loss of miR-125b-1 contributes to head and neck cancer development by dysregulating TACSTD2 and MAPK pathway.* Oncogene, 2014. **33**(6): p. 702-12.
174. Avellini, C., et al., *The trophoblast cell surface antigen 2 and miR-125b axis in urothelial bladder cancer.* Oncotarget, 2017. **8**(35): p. 58642-58653.
175. Ibragimova, I., et al., *Global reactivation of epigenetically silenced genes in prostate cancer.* Cancer Prev Res (Phila), 2010. **3**(9): p. 1084-92.
176. Remsik, J., et al., *Trop-2 plasticity is controlled by epithelial-to-mesenchymal transition.* Carcinogenesis, 2018. **39**(11): p. 1411-1418.
177. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition.* J Clin Invest, 2009. **119**(6): p. 1420-8.
178. Tsai, J.H. and J. Yang, *Epithelial-mesenchymal plasticity in carcinoma metastasis.* Genes Dev, 2013. **27**(20): p. 2192-206.

179. Wang, J., et al., *Loss of Trop2 promotes carcinogenesis and features of epithelial to mesenchymal transition in squamous cell carcinoma.* Mol Cancer Res, 2011. **9**(12): p. 1686-95.
180. Li, X., et al., *TROP2 promotes proliferation, migration and metastasis of gallbladder cancer cells by regulating PI3K/AKT pathway and inducing EMT.* Oncotarget, 2017. **8**(29): p. 47052-47063.
181. Xu, W., Z. Yang, and N. Lu, *A new role for the PI3K/Akt signaling pathway in the epithelial-mesenchymal transition.* Cell Adh Migr, 2015. **9**(4): p. 317-24.
182. Cheng, N., H. Li, and J. Luo, *Trop2 promotes proliferation, invasion and EMT of nasopharyngeal carcinoma cells through the NF-κB pathway.* RSC Advances, 2017. **7**(84): p. 53087-53096.
183. Sun, X., et al., *Knockdown of Trop2 inhibits proliferation and migration and induces apoptosis of endometrial cancer cells via AKT/beta-catenin pathway.* Cell Biochem Funct, 2020. **38**(2): p. 141-148.
184. Zhang, L., et al., *Curcumin inhibits cell proliferation and motility via suppression of TROP2 in bladder cancer cells.* Int J Oncol, 2018. **53**(2): p. 515-526.
185. Wang, J., et al., *Identification of Trop-2 as an oncogene and an attractive therapeutic target in colon cancers.* Mol Cancer Ther, 2008. **7**(2): p. 280-5.
186. Zimmers, S.M., et al., *TROP2 methylation and expression in tamoxifen-resistant breast cancer.* Cancer Cell Int, 2018. **18**: p. 94.
187. Guan, H., et al., *Trop2 enhances invasion of thyroid cancer by inducing MMP2 through ERK and JNK pathways.* BMC Cancer, 2017. **17**(1): p. 486.
188. Sin, S.T.K., et al., *TROP-2 exhibits tumor suppressive functions in cervical cancer by dual inhibition of IGF-1R and ALK signaling.* Gynecologic Oncology, 2019. **152**(1): p. 185-193.
189. Cui, W., et al., *TGFbeta1 inhibits the formation of benign skin tumors, but enhances progression to invasive spindle carcinomas in transgenic mice.* Cell, 1996. **86**(4): p. 531-42.
190. Winslow, M.M., et al., *Suppression of lung adenocarcinoma progression by Nkx2-1.* Nature, 2011. **473**(7345): p. 101-4.
191. Redlich, N., et al., *Anti-Trop2 blockade enhances the therapeutic efficacy of ErbB3 inhibition in head and neck squamous cell carcinoma.* Cell Death Dis, 2018. **9**(1): p. 5.
192. McDougall, A.R., et al., *Trop2: from development to disease.* Dev Dyn, 2015. **244**(2): p. 99-109.
193. Mortazavi, A., et al., *Mapping and quantifying mammalian transcriptomes by RNA-Seq.* Nat Methods, 2008. **5**(7): p. 621-8.
194. Trapnell, C., et al., *Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation.* Nat Biotechnol, 2010. **28**(5): p. 511-5.
195. Graveley, B.R., et al., *The developmental transcriptome of Drosophila melanogaster.* Nature, 2011. **471**(7339): p. 473-9.
196. Guttman, M., et al., *Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals.* Nature, 2009. **458**(7235): p. 223-7.

197. Yaragatti, M., C. Basilico, and L. Dailey, *Identification of active transcriptional regulatory modules by the functional assay of DNA from nucleosome-free regions.* Genome Res, 2008. **18**(6): p. 930-8.
198. Buenrostro, J.D., et al., *ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide.* Curr Protoc Mol Biol, 2015. **109**: p. 21 29 1-21 29 9.
199. Lister, R., et al., *Highly integrated single-base resolution maps of the epigenome in Arabidopsis.* Cell, 2008. **133**(3): p. 523-36.
200. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome.* Cell, 2007. **129**(4): p. 823-37.
201. Johnson, D.S., et al., *Genome-wide mapping of in vivo protein-DNA interactions.* Science, 2007. **316**(5830): p. 1497-502.
202. Jothi, R., et al., *Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data.* Nucleic Acids Res, 2008. **36**(16): p. 5221-31.
203. Lefrancois, P., et al., *Efficient yeast ChIP-Seq using multiplex short-read DNA sequencing.* BMC Genomics, 2009. **10**: p. 37.
204. Robertson, G., et al., *Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing.* Nat Methods, 2007. **4**(8): p. 651-7.
205. Meynert, A.M., et al., *Variant detection sensitivity and biases in whole genome and exome sequencing.* BMC Bioinformatics, 2014. **15**: p. 247.
206. Fang, H., et al., *Reducing INDEL calling errors in whole genome and exome sequencing data.* Genome Med, 2014. **6**(10): p. 89.
207. Ng, S.B., et al., *Targeted capture and massively parallel sequencing of 12 human exomes.* Nature, 2009. **461**(7261): p. 272-6.
208. Chaitankar, V., et al., *Next generation sequencing technology and genomewide data analysis: Perspectives for retinal research.* Prog Retin Eye Res, 2016. **55**: p. 1-31.
209. Dohm, J.C., et al., *Substantial biases in ultra-short read data sets from high-throughput DNA sequencing.* Nucleic Acids Res, 2008. **36**(16): p. e105.
210. DePristo, M.A., et al., *A framework for variation discovery and genotyping using next-generation DNA sequencing data.* Nat Genet, 2011. **43**(5): p. 491-8.
211. Van der Auwera, G.A., et al., *From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline.* Curr Protoc Bioinformatics, 2013. **43**: p. 11 10 1-11 10 33.
212. Raney, B.J., et al., *ENCODE whole-genome data in the UCSC genome browser (2011 update).* Nucleic Acids Res, 2011. **39**(Database issue): p. D871-5.
213. Rosenbloom, K.R., et al., *ENCODE whole-genome data in the UCSC Genome Browser: update 2012.* Nucleic Acids Res, 2012. **40**(Database issue): p. D912-7.
214. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Methods, 2012. **9**(4): p. 357-9.
215. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.
216. Li, H. and R. Durbin, *Fast and accurate long-read alignment with Burrows-Wheeler transform.* Bioinformatics, 2010. **26**(5): p. 589-95.
217. Galinsky, V.L., *YOABS: yet other aligner of biological sequences--an efficient linearly scaling nucleotide aligner.* Bioinformatics, 2012. **28**(8): p. 1070-7.

218. Liu, Y. and B. Schmidt, *Long read alignment based on maximal exact match seeds.* Bioinformatics, 2012. **28**(18): p. i318-i324.
219. Li, R., et al., *SOAP2: an improved ultrafast tool for short read alignment.* Bioinformatics, 2009. **25**(15): p. 1966-7.
220. Luo, R., et al., *SOAP3-dp: fast, accurate and sensitive GPU-based short read aligner.* PLoS One, 2013. **8**(5): p. e65632.
221. Lunter, G. and M. Goodson, *Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads.* Genome Res, 2011. **21**(6): p. 936-9.
222. McKenna, A., et al., *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.* Genome Res, 2010. **20**(9): p. 1297-303.
223. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.
224. Garrison, E. and G. Marth, *Haplotype-based variant detection from short-read sequencing.* arXiv preprint arXiv:1207.3907, 2012.
225. Bao, R., et al., *ExScalibur: A High-Performance Cloud-Enabled Suite for Whole Exome Germline and Somatic Mutation Identification.* PLoS One, 2015. **10**(8): p. e0135800.
226. Li, H., *FermiKit: assembly-based variant calling for Illumina resequencing data.* Bioinformatics, 2015. **31**(22): p. 3694-6.
227. Cantarel, B.L., et al., *BAYSIC: a Bayesian method for combining sets of genome variants with improved specificity and sensitivity.* BMC Bioinformatics, 2014. **15**: p. 104.
228. Pope, B.J., et al., *FAVR (Filtering and Annotation of Variants that are Rare): methods to facilitate the analysis of rare germline genetic variants from massively parallel sequencing datasets.* BMC Bioinformatics, 2013. **14**: p. 65.
229. Lai, Z., et al., *VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research.* Nucleic Acids Res, 2016. **44**(11): p. e108.
230. Eisfeldt, J., et al., *TIDDIT, an efficient and comprehensive structural variant caller for massive parallel sequencing data.* F1000Res, 2017. **6**: p. 664.
231. Chen, X., et al., *Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications.* Bioinformatics, 2016. **32**(8): p. 1220-2.
232. Abyzov, A., et al., *CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing.* Genome Res, 2011. **21**(6): p. 974-84.
233. Plagnol, V., et al., *A robust model for read count data in exome sequencing experiments and implications for copy number variant calling.* Bioinformatics, 2012. **28**(21): p. 2747-54.
234. Fromer, M., et al., *Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth.* Am J Hum Genet, 2012. **91**(4): p. 597-607.
235. Kassahn, K.S., H.S. Scott, and M.C. Caramins, *Integrating massively parallel sequencing into diagnostic workflows and managing the annotation and clinical interpretation challenge.* Hum Mutat, 2014. **35**(4): p. 413-23.
236. Wang, K., M. Li, and H. Hakonarson, *ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data.* Nucleic Acids Res, 2010. **38**(16): p. e164.
237. McLaren, W., et al., *The Ensembl Variant Effect Predictor.* Genome Biol, 2016. **17**(1): p. 122.

238. Cingolani, P., et al., *A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3.* Fly (Austin), 2012. **6**(2): p. 80-92.
239. Xu, Q. and R.L. Dunbrack, Jr., *Assignment of protein sequences to existing domain and family classification systems: Pfam and the PDB.* Bioinformatics, 2012. **28**(21): p. 2763-72.
240. Letunic, I., T. Doerks, and P. Bork, *SMART 7: recent updates to the protein domain annotation resource.* Nucleic Acids Res, 2012. **40**(Database issue): p. D302-5.
241. Arnold, S., et al., *Classifying MLH1 and MSH2 variants using bioinformatic prediction, splicing assays, segregation, and tumor characteristics.* Hum Mutat, 2009. **30**(5): p. 757-70.
242. Jagadeesh, K.A., et al., *S-CAP extends pathogenicity prediction to genetic variants that affect RNA splicing.* Nat Genet, 2019. **51**(4): p. 755-763.
243. Fratkin, E., S. Bercovici, and D.A. Stephan, *The implications of ENCODE for diagnostics.* Nat Biotechnol, 2012. **30**(11): p. 1064-5.
244. Kawaji, H., et al., *The FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation.* Genome Biol, 2009. **10**(4): p. R40.
245. Kawaji, H., et al., *Update of the FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation.* Nucleic Acids Res, 2011. **39**(Database issue): p. D856-60.
246. Kumar, P., S. Henikoff, and P.C. Ng, *Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm.* Nat Protoc, 2009. **4**(7): p. 1073-81.
247. Sim, N.L., et al., *SIFT web server: predicting effects of amino acid substitutions on proteins.* Nucleic Acids Res, 2012. **40**(Web Server issue): p. W452-7.
248. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.
249. Chun, S. and J.C. Fay, *Identification of deleterious mutations within three human genomes.* Genome Res, 2009. **19**(9): p. 1553-61.
250. Schwarz, J.M., et al., *MutationTaster evaluates disease-causing potential of sequence alterations.* Nat Methods, 2010. **7**(8): p. 575-6.
251. Cooper, G.M., et al., *Distribution and intensity of constraint in mammalian genomic sequence.* Genome Res, 2005. **15**(7): p. 901-13.
252. Davydov, E.V., et al., *Identifying a high fraction of the human genome to be under selective constraint using GERP++.* PLoS Comput Biol, 2010. **6**(12): p. e1001025.
253. Kircher, M., et al., *A general framework for estimating the relative pathogenicity of human genetic variants.* Nat Genet, 2014. **46**(3): p. 310-5.
254. Sherry, S.T., et al., *dbSNP: the NCBI database of genetic variation.* Nucleic Acids Res, 2001. **29**(1): p. 308-11.
255. Lappalainen, I., et al., *DbVar and DGVa: public archives for genomic structural variation.* Nucleic Acids Res, 2013. **41**(Database issue): p. D936-41.
256. Mailman, M.D., et al., *The NCBI dbGaP database of genotypes and phenotypes.* Nat Genet, 2007. **39**(10): p. 1181-6.
257. Tryka, K.A., et al., *NCBI's Database of Genotypes and Phenotypes: dbGaP.* Nucleic Acids Res, 2014. **42**(Database issue): p. D975-9.

258. Hamosh, A., et al., *Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders.* Nucleic Acids Res, 2005. **33**(Database issue): p. D514-7.
259. Stenson, P.D., et al., *The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine.* Hum Genet, 2014. **133**(1): p. 1-9.
260. Aartsma-Rus, A., et al., *Entries in the Leiden Duchenne muscular dystrophy mutation database: an overview of mutation types and paradoxical cases that confirm the reading-frame rule.* Muscle Nerve, 2006. **34**(2): p. 135-44.
261. Landrum, M.J., et al., *ClinVar: improving access to variant interpretations and supporting evidence.* Nucleic Acids Res, 2018. **46**(D1): p. D1062-D1067.
262. Thompson, B.A., et al., *Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database.* Nat Genet, 2014. **46**(2): p. 107-115.
263. Forbes, S.A., et al., *COSMIC: somatic cancer genetics at high-resolution.* Nucleic Acids Res, 2017. **45**(D1): p. D777-D783.
264. Lek, M., et al., *Analysis of protein-coding genetic variation in 60,706 humans.* Nature, 2016. **536**(7616): p. 285-91.
265. Spudich, G.M. and X.M. Fernandez-Suarez, *Touring Ensembl: a practical guide to genome browsing.* BMC Genomics, 2010. **11**: p. 295.
266. Kent, W.J., et al., *The human genome browser at UCSC.* Genome Res, 2002. **12**(6): p. 996-1006.
267. Speir, M.L., et al., *The UCSC Genome Browser database: 2016 update.* Nucleic Acids Res, 2016. **44**(D1): p. D717-25.
268. Robinson, J.T., et al., *Integrative genomics viewer.* Nat Biotechnol, 2011. **29**(1): p. 24-6.
269. Garber, M., et al., *Computational methods for transcriptome annotation and quantification using RNA-seq.* Nat Methods, 2011. **8**(6): p. 469-77.
270. Xie, Y., et al., *SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads.* Bioinformatics, 2014. **30**(12): p. 1660-6.
271. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
272. Urnov, F.D., et al., *Genome editing with engineered zinc finger nucleases.* Nat Rev Genet, 2010. **11**(9): p. 636-46.
273. Jinek, M. and J.A. Doudna, *A three-dimensional view of the molecular machinery of RNA interference.* Nature, 2009. **457**(7228): p. 405-12.
274. Dana, H., et al., *Molecular Mechanisms and Biological Functions of siRNA.* Int J Biomed Sci, 2017. **13**(2): p. 48-57.
275. Dell'Arciprete, R., et al., *High-efficiency expression gene cloning by flow cytometry.* J Histochem Cytochem, 1996. **44**(6): p. 629-40.
276. Toji, L.H., et al., *Validation of routine mycoplasma testing by PCR.* In Vitro Cell Dev Biol Anim, 1998. **34**(5): p. 356-8.
277. Orsulic, S., et al., *Induction of ovarian cancer by defined multiple genetic changes in a mouse model system.* Cancer Cell, 2002. **1**(1): p. 53-62.
278. Boyle, J.S. and A.M. Lew, *An inexpensive alternative to glassmilk for DNA purification.* Trends Genet, 1995. **11**(1): p. 8.

279. Alberts, B.J.A.L.J.R.M.R.K.W.P., *Molecular biology of the cell*. 2002, New York: Garland Science.
280. Morikawa, K., et al., *In vivo selection of highly metastatic cells from surgical specimens of different primary human colon carcinomas implanted into nude mice.* Cancer Res, 1988. **48**(7): p. 1943-8.
281. Naquet, P., et al., *Establishment and characterization of mouse thymic epithelial cell lines.* Thymus, 1989. **13**(3-4): p. 217-26.
282. Alberti, S. and L.A. Herzenberg, *DNA methylation prevents transfection of genes for specific surface antigens.* Proc Natl Acad Sci U S A, 1988. **85**(22): p. 8391-4.
283. Alberti, S. and M. Fornaro, *Higher transfection efficiency of genomic DNA purified with a guanidinium thiocyanate-based procedure.* Nucleic Acids Res, 1990. **18**(2): p. 351-3.
284. Elbashir, S.M., et al., *Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells.* Nature, 2001. **411**(6836): p. 494-8.
285. Chalk, A.M., C. Wahlestedt, and E.L. Sonnhammer, *Improved and automated prediction of effective siRNA.* Biochem Biophys Res Commun, 2004. **319**(1): p. 264-74.
286. Brummelkamp, T.R., R. Bernards, and R. Agami, *A system for stable expression of short interfering RNAs in mammalian cells.* Science, 2002. **296**(5567): p. 550-3.
287. Seifert, A., et al., *The metalloproteinase ADAM10 requires its activity to sustain surface expression.* Cell Mol Life Sci, 2021. **78**(2): p. 715-732.
288. Pappin, D.J., P. Hojrup, and A.J. Bleasby, *Rapid identification of proteins by peptide-mass fingerprinting.* Curr Biol, 1993. **3**(6): p. 327-32.
289. Romero-Calvo, I., et al., *Reversible Ponceau staining as a loading control alternative to actin in Western blots.* Anal Biochem, 2010. **401**(2): p. 318-20.
290. Giulietti, A., et al., *An overview of real-time quantitative PCR: applications to quantify cytokine gene expression.* Methods, 2001. **25**(4): p. 386-401.
291. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.
292. Pfaffl, M.W., *A new mathematical model for relative quantification in real-time RT-PCR.* Nucleic Acids Res, 2001. **29**(9): p. e45.
293. Guerra, E., et al., *A bicistronic CYCLIN D1-TROP2 mRNA chimera demonstrates a novel oncogenic mechanism in human cancer.* Cancer Res, 2008. **68**(19): p. 8113-21.
294. Zanna, P., et al., *Trop-1 are conserved growth stimulatory molecules that mark early stages of tumor progression.* Cancer, 2007. **110**(2): p. 452-64.
295. Mamiatis, T., et al., *Molecular cloning–A laboratory manual. New York: Cold Spring Harbor Laboratory. 1982, 545 S., 42 $.* Acta Biotechnologica, 1985. **5**(1): p. 104-104.
296. Guncar, G., et al., *Crystal structure of MHC class II-associated p41 Ii fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S.* EMBO J, 1999. **18**(4): p. 793-803.
297. Sali, A. and T.L. Blundell, *Comparative protein modelling by satisfaction of spatial restraints.* J Mol Biol, 1993. **234**(3): p. 779-815.
298. Merritt, E.A. and D.J. Bacon, *Raster3D: photorealistic molecular graphics.* Methods Enzymol, 1997. **277**: p. 505-24.
299. Zardi, L., et al., *Somatic cell hybrids producing antibodies specific to human fibronectin.* Int J Cancer, 1980. **25**(3): p. 325-9.

300. Carnemolla, B., et al., *Phage antibodies with pan-species recognition of the oncofoetal angiogenesis marker fibronectin ED-B domain.* Int J Cancer, 1996. **68**(3): p. 397-405.
301. Alberti, S., et al., *Biochemical characterization of Trop-2, a cell surface molecule expressed by human carcinomas: formal proof that the monoclonal antibodies T16 and MOv-16 recognize Trop-2.* Hybridoma, 1992. **11**(5): p. 539-45.
302. Alberti, S., M. Nutini, and L.A. Herzenberg, *DNA methylation prevents the amplification of TROP1, a tumor-associated cell surface antigen gene.* Proc Natl Acad Sci U S A, 1994. **91**(13): p. 5833-7.
303. Alberti, S., et al., *Immunofluorescence analysis in flow cytometry: better selection of antibody-labeled cells after fluorescence overcompensation in the red channel.* J Histochem Cytochem, 1991. **39**(5): p. 701-6.
304. Alberti, S., D.R. Parks, and L.A. Herzenberg, *A single laser method for subtraction of cell autofluorescence in flow cytometry.* Cytometry, 1987. **8**(2): p. 114-9.
305. Polishchuk, R.S., et al., *Correlative light-electron microscopy reveals the tubular-saccular ultrastructure of carriers operating between Golgi apparatus and plasma membrane.* J Cell Biol, 2000. **148**(1): p. 45-58.
306. Rossi, C., et al., *Intestinal tumour chemoprevention with the antioxidant lipoic acid stimulates the growth of breast cancer.* Eur J Cancer, 2008. **44**(17): p. 2696-704.
307. Morikawa, K., et al., *Influence of organ environment on the growth, selection, and metastasis of human colon carcinoma cells in nude mice.* Cancer Res, 1988. **48**(23): p. 6863-71.
308. Brattain, M.G., et al., *Heterogeneity of malignant cells from a human colonic carcinoma.* Cancer Res, 1981. **41**(5): p. 1751-6.
309. Scott, R.W., et al., *LIM kinases are required for invasive path generation by tumor and tumor-associated stromal cells.* J Cell Biol, 2010. **191**(1): p. 169-85.
310. van Noort, M., et al., *Wnt signaling controls the phosphorylation status of beta-catenin.* J Biol Chem, 2002. **277**(20): p. 17901-5.
311. Pelech, S., *Tracking cell signaling protein expression and phosphorylation by innovative proteomic solutions.* Curr Pharm Biotechnol, 2004. **5**(1): p. 69-77.
312. Rakha, E.A., et al., *E-cadherin expression in invasive non-lobular carcinoma of the breast and its prognostic significance.* Histopathology, 2005. **46**(6): p. 685-93.
313. Ward Rashidi, M.R., et al., *Engineered 3D Model of Cancer Stem Cell Enrichment and Chemoresistance.* Neoplasia, 2019. **21**(8): p. 822-836.
314. Huang, D.W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nature Protocols, 2009. **4**(1): p. 44-57.
315. Bonasera, V., S. Alberti, and A. Sacchetti, *Protocol for high-sensitivity/long linear-range spectrofluorimetric DNA quantification using ethidium bromide.* Biotechniques, 2007. **43**(2): p. 173-4, 176.
316. Vergara, D., et al., *Comparative proteome profiling of breast tumor cell lines by gel electrophoresis and mass spectrometry reveals an epithelial mesenchymal transition associated protein signature.* Mol Biosyst, 2013. **9**(6): p. 1127-38.
317. Lu, W. and Y. Kang, *Epithelial-Mesenchymal Plasticity in Cancer Progression and Metastasis.* Dev Cell, 2019. **49**(3): p. 361-374.

318. Derynck, R. and R.A. Weinberg, *EMT and Cancer: More Than Meets the Eye.* Dev Cell, 2019. **49**(3): p. 313-316.
319. Brabletz, T., *To differentiate or not--routes towards metastasis.* Nat Rev Cancer, 2012. **12**(6): p. 425-36.
320. Polyak, K. and R.A. Weinberg, *Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits.* Nature Reviews Cancer, 2009. **9**(4): p. 265-273.
321. Dongre, A. and R.A. Weinberg, *New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer.* Nat Rev Mol Cell Biol, 2019. **20**(2): p. 69-84.
322. Querzoli, P., et al., *An immunohistochemically positive E-cadherin status is not always predictive for a good prognosis in human breast cancer.* Br J Cancer, 2010. **103**(12): p. 1835-9.
323. Relli, V., et al., *Distinct lung cancer subtypes associate to distinct drivers of tumor progression.* Oncotarget, 2018. **9**(85): p. 35528-35540.
324. Herbst, A., et al., *Comprehensive analysis of beta-catenin target genes in colorectal carcinoma cell lines with deregulated Wnt/beta-catenin signaling.* BMC Genomics, 2014. **15**: p. 74.
325. Zhao, H., et al., *Identification of beta-catenin target genes in colorectal cancer by interrogating gene fitness screening data.* Oncol Lett, 2019. **18**(4): p. 3769-3777.
326. Aiello, N.M., et al., *EMT Subtype Influences Epithelial Plasticity and Mode of Cell Migration.* Dev Cell, 2018. **45**(6): p. 681-695 e4.
327. Xiong, Y., et al., *Effects of N-Glycosylation on the Structure, Function, and Stability of a Plant-Made Fc-Fusion Anthrax Decoy Protein.* Front Plant Sci, 2019. **10**: p. 768.
328. Klein, C.E., et al., *Expression of 38-kD cell-surface glycoprotein in transformed human keratinocyte cell lines, basal cell carcinomas, and epithelial germs.* J Invest Dermatol, 1990. **95**(1): p. 74-82.
329. Strnad, J., et al., *Molecular cloning and characterization of a human adenocarcinoma/epithelial cell surface antigen complementary DNA.* Cancer Res, 1989. **49**(2): p. 314-7.
330. Trebak, M., et al., *Oligomeric state of the colon carcinoma-associated glycoprotein GA733-2 (Ep-CAM/EGP40) and its role in GA733-mediated homotypic cell-cell adhesion.* J Biol Chem, 2001. **276**(3): p. 2299-309.
331. Wu, C.J., et al., *Matriptase Cleaves EpCAM and TROP2 in Keratinocytes, Destabilizing Both Proteins and Associated Claudins.* Cells, 2020. **9**(4).
332. Page-McCaw, A., A.J. Ewald, and Z. Werb, *Matrix metalloproteinases and the regulation of tissue remodelling.* Nat Rev Mol Cell Biol, 2007. **8**(3): p. 221-33.
333. Seals, D.F. and S.A. Courtneidge, *The ADAMs family of metalloproteases: multidomain proteins with multiple functions.* Genes Dev, 2003. **17**(1): p. 7-30.
334. Overall, C.M. and C.P. Blobel, *In search of partners: linking extracellular proteases to substrates.* Nat Rev Mol Cell Biol, 2007. **8**(3): p. 245-57.
335. Mason, S.D. and J.A. Joyce, *Proteolytic networks in cancer.* Trends Cell Biol, 2011. **21**(4): p. 228-37.

336. Zardi, L., et al., *Transformed human cells produce a new fibronectin isoform by preferential alternative splicing of a previously unobserved exon.* EMBO J, 1987. **6**(8): p. 2337-42.
337. Ikeda, M., et al., *Pr1E11, a novel anti-TROP-2 antibody isolated by adenovirus-based antibody screening, recognizes a unique epitope.* Biochem Biophys Res Commun, 2015. **458**(4): p. 877-82.
338. Truong, A., et al., *AR47A6. 4.2, a functional naked monoclonal antibody targeting Trop-2, demonstrates in vivo efficacy in human pancreatic, colon, breast and prostate cancer models*. 2007, AACR.
339. Thomas, A. and Y. Pommier, *Targeting Topoisomerase I in the Era of Precision Medicine.* Clin Cancer Res, 2019. **25**(22): p. 6581-6589.
340. Bardia, A., et al., *Sacituzumab Govitecan-hziy in Refractory Metastatic Triple-Negative Breast Cancer.* N Engl J Med, 2019. **380**(8): p. 741-751.
341. Tramontano, A., *Homology modeling with low sequence identity.* Methods, 1998. **14**(3): p. 293-300.
342. Crawford, H.C., et al., *ADAM10 as a therapeutic target for cancer and inflammation.* Curr Pharm Des, 2009. **15**(20): p. 2288-99.
343. Duffy, M.J., et al., *Role of ADAMs in cancer formation and progression.* Clin Cancer Res, 2009. **15**(4): p. 1140-4.
344. Nagano, O., et al., *Cell-matrix interaction via CD44 is independently regulated by different metalloproteinases activated in response to extracellular Ca(2+) influx and PKC activation.* J Cell Biol, 2004. **165**(6): p. 893-902.
345. Reiss, K., et al., *ADAM10 cleavage of N-cadherin and regulation of cell-cell adhesion and beta-catenin nuclear signalling.* EMBO J, 2005. **24**(4): p. 742-52.
346. Maretzky, T., et al., *L1 is sequentially processed by two differently activated metalloproteases and presenilin/gamma-secretase and regulates neural cell adhesion, cell migration, and neurite outgrowth.* Mol Cell Biol, 2005. **25**(20): p. 9040-53.
347. Maretzky, T., et al., *ADAM10 mediates E-cadherin shedding and regulates epithelial cell-cell adhesion, migration, and beta-catenin translocation.* Proc Natl Acad Sci U S A, 2005. **102**(26): p. 9182-7.
348. Miller, M.A., R.J. Sullivan, and D.A. Lauffenburger, *Molecular Pathways: Receptor Ectodomain Shedding in Treatment, Resistance, and Monitoring of Cancer.* Clin Cancer Res, 2017. **23**(3): p. 623-629.
349. Caescu, C.I., G.R. Jeschke, and B.E. Turk, *Active-site determinants of substrate recognition by the metalloproteinases TACE and ADAM10.* Biochem J, 2009. **424**(1): p. 79-88.
350. Wang, Y.-Y., et al., *ADAM 10 is associated with gastric cancer progression and prognosis of patients.* Journal of Surgical Oncology, 2011. **103**(2): p. 116-123.
351. Bai, S., et al., *MicroRNA-122 inhibits tumorigenic properties of hepatocellular carcinoma cells and sensitizes these cells to sorafenib.* J Biol Chem, 2009. **284**(46): p. 32015-27.
352. Gavert, N., et al., *L1, a novel target of beta-catenin signaling, transforms cells and is expressed at the invasive front of colon cancers.* J Cell Biol, 2005. **168**(4): p. 633-42.
353. Fogel, M., et al., *L1 expression as a predictor of progression and survival in patients with uterine and ovarian carcinomas.* The Lancet, 2003. **362**(9387): p. 869-875.

354. Ko, S.-Y., et al., *Increase of disintergin metalloprotease 10 (ADAM10) expression in oral squamous cell carcinoma.* Cancer Letters, 2007. **245**(1): p. 33-43.
355. Gaida, M.M., et al., *Expression of A disintegrin and metalloprotease 10 in pancreatic carcinoma.* Int J Mol Med, 2010. **26**(2): p. 281-288.
356. Simeone, P., et al., *A unique four-hub protein cluster associates to glioblastoma progression.* PLoS One, 2014. **9**(7): p. e103030.
357. Relli, V., et al., *Abandoning the Notion of Non-Small Cell Lung Cancer.* Trends Mol Med, 2019. **25**(7): p. 585-594.
358. Gutwein, P., et al., *ADAM10-mediated cleavage of L1 adhesion molecule at the cell surface and in released membrane vesicles.* FASEB J, 2003. **17**(2): p. 292-4.
359. Ludwig, A., et al., *Metalloproteinase inhibitors for the disintegrin-like metalloproteinases ADAM10 and ADAM17 that differentially block constitutive and phorbol ester-inducible shedding of cell surface molecules.* Comb Chem High Throughput Screen, 2005. **8**(2): p. 161-71.
360. Brummer, T., et al., *The metalloprotease ADAM10 (a disintegrin and metalloprotease 10) undergoes rapid, postlysis autocatalytic degradation.* FASEB J, 2018. **32**(7): p. 3560-3573.
361. Sogorb-Esteve, A., et al., *Levels of ADAM10 are reduced in Alzheimer's disease CSF.* J Neuroinflammation, 2018. **15**(1): p. 213.
362. Liu, P.C.C., et al., *Identification of ADAM10 as a major source of HER2 ectodomain sheddase activity in HER2 overexpressing breast cancer cells.* Cancer Biology & Therapy, 2006. **5**(6): p. 657-664.
363. Yan, Y., K. Shirakabe, and Z. Werb, *The metalloprotease Kuzbanian (ADAM10) mediates the transactivation of EGF receptor by G protein-coupled receptors.* J Cell Biol, 2002. **158**(2): p. 221-6.
364. Sahin, U., et al., *Distinct roles for ADAM10 and ADAM17 in ectodomain shedding of six EGFR ligands.* The Journal of cell biology, 2004. **164**(5): p. 769-779.
365. Knösel, T., et al., *Immunoprofiles of 11 biomarkers using tissue microarrays identify prognostic subgroups in colorectal cancer.* Neoplasia (New York, N.Y.), 2005. **7**(8): p. 741-747.
366. Kamble, P.R., et al., *Proteolytic cleavage of Trop2 at Arg87 is mediated by matriptase and regulated by Val194.* FEBS Letters, 2020. **594**(19): p. 3156-3169.
367. Amit, I., et al., *A module of negative feedback regulators defines growth factor signaling.* Nat Genet, 2007. **39**(4): p. 503-12.
368. Trerotola, M., et al., *CD133, Trop-2 and alpha2beta1 integrin surface receptors as markers of putative human prostate cancer stem cells.* Am J Transl Res, 2010. **2**(2): p. 135-44.
369. Bardia, A., et al., *Efficacy and Safety of Anti-Trop-2 Antibody Drug Conjugate Sacituzumab Govitecan (IMMU-132) in Heavily Pretreated Patients With Metastatic Triple-Negative Breast Cancer.* J Clin Oncol, 2017. **35**(19): p. 2141-2148.
370. Bardia, A., et al., *Sacituzumab govitecan, a Trop-2-directed antibody-drug conjugate, for patients with epithelial cancer: final safety and efficacy results from the phase I/II IMMU-132-01 basket trial.* Ann Oncol, 2021. **32**(6): p. 746-756.
371. Zukauskas, A., et al., *TM4SF1: a tetraspanin-like protein necessary for nanopodia formation and endothelial cell migration.* Angiogenesis, 2011. **14**(3): p. 345-54.

372. Allioli, N., et al., *TM4SF1, a novel primary androgen receptor target gene over-expressed in human prostate cancer and involved in cell migration.* Prostate, 2011. **71**(11): p. 1239-50.
373. Song, J., et al., *Sustained survivin expression from OX40 costimulatory signals drives T cell clonal expansion.* Immunity, 2005. **22**(5): p. 621-31.
374. Guba, M., et al., *Overexpression of melanoma inhibitory activity (MIA) enhances extravasation and metastasis of A-mel 3 melanoma cells in vivo.* Br J Cancer, 2000. **83**(9): p. 1216-22.
375. Stoll, R., et al., *The extracellular human melanoma inhibitory activity (MIA) protein adopts an SH3 domain-like fold.* EMBO J, 2001. **20**(3): p. 340-9.
376. Ghoussaini, M., et al., *Genome-wide association analysis identifies three new breast cancer susceptibility loci.* Nat Genet, 2012. **44**(3): p. 312-8.
377. Coulson-Thomas, V.J., et al., *Colorectal cancer desmoplastic reaction up-regulates collagen synthesis and restricts cancer cell invasion.* Cell Tissue Res, 2011. **346**(2): p. 223-36.
378. Barekati, Z., et al., *Methylation signature of lymph node metastases in breast cancer patients.* BMC Cancer, 2012. **12**: p. 244.
379. Shinozaki, M., et al., *Distinct hypermethylation profile of primary breast cancer is associated with sentinel lymph node metastasis.* Clin Cancer Res, 2005. **11**(6): p. 2156-62.
380. Wise, D.R., et al., *Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction.* Proc Natl Acad Sci U S A, 2008. **105**(48): p. 18782-7.
381. Kanou, T., et al., *The transmembrane adaptor Cbp/PAG1 controls the malignant potential of human non-small cell lung cancers that have c-src upregulation.* Mol Cancer Res, 2011. **9**(1): p. 103-14.
382. Salomons, G.S., et al., *X-linked creatine transporter defect: an overview.* J Inherit Metab Dis, 2003. **26**(2-3): p. 309-18.
383. Patra, S., et al., *A short review on creatine-creatine kinase system in relation to cancer and some experimental results on creatine as adjuvant in cancer therapy.* Amino Acids, 2012. **42**(6): p. 2319-30.
384. Li, J.Y., et al., *Ankyrin repeat and SOCS box containing protein 4 (Asb-4) colocalizes with insulin receptor substrate 4 (IRS4) in the hypothalamic neurons and mediates IRS4 degradation.* BMC Neurosci, 2011. **12**: p. 95.
385. Sundberg, B.E., et al., *The evolutionary history and tissue mapping of amino acid transporters belonging to solute carrier families SLC32, SLC36, and SLC38.* J Mol Neurosci, 2008. **35**(2): p. 179-93.
386. Bassett, J.H., et al., *Rapid-throughput skeletal phenotyping of 100 knockout mice identifies 9 new genes that determine bone strength.* PLoS Genet, 2012. **8**(8): p. e1002858.
387. Jung, E.M., et al., *Apoptosis- and endoplasmic reticulum stress-related genes were regulated by estrogen and progesterone in the uteri of calbindin-D(9k) and -D(28k) knockout mice.* J Cell Biochem, 2012. **113**(1): p. 194-203.
388. Aceto, N., et al., *Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis.* Cell, 2014. **158**(5): p. 1110-1122.

389. Cernat, L., et al., *Colorectal cancers mimic structural organization of normal colonic crypts.* PLoS One, 2014. **9**(8): p. e104284.
390. Guerra, E., et al., *p53, cathepsin D, Bcl-2 are joint prognostic indicators of breast cancer metastatic spreading.* BMC Cancer, 2016. **16**: p. 649.
391. Kohn, K.W., et al., *Gene expression correlations in human cancer cell lines define molecular interaction networks for epithelial phenotype.* PLoS One, 2014. **9**(6): p. e99269.
392. Cavallaro, U. and G. Christofori, *Cell adhesion and signalling by cadherins and Ig-CAMs in cancer.* Nat Rev Cancer, 2004. **4**(2): p. 118-32.
393. Scott, R.W., D. Crighton, and M.F. Olson, *Modeling and imaging 3-dimensional collective cell invasion.* J Vis Exp, 2011(58).
394. Solanas, G., et al., *Cleavage of E-cadherin by ADAM10 mediates epithelial cell sorting downstream of EphB signalling.* Nat Cell Biol, 2011. **13**(9): p. 1100-7.
395. Borghi, N., et al., *E-cadherin is under constitutive actomyosin-generated tension that is increased at cell-cell contacts upon externally applied stretch.* Proc Natl Acad Sci U S A, 2012. **109**(31): p. 12568-73.
396. Vivanco, I., et al., *Identification of the JNK signaling pathway as a functional target of the tumor suppressor PTEN.* Cancer Cell, 2007. **11**(6): p. 555-69.
397. Taddei, A., et al., *Endothelial adherens junctions control tight junctions by VE-cadherin-mediated upregulation of claudin-5.* Nat Cell Biol, 2008. **10**(8): p. 923-34.
398. Arima, Y., et al., *Induction of ZEB proteins by inactivation of RB protein is key determinant of mesenchymal phenotype of breast cancer.* J Biol Chem, 2012. **287**(11): p. 7896-906.
399. Garczarczyk, D., et al., *Signal transduction of constitutively active protein kinase C epsilon.* Cell Signal, 2009. **21**(5): p. 745-52.
400. Kato, J., et al., *Direct binding of cyclin D to the retinoblastoma gene product (pRb) and pRb phosphorylation by the cyclin D-dependent kinase CDK4.* Genes Dev, 1993. **7**(3): p. 331-42.
401. Merenda, A., N. Fenderico, and M.M. Maurice, *Wnt Signaling in 3D: Recent Advances in the Applications of Intestinal Organoids.* Trends Cell Biol, 2020. **30**(1): p. 60-73.
402. Brembeck, F.H., M. Rosario, and W. Birchmeier, *Balancing cell adhesion and Wnt signaling, the key role of beta-catenin.* Curr Opin Genet Dev, 2006. **16**(1): p. 51-9.
403. Yang, J., et al., *Guidelines and definitions for research on epithelial-mesenchymal transition.* Nat Rev Mol Cell Biol, 2020. **21**(6): p. 341-352.

X

Khouloud BOUJNAH
DOCTORANT