Intercollegiate Faculty of Biotechnology
of the University of Gdańsk and Medical University of Gdansk

DOCTORAL DISSERTATION

M.Sc., Víctor Urbiola-Salvador

Proteomics characterization of immune responses in inflammation and cancer

Doctoral thesis presented to
Biotechnology Discipline Council of the University of Gdansk
for the award of a doctorate
in the field of science
in the discipline of biotechnology

Promoter: Dr., Zhi Chen
Intercollegiate Faculty of Biotechnology UG&MUG

GDAŃSK 2024

ABSTRACT

The immune system is essential to protect the organism against pathogens, tissue injury, and cancer cells. Any disbalance in this complex immune network can cause multiple pathologies and persistent inflammatory processes can develop in chronic inflammation. Cancer and chronic inflammatory diseases are increasing their incidence and are the most common causes of death. Cancer-associated inflammation has become an important hallmark of cancer, especially in colorectal cancer (CRC) while an imbalance between pro-inflammatory and suppressive immune cells and proteins contributes to abovementioned diseases. Despite the great advances in diagnosis and treatment such as immunotherapies in cancer, most of the patients do not have complete responses and develop drug resistance via alternative immunosuppressive mechanisms. Therefore, deeper understanding of the intricate networks of immune responses involved in diseases is urgently needed. Clinical proteomics development allows for high-throughput quantification of proteins. This thesis focuses on the application of proteomics approaches to characterize immune responses in inflammation and cancer contexts, aiming to identify novel immune regulators and discover potential biomarkers.

SARS-CoV-2 infection results in acute inflammation that can develop in exacerbated immune responses, especially in patients with comorbidities such as chronic inflammatory diseases and cancer. In the first part of the thesis, orthogonal proteomics approaches, mass spectrometry and proximity extension assay, were applied to plasma samples from COVID-19 patients with and without pre-existing comorbidities and corresponding controls to determine plasma protein changes related to SARS-CoV-2 infections, the time of infection and specific anti-SARS-CoV-2 responses. Both technologies showed that COVID-19 patients with comorbidities shared a protein signature characterized by alterations in innate immune proteins including complement cascade and acute-phase proteins such as α-2-antiplasmin, that may support post-COVID-19 clotting perturbations. Key immune proteins were detected including CD4 with associated proteins such as CD28 and anti-microbial BST2. Moreover, indicators of tissue remodeling and damage were detected such as MATN2 and COL6A3 as well as extracellular matrix ECM1 and keratin K22E with potential as novel biomarkers for early detection. In addition, non-previously reported elevated RBP2 and downregulated RNF41 in COVID-19 were found.

CRC diagnosis is mainly based on costly invasive colonoscopy screening programs while CRC prognosis is mainly determined by the tumor stage in the detection time with low survival rates for advanced stages. Therefore, blood-based biomarkers are a promising alternative to improve CRC diagnosis. In the second part of the thesis, previously optimized proteomics approaches were applied to plasma samples from a multi-center CRC cohort and healthy controls to determine protein changes involved in CRC development, progression, and cancer-associated inflammation. MS-based detected protein changes in CRC patients were associated with cholesterol metabolism including APOC2 associated with CRC progression, several SERPIN family members, and the complement cascade including C5, C1QB as well as C4B and C8A both associated with cancer-associated inflammation and CRC progression. Importantly, increased C5 in CRC was validated in an additional cohort. Moreover, increased pro-inflammatory LBP and SAA4 were detected for the first time in CRC while acute-phase reactant LRG1 and ceruloplasmin were linked to cancer-associated inflammation. Proximity extension assay revealed plasma protein changes also associated with inflammation such as MDK, proteins associated with activated Th17, and oncogenic signaling pathways at systemic level. Noteworthy, increased levels of T-cell attractant CXCL9 and CCL23 were discovered and validated in an additional CRC cohort as novel potential diagnostic biomarkers. IFNGγ, IL17C, and IL32 were linked to early CRC stages while ACP6, FLT4, and MANSC1 were linked to late stages, being promising prognostic biomarkers.

In the last part, aiming to determine protein changes from immune cells within the CRC tumor microenvironment (TME), a deep MS-based proteomics analysis of CRC and normal matched tissue enriched with CD4+ T cells and other immune cells. Protein patterns in CRC tissue reflected the ongoing tumorigenic processes and tissue integrity disruption within CRC TME including cell cycle and other hallmarks of cancer such as angiogenesis, apoptosis dysregulation, cancer stemness, and extracellular remodeling. Importantly, a complex network of increased immune proteins in CRC TME was unveiled with innate pro-inflammatory S100A12, S100A8, and S100A9 as well as immunosuppressive mediators such as CD276, PVR, and NT5E. Moreover, protein expression indicated high cell immune heterogeneity with the co-existence of increased levels of FGF2-producing CAFs together with monocyte/macrophage expressing immune checkpoint ICOSL, both of them linked to CRC progression for the first time. Also, higher content of Tregs, activated mast cells, and B cells as well as reduction of IgA plasma cells and CD56 NK cells were predicted within the CRC TME. Interestingly, increased complement cascade within CRC supported findings in CRC plasma analysis which are suggested to

have immunosuppressive properties within the TME. Inferred Treg content was correlated with active MHCII presentation with GILT that may mediate tolerogenic responses and immunosuppressive metabolic reprogramming via tryptophan (KYNU, IDO1, AHR), arginine (ARG1), and taurine (SLC6A6) deprivation. Along the novel potential immune regulators within CRC TME, MCEMP1 may play a relevant role in adhesion and migration of myeloid and T cells, especially Tregs.

In conclusion, this thesis contributed to characterizing proteins associated with immune responses in inflammation and cancer. Novel plasma proteins associated with SARS-CoV-2 infection under pre-existing chronic inflammatory conditions and in CRC provide novel insights into the disease development. The data generated from this thesis work could facilitate the development of novel clinical biomarkers by further validation studies in larger and more diverse cohorts to evaluate their feasibility for clinical usage. Extensive characterization of CRC TME with high immune infiltration highlighted multiple immune-related proteins that may be novel potential immune regulators. Further functional studies may facilitate to unveil underlying molecular mechanisms involved in TME CRC immune responses.

TRESZCZENIE

Układ odpornościowy jest niezbędny do ochrony organizmu przed patogenami, uszkodzeniami tkanek i komórkami nowotworowymi. Wszelkie zaburzenia równowagi w tej złożonej sieci immunologicznej mogą powodować wiele rodzajów patologii, a w skutek utrzymującego się stanu zapaleniu mogą rozwinąć się przewlekłe procesy zapalne. Rak i przewlekłe choroby zapalne zwiększają swoją częstość występowania i są najczęstszymi przyczynami zgonów. Zapalenie związane z rakiem stało się ważną cechą charakterystyczną raka, szczególnie w przypadku raka jelita grubego (CRC), podczas gdy brak równowagi między prozapalnymi i supresyjnymi komórkami odpornościowymi i białkami przyczynia się do wyżej wymienionych chorób. Pomimo dużych postępów w diagnostyce i leczeniu, takich jak immunoterapie w przypadku raka, większość pacjentów nie wykazuje całkowitej odpowiedzi i rozwija lekooporność za pośrednictwem alternatywnych mechanizmów immunosupresyjnych. Dlatego pilnie potrzebne jest głębsze zrozumienie skomplikowanych sieci odpowiedzi immunologicznych zaangażowanych w choroby. Rozwój proteomiki klinicznej umożliwia wysokoprzepustową kwantyfikację białek. Niniejsza rozprawa koncentruje się na zastosowaniu podejść proteomicznych do scharakteryzowania odpowiedzi immunologicznych w kontekście zapalenia i raka, mając na celu zidentyfikowanie nowych regulatorów odpornościowych i odkrycie potencjalnych biomarkerów. Zakażenie SARS-CoV-2 powoduje ostry stan zapalny, który może rozwinąć się w zaostrzonych odpowiedziach immunologicznych, zwłaszcza u pacjentów z chorobami współistniejącymi, takimi jak przewlekłe choroby zapalne i nowotwory. W pierwszej części pracy zastosowano podejścia proteomiki ortogonalnej, spektrometrii masowej i technologii PEA (proximity extension assay) do badania próbek osocza od pacjentów z COVID-19 cierpiących nawcześniejsze choroby współistniejące jak i bez nich oraz odpowiednich kontroli w celu określenia zmian białek osocza związanych z zakażeniami SARS-CoV-2, czasu zakażenia i specyficznych odpowiedzi anty-SARS-CoV-2. Obie technologie wykazały, że pacjenci z COVID-19 cierpiący na choroby współistniejące mieli wspólną sygnaturę białkową charakteryzującą się zmianami w białkach wrodzonej odporności, w tym układu dopełniacza i białek ostrej fazy, takich jak α-2-antyplazmina, które mogą przyczyniać się dokomplikacji procesu krzepnięcia po przebytej chorobieCOVID-19. Wykryto kluczowe białka odpornościowe, w tym CD4 wraz z powiązanymi białkami, takimi jak CD28 i przeciwdrobnoustrojowe BST2. Ponadto białka związane zprzebudową i uszkodzenia tkanek, takie jak MATN2 i COL6A3 oraz białka macierzy zewnątrzkomórkowej ECM1 i keratyny K22E, mogą być nowymi biomarkerami wczesnego wykrywania. Kilka z nich nie zostało wcześniej zgłoszonych, w tym podwyższone RBP2 i obniżona ekspresja RNF41 w COVID-19.

Diagnostyka CRC bazuje głównie na programie badań przesiewowych, które opierają się na kosztownej i inwazyjnej kolonoskopii, podczas gdy prognoza CRC jest głównie określana poprzez stadium guza w w momencie wykrycia przy niskich wskaźnikach przeżycia w zaawansowanych stadiach. Dlatego biomarkery obecne w krwi są obiecującą alternatywą dla poprawy diagnostyki CRC. W drugiej części pracy, wcześniej zoptymalizowane podejścia proteomiczne zostały zastosowane do próbek osocza z wieloośrodkowej kohorty CRC i zdrowych kontroli w celu określenia zmian białkowych zaangażowanych w rozwój CRC, progresję i stan zapalny związany z rakiem. Wykryte na podstawie MS zmiany białkowe u pacjentów z CRC były związane z metabolizmem cholesterolu, w tym APOC2 związanym z postępem CRC, kilkoma członkami rodziny SERPIN i układem dopełniacza, w tym C5, C1QB, a także C4B i C8A, związanymi ze stanem zapalnym związanym z rakiem i postępem CRC. Co ważne, zwiększone stężenie C5 w CRC zostało potwierdzone w dodatkowej kohorcie. Ponadto po raz pierwszy wykryto zwiększone stężenie prozapalnych LBP i SAA4 w CRC, podczas gdy białko ostrej fazy LRG1 i ceruloplazmina były powiązane ze stanem zapalnym związanym z rakiem. Analiza PEA wykazała zmiany białek w osoczu również związane ze stanem zapalnym, takie jak MDK, białka związane z aktywowanymi szlakami sygnałowymi Th17 i onkogennymi na poziomie ogólnoustrojowym. Ponadto, po raz pierwszy wykryto zwiększone stężenie atraktantów limfocytów typu T CXCL9 i CCL23w osoczu CRC co zostało potwierdzone w dodatkowej kohorcie CRC. IFNGγ, IL17C i IL32 były powiązane z wczesnymi stadiami CRC, podczas gdy ACP6, FLT4 i MANSC1 były powiązane z późnymi stadiami, co czyni je obiecującymi biomarkerami prognostycznymi.

W ostatniej części, wysokoprzepustowa analiza proteomiczna oparta o MS była wykorzystana dookreślenia zmian białek komórek odpornościowych w mikrośrodowisku guza (TME) wtkankach CRC i pasujących tkankach niezmienionych nowotworowo wzbogaconych w komórki T CD4 + i inne komórki odpornościowe. Profile ekspresji białek w tkance CRC odzwierciedlały trwające procesy nowotworowe i zaburzenie integralności tkanki w obrębie TME CRC, w tym zaburzenie cyklu komórkowego i innych cech charakterystycznych raka, takie jak angiogeneza, dysregulacja apoptozy, macierzystość komórek rakowych i przebudowa pozakomórkowa. Co ważne, odkryto złożoną sieć zwiększonej ekspresji białek odpornościowych w TME CRC z prozapalnymi

białkami wrodzonej odporności S100A12, S100A8 i S100A9, a także mediatorami immunosupresyjnymi, takimi jak CD276, PVR i NT5E. Co więcej, ekspresja białek wskazywała na wysoką heterogeniczność immunologiczną komórek ze współistnieniem wysokiego poziomu fibroblastów związanych z rakiem (CAF) produkujących FGF2 wraz z ekspresją monocytów/makrofagów prezentujących białko odpornościowe punktu kontrolnego ICOSL, z czego oba zostały powiązane z progresją CRC po raz pierwszy. Ponadto, wyższa zawartość Treg, aktywowanych komórek tucznych i limfocytów typu B, a także redukcja komórek plazmatycznych IgA i komórek NK CD56 została przewidziane w TME CRC. Co ciekawe, zwiększona ekspresja białekkaskady dopełniacza w CRC potwierdziła wyniki analizy osocza CRC, co sugeruje ichwłaściwości immunosupresyjne w TME. Przewidziana ilośćTreg była skorelowana z aktywną prezentacją MHCII z GILT, która może pośredniczyć w odpowiedziach tolerogennych i immunosupresyjnemu przeprogramowaniu metabolicznemu poprzez pozbawienie tryptofanu (KYNU, IDO1, AHR), argininy (ARG1) i tauryny (SLC6A6). Oprócz nowych potencjalnych regulatorów odpornościowych w obrębie CRC TME, MCEMP1 może odgrywać istotną rolę w adhezji i migracji komórek mieloidalnych i limfocytów typu T, zwłaszcza Treg.

Podsumowując, niniejsza praca dyplomowa przyczyniła się do scharakteryzowania białek związanych z odpowiedziami immunologicznymi w stanach zapalnych i nowotworach. Nowe białka osocza związane z zakażeniem SARS-CoV-2 w przypadku przewlekłych stanów zapalnych i w CRC dostarczają nowych spostrzeżeń na temat rozwoju choroby. Dane wygenerowane w ramach niniejszej pracy dyplomowej mogą ułatwić opracowanie nowych biomarkerów klinicznych poprzez dalsze badania walidacyjne w większych i bardziej zróżnicowanych kohortach w celu oceny ich wykonalności w zastosowaniach klinicznych. Szeroka charakterystyka CRC TME z wysoką infiltracją immunologiczną ujawniła wiele białek związanych z komórkami i układem odpornościowym, które mogą być nowymi regulatorami odporności. Dalsze badania funkcjonalne mogą ułatwić określenie podstawowych mechanizmów molekularnych zaangażowanych w odpowiedzi immunologiczne TME CRC.

TABLE OF CONTENTS

ABBREVIATIONS

2-DE: 2-dimensional electrophoresis
A1AT: Alpha-1-antitrypsin
A1BG: Alpha-1B-glycoprotein
A2AP: α-2-antiplasmin
ABCD: Antibody barcoding with cleavable DNA
ACAN: Aggrecan
ACE2: Angiotensin-converting enzyme 2
ACN: Acetonitrile
ACP6: Lysophosphatidic acidphosphatase type 6
AGRP: Agouti-related neuropeptidase
AIDS: Acquired immunodeficiency syndrome
AIRE: Autoimmune regulator
ALN: Axillary lymph node
ALS: Acid labile subunit
ALT: Alternative telomere lengthening
AML: Acute myeloid leukemia
ANGT: Angiotensin II
APBB1IP: Amyloid beta precursor protein binding family B member 1 interacting protein
APC: Antigen Presenting Cell
APOD: Adiponectin
ARG2: Arginase 2
ARHGEF12:Rho guanine nucleotide exchange factor 12
ATRN: Attractin
AUC: Area under the curve
AZU1: Azurocidin 1
BAFF: B-cell activating factor
BCL-2: B-cell lymphoma-2
BCR: B cell receptor
BID: BH3 interacting domain death agonist
BPI: Bactericidal Permeability Increasing Protein
BST2: Tetherin
C4BPA: Component 4 binding protein alpha
C4BPB: Component 4 binding protein beta
CA11: Carbonic anhydrase
CAC: Colitis-associated cancer
CAF: Cancer Associated Fibroblast
CAR: Chimeric Antigen Receptor
CASP: Caspase
CBG: Corticosteroid-binding globulin
CBP: CSK binding protein
CCL: C-C motif chemokine ligand
CCR1: C-C Motif Chemokine Receptor 1
CD: Cluster of differentiation
cDC: Conventional DC
CEA: Carcinoembryonic antigen
CEACAM: CEA Cell Adhesion Molecule
CERU: Ceruloplasmin
CID: Collision-induced dissociation
CIMP: CpG island methylator phenotype
CIN: Chromosomal instability
CITE-seq: Cellular indexing of transcriptomes and epitopes bysequencing
CLEC4G:C-type lectin domain family 4 member G
CLIP: Class II–associated invariant chain peptide
CMS: Consensus Molecular Subtypes
CODEX:CO-Detection by indEXing
COPD: Chronic obstructive pulmonary disease
CP: Patients with comorbidities
CRACC: CD2-like receptor activating cytotoxic cells
CRC: Colorectal cancer
CSF3:colony-stimulating factor 3
CTL: Cytotoxic T lymphocyte
CTLA4:cytotoxic T cell antigen-4
CTSK: Cathepsin K
CXCL:C-X-C motif chemokine ligand
CyTOF: Cytometry by time-of-flight
DAG: Diacylglycerol
DAMP: Damage-associated molecular pattern molecule
DC: Dendritic cell
DC: Disease control
DDA: Data Dependent Acquisition
DEP: Differentially Expressed Protein
DIA: Data Independent Acquisition
DPEP2:Dipeptidase 2
ECM1: Extracellular matrix protein 1
EGFL7: Epidermal growth factor-like protein 7
EGFR: Epithelial cell growth factor receptor
ELISA: Enzyme linked immunosorbent assay
ELN: Elastin
EMT: Epithelial-Mesenchymal Transition
ENPP5: Ectonucleotide pyrophosphatase/phosphodiesterase family member 5
ESI: Electrospray ionization
EV: Extracellular vesicle
FA: Formic acid
FABP1: Fatty acid metabolism
FASL: FAS ligand
FASLG:FAS ligand
FASP: Filter Aided Sample Preparation
FC: Flow cytometry
FCAR: Fc region of IgA
FCGBP: Fc Gamma Binding Protein
FDR: False discovery rate
FFPE: Formalin-fixed paraffin-embedded
FGF: Fibroblast growth factor
FIBA: Fibrinogen alpha chain
FINC: Fibronectin
FIT: Fecal immune test
FLT4:Fms-related tyrosine kinase 4
FOXP3:Forkhead box P3
FSFC: Full Spectrum Dlow Cytometry
FWHM: Full width at half-maximum
GAM: Glioma-associated macrophages
GBM: Glioblastoma
GBP1: Guanylate-binding protein 1
G-CSF: Granulocyte colony-stimulating factor
gFONT: guaiac-based fecal occult blood test
GLUT1: Glucose transporter type 1
GNG4: G protein subunit gamma 4
GO: Gene Ontology
GPRC5A: G protein-coupled receptor, class C, group 5, member A
GPX3: Antioxidant enzyme glutathione peroxidase 3
GSEA: Gene Set Enrichment Analysis
HAGH: Hydroxyacylglutathione hydrolase
HBA: Hemoglobin subunit α
HBB: Hemoglobin subunit β
HC: Healthy control
HCD: High Collision Dissociation
HIV: Human immunodeficiency virus
HLA: Human leukocyte antigen
HNMT: Histamine degradation
IBD: Inflammatory bowel disease
IC: Immune checkpoints
ICAM: Intercellular adhesion molecule
ICU: Intensive care unit
IDO: Indoleamine 2,3-dioxygenase
IFITM3: Interferon Induced Transmembrane Protein 3
IFN: Interferon
IFNG: Interferon γ
IFN-γ: Interferon-γ
IGBP1: Immunoglobulin Binding Protein 1
IgE: Immunoglobulin E
IgM: Immunoglobulin M

IHC: Immunohistochemistry
IL: Interleukin
IL12RB1: Interleukin 12 receptor subunit beta 1
ILC: Innate lymphoid cell
IMC: Imaging mass cytometry
IP3: Inositol trisphosphate
ITAM: Immunoreceptor tyrosine-based activation motif
ITGA11: Integrin subunit alpha 11
ITIM: Immunoreceptor tyrosine-based inhibitory motif
ITK: Inducible T cell kinase
ITLN1: Intelectin 1
iTRAQ: Isobaric tags for relative and absolute quantification
JAK: Janus kinase
K22E: Keratin
KEGG: Kyoto Encyclopedia of Genes and Genomes
KIR: killer cell immunoglobulin receptor
KLKB1: Kallikrein B1
LAG-3: Lymphocyte activation gene-3
LAP: Latent-associated peptide
LARG: Leukemia-associated Rho guanine-nucleotide exchange factor
LAT: Linker for activation of T cells
LCM: Laser capture microdissection
LC-MS/MS: Tandem mass spectrometry coupled with liquid chromatography
LCN2: Lipocalin 2
LFA1: Leukocyte function-associated antigen 1
LOD: Limit of detection
LPA: Lysophosphatidic acid
LTB4R: Leukotriene B4 receptor
LTBP2: Latent Transforming Growth Factor Beta Binding Protein 2
LTF: Lactotransferrin
LYAR: Ly1 Antibody Reactive
mAb: monoclonal antibody
MAC: Membrane attack complex
MACC1: Metastasis-associated in colon cancer protein 1
MAIT: Mucosa-Associated Invariant T
MALDI-TOF: Matrix-assisted laser desorption/ionization time-of-flight
MANSC1:MANSC domain-containing protein 1
MAPK: Mitogen-activated protein kinase
MASP: MBL-associated serine protease
MBL: Mannose-binding lectins
MC: Mass cytometry
MCEMP1: Mast Cell Expressed Membrane Protein 1
mCRC: Metastatic CRC
MDK: Midkine
MDSC: Myeloid-derived suppressor cell
MELC: Multi-Epitope Ligand Cartography
MHC: Major Histocompatibility Complex
MIBI: Multiplexed Ion Beam Imaging
mIF: multiplexed ImmunoFluorescence
MIF: Macrophage Migration Inhibitory Factor
MIST: Multiplexed in situ targeting
MIT: Maximal injection time
MMP: Matrix metalloproteinase
MNDA: Myeloid nuclear differentiation antigen
MPO: Myeloperoxidase
mPOP: Minimal ProteOmic sample Preparation
MRM: Multiple-reaction monitoring
MS: Mass spectrometry
MSI: Microsatellite instability
MZB1: Marginal zone B and B1 cell-specific protein
NAFLD: Non-fatty liver disease
nanoPOTS: Nanodroplet processing in one pot for trace samples
NAPPA: Nucleic Acid Programmable Protein Array
NCAM1: Neural Cell Adhesion Molecule 1
NCE: Normalized Collision Energy
NCF2: Neutrophil cytosolic factor 2
NET: Neutrophil extracellular traps
NFAT: Nuclear factor of activated T cells
NF-κB: Nuclear factor-κB
NGAL: Neutrophil Gelatinase-Associated Lipocalin
NGS: Next generation sequencing
NK: Natural killer
NP: Nucleocapsid protein
NPM3: Nucleophosmin 3
NPX: Normalized protein expression
NSCLC: Non-small-cell lung cancer
NT5E: 5'-Nucleotidase Ecto
PAMP: Pathogen-associated molecular pattern molecule
PBMC: Peripheral blood mononuclear cell
PCA: Principal Component Analysis
PD1: Programmed cell death protein 1
pDC: Plasmacytoid DC
PD-L1: Programmed cell death ligand 1
PEA: Proximity Extension Assay
PF4: Platelet factor 4
PI3K: Phosphatidylinositol 3-kinase
PIP2: Phosphatidylinositol 4, 5-bisphosphate
PKCθ: Protein kinase C
PLA2G4A: Phospholipase A2 Group IVA
PLAYR: Proximity Ligation Assay for RNA
PLCγ1: Phospholipase Cγ1
PON1: Paraoxonase-1
PPY: Pancreatic prohormone
PRDX6:Peroxiredoxin 6
PRM: Parallel-reaction monitoring
PRR: Pattern-recognition receptor
PSM: Peptide-Spectrum Match
PTM: Post-translational modification
PTN: Pleiotrophin
PTPRD: Receptor protein tyrosine phosphatase δ
PYY: Peptide-YY
Q-OT: Quadrupole-orbitrap
Q-TOF: Quadrupole-time of flight
RAGE: Receptor for advanced glycation end products
RBD: Receptor-binding domain
RBP: Retinoic acid binding protein
REAP-seq: RNA expression and protein sequencing
RET: Ret proto-oncogen
RET4: Retinol-binding protein 4
ROI: Region Of Interest
ROS: Reactive oxygen species
RPLC: Reversed-phase liquid chromatography
RPPA: Reverse Phase Protein Arrays
RRM2B: Ribonucleotide reductase regulatory TP53 inducible subunit M2B
RSPO3:R-Spondin 3
RT: Retention Time
RT: Room temperature
S1P1R: Sphingosine 1-phosphate 1 receptor
SAA: Serum amyloid A
SAA4: Serum amyloid A4
sACE2: Soluble ACE2
SAP: SLAM-associated protein
SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2
scATAC-seq: Single-cell sequencing assay for transposase-accessible chromatin
SCBC: Single-cell barcode chip
sCD4: Soluble form of CD4
SCGB1A1: Secretoglobin family 1A member 1
SCoPE-MS:Single Cell ProtEomics by mass spectrometry
SCRN1: Secernin 1
SDBP: Syndecan Binding Protein
SELPLG:selectin P ligand
SERPINA4: Serpin family A member 4

SILAC: Stable isotope labeling using amino acids in cell culture
SIRT: Sirtuin
SISPROT: Simple and integrated spin tip-based proteomics technology
SLAM: Signaling lymphocytic activation molecule
SN: Sentinel Node
SRM: Selective-reaction monitoring
STAGE: STop And Go Extraction
STAT: Signal Transducer And Activator Of Transcription
SUGAR-seq: SUrface-protein Glycan And RNA-seq
SWATH-MS: Sequential Windowed acquisition of All Theoretical fragment ion Mass Spectra
TAM: Tumor-associated macrophage
TAP: Transporter associated with Antigen processing
TCR: T cell receptor
TEM: Central memory T cells
TEMRA: Effector Memory-Expressing CD45RA
TFF: Trefoil Factor
Tfol: Follicular helper T cells
TGFβ: Transforming Growth Factor β
TGFB1: Transforming Growth Factor Beta 1
Th: Helper T cells
TIL: Tumor-infiltrating lymphocyte
TIM3: T cell immunoglobulin and mucin-domain containing-3
TK2: Thymidine kinase 2
TLR: Toll-like receptor
TLR4: Toll-like receptor 4
TLS: Tertiary lymphoid structure
TMB: Tumor Mutational Burden
TME: Tumor MicroEnvironment
TMT: Tandem Mass Tags
TNC: Tenascin C
TNF: Tumor necrosis factor
TNFRSF: TNF Receptor Superfamily
TOF: Time-of-flight
Treg: T regulatory lymphocyte
TYK2: Tyrosine kinase 2
UICC: Union for International Cancer Control
UMAP: Uniform Manifold Approximation and Projection
VCAM: Vascular cell adhesion molecule
VEGF: Vascular endothelial growth factor
VEGFR3:Vascular Endothelial Growth Factor Receptor 3
VLA: Very late antigens
VTDB: Vitamin D-binding protein
VWF: Von Willebrand factor
ZG16: Zymogen 16

Note: References are defined in each chapter according to the published manuscripts, each of them defines the corresponding abbreviations except chapters 2 and 8. Several protein names were not considered as abbreviations to avoid excessive protein listing within the text and/or the protein names were not relevant

LIST OF COMMUNICATIONS

This thesis is based on following original publications and one manuscript:

Publications

Urbiola-Salvador V, Miroszewska D, Jabłońska A, Qureshi T, Chen Z. Proteomics approaches to characterize the immune responses in cancer. *Biochim Biophys Acta Mol Cell Res*. 8, 119266 (2022)

Urbiola-Salvador V, Lima de Souza S, Macur K, Czaplewska P, Chen Z. Plasma Proteomics Elucidated a Protein Signature in COVID-19 Patients with Comorbidities and Early-Diagnosis Biomarkers. *Biomedicines.* 12,840 (2024)

Urbiola-Salvador, V.; Lima de Souza, S.; Grešner, P.; Qureshi, T.; Chen, Z. Plasma Proteomics Unveil Novel Immune Signatures and Biomarkers upon SARS-CoV-2 Infection. *Int. J. Mol. Sci.* 24, 6276 (2023)

Urbiola-Salvador V, Jabłońska A, Miroszewska D, Kamysz W, Duzowska K, Drężek-Chyła K, Baber R, Thieme R, Gockel I, Zdrenka M, Śrutek E, Szylberg Ł, Jankowski M, Bała D, Zegarski W, Nowikiewicz T, Makarewicz W, Adamczyk A, Ambicka A, Przewoźnik M, Harazin-Lechowicz A, Ryś J, Macur K, Czaplewska P, Filipowicz N, Piotrowski A, Dumanski JP, Chen Z. Mass spectrometry proteomics characterization of plasma biomarkers for colorectal cancer associated with inflammation. *Biomark. Insights*. 19, 11772719241257739 (2024)

Urbiola-Salvador V, Jabłońska A, Miroszewska D, Huang Q, Duzowska K, Drężek-Chyła K, Zdrenka M, Śrutek E, Szylberg Ł, Jankowski M, Bała D, Zegarski W, Nowikiewicz T, Makarewicz W, Adamczyk A, Ambicka A, Przewoźnik M, Harazin-Lechowicz A, Ryś J, Filipowicz N, Piotrowski A, Dumanski JP, Li B, Chen Z. Plasma protein changes reflect colorectal cancer development and associated inflammation. *Front. Oncol.* 13:,1158261 (2023)

Urbiola-Salvador V, Miroszewska D, Jabłońska A, Duzowska K, Drężek-Chyła K, Zdrenka M, Śrutek E, Szylberg Ł, Jankowski M, Bała D, Zegarski W, Nowikiewicz T, Makarewicz W, Adamczyk A, Ambicka A, Przewoźnik M, Harazin-Lechowicz A, Ryś J, Filipowicz N, Piotrowski A, Dumanski JP, Chen Z. Deep proteomics characterization of colorectal cancer tumor microenvironment enriched in CD4+ T cells. *Unpublished manuscript*

The original communications have been reproduced with the permission of the copyright holders.

Conferences

Urbiola-Salvador V. Plasma proteomics unveil novel complement and proinflammatory proteins involved in colorectal cancer. Young Science Congress III (Gdańsk, 2023)

Urbiola-Salvador V. Mass spectrometry-based proteomics uncover novel plasma biomarkers for colorectal cancer associated with inflammation. Annual congress of the European Association for Cancer Research 2024 (Rotterdam, 2024)

Awards

Prize of the Dean of Intercollegiate Faculty of Biotechnology UG&MUG 2023

CHAPTER 1. Introduction

1.1. Overview of the immune system in homeostasis and disease

1.1.1. The immune system

The immune system is a complex network of molecules, cells, and tissues that protects organisms from pathogens, harmful objects as well as cancer cells. In homeostasis, the immune system is tightly regulated with a balance between immune activation and suppression controlled by an extensive range of cellular players and biomolecules [1]. In this way, the immune system can successfully recognize and attack pathogens and prevent the tumorigenic transformation of cells, while at the same time, it is capable of timely maintaining the immune reactions and downregulating them to limit tissue damage. A healthy immune system is in balance with constant stimulations and inhibitions that keep it in a steady state in which it can regulate a wide dynamic range of inputs [2,3]. The regulation of immune responses is based on a combination of positive feedback loops and inhibitory control mechanisms that prevent pathologic reactions.

1.1.2. Innate immune responses and inflammation

Innate immune responses are the initial defense against pathogens that can rapidly react to invading pathogens. Innate immunity is mainly composed of physical/chemical barriers such as epithelia and anti-microbial molecules from epithelial surfaces, cellular components including phagocytic cells such as neutrophils and macrophages, mast cells, dendritic cells (DCs), and natural killer (NK) cells among others, from which some of them reside in physical barriers and tissues, as well as circulating proteins such as the complement system and inflammatory mediators. The main protective reaction of the innate immune system is inflammation, but also, anti-viral defense via prevention of viral replication as well as promoting infected cells killing [1].

Acute inflammation is a protective mechanism initiated in response to tissue damage that triggers immune cell infiltration to boost the body's defense process [4]. This process is initiated by pathogen-associated molecular pattern molecules (PAMPs) recognized by Pattern Recognition Receptors (PPR) such as Toll-like receptors (TLRs) from host tissue innate immune cells and epithelial cells or damage-associated molecular pattern molecules (DAMPs) originating from damaged cells [5]. These activated sentinel cells secrete vasodilator molecules such as prostaglandins and histamine that increase capillary permeability, allowing the entrance of plasma proteins such as acute-phase response proteins and complement components to the tissue (Figure 1.1a). The complement cascade is activated in microbial surfaces via classical, lectin and mannose-binding lectins (MBLs), and alternative pathways following a sequential proteolytic cleavage that releases anaphylatoxins C3a and C5a promoting inflammation and opsonize microorganisms while it creates the membrane attack complex (MAC) to destroy them as microorganisms lack host inhibitory signals (Figure 1.1b) [6].

Sentinel cells also produce cytokines such as interleukin (IL)-1 and tumor necrosis factor (TNF) that increase IL6 production and endothelial adhesion. Endothelial cells are induced to ligands such as vascular cell adhesion molecule 1 (VCAM1) and intercellular adhesion molecule 1 (ICAM1) as well as E-selectin. At the same time, these cytokines promote systemic inflammatory responses such as acute-phase protein production in liver, fever, and leukocyte production in bone marrow. At the same time, these cytokines induce leukocyte transendothelial migration to the tissues to kill the pathogens, damaged cell clearance as well as increase inflammation and repair. In this cytokine cascade, TNF and IL1 induce C-X-C Motif Chemokine Ligand (CXCL)-8 production by different cells that recruits neutrophils to destroy pathogens and C-C motif chemokine ligand (CCL)-2 attracts monocytes that are polarized to classic macrophages which release other amplifier inflammatory cytokines that recruits other leukocytes. Neutrophils and macrophages destroy pathogens by phagocytosis. Phagolysosome formation is activated by orchestrated signaling of PPRs such as TLRs, cytokine receptors such as IFNγ receptors, and opsonin receptors such as C3b receptors as well as Cluster of Differentiation (CD)40. Pathogens are destroyed in phagolysosomes with proteolytic enzymes, reactive oxygen species (ROS), and nitric oxide. Moreover, neutrophils also use DNA and granules self-extrusion that trap pathogens and kill them in so-called neutrophil extracellular traps (NETs). Meanwhile, macrophages can use inflammasome-mediated pyroptosis in which the inflammasome is formed by caspase (CASP)-1, NLR sensor proteins, and adaptors to produce IL1 and IL18 as well as gasdermin D membrane pores formation produces pyroptosis, an osmotic cell death. As inflammation causes tissue damage by killing infected cells or collateral damage, innate responses are regulated by inhibitory mechanisms such as macrophages and DCs deactivation by IL10. Meanwhile, alternatively activated macrophages promote tissue repair by Transforming Growth Factor β (TGFβ) secretion, induction of fibroblast collagen production promoting scar tissue formation [1,7].

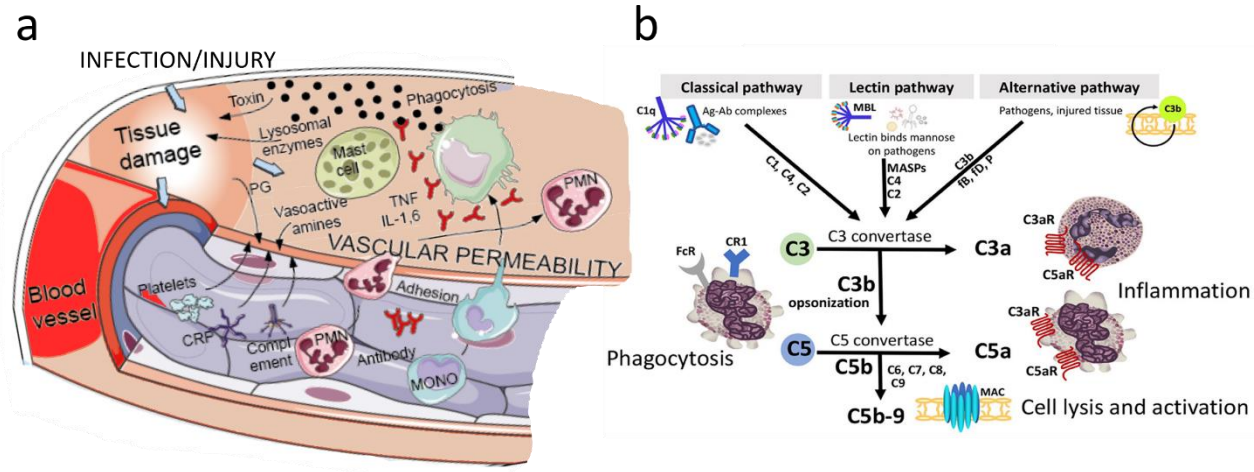

Figure 1.1. Acute inflammation and complement system. (a) Acute inflammation is initiated by tissue damage from infection/injury produces prostaglandins and sentinel cells such as mast cells produces vasoactive amines to increase vascular permeability allowing the entrance of blood components such as complement proteins, antibodies, acute-phase response proteins such as C-reactive protein (CRP), and platelets as well as cellular adhesion of immune cells through TNF, IL1, and IL6 signaling including polymorphonuclear neutrophils (PMNs) and monocytes that are polarized to classic macrophages. Both combat microbes by phagocytosis through toxins and lysosomal enzymes. (b) The complement system is activated via three pathways: in the classical pathway, C1q interacts with antigen/antibody complexes, activating C1s and C1r to cleave C2 and C4, forming the C3 convertase C4b2a. In the lectin pathway, mannose-binding lectin (MBL) recognizes carbohydrate targets followed by C2 and C4 cleavage by MBL-associated serine proteases (MASPs). In the alternative pathway, C3b binds to the pathogen/cell target together with factor B (fB) to form C3bB, and factorD (fD) cleaves fB formingC3-convertase C3bBb stabilized by properdin (P). C3b opsonizes targets for phagocytosis and B-cell activation from C5 convertase and releases anaphylatoxin C3a. Then, C5-convertase cleaves C5 promotes the membrane attack complex (MAC) complex formation binding C5b, C6, C7, C8, and several units of C9and releases anaphylatoxin C5a. Both anaphylatoxins promote inflammation through immune cell activation (adapted from [8] and [9]).

Other innate cells and cytokines are also involved in the inflammatory process. NK cells can kill infected cells regulated by activating/inhibitory receptor combinations including major histocompatibility complex (MHC) I reduction in target cells and secrete IFNγ. Also, innate lymphoid cells (ILCs) are similar in functionality and morphology to T cells but without T Cell Recognition (TCR) receptors clonally distributed. Also, basophils, which are similar to mast cells, and eosinophils, are responsible for anti-parasite protection, while both are involved in allergic inflammation [10]. Meanwhile, previously mentioned DCs are mainly subdivided into conventional DCs (cDCs) that are strong T cell activators and plasmacytoid DCs (pDCs) that can produce IFN against viruses and present antigens to T cells in the spleen. Among other involved cytokines, IL12 induces leukocyte cytotoxicity, IL18 enhances NK cells, and IL15 has similar activities to previous ones. In fact, as the initial line of defense against tissue injury or infection, the innate immune system communicates and acts in concert with the adaptive immune system presented in the next section [1,11].

1.1.3. Adaptive immune responses

Adaptive immune responses are mediated by lymphocytes classified into two main types, B cells and T cells, which are capable of recognizing antigens that are an immense number of substances from microorganisms, host cells, and environment. Lymphocyte clones are antigen-specific to portions of a protein or other biomolecules called epitopes by clonal selection of their antigen receptors. The adaptive immune system memorizes the exposure to foreign antigens and is capable of more efficient responses in secondary exposures while maintains tolerance against self-antigens and other foreign antigens such as from commensal microorganisms. Adaptive immunity can be active when the individual was exposed to the foreign antigen or passive when is transmitted from another individual such as the antibody transfer from placenta to fetus or treatment such as vaccines [1].

T cells are derived from the same bone marrow progenitor as B cells. T cell maturation occurs in the thymus by selection of clones with different T cell receptors (TCRs) by rearrangement of the V(D)J sections of their subunits αβ. Thanks to V(D)J rearrangement the adaptive immune system is capable to generate highly specific antigen receptors for an extremely wide of antigens. When T cells recognize self-antigens are negatively selected that is fundamental for self-tolerance, while after positive selection, T cells migrates to the medulla expressing TCRαβ and adaptor molecules including CD3, CD4 and CD8 where lineage commitment occurs. CD8+ T cells will recognize antigen-MHCI complexes while CD4+ T cells, antigen-MHCII complexes. There are two main T cell types, CD8+ T cells which mainly are cytotoxic T lymphocytes (CTLs) that kills infected and tumoral cells as well as CD4+ T cells, called helper T (Th) cells with several sublineages involved in specific immune responses [12,13].

Peptide antigens are loaded to MHC, internalized and processed for antigen presentation to T cells. Within human leukocyte antigens (HLA), MHCI are theoretically expressed in any nucleated cell, subdivided in classical MHCI (HLA-A,-B,-C) recognized by TCRs of CD8+ T cells and non-classical (HLA-E,-F,-G,-H) with limited peptide antigen diversity recognized

by innate receptors [14]. Meanwhile, MHCII (HLA-DR,-DQ,-DP) are expressed in Antigen Presenting Cells (APCs) including DCs, macrophages, B cells, thymic epithelial cells, and a few others to interact with CD4+ T cells [15]. MHCI processing is initiated by proteasome digestion of cytosolic proteins (including from intracellular pathogens) and transferred to the endoplasmic reticulum by transporter associated with antigen processing (TAP) protein complexes and peptide antigen is assembled to the MHCI-β2-microglobulin complex. Then, stable antigen-MHCI complexes are transported to Golgi complex by chaperones where is encapsuled in exocytic vesicles to the plasma membrane. MHCII processing is mainly performed in APCs but also other cells can after IFNγ stimulation in so-called cross-presentation. Extracellular proteins from lysosomes and endosomes are captured with surface receptors such as lectins, Fc portions of antibodies and receptors for the complement protein C3b, and Ig receptors. MHCII also can present intracellular and membrane proteins from autophagy. MHCII molecules are associated with invariant chain that is proteolyzed in the vesicles by cathepsins and other proteins to class II–associated invariant chain peptide (CLIP). Then, non-classical HLA-DM catalyzes the substitution of CLIP by high-affinity peptides and transported to the plasma membrane, a process regulated by non-classical HLA-DO that inhibits HLA-DM [1] (Figure 1.2).

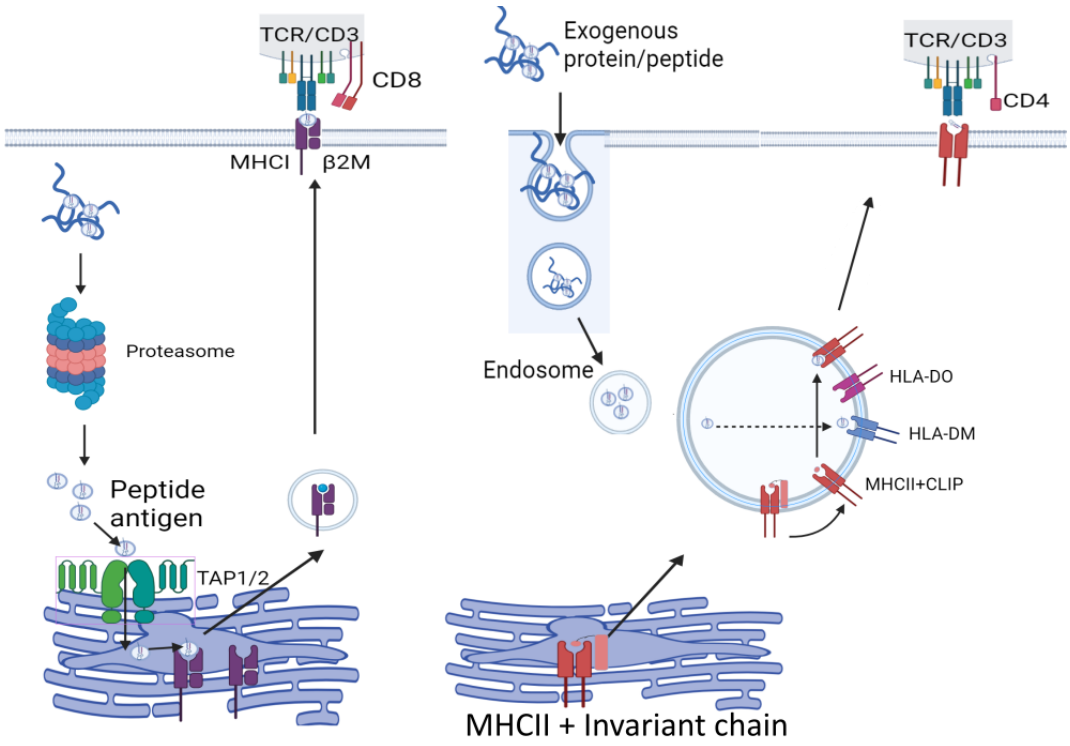

Figure 1.2. Major Histocompatibility Complex I (MHCI) and MHCII antigen processing and presentation. Intracellular antigens such as from viruses and tumors are processed in the proteasome and transported to Golgi apparatus via transporter associated with antigen processing TAP1/TAP2. Then, antigens are associated with MHCI with β2-microglobulin and transported to the membrane to be presented to CD4+ T cells. Meanwhile, exogenous antigens are processed in endosomes and MHCII molecules are processed in vesicles where cathepsins excise the invariant chain to form class II–associated invariant chain peptide (CLIP) and HLA-DM catalyzes the transfer of the antigen to MHCII regulated by HLA-DM. MHCII-antigen complex is transported to the membrane to present the antigen to CD8+ T cells (created with Biorender).

When circulating naïve T cells interact with APCs, activation is tightly regulated [16]. In the immune synapse, TCR recognition of the peptide antigen-MHC complex initiates the signaling transduction in which immunoreceptor tyrosine-based activation motifs (ITAMs) from TCR-CD3 complex are phosphorylated by FYN and LCK kinases. Then, linker for activation of T cells (LAT) signalosome is initiated by LAT, SLP76, and GRB2 phosphorylation via ZAP70. Then, the RAS-mitogen-activated protein kinase (MAPK) pathway is activated resulting in the activation of transcriptional factor AP1 as well as actin polymerization. Moreover, phospholipase Cγ1 (PLCγ1) is phosphorylated by the interaction of several proteins and inducible T cell kinase (ITK) and PLCγ1 hydrolyses phosphatidylinositol 4, 5-bisphosphate (PIP2) into inositol trisphosphate (IP3) and diacylglycerol (DAG). IP3 activates Ca2+ endoplasmic reticulum release that activates calmodulin which induces calcineurin phosphatase, resulting in IL2 and other cytokines expression via the nuclear factor of activated T cells (NFAT). Meanwhile, DAG induces the transcription factor nuclear factor-κB (NF-κB) via protein kinase C (PKCθ) that phosphorylates its adaptor CARMA1 and the signal transducers BCL10 and MALT1. Moreover, CD28 activation induces the PI3K-Akt pathway [17,18].

Then, there are several co-stimulatory and co-inhibitory signals to regulate T cell activation, proliferation, survival, and lineage differentiation including cytokine and membrane ligand-receptor interactions to ensure specific and proportionated

immune responses against infection while maintaining tolerance to self-antigens. A well-established T cell costimulatory molecule is CD28 that recognizes B7-1 (CD80) and B7-2 (CD86) activating signal transduction by phosphatidylinositol 3-kinase (PI3K) that activates NFAT, NF-κB, BCL-XL, glucose transporter type 1 (GLUT1), and IL2 mediated clonal expansion among others. Other co-stimulatory signals, also present in B cells and NK cells, include signaling lymphocytic activation molecule (SLAM) family with SLAMF1, CD244, LY9, CD84, and CD2-like receptor activating cytotoxic cells (CRACC) among others [19]. T cell activation induces a metabolic switch from oxidative phosphorylation to glycolysis and amino acid metabolism while mTOR signaling regulates protein translation to increase the production capacity of cytokines such as IL2 and IFNs among other immune biomolecules [20].

To regulate the immune response, co-inhibitory receptors deactivate T cells, NK cells, and B cells. Most of them contain immunoreceptor tyrosine-based inhibitory motifs (ITIMs) to recruit SH2-domain–containing phosphatases. Activator SCR kinases phosphorylate ITIMs that recruit phosphatases SPH1 in NK cells, SPH2 in B cell receptor (BCR) and TCR signaling, while SHIP deactivates PIP3 in adaptive and innate cells. For instance, the competitive inhibitor of B7-CD28 interactions, cytotoxic T cell antigen-4 (CTLA4), has higher affinity than CD28 for B7 proteins, regulating T cell activation. Multiple co-inhibitory receptors, also called immune checkpoints, are involved in T cell regulation such as PD1, TIM3, LAG3, and TIGIT that brought the development of immunotherapies [21]. In NK cells, killer cell immunoglobulin receptors (KIRs) recognize subsets of MHCI, some with ITIMs as inhibitory receptor CD94/NKG2A that recognizes HLA-E [22].

CD8+ T cells require antigen presentation by classical DCs, especially cDC1, and CD4+ T cell stimulation of APCs by CD40L promotes CD8+ T cell differentiation to effector CTLs and memory cells. Effector CD8+ T cells recognize target cells expressing the antigen-MHCI complex with TCR together with leukocyte function-associated antigen 1 (LFA1) that interacts with ICAM1 in the target cell. This interaction activates the release of perforins, granzymes, and FAS ligand (FASL) that binds to the death receptor FAS, inducing apoptosis by caspase activation of the target cell. Moreover, activated CTLs secrete IFNγ that activate macrophages [12,23]. Meanwhile, CD4+ T cells are essential moderators of the immune response that act as helper T cells supporting antigen-mediated responses. Circulating naïve T cells migrate to secondary lymphoid organs where non-specific antigen CD4+ T cells may die or return to circulation while CD4+ T cells with TCR specific to the antigen are activated by TCR-MHCII complex and co-stimulatory signals inducing specific cytokine receptor expression, integrins such as very late antigens 4 and5 (VLA4, VLA5) that binds fibronectin, and CD44 that adheres to hyaluronan, increasing the migration capacity in tissue. Depending on the microbe antigen and cytokine signaling from APCs, CD4+ T cells can differentiate into different subsets defined by the expression of transcriptional factors, epigenetic modifications, and cytokines. These cytokines reinforce CD4+ T cell differentiation and can inhibit other subsets to potentiate specific polarizations toward alternative subsets. There are three main effector subsets, Th1, Th2, and Th17 [13] (Figure 1.3).

Among them, Th1 cells are mainly responsible for fighting against viral and intracellular bacterial infection by classical macrophage activation mediated by CD40-CD40L via AP1 and IFNγ signaling via Signal Transducer And Activator Of Transcription (STAT)-1 in macrophages. T cell activation together with IL12-induced STAT4 and IFNγ induces STAT1 and T-BET that are essential for Th1 commitment and IFNγ production. Th1 cells also secrete TNF to increase leukocyte recruitment and inflammation as well as other cytokines such as IL10 acting as a negative feedback loop by APC deactivation. Th2 cells are responsible for helminthic infection responses, allergic reactions, and tissue repair. Th2 are induced by IL4 via activation of STAT6 and consecutive GATA3 expression that amplifies IL4 secretion and signaling. Th2-derived IL4 and IL13 induce helminthic specific IgE switch in B cells, alternative macrophage activation in which macrophages induce tissue repair via growth and angiogenic factors as well as cytokines, and promote epithelial barrier defense via intestinal peristalsis and mucus production. At the same time, IL5 promotes eosinophil proliferation for helminth clearance and mature neutrophil activation. Th17 cells are responsible for extracellular bacterial and fungi infection. IL1 and IL6 are the first signals to induce Th17 via STAT3 and RORγt while IL23 maintains the phenotype and proliferation. TGFβ can also induce Th17 differentiation under the presence of other determining cytokines although it also has anti-inflammatory capacities. These cells mainly produce IL17A and IL17F that induce anti-microbial biomolecules such as defensins and attract neutrophils to inflammation sites to promote their proliferation via granulocyte colony-stimulating factor (G-CSF). Moreover, Th17-derived IL22 promotes anti-microbial peptides and epithelial barrier integrity. There are other less characterized subsets such as Th9 cells that are involved in numerous allergic and infection processes mainly producing IL9 after induction by TGFβ and IL4 or Th22 characterized by IL22 production but without IL17 co-expression [13] (Figure 1.3).

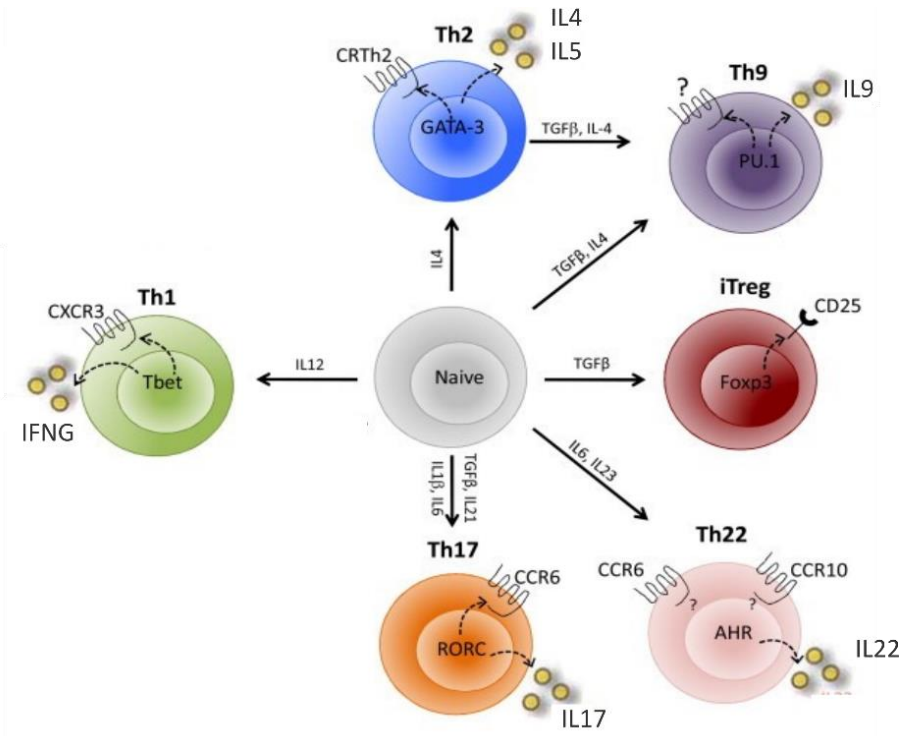

Figure 1.3. Schematic representation of different CD4 T cell subsets. Naïve CD4+ T cells can differentiate in different subsets via induction with specific cytokines indicated in the arrows from naïve to differentiated subsets. Each subset is characterized by specific transcriptional factors that regulate the expression of specific cytokine receptors as well as production and secretion of cytokines (adapted from [24]).

Another CD4+ T cell subset is follicular helper T cells (Tfol) that are responsible for germinal center formation where B cell development occurs. They are initiated by strong TCR-MHC interactions with DCs interactions together with ICOS and IL6 signaling. These results in BCL6 induction which reduces IL2Rα expression to inhibit differentiation to other Th subsets, reduces CCR7 and promotes CXCR5 expression. Consecutively, these cells mainly migrate to lymph node T and B zones and spleen. Then, the interaction with antigen-specific B cells induces SLAM-associated protein (SAP) that stabilizes BCL6 and other transcriptional regulators that prevent SLAMF6-mediated inhibition. Tfol is characterized by expression of CXCR4, SLAM, and CD86 as well as low PSGL1, sphingosine 1-phosphate 1 receptor (S1P1R), and loss of EBI2 [25].

Importantly, regulatory T cells (Treg) are responsible for dampening effector T cells to control the immune response and avoid excessive tissue damage. Moreover, Tregs play a central role in immune tolerance by suppression of immune responses against self-antigens and other antigens. Treg differentiation requires external IL2 to interact with CD25 that activates STAT5 and may enhance Forkhead box P3 (FOXP3), the Treg transcriptional factor master, as well as TGFβ.Thymocyte Tregs originate in thymus after exposition to tissue-restricted antigens regulated by autoimmune regulator (AIRE) from medullary thymic epithelial cells. Meanwhile, peripheral Tregs are generated from circulating naïve CD4+ T cells under inflammatory processes and antigen recognitions without strong innate immune responses. Peripheral Treg differentiation is mainly induced by TGFβ that induces FOXP3 expression. Treg can suppress other T cell subsets, B cells, NK cells, DCs, and macrophages [1]. Tregs use several immunosuppressive mechanisms to exert their functions. Mainly, Tregs expressing high CTLA4 can interact and convert DCs to antigen-specific tolerogenic DCs that can lose the antigen presentation or reduce effector T cell proliferation by DC production of Indoleamine 2,3-dioxygenase (IDO) that limits tryptophan metabolism. Tregs employ other immune checkpoints such as PD1. Once Tregs are differentiated via TGFβ stimulation. Another immunosuppressive mediator is TGFβ which limits effector T cell function, macrophage activation, and Th1 and Th2 differentiation while promoting tissue repair and IgA production by B cells. Meanwhile, Tregs can also secrete IL10, especially in intestinal tissue, that inhibits TCR co-stimulators as well as IL12 required for IFNγ production on DCs and macrophages. Multiple immunosuppressive mechanisms of Treg have been identified such as Ca2+ disruption in effector T cells, extracellular generation of adenosine by CD39/CD73 axis to induce tolerogenic DCs and inhibit effector T cells proliferation, perforin-granzyme cytolysis of effector T cells, apoptosis induction by interaction of Treg TRAIL with DR5 in effector T cells or IL2 consumption by high CD25 expression required for CD8+ T cell proliferation [26]. At the same time, multiple Treg subsets have been identified according to their maturation and location. For instance, follicular BCL6+CXCR5+Tregs with low CD25 regulate B cell differentiation process [27].

Noteworthy, CD4+ T cells change cytokine expression patterns according to stimulus and cell fate with high phenotypic plasticity. Once activated and differentiated, CD4+ T cells can change the phenotype from effector to regulatory or vice versa. For instance, there are T cell subsets expressing Th1 IFNγ and Th2 IL4 cytokines defined as Th1/Th2 and Th2 cells can

derived into Tfol against helminthic infections or to Th9. Importantly, Th17 and Treg cells are heterogeneous with multiple subsets expressing cytokines from other CD4+ T cells such as Th1, Th2, and Th9 as well as high plasticity between them [28,29].

Independently of the T cell subset, after CD45RA naïve T cell differentiation, CD45RO memory T cells are generated to fight against the next antigen challenge and are mainly subdivided according to their homing and functional properties. Central memory T cells (TCM) mainly reside in lymph nodes by L-selectin and CCR7 expression and are specialized in high proliferation but with limited effector capacity. Effector memory T cells (TEM) reside in peripheral tissues, especially mucosa being capable of performing rapid cytotoxic responses in antigen challenge but limited proliferation. A less characterized subtype are Effector Memory-Expressing CD45RA (TEMRA) cells that expresses the naïve marker CD45RA, high cytotoxic capacities, and senescence markers (CD57, KLRG1) [30].

Apart from CD4+ and CD8+ T cells, other unconventional T cells are involved in protection and early defense in epithelial barriers, damaged and tumorigenic cell removal, and cytokine production to improve later adaptive responses. γδ T cells, characterized by γδ TCR expression with limited diversity to bind peptides and others such as lipids, alkyl amines, and phosphorylated molecules [31]. NKT cells contain limited diversity of conventional TCR that recognize glycolipid antigens presented by CD1, NK markers such as CD56 and produce cytokines involved in Th1 and Th2 differentiation [32]. Mucosa-Associated Invariant T (MAIT) cells express an invariant TCRαβ Vα7.2-Jα33 that recognizes bacterial and fungal riboflavin metabolites presented by MR1 (MHCI–related protein 1) and are mainly located in blood, gastrointestinal tract, and are around half of T cells in liver where may protect against gut microbiota infiltrated in blood. After activation by TCR-MR1 interaction or cytokines IL12 and IL18, MAIT cells become cytotoxic producing TNF and IFNγ [33].

Briefly, the other lymphocyte subtype, so-called B cells, is responsible for humoral immunity by the production of antigen-specific antibodies that recognize and bind pathogen protein and non-protein antigens to neutralize them by opsonization and stimulate their elimination by phagocytosis and the complement system in blood, and respiratory and gastrointestinal tracts. Transitional B cells from bone marrow express low-affinity antibodies immunoglobulin M (IgM) and are negatively selected in the spleen and other secondary lymph organs, selected cells become follicular B cells with IgM and IgD that populate secondary lymphoid organs searching for BCR stimulation. Native proteins are recognized by antigen-specific BCRs Igα and Igβ and once are cross-linked by IgM and IgD and ITAMs are phosphorylated by LYN, FYN, and BLK kinases that recruit SYK with similar functions to ZAP70 in T cells. Co-stimulators include CR2, TLRs, proliferation inducer APRIL, and BAFF, especially in T cell-independent B cell stimulations [34,35].

In T cell-dependent B cell stimulation, the protein antigen-BCR is internalized, processed, and presented the antigen by MHCII to CD4+ helper T cells in follicle margins or extrafollicular sites, that are also activated with the same antigen. In the follicles, CCR7+ naïve Th cells as well reside in the T cell zone attracted by CCL19 and CCL21 while naïve B cells express CXCR5 and are attracted by CXCL13 produced by follicular DCs. When activated by antigen BCR recognition, B cells reduce expression of CXCR5 and increase CCR7 and EBI2 that recognize oxysterols produced in the T cell zone to encounter the corresponding activated CD4+ T cells. The B cell-T cell antigen presentation and cytokine inter-communication promoted by CD40 in T cells and its ligand CD40L in B cells induce B cell proliferation and differentiation by similar signal transduction pathways as T cells including activation of transcriptional factors AP1 and NF-κB. When the interaction is in extrafollicular sites, isotype switching occurs in the foci generating short-lived plasma cells. In follicles, Tfol IL21 secretion promotes activated B cells to form germinal centers in which somatic hypermutation produces antibody affinity maturation [36]. In the maturation and selection process, additional isotype switching by cytokine T cell signaling produces high affinity antibodies IgG subdivided into four subclasses (IgG1-4) according to their receptors affinity, IgE involved in helminthic infections and hypersensitivity, or IgA involved in the prevention of mucosal infections. Then, memory B cells are generated that can respond rapidly to the next infections and long-lived plasma cells that migrate to bone marrow. Meanwhile, antibodies against non-protein antigens are T cell-independent in B cell subsets called B-1 and marginal zone B cells originated from fetal liver–derived stem cells. Their isotype switching is limited to low-affinity IgG and IgA antibodies, especially located in the peritoneum and mucosa from where can differentiate into IgA plasma cells [37]. Antibody production is regulated by a negative feedback loop by antigen-antibody complex interaction with membrane IgGand Fc portion receptor FcγRIIB among others [38].

1.1.4. Immune-related diseases and chronic inflammation

Disruptions of this intricate immune network by genetic alterations and/or environmental factors cause a plethora of diseases. From one side, immunodeficiency diseases result in increased infection susceptibility and risk of cancer development by hereditary or acquired defects in the innate system such as phagocyte function, complement system, TLR signaling, or leukocyte adhesion, in T cells such as deficiencies in MHC-TCR presentation, cytokine signaling, abnormal purine metabolism, and in B cells with defect in maturation and T cell interaction. Acquired immunodeficiencies can be caused by infections such as acquired immunodeficiency syndrome (AIDS) from human immunodeficiency virus (HIV) infection that targets CD4+ T cells and others andby immunosuppressive and chemo/radiotherapy treatments, malnutrition, or other diseases such as cancer and autoimmune diseases [1,39].

Autoimmune diseases result from immune reactions against self-antigens and foreign antigens damaging self-cells and tissues due to genetic alterations in MHC genes related to antigen processing. Importantly, tissue alterations by inflammation or injury can result in autoimmune reactions while commensal microbiota and infections may also affect autoimmunity development. Abnormal immune responses to self-antigens and foreign antigens are considered hypersensitivity diseases classified into four subclasses. Immediate hypersensitivity type I includes allergies such as asthma or atopic dermatitis in

which B cells produce IgE after antigen exposure and induction by IL4 and IL13 from Tfol. Foreign antigens-IgE complexes are recognized by FcεRI that activate mast cells (also expressed in basophils) to release vasodilators, leukotrienes, prostaglandins, platelet-activating factors, TNF, IL4, IL13, and IL5. Then, Th2 cells are recruited that promote late inflammation by neutrophil and eosinophil recruitment aggravating tissue damage [40].Antibody-mediated hypersensitivity type II is activated by IgG and IgM antibodies specific to membrane or extracellular matrix (ECM) antigens causing complement activation and consequent inflammation and tissue damage or by antibodies specific to hormones and metabolites needed for normal tissue function. In immune complex-mediated hypersensitivity type III, antibodies target soluble circulating antigens and antigen-antibody complexes can deposit in blood vessel walls activating the complement and consequent inflammation in multiple tissues as in systemic lupus erythematosus. Lastly, T cell-mediated hypersensitivity type IV includes multiple autoimmune diseases such as multiple sclerosis, rheumatoid arthritis, type 1 diabetes, and psoriasis in which CD4+ T cells, mainly Th1 and Th17, produce cytokines to promote inflammation and tissue injury against self-antigens or pathogen antigens such as *Mycobacterium tuberculosis* or SARS-CoV-2 that promote an exacerbated immune response. In fact, in viral infections, CTL cytotoxicity also causes tissue damage, while chemicals, metals, and plant metabolites can also promote these hypersensitive reactions [41,42].

As autoimmune diseases exemplify, when the inflammation persists along long periods of time and becomes chronic, it causes health disorders. In fact, chronic inflammation is the main contributor of diseases that are the most significant cause of death in the world including abovementioned autoimmune diseases, cardiovascular diseases, arthritis and joint diseases, chronic obstructive pulmonary disease (COPD), Alzheimer's disease, chronic kidney disease, and inflammatory bowel disease (IBD) [43]. Importantly, chronic inflammation is an essential player in the tumorigenic process being a hallmark of cancer [44].Moreover, these diseases are considered risk factors for developing clinical complications for multiple bacterial and viral infections including SARS-CoV-2 infection (further developed in sections 2.2 and 3.2) [45].

1.1.5. Cancer and tumor immunity

Cancer is a group of malignant diseases characterized by genetic, epigenetic, metabolic, and signaling alterations that result in uncontrolled growth by malignant transformation, apoptotic death resistance, invasive capacities that spread through normal tissues, and distant site metastasis [46]. Immune surveillance is the continuous recognition of transformed cells by the immune system to avoid tumor formation. Mainly, CD8+ CTLs together with APCs, are responsible for immune surveillance by tumor antigen recognition, including neoantigens from genetic alterations, oncogenic viruses, abnormally overexpressed proteins, and glycoproteins such as cancer-testis, lineage-restricted, and oncofetal antigens [47]. Moreover, CD4+ T cells are also involved in anti-tumor immunity with high Th1 and CTLs levels associated with good prognosis. NK cells recognize tumor cells with MHCI downregulation or with NK cell co-stimulators while classic M1 macrophages can also exert anti-tumorigenic activities by IFNγ activation from tumor-specific T and NK cells. However, tumors evolve by selective pressure to evade or resist anti-tumor immune responses by multiple mechanisms such as exploitation of co-inhibitory signaling by PD1, CTLA4, and other immune checkpoints. In fact, immune cells also contribute to tumor development and immune evasion such as alternatively activated M2 macrophages that promote tissue remodeling and angiogenesis supporting tumor spreading and Tregs dampening immune responses with previously presented tolerogenic mechanisms [48].

Immunotherapy is based on the activation of anti-tumor effector immune cells. At first,passive immunotherapy was developed based on passive immunization mainly with antibodies against tumor antigens such as first FDA-approved anti-CD20 for B-cell lymphomas. Passive immunotherapy is limited to the lifetime of the antibody and acquired resistance by loss of antigen expression within the tumor due to the applied selective pressure [49]. These antibodies can be conjugated with radioisotopes or chemotherapeutic compounds for targeted radio/chemotherapy or bispecific to target tumor cells and T cells to induce T cell anti-tumorigenic activity [50,51]. Immune checkpoint blockade is another approach that targets co-inhibitory signaling by PD1, CTLA4, and other immune checkpoints to re-activate effector T cells and dampen immunosuppression. However, only a minority of patients benefit from immune checkpoint blockade because tumors have multiple immune evasion mechanisms and acquired resistance by selective pressure [52]. Other immunotherapy includes adoptive cellular therapy, Chimeric Antigen Receptor (CAR)-T cell therapy, tumor-antigen vaccines, oncolytic viruses, and cytokine therapies. Adoptive cellular therapy consists of infusion of autologous or allogeneic T cells such as cultured tumor-infiltrating lymphocytes (TILs) while CAR-T cells are genetically engineered with activating TCR signals to circumvent tumor immunosuppression. These alternative therapies are in development and clinical trials with promising potential in cancer therapy, however, similar outcomes to immune checkpoint blockade are observed. [53].

1.2. Colorectal cancer (CRC)

Worldwide, CRC is the third most common cancer and the second leading cause of cancer death. However, global distribution is not homogeneous with higher incidence and mortality in more-developed countries due to the aging population and preponderance of poor dietary habits, smoking, low physical activity, and obesity among other environmental risk factors [54]. CRC screening populations are mainly based on colonoscopy to detect with a 61% reduction of mortality. However, this invasive procedure conveys risks and high costs [55]. Non-invasive methods such as occult blood testing and blood-based biomarkers are potential alternatives but current biomarkers do not reach enough sensitivity and specificity to substitute the golden standard colonoscopy (more details in sections 4.1 and 5.1). CRC is often successfully

treated with curative surgery and adjuvant chemotherapy with a 5-year survival rate of up to 90%. However, approximately 60% of patients are diagnosed with metastatic CRC (mCRC) with a poor 5-year survival rate [56].

From histological classification, over 90% of CRCs are adenocarcinoma originating from epithelial cell polyps in the adenoma-carcinoma sequence with clinicopathological and genetic characteristics [57]. There are other CRC types with different histological characteristics such as mucinous adenocarcinoma with high accumulation of mucin, signet ring cell carcinoma with high intracytoplasmic mucin, squamous cell carcinoma is a relatively rare subtype or undifferentiated carcinoma [58]. CRC results from the progressive accumulation of genetic mutation and epigenetic alterations which activate oncogenes and inactivate tumor suppressor genes that regulate cancer control hallmark pathways [59,60]. Regarding genetic factors, more than 70 % of CRC patients are sporadic cases that are characterized by mutations in KRAS, APC, and TP53 genes and chromosomal instability (CIN) characterized by aneuploidy and loss of heterozygosity. In contrast, around 15% of sporadic cases and almost all hereditary colorectal syndromes, that only represent around 5-10% of patients, present mutations in the DNA mismatch repair pathway. Lynch syndrome is the most common with mutations in MLH1, MSH2, MSH6, and PMS2 which generate microsatellite instability (MSI) and, frequently, CpG island methylator phenotype (CIMP) with high methylation levels of the genome. MSI results in high Tumor Mutational Burden (TMB) with increased neoantigen production and high immune infiltration [61]. Further transcriptomics characterization classified CRC tumors into four Consensus Molecular Subtypes (CMS). CMS1(immune) with high immune infiltration and activation with Th1 and CTLs as well as MSI[62], CMS2 (canonical) with CIN and alterations in WNT, MYC, and epithelial cell growth factor receptor (EGFR) pathway, CMS3 (metabolic) with metabolic dysregulation and upregulated glycolysis as well as CIN and high CIMP and KRAS signaling activation, and CMS4 (mesenchymal)with epithelial-mesenchymal transition (EMT), angiogenesis, complement and TGFβ signaling activation, angiogenesis, and matrix remodeling pathways [63].

1.2.1. Cancer-associated inflammation and CRC

A well-established connection between chronic inflammation and cancer is the transition from inflammatory bowel disease (IBD), including Ulcerative Colitis and Crohn's Disease, to CRC. IBD patients show increased risk of CRC development, being the cause of death for 10% of IBD patients [64]. IBD is a multifactorial disease such as genetic, environmental, microbiota, and immune factors characterized by relapsing chronic intestinal inflammation [65]. IBD immunopathogenesis is initiated by microbiota and food antigen recognition by APCs that activate the inflammatory cascade. Then, effector CD4+ T cells subsets are recruited by IL12, IL23, and TNF, especially Th2 and Th17. Th2 cells produce IL13-mediated apoptosis of epithelial cells and Th17 cells secrete IL17 resulting in exacerbated and prolonged inflammation, while Treg levels are reduced [66,67]. In chronic inflammation, excessive production of ROS and NOS induces DNA damage, molecular alterations, and epigenetic changes in epithelial cells. Moreover, inflammation promotes epithelial proliferation via inflammatory cytokines such as IL6 and IL22 and acts as a selective pressure to develop stem epithelial cells with mutations [68]. Another consequence of inflammation is gut dysbiosis that produces carcinogens and reactive metabolites. Excessive epithelial damage causes epithelial barrier disruption and intestinal permeability that continuously activate immune responses as a damaging positive feedback loop [69]. CRC tumorigenesis derived from IBD involves multiple immune signaling pathways such as an exacerbated Th17-dependent IL17 signaling that can influence other T cell subsets. In fact, IBD-derived CRC patients contain high levels of Th17-like Treg that can activate IFNγ, IL17, and TNF production [70]. Importantly, the same cytokines such as IL17 can be involved in CRC development inducing angiogenesis and vascular endothelial growth factor (VEGF) production in endothelial cells [71].

1.2.2. The immune Tumor MicroEnvironment (TME) in CRC

The immune TME plays a key role in tumorigenesis, CRC progression, metabolic rewiring, and drug resistance. TILs can have a prognostic value in CRC demonstrated by an Immunoscore that is based on densities of CD3+ and cytotoxic CD8+ T cells or the good prognostic value of Th1 and CTL infiltration [72]. Moreover, MSI CRC tumors with high immune infiltration have partially benefited from immune checkpoint therapies however only a minority of patients had complete response [73]. Within CRC TME, tumor cells can secrete IL10 and TGFβ that suppress DCs to avoid immune surveillance and induced tolerogenic DCs can secrete more IL10 while reduced IL12 and CXCL1 [74]. CD8+ T cell exclusion correlates with worse CRC prognosis and CD8+ T cells are excluded from tumors by reduced CXCL9 and CXCL10 tumor expression, limited ICAM1 and VCAM1 endothelial expression, and dense ECM barriers formed by cancer-associated fibroblasts (CAFs). Instead of cytotoxic T cells, immune TME of CRC is enriched in immunosuppressive innate cells such as M2 macrophages, CAFs, and T cells that impair T cell cytotoxic activity expressing multiple immune checkpoints and immunosuppressive cytokines. After repeated stimulations with co-inhibitory signals, T cells present exhaustion phenotypes characterized by activation without cytotoxic activity [75]. Tregs play a central role in CRC TME with high FOXP3 induction associated with STAT5 and TET2 in CD4+ T cells from CRC. Treg recruitment is mainly by CCR4 interaction with CCL22 and CCL17 from CRC tumor cells and macrophages as well as specific CCR8 and CCR6 Treg interaction with myeloid-secreted CCL1 and CCL20, respectively. Also, Treg induction by metabolic tryptophan metabolism in which IDO converts tryptophan to kynureine that promotes FOXP3 expression via AHR. CRC-infiltrated Treg can present highly immunosuppressive phenotypes with co-expression of immune checkpoints (CTLA4, PD1, TIM3) and co-stimulators ICOS and OX40 together with high CD39/CD73 expression with inhibitory effects in effector T cells via adenosine receptors among others (Figure 1.4). Moreover, Treg metabolic reprogramming to oxidative

phosphorylation, fatty acid oxidation, and lactate metabolism favor their adaption within hypoxic TME enriched in lactate [76,77].

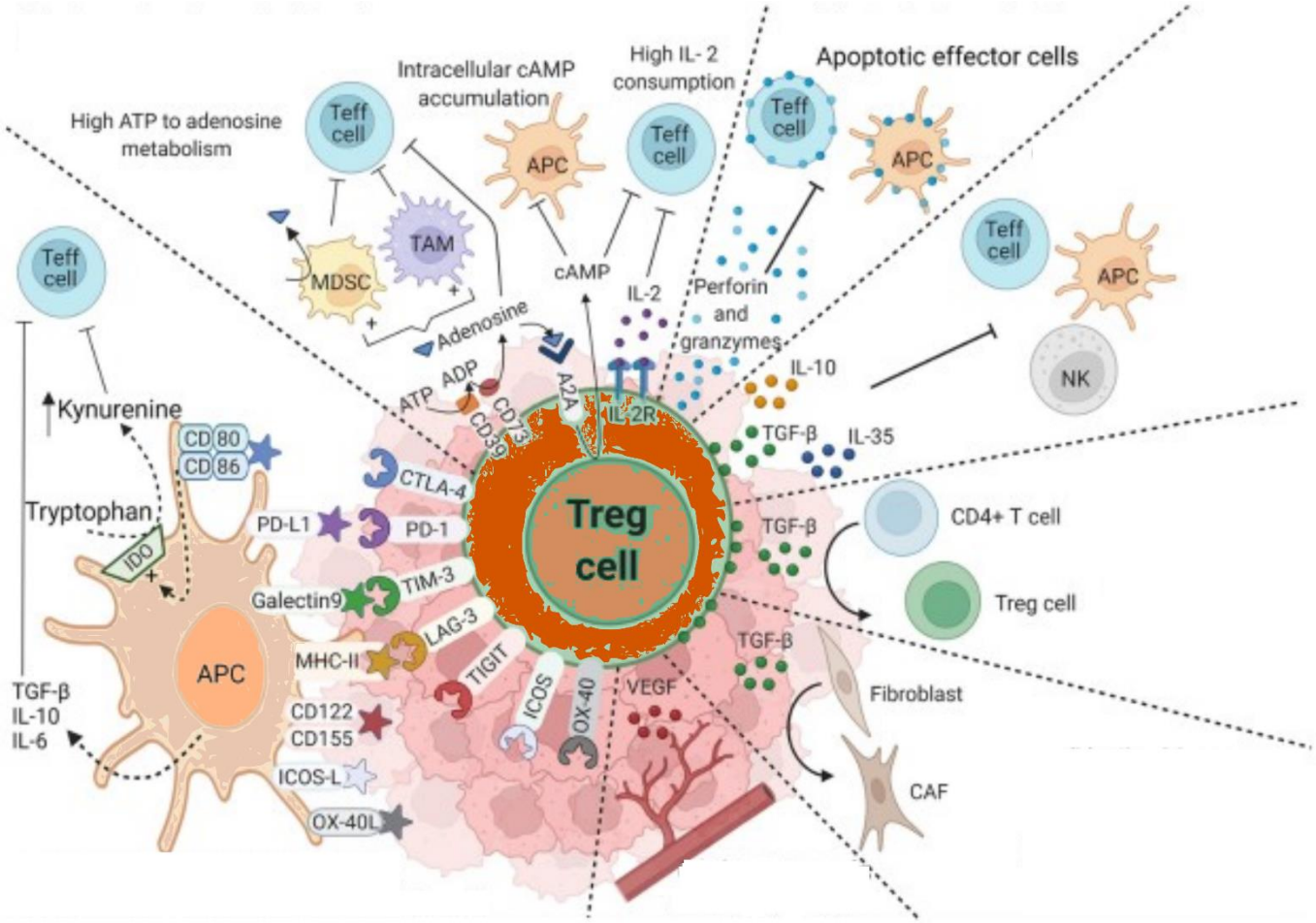

Figure 1.4. Representation of multiple immunosuppressive mechanisms exerted by Treg in CRC. Treg can suppress other immune cells by metabolic alterations, high IL2 consumption, effector cell killing, secretion of immunosuppressive cytokines inducing CD4+ T cell conversion to Tregs, and express multiple co-inhibitory signals inducing tolerogenic APCs that also can inhibit effector T cells by cytokine and tryptophan consumption. Tregs also can promote angiogenesis and CAFs (adapted from [77]).

1.3. Proteomics approaches to characterize the immune responses in cancer

In this section, the introduction briefly presents CD4+ T cell's role in the tumor microenvironment and the potential reasons for low responses to immunotherapy among cancer patients. Then, the main concepts of proteomics approaches are briefly presented. This review aims to collect the current state-of-the-art proteomics approaches to characterize the immune responses and immunosuppression within the tumor microenvironment.

This review article was originally published in **Biochimica et Biophysica Acta (BBA) - Molecular Cell Research**and the content is presented with minor modifications:
***Urbiola-Salvador V*, Miroszewska D, Jabłońska A, Qureshi T, Chen Z. Proteomics approaches to characterize the immune responses in cancer. *Biochim Biophys Acta Mol Cell Res*. 8, 119266 (2022).**

1.3.1. Introduction

In 2020, according to the International Agency for Research on Cancer, over 19 million new cases and 10 million deaths caused by cancer were estimated to occur worldwide. Breast, lung, and colorectal cancer (CRC) were assigned as the most commonly occurring types of cancer [78]. There are many known risk factors of cancer, both independent from lifestyle e.g., genetic predisposition or random DNA mutation, and lifestyle dependents such as tobacco smoking habits, lack of exercise and obesity, exposure to radiation, or poor diet [79]. Despite some differences in the mortality rate due to cancer between developed and developing countries, undeniably this issue concerns the global population [78]. For some types of cancers, inflammation is associated with tumor development, either as a cause or a consequence of ongoing tumor growth.

Regardless of the origin, the inflammation and immune cells in the tumor microenvironment (TME) play an important role in cancer development [80,81]. Helper T (Th) cells, essential moderators of the immune response, exhibit a dual role in cancer progression and immunity. The cluster of differentiation (CD)4+ T cells orchestrate immune responses against tumors and can differentiate into different subsets within TME [82]. Th1 lymphocytes, as the main producers of interferonγ (IFNγ), play the major role in anti-tumor response by activating innate immune cells such as macrophages and natural killer (NK) cells, promoting proinflammatory phenotype of macrophages, and inducing expression of major histocompatibility complex (MHC) class II on the surface of antigen-presenting cells (APCs). In addition, Th1, via the production of IFNγ, induce the differentiation of cytotoxic CD8+ T cells and inhibit T regulatory lymphocytes (Tregs) function [83]. Th2 lymphocytes are the key players in host immunity and tissue repair signaling. Signatory cytokines produced by Th2 cells, interleukin (IL)-4, IL5, IL9, and IL13, participate in B cell proliferation and immunoglobulin E (IgE) production. They are also associated with the pathological states of chronic inflammation e.g., asthma [84]. Their role in cancer clearance has been linked with the recruitment of eosinophils, neutrophils, and macrophages at tumor sites via IL4 signaling [85].

Another subset of CD4+ T cells, Th17, are the main producers of IL17 and play a key role in the host defense against pathogens, especially in the gut [86]. Th17 cells have been linked with the induction of a protumor environment [87], however, preclinical and clinical studies demonstrate that Th17 cells contribute to the recruitment of effector cells such as neutrophils to TME [71]. Therefore, the role of Th17 in cancer progression remains controversial and requires further studies [88]. On the other hand, Treg cells are a subpopulation of T cells that are engaged in sustaining immunological self-tolerance and homeostasis. They can suppress and downregulate the immune response, as such, participate in promoting the tumor favorable conditions [86,89]. Moreover, Treg cells' phenotypic plasticity facilitates the conversion to different subsets with superior immunosuppressive activity such as IL17 producing Treg and latent-associated peptide (LAP)+ Treg cells [90]. More recently, other novel T cell subsets such as Th9, Th22, and follicular Th cells have been suggested to affect the TME with controversial effects, regarding their anti-tumor or protumor activity [91,92]. Despite the great advance in cancer immunology in the last few years, a better understanding of the TME heterogeneity and the complexity of immune cell interactions are needed.

Cancer immunotherapy with monoclonal antibodies (mAbs) that block the interaction of programmed cell death protein 1 (PD1) with its ligand PD-L1 has shown clinical response in a wide range of solid and hematological cancers [93]. However, only a minority of patients exhibit dramatic positive responses. The low response can be linked to other immunosuppressive mechanisms and an array of factors affecting immunotherapy effectiveness, such as tumor genomic instability, immune phenotype, level of inflammation, microbiome, T cell memory, or even sunlight exposure [48]. Therefore, a comprehensive understanding of the role of T cells in TME is needed to discover novel targets and biomarkers for the effective treatment of cancer.

High-dimensional and high-throughput techniques are promising tools in unraveling this issue [94]. Omics-based strategies such as transcriptomics have been applied to uncover the immune surveillance mechanisms and immune profiling in various cancer types [95–98]. However, the knowledge about the mechanism of gene regulation at the posttranscriptional, translational, and posttranslational levels is still limited. Poor levels of concordance between changes in protein abundance and mRNA expression have been reported, especially in CD4+ T cells [99,100]. Therefore, with steady progress in proteomics technology, proteomics analyses can provide a more comprehensive view of T cells' fate in cancer progression through simultaneous detection, identification, and quantification of thousands of proteins in a single study. In particular, tandem mass spectrometry (MS) coupled with liquid chromatography (LC-MS/MS) provides an integrated system for proteomics analysis with improved sensitivity and moderate throughput [101,102].

Nowadays, two basic proteomics strategies are commonly used in cancer study: MS-based and antibody-based. Bottom-up proteomics is currently the predominant MS-based strategy, which is applied to discovery research aiming at the deep identification of a given proteome in an exploratory and unbiased manner. In contrast, antibody-based strategies are widely used in targeted approaches, which can detect preselected proteins from a given sample, ideally, with high sensitivity, selectivity, quantitative accuracy, and reproducibility. However, antibody-based approaches are limited by the number of proteins detected and the availability of antibodies. MS-based strategies can potentially detect hundreds or thousands of proteins to establish novel biomarkers, potential drug targets, and other research efforts [103]. So far, neither of the two strategies has achieved the detection of the whole proteome. In this review, we focus on different proteomics approaches, including antibody-based and MS-based strategies, for immune characterization of cancer states with an emphasis on CD4+ T cells. Finally, we will present novel single-cell proteomics approaches with great potential in cancer immunology.

1.3.2. A brief overview of proteomics

Proteomics is a large-scale analysis of the sum of proteins from an organism, tissue, cell, or biofluid [104]. Clinical proteomics aims at understanding how their abundance, expression, localization, posttranslational modifications (PTMs), and molecular interactions cause disease to improve patient care [105]. Various protein identification techniques have been applied to study proteins involved in cancer formation and progression such as flow cytometry (FC), mass cytometry (MC or CyTOF; cytometry by time-of-flight) [106,107], and immunohistochemistry (IHC) [108]. However, these strategies are limited by their multiplexing capacity and the availability and quality of specific antibodies [102].

Bottom-up proteomics is currently a predominant strategy that utilizes protein digestion before MS analysis. The general sample preparation workflow in bottom-up proteomics (Figure 1.5) consists of protein extraction, solubilization with detergents, reduction of disulfide bonds, alkylation of free cysteines, and lastly enzymatic digestion (normally trypsin)

conducted in-solution or filter-aided. Then, obtained peptides are desalted with reversed phase C18 tips [109,110]. This workflow can be combined with fractionation steps at protein and peptide levels with different biochemical approaches such as two-dimensional electrophoresis (2-DE), strong cation exchange, or enrichment of peptides with PTMs (e.g., phosphorylation, acetylation, glycosylation) [111]. The resulting mixtures of peptides are identified and quantified in the mass spectrometer by the analysis of mass-to-charge ratios of molecular ions.

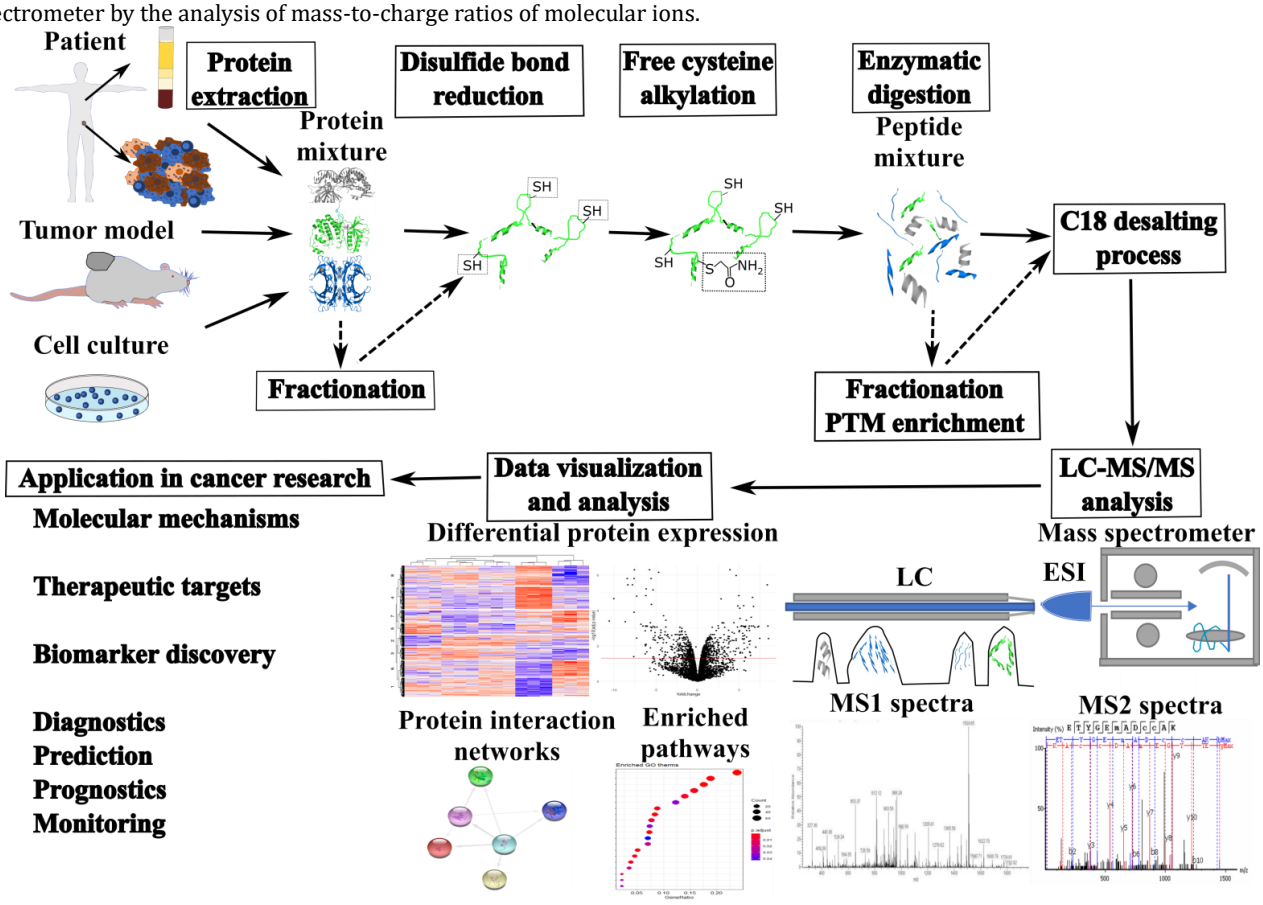

Figure 1.5. Bottom-up proteomics workflow. Protein mixtures are extracted from patient samples, tumor model samples, or cell culture. Proteins are solubilized, disulfide bonds are reduced, free cysteines are alkylated, and proteins are digested with enzymes. Alternatively, proteins and peptides can be fractionated or enriched in posttranslational modifications (PTMs). Peptide mixture is desalted with reversed phase C18 tips and prepared for tandem mass spectrometry coupled with liquid chromatography (LC-MS/MS) analysis. LC separates peptides that are ionized by electrospray ionization (ESI) and analyzed in the mass spectrometer, generating MS1 and MS2 spectra. Data visualization and analysis allow the identification and quantification of differentially expressed proteins as well as the identification of enriched pathways and protein interaction networks. Proteomics analysis has several applications in cancer research such as the discovery of underlying molecular mechanisms, therapeutic targets, and biomarkers as well as improvement of diagnostics, prediction, prognostic, and therapy monitoring.

LC-MS/MS has revolutionized proteomics because of the great advances in reproducibility, high resolution, high mass accuracy, improvement of scanning modes, and excellent sensitivity. The combination of nano-LC technology or capillary electrophoresis with electrospray ionization (ESI) enables the identification and quantification of thousands of proteins from one single injection in high-resolution mass spectrometers [102,112,113]. This progress in clinical proteomics accelerates the study of the underlying mechanisms of cancer as well as biomarkers discovery and, at the same time, improves diagnostic, prediction, prognostic, and monitoring efficacy of novel immunotherapies [101,114,115].

MS Imaging is a cutting-edge technology that incorporates matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) with micrometer laser beams that shed on frozen or Formalin Fixed Paraffin-Embedded (FFPE) tissue samples. Each laser-excited spot generates ionized proteins/peptides which are generally identified by MALDI-TOF. Thus, tissue images are generated via a raster scan in which each spot is associated with its mass spectrum, providing the spatial distribution and relative abundance of the analytes over the entire tissue section [116]. MS Imaging is mostly non-destructive and can be combined with histological staining to study regions of interest or digital PCR [117,118]. This technique can resolve the complexity of spatial protein patterns and other biomolecules (lipids, glycans, and metabolites) within the TME in an untargeted manner [119–122]. Interestingly, recent technical advances in laser resolution enable the measurement of analytes at the single-cell level [123]. However, its wider application is currently limited by the required heavy instrumentation, non-standardized workflows, and its suboptimal quantification capability [124].

Another approach is top-down proteomics that identifies intact proteins by different protein separation techniques with LC-MS/MS, where the proteins are ionized and subsequently fragmented. However, the sensitivity is about 100-fold lower than bottom-up proteomics with lesser proteomic coverage and throughput due to its lower efficiency in fragmenting intact proteins [125,126].

1.3.3. MS-based proteomics approaches applied to study immune responses in cancer

Upregulation of immune checkpoints (IC) such as cytotoxic T cell antigen-4 (CTLA4) and PD1 molecules within the TME is considered the major immunosuppressive mechanism that inhibits effector T cell functions [127]. Apart from that, the TME is enriched in soluble factors such as tumor growth factor-β (TGFβ), IL10, and CD73-derived adenosine which potently suppress T cell anti-tumor functions and promote the conversion of naïve CD4+ T cells into Tregs [128,129]. Moreover, metabolic restriction of T cells by nutrient competition from tumor cells inhibits effector T cell anti-tumor functions [130]. MS-based discovery proteomics can contribute to elucidating the most relevant proteins, molecular mechanisms, and pathways involved in immunosuppression, which will lead to the identification of novel targets for potential immunotherapy. This section describes various MS-based proteomics approaches and their application in the analysis of the immune responses in cancer by characterization of T cells, the tumor-infiltrating lymphocytes (TILs) as well as biofluids in mice models and clinics.

1.3.3.1. The potential of MS-based proteomics approaches in preclinical cancer model studies for discovery research

Preclinical studies in mice models are an essential milestone towards novel therapeutic strategies in humans as well as uncovering molecular mechanisms involved in disease progression. Despite the great potential of proteomics to discover novel therapeutic targets, proteomics analysis has not been broadly applied in mice models in the research field of cancer immunology. Interestingly, a few bottom-up proteomics studies exemplify its ability to characterize T cells originating from spleen and lymph nodes in cancer mice models, providing novel insights into this field. For instance, proteomics analysis of T cells in a mice model of colitis-associated colorectal cancer (CAC) demonstrated that sirtuin 5 (SIRT5) downregulates numerous proteins related to the T cell receptor signaling pathway and enhances immunosuppressive Treg cell differentiation. However, further studies are needed to evaluate the broader role of SIRT5 in cancer immunotherapy [131]. In addition, bottom-up proteomics analysis can be applied to reveal PTMs involved in tumor immunosuppression. MS-based proteomic analysis of SIRT2-immunoprecipitated proteins as well as acetyl-lysine peptides demonstrated that SIRT2 suppresses key metabolic enzymes by deacetylation in T cells, promoting a T cell exhausted phenotype. These findings were validated in melanoma and lung cancer mice models as well as in T cells originating from healthy donors and TILs isolated from non-small cell lung cancer patients *in vitro*, which revealed that pharmacologic inhibition of SIRT2 can enhance cancer immunotherapies [132]. Interestingly, the sirtuins family has been associated with cancer progression and metastasis through different mechanisms [133–135]. Application of bottom-up proteomics in an arginase 2 (*Arg2*)$^{-/-}$ T-cell-specific knock-out in CRC and melanoma xenograft models discovered the immunosuppressive function of mitochondrial ARG2 in CD8+ T cells. *Arg2*-deficient CD8+ T cells were synergized with PD1 blockade, unveiling the potential application of ARG2 inhibition as a novel immunotherapy [136]. Bottom-up proteomics has also been applied to study the immune response to treatment in a breast cancer mice model. Shotgun MS analysis of mice serum revealed that cryo-thermal therapy induces acute phase response with IL6 activation, promoting Th1 anti-tumor activity [137].

Application of shotgun proteomics in hyperactive platelets derived from CAC mice revealed an increased level of protumor serum amyloid A (SAA) proteins, suggesting a novel target to treat CAC patients at early clinical stages, or even to prevent cancer development [138]. Also, bottom-up proteomics analyzed extracellular vesicles (EVs) from tumor-associated macrophages (TAMs) derived from a CRC mouse model. Surprisingly, TAM-EVs possessed a proteomic signature that was associated with inflammation and immune response through Th1/M1 macrophage polarization [139]. Both studies show the broad application of MS-based proteomics in the analysis of innate immune cells which influence the cancer immune response.

The abovementioned studies show the potential of MS-based proteomics in preclinical cancer mice models to understand molecular mechanisms involved in immunosuppression, the effect of therapies at the protein level as well as to discover novel therapeutic targets for immunotherapy. However, instead of inferring their activity from peripheral blood, further proteomics analysis of TILs will provide more valuable information on T cell functions within the TME.

1.3.3.2. MS-based proteomics application in clinical studies to characterize cancer immune responses

The advancement of shotgun MS-proteomics enables better characterization of TILs in clinical samples. The first step towards this goal was the development of the simple and integrated spin tip-based proteomics technology (termed SISPROT) combined with laser-capture microdissection technology (LCM) [140]. LCM-SISPROT provided spatial proteome profiling of cancer cells, enterocytes, lymphocytes, and smooth muscle cells of both normal and CRC tissue obtained from the same patient. Each cell type possessed an individual proteomic signature such as immune processes enrichment in lymphocytes. Interestingly, the spatial proteomic composition from the same cell type showed expression fluctuations across micrometer spatial distance which highlights the heterogeneity of TME [140].This proof-of-concept study demonstrates the technical advancement towards high-throughput proteomics characterization of TILs. The next step is the application of LCM combined with shotgun proteomics in studies of clinical importance. For instance, this approach has been recently applied to compare the proteomes of microdissected TILs from 3 metastatic melanoma patient samples (IFN-γ-high, lymphocyte

activation gene-3 (LAG-3)-high, and none), showing that only the IFN-γ-high sample was enriched in different inflammatory pathways [141].

It is well known that tumor-secreted factors and exosomes enrich immunosuppressive cells within the tumor-draining lymph nodes, leading to defective local T cell priming [142,143]. Further characterization of the tumor-draining lymph node cellular and protein composition is needed to release T cell inhibition and to develop potential immunotherapy. MS-based proteomics has been recently applied to characterize the pathophysiology of perfused breast cancer patient-derived axillary lymph nodes (ALNs) sustained *ex vivo* using normothermic perfusion [144]. Neutrophil degranulation and extracellular matrix degradation pathways were enriched in metastatic ALNs compared to reactive ALNs. Similar results of enriched pathways were observed in metastatic lymph nodes from pancreatic ductal adenocarcinoma and prostate cancer [145,146]. These studies demonstrate that MS-based proteomics is a powerful tool to characterize biofluids such as perfusates from tissue, facilitating the protein characterization of lymph nodes. MS-based shotgun proteomics analysis has also been applied to study the cellular composition of draining lymph nodes, such as Treg cells from Sentinel Nodes (SN) compared to non-SN Tregs in bladder cancer patients [147]. It was found that SN-resident Tregs were enriched in growth and immune signaling pathways with IL16 playing a central role. Moreover, Treg cells *in vitro* exposition to tumor secretome increased the IL16 processing into its bioactive form through caspase-3 activation, reinforcing Treg suppressive capacity [147].

Currently, MS imaging has been applied to study the protein heterogeneity as well as spatial patterns in multiple solid tumors, focusing on sub-histological classification as well as the discovery of new candidate biomarkers [148–151]. In breast cancer patients' samples, using MS imaging revealed a correlation between high intra-tumor heterogeneity, high levels of TILs, and better prognosis. [152]. These findings suggest that unveiling the proteome heterogeneity is crucial for defining the extent of cellular heterogeneity within the TME. In recent years, MS imaging has been approved as a powerful tool to characterize immune cell population changes and to identify protein signatures in response to immunotherapy. Berghmans et al. [153] used MS imaging to measure anti-PD-L1 immunotherapy response in non-small cell lung cancer patients. Downstream analysis and IHC validation demonstrated that neutrophil defensins-1, -2, and -3 are predictive biomarkers associated with a positive immunotherapy response. Indeed, *in vitro* experiments showed that these defensins activate immune cells against cancer cells. Importantly, MS imaging can be combined with LCM and subsequent bottom-up/top-down proteomics to facilitate the identification of putative proteins within the TME [154,155]. This combination revealed that the proteomes from TME cell subpopulations are associated with unique molecular signatures in breast cancer [154]. This proof-of-concept study demonstrates that the combination of proteomics approaches can reveal TME proteomics heterogeneity.

Top-down proteomics has not been widely applied to cancer immunological research, but several studies exemplify the potential of this technique. Generally, top-down proteomics is combined with bottom-up proteomics or MS imaging. On one hand, top-down/bottom-up proteomics has been used to identify potential biomarkers in prostate cancer [156] and pediatric brain cancers [157–159] as well as to investigate the proteome landscape of breast cancer patient-derived mouse xenograft models [160]. Bottom-up proteomics has a higher coverage of the proteome, while top-down facilitates the identification of proteoforms with specific PTMs. These studies highlight the benefit of the integration of both approaches. On the other hand, a combination of top-down proteomics and MS imaging can identify the spatial patterns of protein products from alternative Open Reading Frames within the TME. This integrative approach can detect potential biomarkers that were not considered before. Interestingly, top-down proteomics also facilitates the identification of protein complexes [161], novel quaternary structures [162], and tumor mutant proteoforms [163].

In summary, MS-based proteomics has been widely applied in cancer immunology research. Studies have approved that novel insights into the current understanding of tumor-mediated immunosuppression have been gained by using these technologies. Systematic untargeted proteome characterization of different T cell subsets and other cell subtypes within the TME, and biofluids will facilitate the discovery of novel biomarkers and therapeutic targets to overcome tumor-mediated suppression of effector T cell activation.

Despite these great advances, several technical challenges must be addressed. MS-proteomics does not provide the full sequence of a protein but rather relies on the identification of unique peptides from a protein. Its sensitivity is limited by the number of acquired spectra to identify a specific peptide [164]. However, an average of 75% of collected spectra can remain unidentified [165]. This lack of sensitivity limits the dynamic range of mass spectrometers as well as the identification of low-abundant proteins, especially in clinical samples, such as serum, in which the dynamic range can overpass 10 orders of magnitude [166]. Once a peptide is correctly identified, another challenge is the identification of different isoforms of the protein, called proteoforms. These proteoforms are generated by posttranscriptional processing and PTMs, yielding multiple proteoforms from the same canonical amino acid sequence [167]. Despite the development of PTM enrichment strategies, the identification of modified peptides arises more complications due to their lower abundance, lower ionization and fragmentation efficiency, inaccurate mass determination, confusion with the assignment of residue substitutions, and uncertainty in the PTM site assignment [168,169]. Lastly, the high cost of MS instrumentation as well as the level of expertise required to perform MS-proteomics hinders its wider usage.

1.3.4. Antibody-based technologies to characterize immune responses in cancer

MS-based proteomics is widely used in discovery proteomics while antibody-based approaches are the most widely chosen for targeted proteomics, although the number of detected proteins is limited. One of the main challenges in cancer immunology is to find novel biomarkers to guide the choice of therapeutic strategies to maximize patient benefit. Predictive

biomarkers for immunotherapy require a more holistic approach with panels of biomarkers to identify the underlying biology and complexity of the tumor immune response [170]. Recently developed antibody-based detection techniques can detect from tens to hundreds of proteins simultaneously, being a powerful tool to identify these panels of biomarkers.

Multiplex immunoassays utilize antibodies as anchors that are immobilized on a solid surface or the surface of beads. In both, the protein of interest is bound to the specific antibody. The technology enables simultaneous detection and quantitation of tens of proteins. It is a powerful tool, especially for the detection of secreted proteins, such as cytokines and growth factors from a limited amount of biological and clinical materials. For example, this technique was applied to study the correlation between 59 serum-derived proteins and response to immunotherapy in gastrointestinal cancers. As a result, protein signatures characterized by higher levels of IC molecules, namely PD-L1, CD28, immunoglobulin and mucin domain 3 (TIM-3), LAG-3, and CTLA4, correlated with better prognosis and higher response, being a promising panel of predictive biomarkers [171]. In addition to detecting proteins from serum or plasma samples, recently, this technique has been applied to characterize inflammation-involved proteins in CRC tumors and matched normal tissues, providing a panel of 32 biomarkers differentially expressed in CRC tumors [172].

Another antibody-based technology, **Proximity Extension Assay** (**PEA**) further extends the number of detected proteins from tens to hundreds and even thousands. The technology is based on target-specific antibodies conjugated with unique complementary DNA. The antibody pairs targeting one protein bind to the target and a barcoded DNA duplex is formed, which is amplified by qPCR or next-generation sequencing (NGS), allowing quantification of up to 3072 proteins [173,174]. In a recent study, the oncology panel of PEA with 92 cancer-related proteins was utilized to identify potential circulating tumor biomarkers for meningioma. The pathway analysis revealed upregulation of immunomodulatory proteins such as CD69, C-C motif chemokine 24 (CCL24), IL24, CCL9, and B-cell activating factor (BAFF) [175]. In another study, the PEA immune-oncology panel was applied to study the serum/plasma proteomic profiles of pancreatic neuroendocrine neoplasm patients. Many well-known immune regulators, such as CCL3, IL7, IL10, CCL20, were significantly elevated in patients compared to healthy controls, whereas FAS ligand (FASLG) was downregulated [176]. The PEA technology has shown a promising potential to detect chemokine variability within metastatic melanoma patients subjected to anti-PD1 therapy [177]. Likewise, it has also been used to assess the immune profile of chronic lymphocytic leukemia patients undergoing different treatments. [178]. PEA analysis of 29 CRC tumors using the immune-oncology panel resulted in only 9 tumors clustered together in unsupervised hierarchical clustering, which revealed the intra-tumor TME heterogeneity [179]. PEA technology possesses a validated specificity and sensitivity (sub-pg/ml) which allows multiplexed protein detection, consuming a minimal amount of sample. Further progress will have a powerful impact on the discovery of new diagnostic, predictive, prognostic, and monitoring biomarkers as well as on the understanding of the proteome of cancer patients [180].

Moreover, other antibody-based proteomics techniques, such as **Reverse Phase Protein Arrays** (**RPPA**) [181] and chip array cDNA-based **Nucleic Acid Programmable Protein Array** (**NAPPA**) [182] have been applied in cancer immunology research. RPPA has been used to correlate the tumor heterogeneity and immune response in melanoma patients [183], while NAPPA to analyze tumor autoantibodies in CRC patients [184]. However, antibody-based approaches are limited by the availability and the specificity of antibodies that implies cross-reactivity. Another disadvantage is the variability between batches, especially when the antibody is produced in a new population of antibody-producing animals [185]. Most importantly, these approaches only detect limited numbers of preselected proteins.

1.3.5. Emerging single-cell proteomics applied to characterize the immune TME

The interplay between cancer cells and their microenvironment plays an important role in many cancer-related biological processes, including progression, metastasis, drug resistance as well as immune response. These complex cellular interactions of the TME and cancer cells are driven by cell heterogeneity [186,187]. Therefore, to develop more effective immune therapies, it is fundamental to understand the interaction between immune and cancer cells. Single-cell protein measurements rather than a conventional bulk analysis can provide more precise information on this heterogeneity. This section reviews the different single-cell proteomics strategies applied or with potential application in cancer immunity and immune cell characterization. The following section includes a short description of antibody-based approaches, MS-based approaches, and multi-omics strategies applied to cancer immunity at the single-cell level.

1.3.5.1. Antibody-based approaches

For the past 30 years, FC has become the 'gold standard' in marker analysis at the single-cell level. Despite its popularity, this method is limited to a low number of markers for simultaneous analysis due to overlapping fluorescence spectra [188,189]. A recently developed modification of traditional FC, **full spectrum flow cytometry (FSFC)** overcomes the issue of overlapping fluorescence spectra of fluorophore-conjugated antibodies, as the detection and measurement include an entire fluorescence spectrum. This enables the simultaneous detection of up to 64 proteins [190]. This technique has been applied to characterize specific cells populations within the TME. For instance, FSFC with over 30 markers found a tumor favorable environment formation caused by arginine-metabolizing myeloid cells co-localized with CD4+ T cells of unconventional phenotype in neuroblastoma mice models [191]. FSFC was applied to characterize the immune cells populations in syngeneic melanoma, breast, ovarian, and CRC cancer models with the focus on Tim-3 as a focal molecule [192]. Comparable higher cytolytic activity of Tim-3+PD1+CD8+ TILs lead researchers to conduct the validation of combined treatment of Tim-3/PD-1 mAbs with indication of an enhanced anti-tumor effect [192].

By the combination of features of FC and MS, **MC (CyTOF)** has been developed to overcome the limitations of simultaneous analysis of up to 100 proteins at the single-cell level. In this method, cells are stained with metal isotope-tagged antibodies and separated in a mass cytometer, followed by TOF analysis of isotopes mass ratio in the analyzed samples. MC has been successfully applied in the study of the immune signature and immune response in cancer and exhibits potential in the discovery of novel cell populations in different types of cancer [193–201]. . For example, MC and RNA-seq analysis of tumor and peripheral blood mononuclear cells (PBMC) of CRC patients revealed that exhausted T cells are induced and recruited by the TME at all stages of the tumor development, demonstrating the link between immunosuppressive TME and the lack of immunotherapy response [193] This study demonstrated the superiority of MC analysis of TME over RNA-seq to characterize the single-cell proteome state. Interestingly, another CyTOF study identified a novel specific population of effector Tregs with protective function in CRC tumors [194]. In glioblastoma (GBM), MC provided data confirming the inter- and intra-tumor heterogeneity of glioma-associated macrophages (GAM). Moreover, the proportion of GAMs was decreased and exhausted T cells and Tregs were increased in recurrent tumors, contributing to an immunosuppressive environment [195].In xenografts GBM models, MC was utilized as a comparative tool of immune landscape between tumor-silent and tumor-active models revealing distinct differences in the cells profiles [196]. Additionally, using cell barcoding in MC enables sample multiplexing which is a very useful option when dealing with valuable clinical samples and low amounts of murine tissue samples. Recently, MC has been successfully applied in high-throughput clinical analysis, where multiple samples have been analyzed with more than 35+ isotope tags [197].

Further advances in antibody-based proteomics utilize the combination of already established antibodies properties and application with microchips or microfluidics to perform proteomic analysis in isolated single cells. **Single-cell barcode chips** (**SCBC**) separates single cells in microchambers and secreted or intracellular proteins are captured on an antibody array. Then, captured proteins undergo the staining and quantification with the corresponding biotinylated antibodies and fluorescent streptavidin [202]. Advances in this technology led to the development of a commercial platform that quantifies a panel of 40 key secreted proteins from a single, viable cell [203]. Among other applications, this platform was used to study the heterogeneity of CD8+ TILs in metastatic melanoma patients [204].

Multiplexed *in situ* targeting (**MIST**) technology uses microbeads hybridized with antibodies conjugated to single-stranded DNA. Once the secreted target proteins are captured, an ELISA assay with the usage of a second, complementary DNA-conjugated antibody is performed [205]. Both technologies, SCBC and MIST, have to compromise the multiplex capacity and detection sensitivity, i.e. increasing the number of different antibodies can increase the multiplexing capacity but, in parallel, decrease the amounts of particular antibodies used, decreasing the sensitivity [206]. **Antibody barcoding with cleavable DNA** (**ABCD**) is the next technology that improves multiplexing capacity by utilizing antibodies linked to a unique DNA barcode via a photocleavable linker. DNA barcodes are released after incubation by UV exposition and are quantified by fluorescence hybridization [207]. Moreover, ABCD allows simultaneous analysis of hundreds of proteins from cancer cells and it was applied to characterize lung cancer cells from minimally invasive fine-needle aspirates [208].

TME heterogeneity does not only rely on the different cell types but also their spatial distribution and cell-cell interactions [209]. Whereas previous techniques analyze proteins in isolated single cells, the next antibody-based strategies are focused on comprehensive protein profiling in their natural spatial contexts. Multiplex immunofluorescence (**mIF)** is based on cycles of antibody staining, imaging, and antibody removal in tissue slides. This method allows the simultaneous identification of several immune markers in the same cell providing data about both the expression and location of target proteins (Figure 1.6a). A combination of tissue microarrays with mIF has been optimized (e.g., for TME immune profiling) [210]. Gerdes et al. [108] applied mIF to analyze 61 proteins in CRC, revealing extensive tumor heterogeneity. Recently, mIF has been used to unveil the immune heterogeneity within the TME of melanoma and breast cancer ALNs [141,144].

Since the specific intracellular localization of the proteins is essential to performing their biological function(s), whilst localization abnormality may severely disrupt biological processes causing disease, characterization of protein expression, as well as its localization in a high resolution, is needed. Single-cell spatial proteomics aims at solving this problem in a comprehensive manner (reviewed in [211] and [212]). An mIF technique called **Multi-Epitope Ligand Cartography** (**MELC**) uses an automated microscopic robot that allows multiplexed protein characterization at subcellular level. In a pioneering work, MELC was applied to identify changes in key immune function-related proteins in CRC tissue at subcellular level [213]. In this study, 1,930 clusters of proteins distinguished CRC from healthy tissue, and CRC tissue was enriched in T cells with altered T cell adhesion and NK cells with high nuclear factor-κB (NF-κB) expression. Later, Bhattacharya et al. [214] used **Toponome Imaging System**, a similar mIF strategy, to compare CRC with a normal colon. 5,708 clusters of proteins that are specific to colon cancer were identified, showing that CRC has a unique higher-order toponomy signature.

Since the application of mIF techniques carries a risk of damaging the epitopes' integrity, oligonucleotide conjugated antibodies alternatives have been explored [215–217]. **CO-Detection by indEXing** (**CODEX**) iteratively visualizes targets through in situ polymerization-based indexing procedure with oligonucleotide-conjugated barcodes and dNTPs analogs tethered to fluorophores (Figure 1.6b) [218]. CODEX has been applied to study the immune TME of CRC with 56 markers, showing the importance of the spatial distribution and cell neighborhoods in CRC [219]. Despite the recent advances in multiplexed analysis, it was found that oligonucleotides negatively affect the specificity and the binding affinity of antibodies. To avoid this interference, other alternatives are used e.g., removable antibodies with fluorophores linked by an azido group [220].

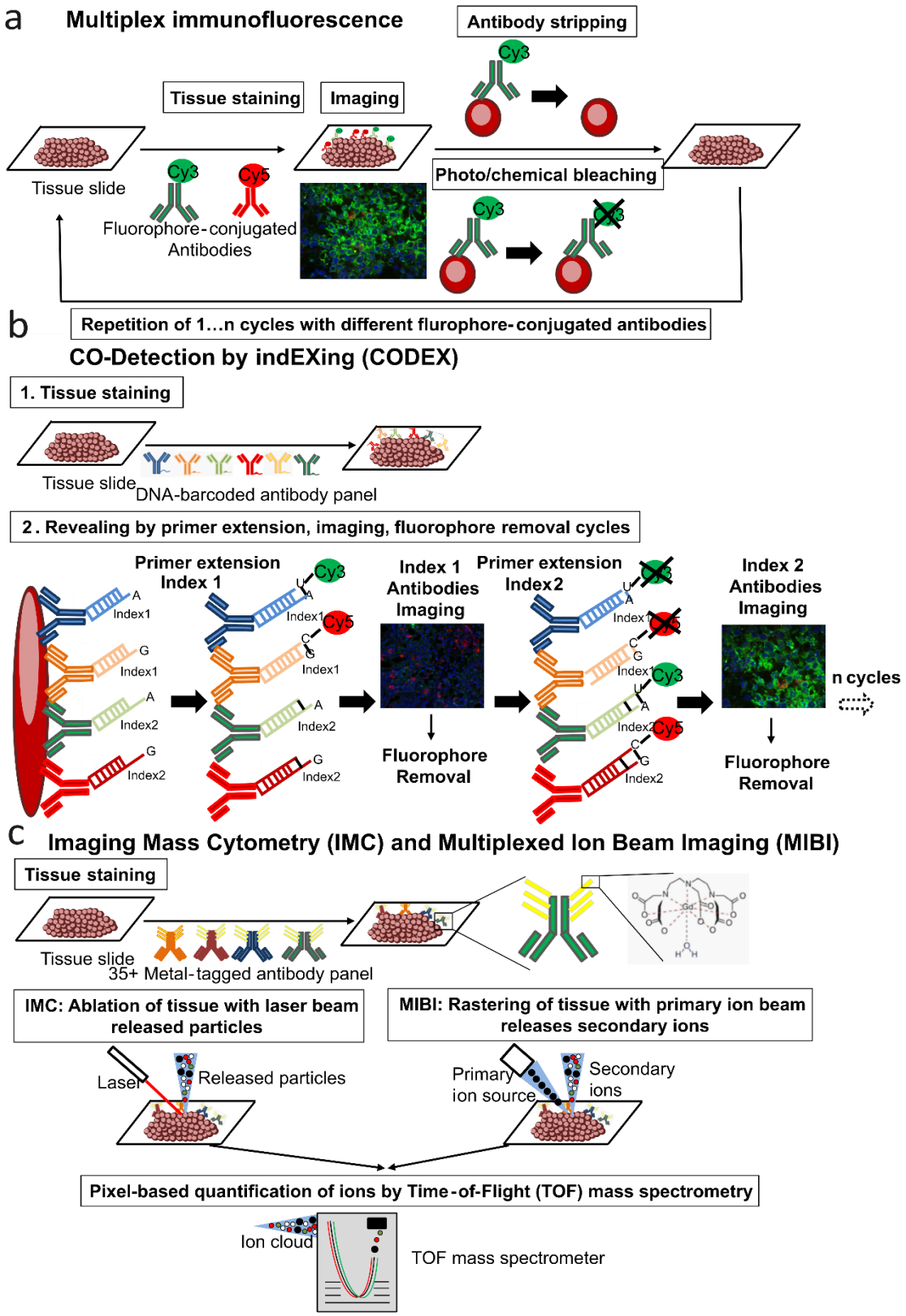

Figure 1.6. Schematic representation of single-cell spatial proteomics approaches. (a) Multiplex immunofluorescence (mIF), (b) CO-Detection by indEXing (CODEX), (c) Imaging Mass Cytometry (IMC) and Multiplexed Ion Beam Imaging (MIBI).

In the context of cancer immunology, **imaging mass cytometry (IMC)** and **multiplexed ion beam imaging (MIBI)** are powerful tools to assess the complexity of the TME and networks of cell-cell interactions in their spatial context within the tissue. IMC is a technology that combines CyTOF (MC) and imaging to analyze proteins *in situ* (Figure 1.6c). First, the tissue slide is stained with a panel of metal conjugated antibodies and then the stained tissue is converted to a stream of particles pixel-to-pixel by a laser. Next, the mass spectrometer determines and quantifies the metal isotopes linked to the antibodies in each particle and, finally, a computational algorithm combines the MS data of each pixel with its coordination information to generate a two-dimensional image [221]. IMC not only provides information on single-cell proteomics but also on the localization of the particular protein in the tissue and construct the cellular interaction within the TME. This methodology gives additional data potentially relevant in the context of prognosis or treatment. IMC analysis with 35 biomarkers of patients' breast tumors samples, together with available survival data, yielded high-dimensional images providing information on the complexity of organization of tumor and stromal cells, their location within the tissue, and distinct phenotypes of tumor cells. This study led to the proposal of novel breast cancer subgroups closely related to the particular patient's prognosis [222]. IMC was also used to explore the TME of different cancer types including Hodgkin lymphoma, and CRC, in which tertiary lymphoid structures in CRC were found to have abundant forkhead box P3 (FoxP3)+ Treg expression, demonstrating its potential for immune profiling in tumors [223].

The second technology, MIBI is a variation of IMC which operates an ion beam to release metal ion reporters, therefore increasing its multiplexing capacity to more than 100 targets at once [224]. An interesting application of MIBI is single-cell metabolic regulome profiling, which enables to study the composition of the metabolic regulome in combination with phenotypic identity with more than 110 antibodies against metabolite transporters, metabolic enzymes, or regulatory modifications. The study revealed the metabolic heterogeneity and spatial organization of CD8+ T cells in CRC, including subsets expressing the T cell exhaustion-associated molecules CD39 and PD-1, indicating their exclusion from the tumor-immune boundary [225]. Undeniably, IMC and MIBI are superior methods to fluorescence-based technologies because they detect simultaneously targeted proteins with a higher dynamic range, avoiding staining/stripping cycles that can compromise epitope integrity [226]. However, their disadvantage is the availability of the number of antibodies conjugated with metal isotopes suitable for FFPE and fresh frozen tissue staining [227]. In summary, the bottleneck of single-cell measurements with antibodies is the limit of sensitivity, which stems from the molecular shot noise, limiting accurate quantification to the low attomolar (aM) range, as well as the quality of the antibody [228].

1.3.5.2. Single-cell MS-based approaches

Unbiased single-cell MS-based proteomics approaches are currently in development, being a promising alternative that can overcome the limitation of antibody-based approaches, potentially leading to an increased number of detected proteins [229]. However, single-cell MS analysis must overcome additional challenges apart from the abovementioned for bulk MS proteomics. Proteins cannot be amplified as nucleic acids. Thus, one of the major challenges is the delivery of peptides to the mass spectrometer taking into account the low protein content of a single cell. Single-cell sample preparation requires miniaturization and automation to reduce protein losses and increases their concentration [230]. Single cells are separated by FACS or other alternative techniques and subsequently, protein extraction and digestion are performed in reduced volumes (1µl/cell or lower). Different strategies of sample preparation have been successfully developed such as nanodroplet processing in one pot for trace samples (nanoPOTS) [231], oILair droplets [232], or Minimal ProteOmic sample Preparation (mPOP) based on freeze-heat cycles [233]. Moreover, peptide separation in the LC column and its corresponding ESI must be miniaturized with flow rates at low-nanoliter-per-minute or even picoliter-per-minute range. Therefore, the inner diameter of nanoLC columns is reduced from 75 µm to 30 µm which in consequence improves single-cell proteome coverage [234].

Importantly, single-cell MS analysis needs an increase in peptide sequence identification as well as its multiplexing capacity to analyze the proteome from thousands of cells at an affordable cost [229]. A great advance has recently been achieved with an approach called **Single Cell ProtEomics by mass spectrometry** (**SCoPE-MS**) [235]. SCoPE-MS prepares the sample by mPOP and adds an isobarically labeled carrier (e.g., the proteome of 100 cells) with tandem mass tags [236]. The usage of a proteomic carrier mitigates sample losses, facilitates peptide sequence identification, and increases the multiplexing capacity with a limit of 12 single-cell proteomes in one run due to the limited tandem mass tags available. With such technological development, SCoPE-MS found its application in heterogeneity studies. SCoPE-MS quantified 3,042 proteins in 1,490 single monocytes and macrophages, suggesting that heterogeneity of macrophages may emerge without the participation of polarizing cytokines [237]. Moreover, SCoPE-MS quantified 1,500 proteins from 152 cells from three acute myeloid leukemia (AML) cell lines, revealing functionally distinct differences between the three cell clusters [238,239]. Moreover, the combination of nanoPOTS and SCoPE-MS quantified around 1,000 proteins per cell of 3,000 FACS-sorted cells from an AML culture model. It allowed resolving AML heterogeneity at a single-cell level along different hierarchical stages of differentiation [240].

Further improvements will be achieved through innovations in sample preparation and peptide separation, hardware advances of mass spectrometers as well as innovative acquisition and interpretation methods. These improvements will facilitate increased coverage of single-cell proteomes as well as the sensitivity and confidence of peptide sequence identification, revolutionizing cancer immunology [241].

1.3.6.5. Single-cell multi-omics strategies

For precision oncology, to deeply and comprehensively understand the complexity of the TME, in addition to proteomics, an integration of multi-omics data at the individual cell level with the molecular landscape of each cell is needed [242,243].

Proteogenomics approaches combine bulk MS-based proteomics with genomics and transcriptomics. This strategy has been applied to several cancer types, providing novel insights into somatic mutation consequences at the protein level as well as neoantigens discovery for immunotherapy [244–247]. However, the genomic and proteomic data integration at the single-cell level is currently in development. Recently, a pioneering study designed DAb-seq which allows analysis of 49 DNA targets and 23 protein markers by the combination of DNA barcodes conjugated to antibodies and multiplex PCR. Although this technology requires an increase in its multiplexing capacity, it demonstrated the heterogeneous interactions of somatic mutations and protein expression in AML single cells [248].

On the other hand, there are some techniques designed to link mRNA and antibody based protein analysis in single cells approaches. **Proximity Ligation Assay for RNA** (**PLAYR**) is a method that uses FC/MC for simultaneous analysis of target proteins stained with antibodies and RNA. PLAYR probe pairs hybridize their targets and then the insert and backbone are hybridized and ligated to the probes. After rolling circle amplification, labeled oligonucleotides bind the insert regions for detection and quantification [249]. Recently, this method has been used to demonstrate intra-clonal heterogeneity in chronic lymphocytic leukemia cells [250]. **Cellular indexing of transcriptomes and epitopes by sequencing** (**CITE-seq**) and its sister technology **RNA expression and protein sequencing** (**REAP-seq**) combine DNA-conjugated antibodies with scRNA-seq [251,252]. The difference is that CITE-seq uses biotinylated antibodies whereas REAP-seq uses antibodies covalently bonded to aminated DNA sequences. These methods integrate cellular surface protein and transcriptome measurements into single-cell readout. CITE-seq provides a more detailed characterization of cellular phenotypes compared to scRNA-seq alone and allows simultaneous protein expression and transcriptome profiling of thousands of single cells (Figure 1.7). CITE-seq may also show quantitative differences in marker expression between subsets e.g., expression difference of CD8a between NK and T cells [252]. A CITE-seq panel of 157 antibodies was applied to immunophenotype breast cancer patients. 18 clusters of T cells and innate lymphoid cells (ILCs) were found with different proportions among clinical subtypes. Interestingly, IC molecules were also differentially expressed among breast cancer subtypes. These findings may lead to personalized immunotherapy strategies for each subtype [253]. Moreover, it was found that CITE-seq can be combined with single-cell sequencing assay for transposase-accessible chromatin (scATAC-seq) and used to study the RNA expression, surface proteins, and chromatin accessibility at the single-cell level. Granja et al. [254] applied such a strategy to find distinct and shared molecular mechanisms of leukemia. Among the challenges for both technologies (CITE-seq and REAP-seq), the efficiency of cell captures must be increased, the system requires total automation, and the multiplex detection must be extended to intracellular proteins which is currently limited to a reduced number of proteins [255,256]. Recently, **SUrface-protein Glycan And RNA-seq** (**SUGAR-seq**) has been designed to enable the detection and analysis of N-linked glycosylation, extracellular epitopes, and the transcriptome at the single-cell level. SUGAR-seq is an extension of CITE-seq in which glycans are captured with a biotinylated lectin and subsequently detected using an anti-biotin mAb conjugated to a DNA-barcode. Integrated SUGAR-seq and glycoproteome analysis identified TILs with unique N-glycan profiles as cellular T cell subsets with the altered epigenetic and functional state in CRC and melanoma mice models [257].

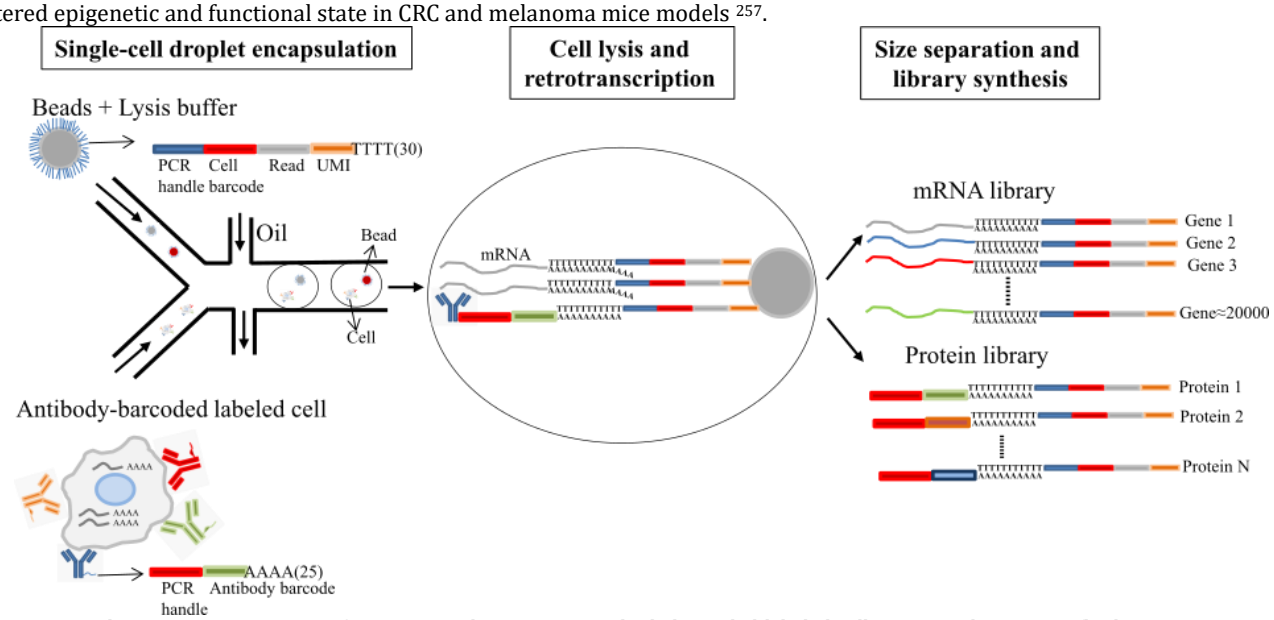

Figure 1.7. Schematic representation of CITE-seq and REAP-seq. Antibody-barcoded labeled cells are mixed in a microfluidic system in which each droplet contains a cell, beads with the PCR adapters with the corresponding cell barcodes, and lysis buffer. After cell lysis within the droplet, mRNA and DNA barcodes from antibodies are hybridized with PCR adapters. Subsequent retrotranscription generates cDNAs and droplets are disrupted. Upon disruption, the respective cDNAs for mRNAs and proteins are separated by size. These synthesized libraries are sequenced, providing the single-cell expression profiles of mRNA and targeted proteins.

Zhang et al. [258] combined scRNA-seq and mIF to study the immune TME of CRC patients. They found that TILs showed an exhausted phenotype compared to T cells originating from normal tissue and peripheral blood. Moreover, they identified a population of Th1-like cells that were enriched in microsatellite instability (MSI) CRC, providing a possible explanation for

MSI patients' good response to anti-PD-1 immunotherapy. Finally, de Vries et al. [259] combined MC with 36 markers, FC, scRNA-seq, and mIF to analyze T cells from CRC, matched associated lymph nodes, healthy mucosa, and peripheral blood. Different phenotypes of CD8+/γδ T cell and CD4+ memory T cells were observed in each examined tissue. Interestingly, an innate lymphoid cell (ILC) population was enriched in CRC tissues with high expression of cytotoxic molecules. Additionally, this ILC population correlated with the presence of tumor-resident cytotoxic, helper, and γδ T cells with similar activated profiles. This study not only sheds some light on the complexity of lymphocytes composition dependent on the sample type but also demonstrates that multi-omics data integration provides much more data and in-depth analysis, which otherwise would not be obtained.

1.3.6. Conclusions and future perspectives

Despite the great advances in cancer immunology and the development of immunotherapy, the patients' response rate remains a clinical challenge. Understanding the complexity of TME and immunosuppression mechanisms may lead to design of more effective cancer immunotherapies. Proteomics is a powerful approach to accelerate the studies on immune responses in cancer. MS-based proteomics can uncover novel insights into molecular mechanisms and potential therapeutic targets, while the application of antibody-based proteomics approaches does not require specialized expertise as in MS and is widely applied as a tool to characterize selected proteins and discover new clinical biomarkers. However, both approaches possess limitations and technical challenges that complicate the characterization of the whole proteome of biological systems, especially to differentiate between proteoforms.

Emerging single-cell proteomics approaches will revolutionize our understanding of the complex cellular networks within the TME and interactions between cancer and immune cells. Several technologies have been recently developed with the potential for comprehensive proteomic characterization that facilitates the deep profiling of immune responses in cancer at the single-cell level. Novel technical solutions will provide higher sensitivity and higher resolution at the subcellular and molecular level [260–262]. Importantly, a new era in proteomics was born with single-molecule protein sequencing based on fluorescence-mediated *in situ* protein identification [263,264] as well as nanopores [265,266]. Further technical development of these next-generation proteomics approaches will ideally enable the whole proteome characterization and unveil the distribution of proteoforms at the single-cell level.

In summary, together with the technological advancements in single-cell analysis, progress in a holistic system of multi-omics data analysis and discovery is needed. To date, it was found that a combination of different 'omics' data with single-cell proteomics, may provide information on cancer origin, progression, and prognosis, which could remain undiscovered if were analyzed separately. It is well-recognized that a comprehensive approach to TME composition is crucial in personalized therapy and efficient treatment. In this review, we have discussed examples of immune heterogeneity studies of TME in cancer, focusing on both MS-based bulk/antibody-based and single-cell analysis. Moreover, we reviewed emerging single-cell proteomic analysis methods with examples of the combination of multi-omics studies, which we believe become widely applied in cancer research in the future.

1.4. LC-MS/MS-based label-free shotgun proteomics analysis

As previously presented in section 1.2.3, there are multiple antibody and MS-based proteomics technologies. Within MS-based technologies, this section will focus on the technical aspects including MS instrumentation and label-free quantitative strategies from shotgun proteomics analysis by LC-MS/MS applied in this thesis. In biomedical research, sample preparation of proteins extracted from different sources such as cells, tissues, or biofluids, resulting in digested peptides is followed by LC-MS/MS analysis.

1.4.1. LC-MS/MS

Several separation methods of peptides are applied in proteomics analysis including LC and capillary electrophoresis. Peptide mixtures are injected into the LC system in which peptides are separated by their molecular properties. The most widely used is reversed-phase liquid chromatography (RPLC) in which peptide mixtures are separated based on their hydrophobicity in C18 columns as stationary phase by increasing gradients of acetonitrile in acidic pH as mobile phase at micro/nanoflow. In this way, peptides are separated to minimize their co-elution and directly injected into the mass spectrometer [267].

Mass spectrometry analysis is based on the measurement of mass-to-charge ratios (m/z) of ions in gas phase. Mass spectrometers are composed of three main sections, ion source, mass analyzer, and detector [268]. In LC-MS/MS, electrospray ionization (ESI) is the most common ion source in which high voltage produces dispersion of charge droplets that decrease their size by evaporation until gas-phase ions are ejected and transferred to the mass spectrometer [269].Several types of mass analyzers are available and mainly separate ions according to their m/z ratios by different techniques. These separated ions are transferred to the detector, such as electron multipliers, that capture and amplify the ion current signals. In tandem mass spectrometry (MS/MS), the precursor ions introduced in the mass spectrometer are selected and consecutively fragmented into product ions that are analyzed [270]. Two main hybrid high-resolution accurate-mass mass spectrometers, quadrupole-time of flight (Q-TOF) and quadrupole-orbitrap (Q-OT) are mainly used in untargeted LC-MS/MS proteomics, that were applied in this thesis, although there are other combinations and mass spectrometers [271]. In both instruments, quadrupoles consist of four parallel cylindrical rods of hyperbolic or cylindrical cross-section in which precursor ions are selected

according to their m/z ratio through the central axis of quadrupoles. Selected ions achieve a stable trajectory through this axis while other ions are neutralized by hitting the rods by application of sets of direct current and radiofrequency within parallel rods while adjacent rods have an opposite radiofrequency [272].

In Q-TOF, the first quadrupole provides improved quality of precursor ions by collisional damping, the second selects ions, and the third quadrupole performs collision-induced dissociation (CID) with neutral gases to fragment precursor ions. Resulting product ions and remaining precursor ions are transferredto a reflectron TOF mass analyzer in which ions are separated by their velocity along TOF and kinetic energy by the ion mirror allowing the determination of both precursor and fragment ions required to provide peptide amino acid sequences [273]. In Q-OT, ESI-derived precursor ions are selected in a quadrupole after collisional damping and transferred to a C-trap that is connected to a High Collision Dissociation (HCD) cell and the orbitrap mass analyzer. Selected ions are fragmented in the HCD cell and product ions are sent to the Orbitrap by the C-trap. In peptide fragmentation by CID and HCD, precursor peptide ions are fragmented in different amide bonds, resulting mainly in y- (C-terminal charged) and b-ions (N-terminal charged) that allow to determine the peptide sequence (next section) [268]. Orbitrap analyzer is an ion trapping based on electrostatic fields applied through its outer and inner spindle shape electrodes [274]. Then ions are detected and converted by Fourier transformation into the frequency domain and then mass spectra [275]. Another mass spectrometer is the tribrid orbitrap that combines the Q-OT structure with an additional linear ion trap for parallelization of mass spectra acquisition, increasing scanning rates of MS1 spectra from precursor ions and MS2 spectra from product ions [276]. Last technological and software acquisition advances resulted in higher resolving power with a full width at half-maximum (FWHM) of 500000 at 200 m/z, a measurement of peak resolution, and high mass accuracy (<1 ppm) [277].

1.4.2. Data Dependent Acquisition (DDA) and Data Independent Acquisition (DIA)

Untargeted LC-MS/MS proteomics analysis is divided into two main quantification strategies: Data Dependent Acquisition (DDA) and Data Independent Acquisition (DIA). Meanwhile, targeted proteomic analysis includes selective/multiple-reaction monitoring (SRM/MRM) in Q-TOF and parallel-reaction monitoring (PRM) in Q-OT with high reproducibility, selectivity, and sensitivity for limited numbers of peptides [278]. In this thesis, DDA and DIA label-free strategies were applied, however, there are also labeling quantification strategies based on stable isotope–enriched labels such as Tandem Mass Tags (TMT) and isobaric tags for relative and absolute quantification (iTRAQ) or metabolic labeling such as Stable isotope labeling using amino acids in cell culture (SILAC) to improve quantification reproducibility and increase multiplexing capabilities [279].

DDA consists of iterative acquisition cycles of MS1 and MS2 spectra from selected peptides in real time for each MS survey, normally the top N most abundant ions, together with dynamic exclusion of already selected precursor ions for a certain time. Selected peptides for fragmentation are filtered in a narrow m/z window to record the spectrum from individual peptides instead of co-eluted peptides. This strategy results in large amounts of high-quality MS2 spectra allowing to the identification of thousands of peptides in a single run. However, the selection of precursor ions is stochastic and detection of low precursor ions do not reach the mass spectrometer limit of detection [280]. In general, DDA data is analyzed by spectrum-centric analysis using sequence database searching software such as MaxQuant [281] or SEQUEST [282] in which algorithms evaluate matches between *in silico* MS2 spectra of the theoretically peptide fragments from a proteome sequence database and each recorded spectrum that generates Peptide-Spectrum Matches (PSMs). Database search is including artificial decoy peptide sequences that are followed by control of False Discovery Rate (FDR) calculated as the ratio between decoy matches and reported target matches [283]. Other strategies include *de novo* sequencing that predicts peptide identity without database support and MS library searches in which recorded spectra are matched to previously recorded experimental spectra as well as combinations of them with database search [284].

DIA strategy, also called Sequential Windowed acquisition of All THeoretical fragment ion Mass Spectra (SWATH-MS) in Q-TOF analysis, combines the high-throughput from DDA and robust quantification of targeted approaches with a different acquisition scheme. In iterative cycles, precursor ions are separated in successive m/z windows of 5-25 m/z ranges in which all the co-isolated peptide ions are fragmented, resulting in complex MS2 spectra [278]. Consequently, all the peptides above the detection threshold are theoretically identifiable but MS2 spectra interpretation is complicated. Peptide identification is based on peptide-centric analysis in which bona fide MS2 spectral libraries containing m/z and intensities of fragment ions and LC Retention Times (RT) for each peptide are used to extract fragment ion traces in target MS2 data and assess their coelution quality to infer peptide identification [284]. Similarly to spectral-centric analysis, target-decoy FDR correction is performed by scoring the chromatographic peaks of identified fragmented peptides. Primarily, MS2 spectral libraries are built by extensive pre-fractionation of pooled queried samples in DDA to guarantee the library quality, which is essential for DIA [284]. Recent advances in DIA informatics analysis allow for "library-free" approaches in which *in silico* libraries are generated from proteome databases by deep learning approaches such as DIA-NN as well as DIA MS2 spectral library creation and consequently hybrid DIA+DDA libraries such as FragPipe to improve protein identification and quantification [285–287].

CHAPTER 2. Aims of the thesis

Inflammation is the most relevant contributor to several diseases. Chronic inflammatory diseases, and cancer, are considered the most significant causes of death and their prevalence is increasing. Recently, there were great advances in treatment and diagnosis such as cancer immunotherapies that are a milestone in cancer treatment However, a majority of patients exhibit low responses due to the influence of other non-targeted inflammatory/immunosuppressive mechanisms. Therefore, deeper understanding of pathogenic molecular mechanisms underlying the immune responses is urgently needed. An imbalance between regulatory and inflammatory CD4+ T-cell populations and other immune cells plays an essential role in chronic inflammatory diseases and cancer [288]. Although several studies have applied transcriptomics with remarkable contributions to biomedical research, poor levels of concordance between changes in protein abundance and mRNA expression have been reported [99]. Proteins are well-known critical cell effectors and proteomics is revolutionizing molecular biology and clinical research with robust protein expression profiling considering post-translational regulation, interactions, and subcellular localization [289]. This thesis aims to identify and characterize protein changes associated with chronic inflammation and cancer contexts by proteomics approaches, especially focused on immune-related proteins related to CD4+ T cell subsets and other immune cells. We speculate that proteomic characterization will facilitate biomarkers discovery as well as the identification of novel regulators of T-cell-driven immune responses. These novel regulators may be potential immunotherapy targets for chronic inflammatory diseases and cancer treatment, diagnosis, and prognosis.

SARS-CoV-2 infection causes an exacerbated immune response and acute inflammation with aggravating effects in COVID-19 pathogenesis. Patients with chronic inflammatory diseases have a higher risk of developing severe symptoms after SARS-CoV-2 infection and increased mortality. Despite the great advances in SARS-CoV-2 immunopathogenesis since the COVID-19 outbreak, further understanding of SARS-CoV-2 immune responses and biomarkers are needed. In collaboration with Dr. Chen team from University of Oulu, I had the opportunity to take part in a SARS-CoV-2 project that resulted in two publications (Chapters 3 and 4). The aims of this first part of the thesis are:

1. Set up and optimization of sample preparation protocols for LC-MS/MS proteomics analysis and bioinformatics analysis of LC-MS/MS and PEA data.
2. Plasma proteomics characterization of COVID-19 patients with/without pre-existing chronic inflammatory diseases by two orthogonal technologies (LC-MS/MS and PEA)
3. Determine plasma protein changes associated with COVID-19 infection and protein signatures linked to COVID-19 response in patients with comorbidities
4. Characterize plasma protein changes related to time of infection and generation of SARS-CoV-2-specific antibodies

Immune-related proteins derived from CD4+ T cells and other immune cells are essential players in cancer-associated inflammation and in CRC tumorigenesis, progression, and therapy resistance.Once sample preparation and bioinformatic analysis were established in the COVID-19 study, the second and main part of this thesis aims to:

1. Plasma proteomics characterization of CRC patients and healthy controls by LC-MS/MS and PEA (Chapter 5 and 6)
2. Determine plasma protein changes caused by cancer-associated inflammation and cancer progression
3. Validation of potential biomarkers in larger CRC cohorts
4. Proteomics characterization of CRC and normal-matched tissue enriched in CD4+ T cell infiltration and other immune cells (Chapter 7)
5. Determine protein changes linked to CRC progression and infiltration
6. Validate potential immune regulators in public datasets

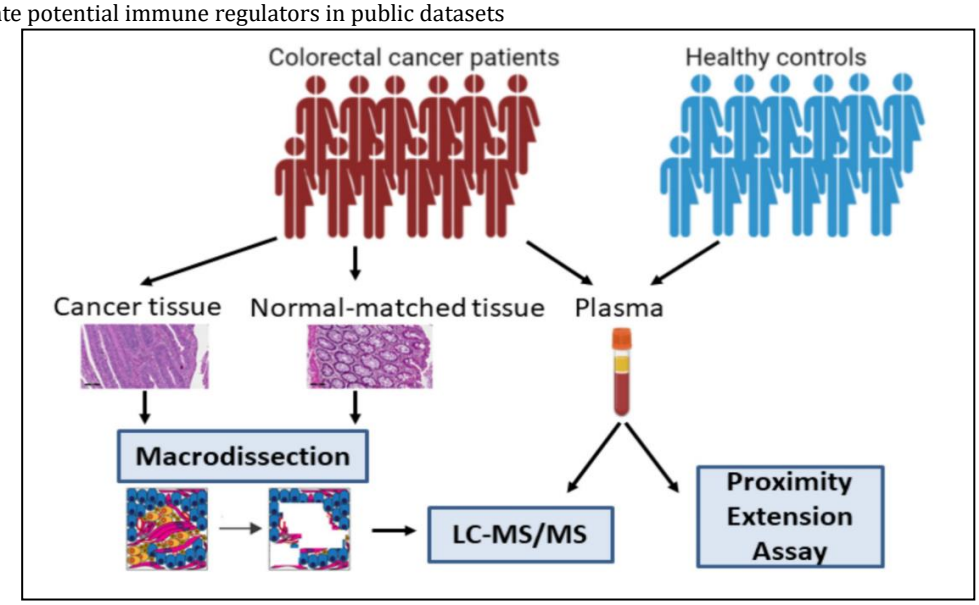

Figure 2.1. General overview of the second main part of the thesis.

CHAPTER 3. Plasma Proteomics Elucidated a Protein Signature in COVID-19 Patients with Comorbidities and Early-Diagnosis Biomarkers

SARS-CoV-2 infection can produce severe symptoms characterized by systemic inflammation and an exacerbated immune response. Moreover, COVID-19 patients with pre-existing chronic inflammatory diseases have a higher risk of developing pathological complications. In this chapter, untargeted LC-MS/MS proteomics analysis was applied to plasma samples from SARS-CoV-2 infected patients with and without pre-existing comorbidities together with their age-and-sex matched HCs and disease controls to characterize the protein changes caused by SARS-CoV-2 infection. This study resulted in an article published in **Biomedicines**in collaboration with the University of Oulu and the IFB laboratory of mass spectrometry and is presented with minor modifications:

***Urbiola-Salvador V*, Lima de Souza S, Macur K, Czaplewska P, Chen Z. Plasma Proteomics Elucidated a Protein Signature in COVID-19 Patients with Comorbidities and Early-Diagnosis Biomarkers. *Biomedicines.* 12, 840 (2024).**

3.1. Introduction

Since the COVID-19 outbreak, extensive research efforts have improved our understanding of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection pathogenesis as well as diagnostics, treatment and prevention with the effective design of vaccines [290]. The pathophysiology of SARS-CoV-2 infection is closely interconnected with the immune response resulting in diverse clinical presentations from asymptomatic/mild to severe patients with high mortality rates. In fact, SARS-CoV-2-induced tissue damage recruits immune cells causing a local and systemic inflammatory response, also called cytokine storm, which in severe cases lead to pneumonia, microthrombi deposition, systemic symptoms and multi-organ failure in fatal cases .With the virus evolution and new emerging variants, continuous research must characterize the interaction between SARS-CoV-2 infection and the immune response [291].

The heterogeneity in the COVID-19 response is caused by the clinical variability including multiple factors such as sex, age, and ethnicity as well as pre-existing comorbidities such as cardiovascular diseases, cancer, diabetes, and obesity among others with higher severity and mortality rates [45]. Despite the great advances in COVID-19 research, deeper understanding of COVID-19 immunopathology especially in patients with pre-existing comorbidities and the determination of specific COVID-19 biomarkers are urgently needed [292]. Importantly, proteomics approaches can provide novel insights into the protein changes caused by SARS-CoV-2 infection at cellular and systemic level to identify the drivers of the pathogenesis [293]. In fact, proteomics characterization of COVID-19 patients' plasma is a non-invasive strategy that can reflect the disease status and determine potential biomarkers associated with the pathophysiological mechanisms of SARS-CoV-2 infection.

In this study, we applied tandem mass spectrometry coupled with liquid chromatography (LC-MS/MS) proteomics analysis to plasma samples from 28 SARS-CoV-2 infected patients with and without pre-existing comorbidities, age-and-sex-matched healthy controls (HCs) and disease controls (DCs) from the Finnish Clinical Biobank. The main aims of this study was to characterize plasma protein changes associated with or without the presence of comorbidities and time of infection. We found a common protein signature among COVID-19 patients with comorbidities characterized by protein alterations involved in the coagulation and complement pathways, tissue damage and remodeling, acute-phase reaction as well as cholesterol metabolism among others. Moreover, novel potential diagnostic biomarkers of early SARS-CoV-2 infection were detected including the keratin K22E, the extracellular matrix protein ECM1, and the acute-phase response protein α-2-antiplasmin (A2AP) among others. This study determined novel insights into the plasma protein changes caused by SARS-CoV-2 infection that may lead to the validation of these potential biomarkers in further studies.

3.2. Materials and Methods

3.2.1. Study Cohort

This retrospective study included 28 plasma samples collected from SARS-CoV-2 virus-infected patients (25% males, age range 19-79). Nineteen of them had other pre-existing diseases, while the remaining nine were without comorbidities. Two control groups included 28 SARS-CoV-2 virus-negative healthy subjects without any diseases and the disease control group including 20 SARS-CoV-2 virus-negative subjects with additional diseases. The 19 SARS-CoV-2 patients with additional diseases were age-and-sex-matched with 19 healthy controls as well as with 20 disease controls, in the case of which the matching considered the major disease(s) as well. Two patients with comorbidities were age-and-sex-matched to two disease controls with the same clinical conditions, while for another patient, a disease control was not found. The remaining 9 comorbidities-free SARS-CoV-2 patients were age-and-sex-matched to their corresponding 9 healthy controls (Table 3.1). Plasma samples and clinical information were obtained from the Finnish Clinical Biobank, Tampere. The study was conducted in accordance with the Declaration of Helsinki and was approved by the HUS ethics committee. All the participants provided informed consent.

3.2.2. Sample Preparation for Mass Spectrometry

Proteins were extracted from plasma samples with lysis buffer containing 1% SDS, 50 mM DTT in 100 mM Tris-HCl pH 8.0 with protease and phosphatase inhibitors. Samples were incubated at 95 °C for 10 min and protein concentrations were

determined in NanoDrop 2000 at 280 nm. Samples were processed according to the Filter Aided Sample Preparation protocol [110]. Briefly, 100 µg of proteins were transferred to Microcon 10 kDa filters (Merck) and were washed three times with 200 µl of urea buffer (8 M urea in 100 mM Tris-HCl pH 8.5) at 10000 rcf for 20 min at room temperature. Free cysteines were alkylated by incubation in the darkness for 20 min at room temperature with 100 µl of 55 mM iodoacetamide in urea buffer. Samples were centrifuged at 10000 rcf for 15 minutes followed by three washes with 100 µl of urea and two final washes with 100 µl of digestion buffer (50 mM Tris-HCl pH 8.5). After washing steps, proteins were digested by incubation at 37°C overnight with Sequencing Grade Modified Trypsin (Promega) at a trypsin:protein ratio 1:100 in 60 µl of digestion buffer. Peptides were eluted in a new collection tube by centrifugation and with two additional elutions with 125 and 100 µl of digestion buffer. Next, trypsin activity was quenched with a final concentration of 0.1% trifluoroacetic acid. Peptide concentrations were measured in Nanodrop at 280 nm and 10 µg of peptides were desalted via the STop And Go Extraction (STAGE) Tips protocol [294] using Empore C18 extraction disks (3M) with elution by 60% acetonitrile/1% acetic acid solution. Samples were dried using SpeedVac and stored at -20°C until analysis.

3.2.3. Mass Spectrometry Analysis

LC-MS/MS analysis was performed in the positive ion mode using a TripleTOF 5600+ mass spectrometer equipped with TurboV Ion Source (SCIEX, Framingham, MA) and coupled with the EkspertMicroLC 200 Plus System (Eksigent, Redwood City, CA). SCIEX Analyst TF 1.8.1 software controlled the microLC-MS/MS system. Two µg of peptides were injected per technical replicate and the chromatographic separations were performed with a 5 µl/min flow for 60 min on a ChromXP C18CL column (3 µm, 120 Å, 150 × 0.3 mm) placed in a column oven at 35 °C. Peptides were separated with a gradient from 11% to 35% of acetonitrile in 0.1% formic acid. The mass spectrometer operated in data-dependent acquisition mode and the m/z range of 400-1200 Da was applied for the TOF MS survey scan with an accumulation time of 250 ms. A maximum of top 20 precursor ions with charges between +2 and +5 were selected for collision-induced dissociation (CID) fragmentation with rolling collection energy. Precursor ions were excluded from reselection for 5 s after two occurrences. Product ions spectra were acquired in the range of 100-1800 Da within an accumulation time of 50 ms.

3.2.4. Mass Spectrometry Data Analysis

Acquired raw files were converted with MSConvertGUI 3.0 to mzML format to use as input for protein identification and quantification analysis using PeaksStudio Xpro 10.6 software. Peptide sequences were searched against *Homo sapiens* UniProtKB/Swiss-Prot database (release 2022_03) for peptides with specific trypsin digestion and a maximum of 3 missed cleavages per peptide. Carbamidomethylation was set as fixed post-translational modification (PTM), whereas N-terminal acetylation and methionine oxidation as variable PTMs. Peptides and proteins were identified with a < 1% false discovery rate (FDR) and proteins were considered identified with at least 1 significant unique peptide. Label-free quantification was performed based on the integration of the area under the curve (AUC) of peptides with the use of label-free quantification feature available in PeaksStudio Xpro 10.6 software.

3.2.5. Proteomics Data and Statistical Analysis

Statistical analysis was performed in RStudio (version 1.3.1093) using R (version 4.0.3) [295,296]. Peptide results from PeaksStudio software were used for data preprocessing with the "SummarizedExperiment" (version 1.28.0) and "QFeatures" (version 1.8.0) R packages. Peptides that were only detected twice across all samples were removed. Data preprocessing was performed by logarithmic transformation and quantile normalization. Peptides were aggregated in proteins by robust summarization [297]. Quantification reproducibility among technical replicates was evaluated by Pearson correlation. Between-group differences in protein expression levels were analyzed by means of the general linear model regression approach with analysis of contrasts using the "emmeans" R package (version 1.6.2.1). First, 19 SARS-CoV 2 virus-infected patients with comorbidities were involved to search for differential protein expression due to virus infection and/or coexistent comorbidities by comparing the protein expression levels in virus-infected patients to respective healthy and disease control subjects. The models used in this analysis comprised the presence of comorbidities as nested confounding factor. While the remaining nine comorbidities-free patients were compared to their age-and-sex-matched healthy controls to determine protein expression changes due to the virus infection itself without additional confounders in these models. In the last regression modeling, all 28 SARS-CoV-2 virus-positive patients' samples were analyzed together to search for differential protein expression due to infection time including the presence of comorbidities as confounding factor. False Discovery Rate in contrast analyses was controlled by using the Benjamini & Hochberg correction [298]. Proteins were considered differentially expressed with an FDR adjusted p-value < 0.05. Volcano plots were generated with the R package "ggplot2" (version 3.3.5) and the heatmap with the R package "ComplexHeatmap" (version 2.6.2). KEGG pathway enrichment analysis via active subnetworks from the STRING database was performed with the R package "pathfindR" (version 1.6.3) with FDR correction [299]. No custom code was used in this analysis.

3.3. Results

3.3.1. Proteomic Profiles of Plasma from COVID-19 Patients and Their Controls

We applied LC-MS/MS proteomics analysis of plasma samples from 28 COVID-19 patients and their corresponding sex-and-age-matched HCs to elucidate the protein changes associated with SARS-CoV-2 infection at systemic level. As 19 of these patients presented diverse pre-existing comorbidities (CP) including cancer, type 2 diabetes mellitus, asthma, type 1 diabetes, and other autoimmune diseases such as rheumatoid arthritis, and psoriasis among others, we also analyzed plasma samples from 20 sex-and-age-matched DCs with the same pathological conditions. With this experimental design, this study aims to identify common protein changes along SARS-CoV-2 patients with comorbidities. Across the plasma samples from this cohort of 76 individuals, 235 circulating proteins were quantified with a FDR < 1%, from which 208 proteins were commonly identified among the three clinical groups (Figure 3.1a).

High-resolution LC-MS/MS mass spectrometry analysis allowed us to quantify proteins that span a concentration range of nine orders of magnitude reflected in the high level of coverage of quantified areas in the protein rank plot (Figure 3.1b). The plasma protein levels of high abundant proteins were similar between clinical groups while proteins with lower quantification levels showed high variation among these groups. To evaluate the reproducibility of LC-MS/MS measurements, correlation analysis for random chosen matched samples demonstrated the high reproducibility among technical replicates with high correlation coefficients (Figure 3.1c).

Taken together, LC-MS/MS analysis of plasma from SARS-CoV-2 infected patients and their corresponding controls allow the quantification of 235 proteins within a high range of concentrations with high reproducibility among technical replicates.

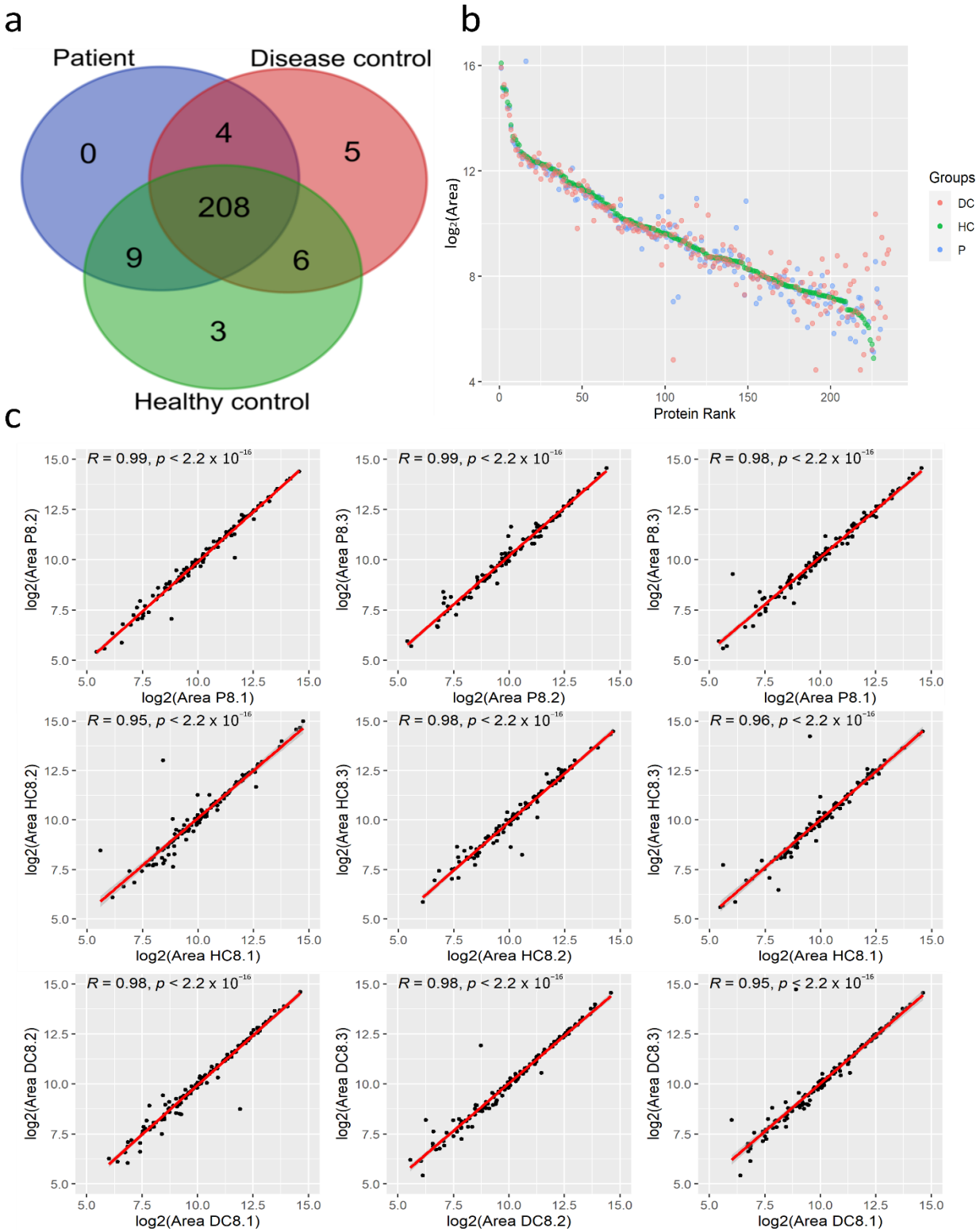

Figure 3.1. Quantification of plasma proteins from COVID-19 patients and their controls by tandem mass spectrometry coupled with liquid chromatography (LC-MS/MS). (a) Venn diagram with identified proteins in the three clinical groups. (b) Protein rank plot with the mean of the areas of the 235 identified proteins from each clinical group. (c) Correlation analysis of the protein quantified areas (after log2 transformation) of three technical replicates from a representative patient and the corresponding controls with the R coefficients and p-values.

3.3.2. COVID-19 Patients with Comorbidities Share a Common Plasma Protein Signature

Firstly, to investigate proteomic alterations caused by SARS-CoV-2 infection under pre-existing pathological conditions, plasma protein levels between SARS-CoV-2 infected patients with comorbidities (CPs) and their sex-and-age-matched HCs were compared. Among the 235 quantified proteins, the levels of 25 proteins were significantly changed, among which levels of fibronectin (FINC), keratins K1C10 and K22E, SHBG, and immunoglobulin variable chains HVD82 and LV39 were elevated in CPs (Figure 3.2a, Appendix I Table S1). The elevated level of FINC, a mediator of blood clotting was previously reported, suggesting the reliability of our results [300]. As a key extracellular matrix (ECM) component, FINC can also be considered as an indicator of tissue remodeling after damage caused by SARS-CoV-2 infection [301]. Moreover, for the first time, the increased levels of these two cytoskeletal keratins, K1C10 and K22E were revealed that may indicate the SARS-CoV-2-induced damage of epithelial cells.

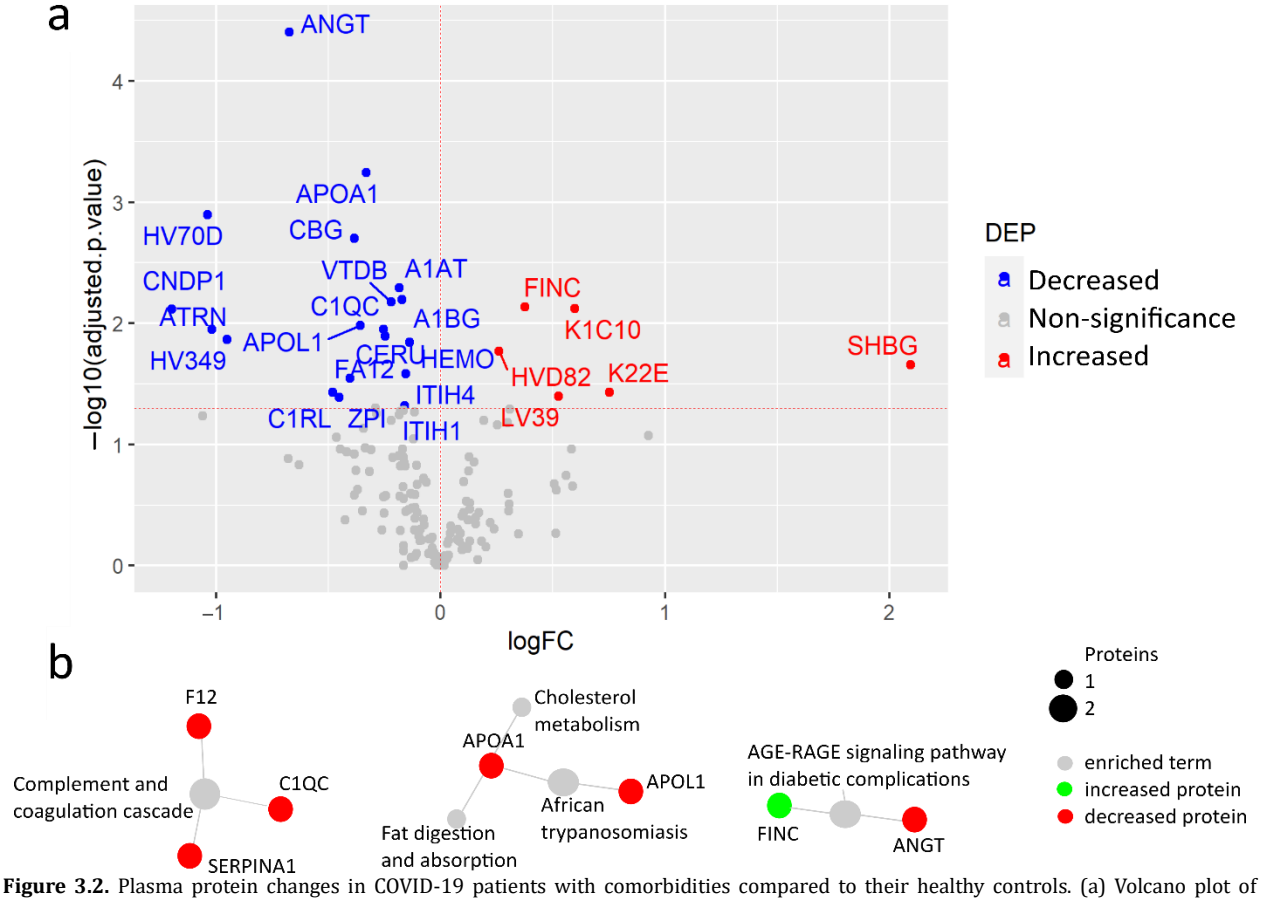

Figure 3.2. Plasma protein changes in COVID-19 patients with comorbidities compared to their healthy controls. (a) Volcano plot of differential expression analysis between COVID-19 patients with comorbidities and sex-and-age-matched healthy controls. (b) Network of selected pathways from the differentially expressed proteins (DEPs) with KEGG pathway enrichment analysis via active subnetworks from STRING database.

Whereas proteins with lower plasma levels in CPs compared to HCs included angiotensin II (ANGT), apolipoproteins APOA1 and APOL1, vitamin D-binding protein (VTDB), the protease inhibitors alpha-1-antitrypsin (A1AT) and ZPI, corticosteroid-binding globulin (CBG), alpha-1B-glycoprotein (A1BG), and ceruloplasmin (CERU) among others. Importantly, angiotensin II (ANGT) showed the most significant change. ANGT is a regulator of blood pressure and cardiac function that is proteolyzed by angiotensin-converting enzyme 2 (ACE2), the receptor of SARS-CoV-2. The reduced levels of ANGT may be caused by the alteration of ACE2 due to SARS-CoV-2 binding [302]. Noteworthy, two negative acute-phase reactants, CBG and ZPI, showed lower levels in patients in COVID-19 studies. The reduced plasma levels of CBG and ZPI in COVID-19 patients have not been previously reported.

KEGG pathway enrichment analysis of DEPs revealed that SARS-CoV-2 infection altered the complement and coagulation pathways, cholesterol and fat metabolism including APOA1 and APOL1, and the RAGE-AGE signaling pathway (Figure 3.2b, Appendix I Table S2). APOA1 has anti-viral activity and was previously detected at low levels in COVID-19 patients as a strong predictive factor of COVID-19 severity [303,304]. Moreover, RAGE-AGE signaling pathway plays a key role in pulmonary inflammatory responses including viral infection as well as in diabetes by NF-κB activation and pro-inflammatory cytokine release [305,306].

To identify specific plasma proteins that are differentially regulated due to the SARS-CoV-2 infection itself in COVID-19 patients with pre-existing comorbidities, we compared CPs versus their disease controls (DCs). From the 26 detected DEPs,

18 were increased in CPs including hemoglobin subunits α and β (HBA, HBB), paraoxonase-1 (PON1), and α-2-antiplasmin (A2AP) among others, while lower levels of 8 proteins were observed, including the antioxidant enzyme glutathione peroxidase 3 (GPX3), immunoglobulin IGG1, and variable immunoglobulin chains (LV310, LV211, HV349, and HC70D) (Figure 3.3a, Appendix I Table S3). Importantly, this comparison confirmed that increased FINC together with the keratins K1C10, K22E and additional K2C1, not detected in the previous comparison, are indicators of tissue damage by SARS-CoV-2 infection not from the underlying comorbidities. Moreover, elevated levels of HBA and HBB may result from the SARS-CoV-2 infection of ACE2-expressing erythrocytes in which SARS-CoV-2 interacts with HBA and HBB changing to a defective conformation and consequent hemolysis [307]. Another elevated protein in CPs is PON1, which has antioxidant activity by the hydrolysis of lipoperoxides, participating in the innate immune response to infections and oxidative stress [308]. Of particular interest, there is a substantial increase in the acute phase inflammatory protein A2AP which is part of the plasmin-antiplasmin system that plays a key role in blood coagulation and fibrinolysis [309].

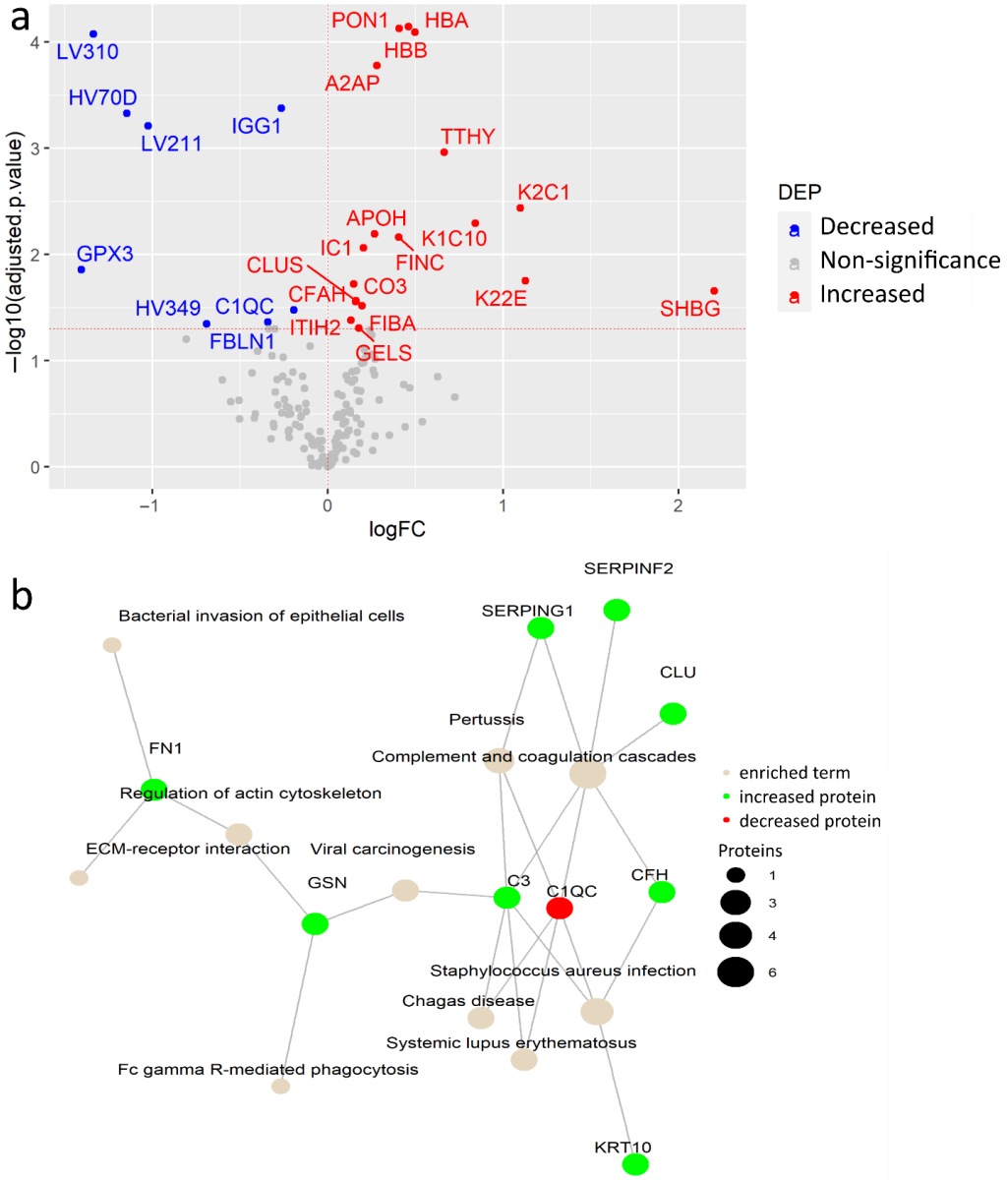

Figure 3.3. Plasma protein changes caused by SARS-CoV-2 infection in patients with comorbidities. (a) Volcano plot of differential expression analysis between COVID-19 patients with comorbidities and sex-and-age-matched disease controls. (b) Network of KEGG pathway enrichment analysis of DEPs via active subnetworks using STRING database.

The network of KEGG enriched pathways in DEPs compared to their DCs showed that elevated level of proteins in CPs were involved in complement and coagulation cascades, viral carcinogenesis, and ECM-receptor interaction among others (Figure 3.3b, Appendix I Table S4). In fact, the complement system is a fundamental player in the anti-microbial innate immune response, including response against SARS-CoV-2, but also, its hyperactivation contributes to the exacerbated

immune response of severe COVID-19 cases, suggesting a dual role in SARS-CoV-2 infection [310]. Among them, the cleavage of C3 (protein CO3) releases C3a anaphylatoxin that contributes to the hyperinflammatory state of COVID-19 patients [311].

Collectively, COVID-19 patients with different pre-existing comorbidities shared a common plasma protein signature. In addition, the synergistic effect of SARS-CoV-2 infection and pre-existing comorbidities causes plasma protein changes that are associated with metabolic alterations, coagulation and innate immune responses.

3.3.3. Coagulation and Cholesterol Metabolism are Altered in COVID-19 Patients without Comorbidities

To determine the plasma protein changes in the COVID-19 patients without comorbidities, we quantified and compared plasma protein levels of COVID-19 patients without comorbidities from this cohort to sex-age-matched HCs. Among the significantly altered protein levels, 8 proteins were elevated in COVID-19 patients while 40 were with decreased levels (Figure 3.4a, Appendix I Table S5). Among the increased proteins, FINC was also elevated in patients without comorbidities, suggesting that it is caused by SARS-CoV-2 infection independently of the pre-existing comorbidities. Meanwhile, the negative acute-phase reactant FETUA and APOM were reduced in COVID-19 patients suggesting that their reduction is caused by the exacerbated innate immune response to SARS-CoV-2 infection [312,313].

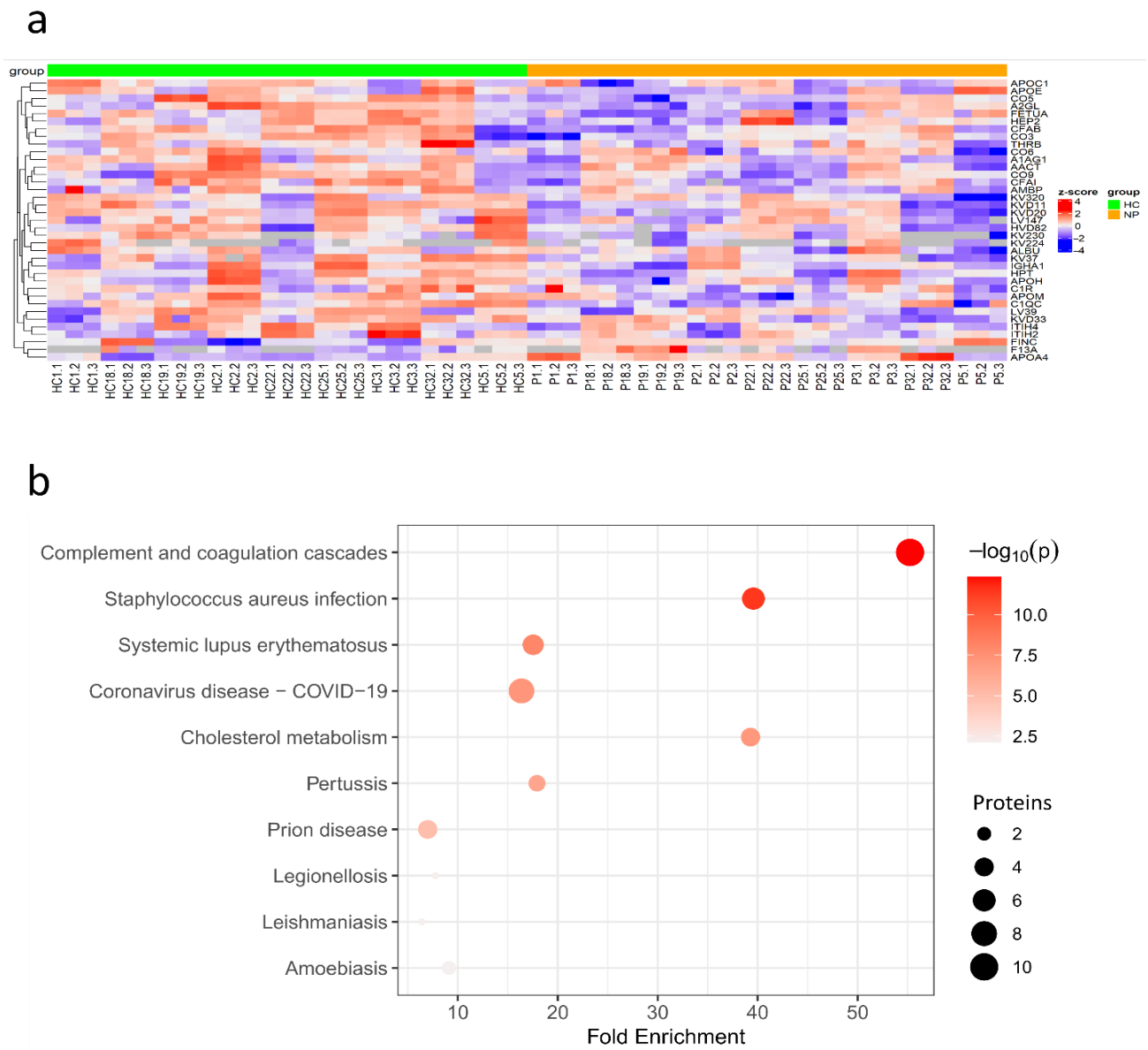

Figure 3.4. Protein changes in SARS-CoV-2 infected patients without comorbidities compared to sex-an-age-matched healthy controls. (a) Heatmap of hierarchical clustering of selected DEPs after z-score normalization. (b) Bubble plot of KEGG enriched terms from DEPs between patients without comorbidities and their healthy controls. p, p-value.

KEGG pathway enrichment analysis revealed that these DEPs were mainly associated with the complement and coagulation cascades, cholesterol metabolism as well as the expected COVID-19 KEGG term itself and other diseases in which the complement cascade is involved in the innate immune response such as *Staphylococcus aureus* infection (Figure 3.4b, Appendix I Table S6). Interestingly, the coagulation factor XIII (F13A) which stabilizes the fibrin clot in the coagulation cascade, was found elevated in COVID-19 patients. In fact, coagulopathy plays an essential role in COVID-19 morbidity and F13A was previously reported elevated in COVID-19 patients with less severe symptoms, indicating the capacity to identify COVID-19 patients with mild symptoms [314]. The majority of proteins associated with cholesterol metabolism including APOC1, APOH, and APOE, were found decreased except APOA4 that was elevated in COVID-19 patients which is supported by a previous study that demonstrated its elevation in COVID-19 patients with mild symptoms versus severe patients [315].

Taken together, SARS-CoV-2 infection in patients without comorbidities causes the alteration of proteins related to the coagulation and cholesterol metabolism including negative acute-phase reactants that could counteract the innate immune response against the virus.

3.3.4. Early SARS-CoV-2 Infection is Associated with Immune Protein Changes and Tissue Remodeling

To elucidate the plasma protein changes in the early SARS-CoV-2 infection, we compared between patients with collected plasma samples in the early infection and late infection (after 3 months). This analysis revealed 36 proteins increased in patients at early stage of SARS-CoV-2 infection, including the antioxidant PON1; whereas, 4 proteins with elevated levels in late infection (Figure 3.5a, Appendix I Table S7). Interestingly, we detected increased levels of attractin (ATRN) in patients with early infection. ATRN is involved in the initial immune cell clustering during inflammatory responses, suggesting ATRN involvement in the initial response to SARS-CoV-2 infection [316]. Moreover, the keratin (K22E) and ECM1 were with higher levels in early infected patients indicating active tissue damage and remodeling due to SARS-CoV-2-infection. In fact, ECM1 is a regulator of differentiation of several subsets of helper T cells, being a potential link between the tissue damage and the immune response against SARS-CoV-2 infection [317]. Pathway enrichment analysis showed elevated proteins in early infection were associated with up-regulated pathways including the complement and coagulation pathways as well as thyroid hormone synthesis (Figure 3.5b, Appendix I Table S8). Notably, high levels of VWF in COVID-19 patients are supported by previous studies [318], while our study indicates that VWF is specifically elevated in the early SARS-CoV-2 infection. Additionally, increased A2AP, also called SERPINF2, was determined as a potential biomarker of early COVID-19 infection

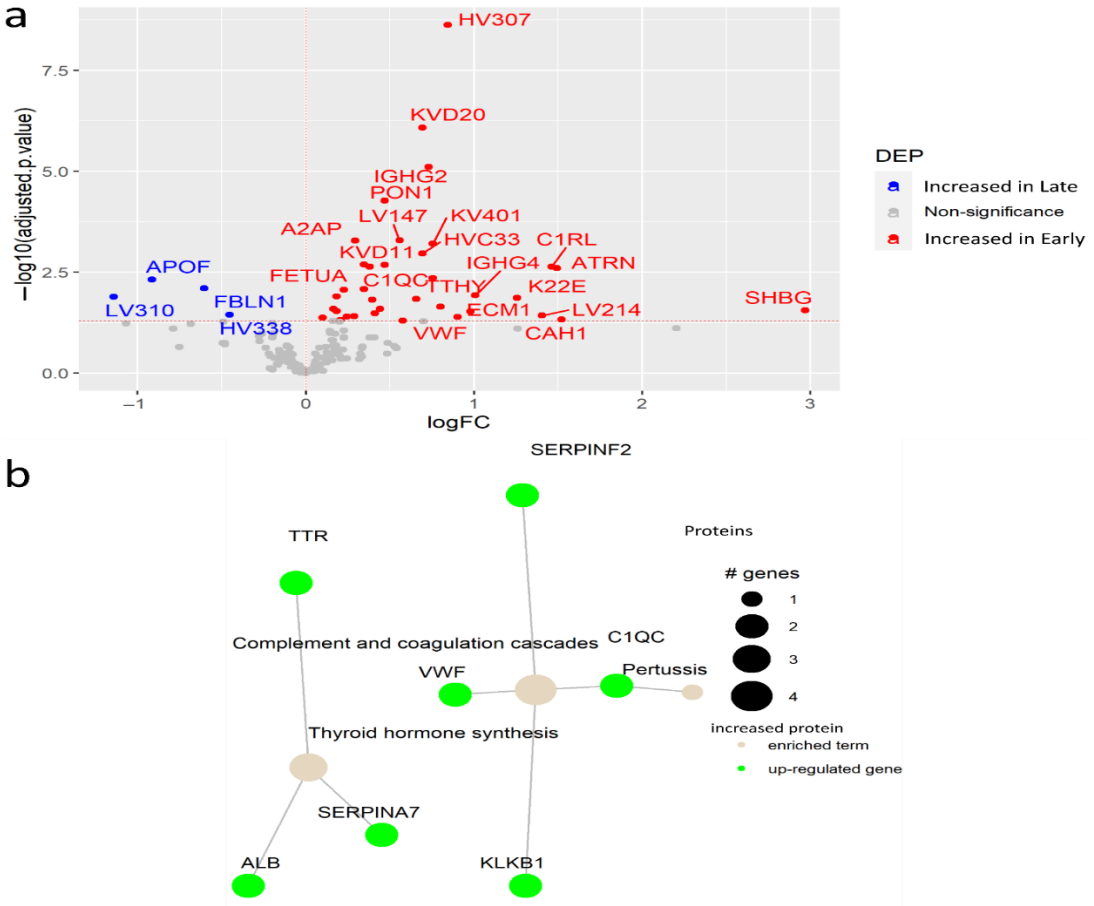

Figure 3.5. Plasma protein changes in early SARS-CoV-2 infection. (a) Volcano plot of DEPs between early SARS-CoV-2 virus-infected patients and patients with late infection. (b) Network of KEGG pathway enrichment analysis with DEPs via active subnetworks from STRING protein-protein interaction database.

3.4. Discussion

Current understanding of COVID-19 pathophysiology needs further development to determine the role of the immune response against SARS-CoV-2 infection and clinical symptoms variability. Here, we performed LC-MS/MS proteomics analysis of plasma from COVID-19 patients and their corresponding controls to characterize the systemic protein changes underlying SARS-CoV-2 infection. In this study, changes in plasma levels of proteins involved in tissue damage and remodeling (K1C10, K22E, and ECM1), coagulation (FINC, F13A, ANGT, and VWF), inflammation (A2AP, ZPI, and CBG), complement activation (C3, C1QC), and cholesterol metabolism (APOA1, APOL1, APOE, APOA4) as well as levels of antioxidant enzymes PON1 and GPX3 among others were associated with SARS-CoV-2 infection.

Our analysis detected for the first time the elevation of the keratins K1C10, K22E in COVID-19 patients with pre-existing comorbidities while previous studies found an enrichment of keratinization and increased plasma levels of KRT19 in COVID-19 patients as well as the association of KRT7 with severity [319,320]. Moreover, Cytoskeletal remodeling of these keratins may be induced by SARS-CoV-2 infection to facilitate its spread between epithelial cells as K1C10 showed interaction with several SARS-CoV-2 proteins [321]. Then, these keratins are released with the cell content to the bloodstream and can be indicators of epithelial tissue damage [320]. Interestingly, K22E and ECM1 were also increased in early SARS-CoV-2 infection, being potential diagnostic biomarkers of early SARS-CoV-2 infection. We confirmed that FINC is increased in COVID-19 patients despite the presence of pre-existing comorbidities, being a potential biomarker of tissue damage by SARS-CoV-2 infection [300].

Between the enriched pathways of elevated plasma proteins in COVID-19 patients, complement and coagulation cascade were present in all the comparisons, demonstrating their active role in SARS-CoV-2 infection. Among them, elevated C3 in COVID-19 patients with comorbidities was revealed while previous reports demonstrated that C3 is associated with COVID-19 severity and was proposed as potential treatment to diminish inflammatory symptoms [311,322,323]. Meanwhile, the acute phase reactant A2AP was elevated in COVID-19 patients with comorbidities as well as in the early infection. Recently, A2AP was also found elevated in persistent circulating plasma microclots of long-COVID-19 patients that highlights A2AP contribution to the multiple coagulation/fibrinolysis pathophysiology of SARS-CoV-2 infection [324]. In contrast, the negative acute-phase reactants ZPI and CBG were found for the first time reduced in SARS-CoV-2 infected patients with comorbidities. In fact, both proteins have a role in dampening the excessive inflammatory response [325,326]. Meanwhile, other two negative acute-phase reactants, FETUA and APOM, were reduced in COVID-19 patients without comorbidities which are supported by previous studies [312,313].

There are some controversies in the literature regarding ANGT levels in COVID-19 patients [327]. However, low levels of the blood regulator ANGT together with high soluble ACE2 in severe COVID-19 patients compared to healthy subjects were found in a previous study by LC-MS/MS [328]. Interestingly, our previous study showed also elevated levels of soluble ACE2 in the same cohort of COVID-19 patients with comorbidities [329]. These complementary results suggest that SARS-CoV-2 infection alters the renin-angiotensin system, affecting the cardiovascular and inflammatory stability in patients with pre-existing comorbidities. Additionally, as ANGT is involved in stimulating TNF-α production and T cell activation, its reduction may negatively affect to the adaptive immune response against SARS-CoV-2 [330].

Among the dysregulated metabolic enzymes, high levels of the antioxidant protein PON1 was found in COVID-19 patients with comorbidities as well as in the early infected patients; whereas previous studies showed that PON1 activity is reduced in COVID-19 patients which is attributed to excessive oxidative stress and lipoprotein alterations secondary to infection [331]. Supporting our results, PON1 levels were found increasing along the first month of infection as well as decreased with severity by LC-MS/MS while COVID-19 patients from our study were with comorbidities [315,332]. In contrast, GPX3 was with lower levels in COVID-19 patients with comorbidities while previous reports revealed lower glutathione peroxidase activity due to SARS-CoV-2 infection [333] as well as Epstein–Barr virus infection in diabetic patients [334]. In fact, SARS-CoV-2 triggers an oxidative stress reinforcing inflammation and leading to a weakened antioxidant system, especially in patients with pre-existing comorbidities that generate oxidative stress [335].

The SARS-CoV-2 infection of erythrocytes is a well-established phenomenon that produces hemolysis. Consequently, patients develop anemia together with hyperferritinemia and systemic hypoxia that can result in multi-organ failure [307]. Interestingly, HBB participates in the anti-viral innate immune response against RNA viruses [336]. While previous studies demonstrated elevated levels of HBA and HBB in the airway mucus of severe COVID-19 patients as well as in serum of patients with high IL6 compared to healthy subjects [337,338], we also detected high plasma levels of HBA and HBB in patients with comorbidities compared to their disease controls.

The limited number of included patients from the Finnish Clinical Biobank complicates the results inference at population level and all of them were with mild symptoms impeding the assessment of protein changes association with COVID-19 severity. Therefore, further validation in multi-center studies including more and diverse ethnic groups will facilitate the confirmation of the findings in specific comorbidities and the final application of the discovered novel biomarkers.

To conclude, this study determined a plasma protein signature shared by COVID-19 patients with pre-existing comorbidities characterized by alterations in acute-phase reactant proteins, coagulation and complement cascades, innate

immune responses as well as cholesterol and redox enzymes. Moreover, several potential biomarkers of early SARS-CoV-2 infection were identified that further research can establish its applicability in clinics.

CHAPTER 4. Plasma proteomics unveil novel immune signatures and biomarkers upon SARS-CoV-2 infection

Instead of untargeted proteomics analysis as in the previous chapter, in this chapter, targeted analysis by proximity extension assay was applied to plasma samples from the same COVID-19 patients. This study aims to determine plasma protein associated with SARS-CoV-2 infection as in the previous chapter with the advantage that specific immune and other proteins are quantified. This study resulted in an article published in **International Journal of Molecular Sciences**in collaboration with the University of Oulu and the IFB laboratory of mass spectrometry and is presented with minor modifications:

***Urbiola-Salvador V*, Lima de Souza S, Grešner P, Qureshi T, Chen Z. Plasma Proteomics Unveil Novel Immune Signatures and Biomarkers upon SARS-CoV-2 Infection. *Int. J. Mol. Sci.* 24, 6276 (2023).**

4.1. Introduction

The ongoing COVID-19 pandemic, which is caused by SARS-CoV-2, is the most significant catastrophe that humanity has faced in the 21st century. With several widespread variants, more than 6 million COVID-19 related deaths, and over 600 million cases [339], COVID-19 is ranked as one of the deadliest pandemics of recorded human history [340]. Despite all of the advancements in medical sciences, preventive measures, and efficient vaccination programs, the COVID-19 pandemic still persists and thousands of COVID-19 related deaths are reported every day [339].

Regarding the immune reaction upon SARS-CoV-2 infection, a large variability is observed among the human population and current studies show that immunopathology is largely responsible for COVID-19 pathogenesis and the related mortality [341–343]. The release of large quantities of pro-inflammatory cytokines, which is known as a cytokine storm, is considered as an underlying reason for the hyperactive immune response against SARS-CoV-2 infection and is correlated with COVID-19 disease severity [344]. The ample production of cytokines attracts immune cells to the site of infection and causes tissue damage that may lead to pneumonia, lung injury and multi-organ failure, which are complications commonly seen in critical and deceased COVID-19 patients [345].

Although vaccines have shown a promising outcome to curb the spread of COVID-19 infection, the rate of mutation in the SARS-CoV-2 single-stranded RNA-based genome will likely result in a greater landscape of variants. The outright eradication of COVID-19 seems improbable in the near future [346–348]. Importantly, despite that the majority of COVID-19 patients do not develop severe symptoms that require intensive care unit (ICU) admission, several patients contract post-COVID-19 syndrome characterized by multi-organ symptoms that persist for months after acute COVID-19, independently of the disease severity [349]. A remarkable number of studies have been published on COVID-19 since the inception of the pandemic. Nevertheless, our current knowledge of the immunological changes upon SARS-CoV-2 infection is still incomplete. There is a pressing need to broaden our understanding of the immune dynamics that occur upon this viral infection as well as its long-term effects of post-COVID-19 syndrome.

COVID-19 progression and severity are determined by age, sex, ethnicity, comorbidities, and some risk genetic mutations carried by patients [350–353], which makes it rather difficult to pinpoint the immune signatures that could be used to anticipate a hyperactive immune response. Importantly, comorbidities, such as cancer, diabetes, cardiovascular, and neurological diseases, are reported to be associated with higher severity and increased risk of death in patients with COVID-19 [354–357]. Albeit interpatient heterogeneity of medical conditions, there could be key signature proteins that may serve to determine the underlying immune and physiological responses in SARS-CoV-2 infected patients with comorbidities. Recent studies show that around 20% of people infected by SARS-CoV-2 may continue develop symptoms diagnosed as post-COVID19 condition (also known as long COVID), a condition there still remains limited information [349,358].

In this study, we investigated the plasma proteomic profiles of COVID-19 patients by Proximity Extension Assay (PEA) with reference to age and sex-matched disease controls (DC) and healthy controls (HC) collected in Finnish Clinical Biobank Tampere to understand how and which proteins are diverted from their normal expression patterns that may lead to disruption in immune features and cellular functionality. Among the identified protein changes, many key players in immune regulation and inflammation were upregulated, especially in patients with comorbidities. Moreover, several protein changes were associated with the generation of specific SARS-CoV-2 antibodies and time after infection that reflects the long-term effects of post-COVID-19 condition.

Our analyses provide deeper insights into the plasma proteomic changes caused by a SARS-CoV-2 infection that modulate the immunological and physiological response in the Finnish population. Several novel plasma proteins involved in innate and adaptive immunity, T cell activation/cycling, and extracellular matrix (EMC) remodeling were associated with a SARS-CoV-2 infection that may serve as potential diagnostic and prognostic biomarkers as well as potential therapeutic targets for COVID-19 patients.

4.2. Materials and Methods

4.2.1. Study cohort

The study involved 28 SARS-CoV-2 virus-infected patients (21 females and 7 males, age range 19-79) collected between April 2020 and November 2020. Nineteen of them were found to have other diseases, while the remaining nine were comorbidities-free. The control group consisted of 28 SARS-CoV-2 virus-negative subjects without any diseases (the so-called healthy controls) and additional 20 SARS-CoV-2 virus-negative subjects with additional diseases (the so-called disease control). The 19 SARS-CoV-2 patients with additional diseases were age-and-sex-matched with 19 healthy controls as well as with 20 disease controls, in the case of which the matching considered the major disease(s), as well. The remaining 9 SARS-CoV-2 virus-infected patients were age-and-sex-matched to remaining 9 healthy controls only. For one COVID-19 patient with chronic diseases, no age-and-sex-matched disease control was found, while two others needed to have two age-and-sex-matched disease controls owing to his clinical conditions (see Appendix I Table S1 for details). Plasma samples of all subjects (patients and controls) together with respective clinical information were obtained from the Finnish Clinical Biobank, Tampere.

4.2.2. Proximity extension assay

The plasma samples were pipetted to a 96-well plate in a randomized order to circumvent the effects of experimental variables. The last column of the plate was left empty for two Olink samples, three negative and three inter-plate controls, which are used to calculate the intra-assay coefficient of variance, monitor background noise for the limit of detection calculation and compensate for potential variation between the runs, respectively. The sample plate was dispatched to Olink Proteomics (Uppsala, Sweden).

The PEA chemistry is based on antibody-antigen complexes. Two antibodies carrying unique and complementary DNA sequences bind to one specific protein in two different epitopes and the DNA-tags hybridize and are amplified to generate a library of DNA fragments. The Olink Explorer collection is sequenced by next generation sequencing (NGS) that identifies each protein from a different sample using adaptors and unique barcode sequences [174]. Then, the number of reads can be translated to protein concentration using normalized protein expression (NPX) values that are represented in log2 scale. The higher NPX value corresponds to higher protein concentration and vice versa. The Olink Explorer collection consisting of 1,472 proteins (1,463 unique proteins) encompassing the cardiometabolic, neurology, inflammation, and oncology panels (369, 367, 368, and 368 proteins, respectively) was run for the extensive proteome profiling of the samples.

4.2.3. Statistical analyses

Statistical analysis was performed in RStudio (version 1.3.1093) using R (version 4.0.3) [295,296]. First, proteins and samples were filtered when the quality control was negative or the quantification led to values below the respective protein limit of detection (LOD) in at least 50% of the samples. The remaining NPX values were imputed with the respective $LOD/\sqrt{2}$ value. Between-group differences in protein expression levels were analyzed by means of the general linear model regression approach with analysis of contrasts using the "emmeans" R package (version 1.6.2.1). The whole regression modeling was divided into three main steps. In the first step, 19 SARS-CoV 2 virus-infected patients with comorbidities were involved to search for differential protein expression due to virus infection and/or coexistent comorbidities by comparing the protein expression levels in virus-infected patients to respective healthy and disease control subjects. Models used in this analysis comprised the presence of comorbidities as nested confounding factor. In the second step of analysis, the remaining nine SARS-CoV-2 virus-infected comorbidities-free patients were involved to search for DEPs due to virus infection itself in comorbidities-free virus-infected group of patients by comparing the protein expression levels to those in age-and-sex-matched healthy controls. No additional confounders were assumed in this analysis. In the last, third step of regression analysis, all 28 SARS-CoV-2 virus-positive patients were analyzed together to search for differential protein expression due to time of sampling (early vs. late) and presence/absence of antibodies, by simple comparing of respective groups. Models used in this step comprised the presence of comorbidities as confounding factor, as well. The False Discovery Rate (FDR) in contrast analyses was controlled by means of the Benjamini & Hochberg correction [298]. Proteins were considered differentially expressed with an FDR adjusted p-value < 0.05. KEGG and Gene Ontology (GO) pathway enrichment analysis via active subnetworks from STRING database was performed with the R package "pathfindR" (version 1.6.3) with FDR correction [299]. Graphics were generated with the R package "ggplot2" (version 3.3.5), excepting heatmaps that were generated with the R package "ComplexHeatmap" (version 2.6.2) and box and whisker plot that were generated with GraphPrism (version 9.3.1). Hierarchical clustering of differentially expressed proteins was performed after transformation of the NPX values or means from each group of samples into z-score values using the Euclidean distances. The heatmap with means from each group of samples is split by k-means clustering. Protein networks from STRING database were constructed in Cytoscape (version 3.8.2) with a confidence cut-off 0.7.

4.2.4. Enzyme linked immunosorbent assay (ELISA)

ELISAs were carried out to detect serum IgM and IgG antibodies against SARS-CoV-2 receptor-binding domain (RBD), a subunit of the Spike S1 protein, and the nucleocapsid protein (NP). The assays were performed with kits from TestLine

(Brno, Czech Republic) specific for IgM or IgG against the RBD (CoRM96 and CoRG96, respectively) or NP (CoNM96 and CoNG96, respectively). The IgG for both spike and NP all together was detected by using a kit from Vircell Microbiologists (Granada, Spain, G1032). All of the assays were performed according to the manufacturer's instructions.

4.3. Results

4.3.1. SARS-CoV-2 infection in patients with comorbidities causes plasma protein changes with enhanced soluble CD4 and associated proteins.

To determine the changes of the protein profiles in peripheral blood caused by a SARS-CoV-2 infection, we performed plasma protein analysis by using Proximity Extension Assay (PEA) technology. Out of the total 1,472 proteins from the entire Explore panels, after removing repetitions in the panels and filtration of proteins with low detection rates among the samples, 1,387 proteins were quantified in our samples. In our cohort, 19 out of the 28 COVID-19 patients have comorbidities in addition to a SARS-CoV-2 infection, such as type 2 diabetes mellitus, asthma, cancer, multiple sclerosis, and rheumatoid arthritis. Firstly, we compared the plasma protein changes between the samples from the SARS-CoV-2 virus infected patients with comorbidities versus their HCs. Among all of the 1,387 quantified proteins, the expression of 116 proteins was significantly changed (Figure 4.1a). Among these differently expressed proteins (DEP), 106 were upregulated in patients with comorbidities, such as ACE2, FOLR2 and AGRN, and 10 were downregulated, such as RNF41 and TRAF2. The membrane-bound ACE2 is essential to facilitate the entry of SARS-CoV-2 [359]. Importantly, ACE2 was among the elevated proteins detected in this group of COVID-19 patients. (Figures 3.1a, Appendix II Figure S1a and Table S2). A recent study also indicates that the interaction of the SARS-CoV-2 spike protein with soluble ACE2 (sACE2) or with a complex of sACE2 and vasopressin leads to receptor-mediated endocytosis of the virus [302]. Next, in order to further identify their expression patterns, we clustered the 116 identified DEPs (Appendix II Figure S1a). Hierarchical clustering clearly distinguishes patients with comorbidities with more upregulated and less downregulated proteins observed versus HCs. Notably, the helper T cell surface marker CD4, the co-stimulatory molecule of T cells CD28 and the B-cell activation protein CD83 were found in this group, indicating the active involvement of immune cells in response to a SARS-CoV-2 infection. To distinguish which of the protein changes were due to an altered secretion from a certain type of cells and which were a result of destructed tissues or cells due to the viral infection, from the 116 DEPs, 30 proteins were identified in the human blood secretome, such as FOLR2, GRN, EFNA4, TNFSF13B, CCL26, CCL21, PILRA, VWF, LTBP3, LILRA5, and TGFA (Appendix II Table S2).

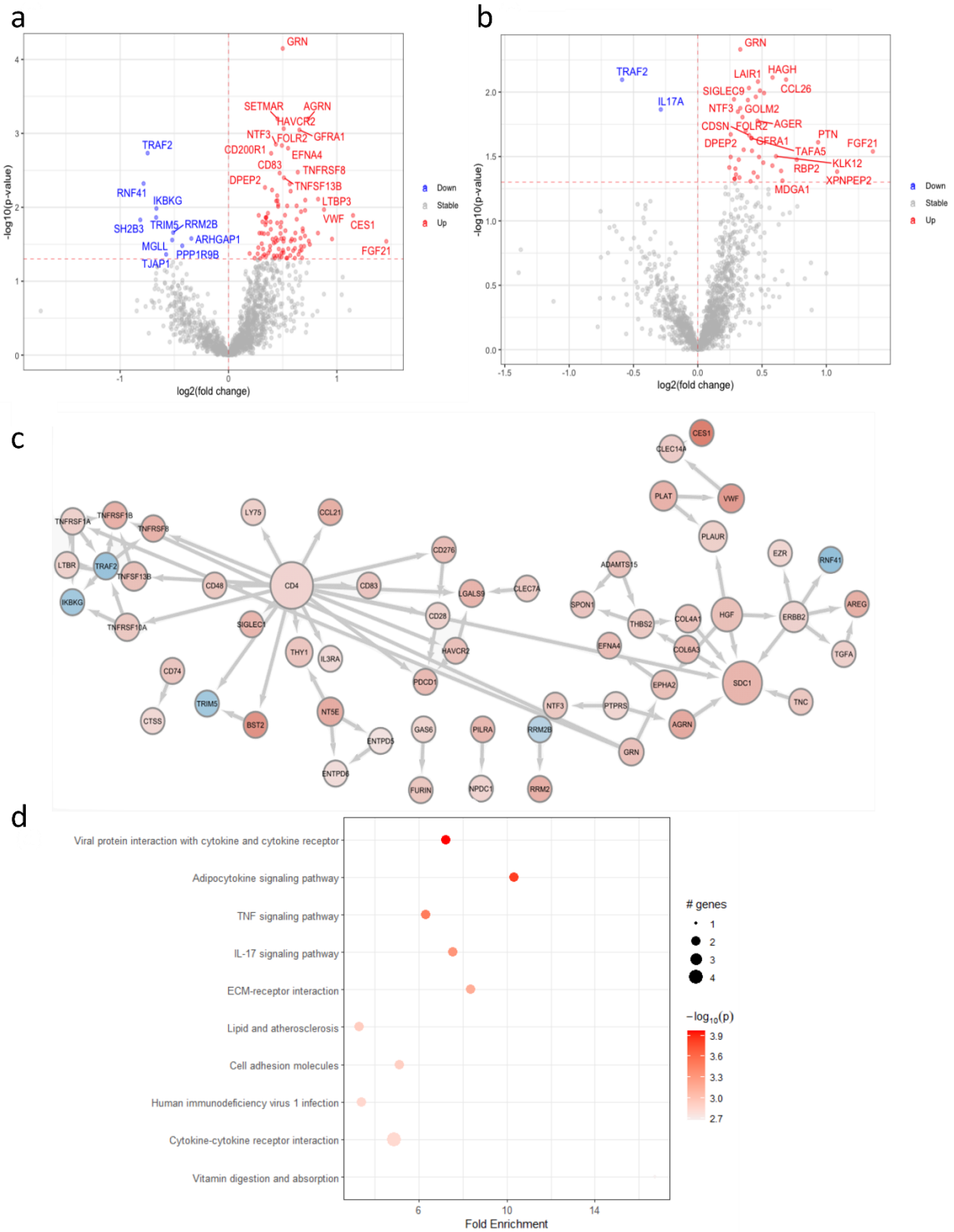

Figure 4.1. SARS-CoV-2 infection in patients with comorbidities causes plasma protein changes with enhanced soluble CD4 and associated proteins. (a) Volcano plot of statistical significance against fold-change of proteins between SARS-CoV-2 virus infected patients with comorbidities and healthy controls. Colored dots indicate statistically differentially expressed proteins (DEPs). (b) Protein-protein interaction network of DEPs between SARS-CoV-2 virus infected patients with comorbidities and healthy controls with organic layout from

STRING database query with a 0.7 confidence cut-off. The size of nodes indicates the degree of connectivity of the nodes. (c) Volcano plot of statistical significance against fold-change of proteins between SARS-CoV-2 virus infected patients with comorbidities and paired disease controls. Dots indicate statistically DEPs.; (d) Dot plot of KEGG pathway enrichment combined with STRING protein-protein interaction network analysis from DEPs between patients and disease controls. (a-c) The red and blue dots represent up-regulation and down-regulation in patients, respectively.

To understand the highly complex interaction patterns of the DEPs, we constructed a protein-protein interaction network using STRING *Homo Sapiens* database. CD4 is highly expressed on the surface of helper T cells [360]. Previously, the soluble form of CD4 (sCD4) has been reported to be detected in patient serum with an HIV infection and the sCD4-to-CD4 lymphocyte ratio increases with disease severity [361]. Here, we observed an increased level of CD4 in the SARS-CoV-2 infected plasma samples compared to plasma from healthy donors. The network analysis showed the association of CD4 with multiple identified DEPs, including CD28, CD48, CD83, and PDCD1 (also named PD1), which were also elevated in patients' plasma compared to their HCs (Figure 4.1b), suggesting the potential damage of helper T cells in response to a SARS-CoV-2 infection. Furthermore, a subgroup of TNF and TNF Receptor Superfamily Members (TNFRSF), including TNFSF13B, TNFRSF8, TNFRSF6B, TNFRSF1B, TNFRSF1A, TNFRSF10A was elevated in patients' plasma and formed a small network (Figure 4.1b). The elevated level of these proteins in SARS-CoV-2 infected plasma suggests the active regulation of this pathway in blood cells in response to this viral infection. Moreover, pathway enrichment analysis showed the upregulation of proteins related to extracellular matrix (ECM)-receptor interaction that are involved in ECM remodeling, tissue damage, and repairing (Appendix II Figure S1b). Viral infection can activate innate and adaptive immune responses, in which the NF-κB signaling pathway plays an important role; on the other hand, viruses may suppress NF-κB pathway activation to dampen the host immune responses [362,363]. Not surprisingly, pathway analysis of the DEPs in a comparison of 19 patients with comorbidities versus HCs showed the enriched positive regulation of I-kappaB kinase/NF-κB signaling pathway and positive regulation of viral protein interaction with cytokine and cytokine receptor (Appendix II Figure S1b). Collectively, SARS-CoV-2 infection clearly causes changes of proteins in peripheral blood that are associated with immune responses, and especially might induce T cell death as indicated by the enhanced soluble CD4 and its associated proteins.

The clustering analysis showed that some of the protein expression changes identified in SARS-CoV-2 infected patients were also observed in their age and sex-matched DCs, such as proteins in cluster 3 in Appendix II Figure S2a. While out of the 116 DEPs in patients with comorbidities, 16 were also changed when we compared their DCs with HCs, indicating changes of these plasma proteins may be caused by their comorbidities (Appendix II Figure S2b).

To dissect what proteins were indeed affected by COVID-19 but not due to their comorbidities, we compared the group with COVID-19 and comorbidities to their matched DCs but without SARS-CoV-2 infection. From the 44 DEPs, all of the proteins were upregulated in patients with COVID-19 compared with their DCs, except TRAF2 and IL17A (Figure 4.1c, Appendix II Table S3).

KEGG pathway enrichment combined with STRING protein-protein interaction network analysis of the DEPs revealed a reduced IL17 signaling pathway (Figure 4.1d). On the other hand, among the 42 elevated proteins in patients with comorbidities compared with their DCs, 28 plasma proteins, including CD28 were found also elevated when compared with their HC, indicating that the elevated plasma level of these 28 proteins is most likely caused by COVID-19 infection but not the comorbidities. FGF21 and NTF3, associated with the MAPK signaling pathway, TNFRSF1B, and CCL26 associated viral protein interaction with cytokine and cytokine receptor pathway were upregulated in plasma proteins from COVID-19 patients with comorbidities compared to their DCs, indicating that COVID-19 caused changes of these pathways. Collectively, these results demonstrate the shared plasma protein signatures in COVID-19 patients with comorbidities, despite the heterogeneity of their medical conditions.

4.3.2. Reduced RNF41 in plasma is associated with a SARS-CoV-2 infection

Next, we performed one more analysis to compare the patient samples without chronic diseases to their healthy controls to define the proteins related to the SARS-CoV-2 infection itself. From this analysis, 21 DEPs were found, 7 proteins were detected higher in patients, and 14 proteins were higher in the HCs (Figure 4.2a, Appendix II Table S4). It is noteworthy that this comparison resulted in a lower number of DEPs. RNF41, an E3 ubiquitin-protein ligase, is among the 14 proteins with higher level in HCs than that of COVID-19 patients. And it is the only protein with reduced level in COVID-19 patients with and without comorbidities compared with their HCs (Figure 4.1a, 4.2a-b). RNF41 is involved in type 1 cytokine receptor signaling, regulating JAK2-associated cytokine receptor surface by degradation and ectodomain shedding [364]. RNF41, as in the other RNFs, can inhibit antiviral responses; nevertheless, with a regulation of IFN signaling, it is still unclear [365]. RNF41 expression is with low tissue specificity and is mainly involved in the B-cell immune response. It has been reported that SARS-CoV-2 NSP15 protein targets RNF41 [366], and this interaction may be involved in antagonizing the induction of IFN-I [367]. In murine dendritic cells, RNF41 negatively regulates the cross-presentation of dead cell-derived antigens by the ubiquitination of CLEC9A [368]. Now, we observed a reduced plasma level of RNF41 in COVID-19 patients, suggesting a potential function of RNF41 in a SARS-CoV-2 infection. Therefore, the analysis of plasma proteomics revealed novel potential diagnostic biomarkers for COVID-19.

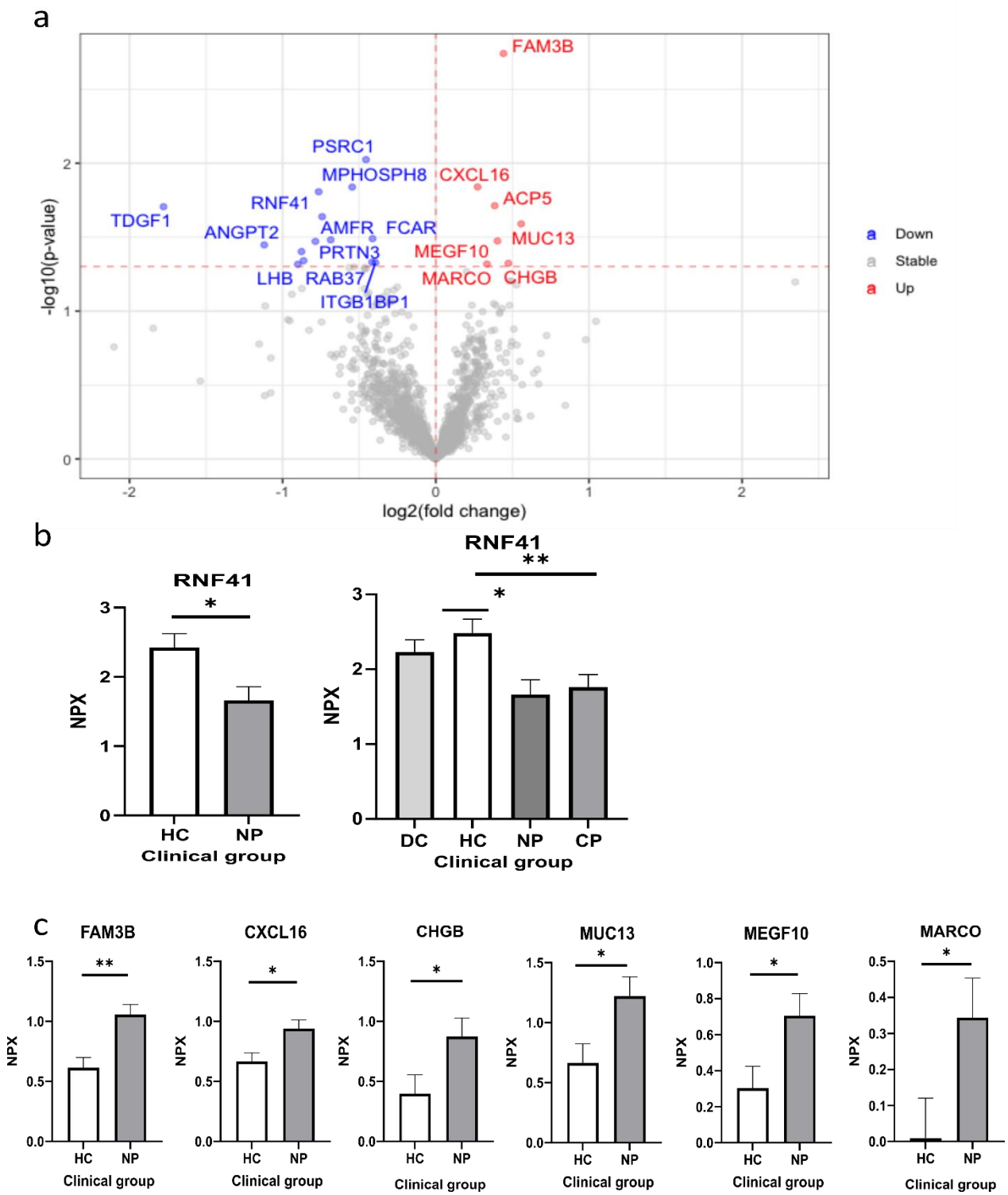

Figure 4.2. Reduced plasmas level of RNF41 is associated with SARS-CoV-2 infection. (a) Volcano plot of statistical significance against fold-change of proteins between patients without comorbidities and healthy controls. The red (upregulated in patients without comorbidities) and blue (downregulated in patients without comorbidities) dots indicate statistically DEPs. (b) Bar plots with normalized protein expression of RNF41 among different clinical groups. (c) Bar plots with normalized protein expression of FAM3B, CXCL16, CHGB, MUC13, MEGF10 and MARCO in patients without comorbidities and their respective healthy controls. * and ** indicate statistically significant with an adjusted p-value < 0.05 and < 0.01, respectively. CP, Comorbidities Patient; DC, Disease Control; HC; Healthy Control; NP, Non-comorbidities Patient; NPX, Normalized Protein eXpression.

A group of plasma proteins were found elevated in patients without comorbidities compared with their HCs, such as FAM3B, CXCL16, CHGB, MUC13, MEGF10 and MARCO (Figure 4.2a, c).

4.3.3. Characterization of early and long-term plasma protein responses associated with SARS-CoV-2 infection

To characterize the plasma protein response to early and late SARS-CoV-2 infection, we compared the 14 plasma samples collected within 3 months after infection with the 14 samples collected after 3 months of infection. As shown in Figure 4.3a, the longer period after infection caused more elevated plasma proteins than early infection. Only Receptor for the Fc region of IgA (FCAR) was significantly elevated in early infected patients. FCAR interacts with IgA-opsonized targets and triggers several immunologic defense processes. However, no association of its plasma level with COVID-19 has been reported yet. 67 plasma protein levels were higher in patients infected longer than 3 months compared to recent infected patients, including a group of cytokines and chemokines, such as IL6ST, TGFB1, IL32, IL19, CXCL8, CCL2, CCL11 and CCL14 (Figure 4.3, Appendix II Table S5). These elevated cytokines indicate the pathologic inflammation that may drive post-COVID-19 condition [349]. Among the elevated chemokines, the eosinophILattracting chemokine CCL11 was previously found elevated in late post-COVID-19 syndrome patients which may be associated with cognitive symptoms [369]. Meanwhile, CCN3 and PLTP were upregulated in patients with longer period of infection which are involved in the negative regulation of inflammation [370,371]. They may attenuate COVID-19 pro-inflammatory conditions in order to re-establish homeostasis. Moreover, high plasma levels of aggrecan (ACAN), the main component of cartilaginous ECM, may indicate degradation of cartilaginous tissues after COVID-19 pro-inflammatory conditions. In fact, joint aged joint and arthritis may be associated with post-COVID-19 condition [372]. Taken together, the characterization of early and long-term effect of plasma protein responses to SARS-CoV-2 infections resulted in the identification of novel potential biomarkers of post-COVID-19 condition.

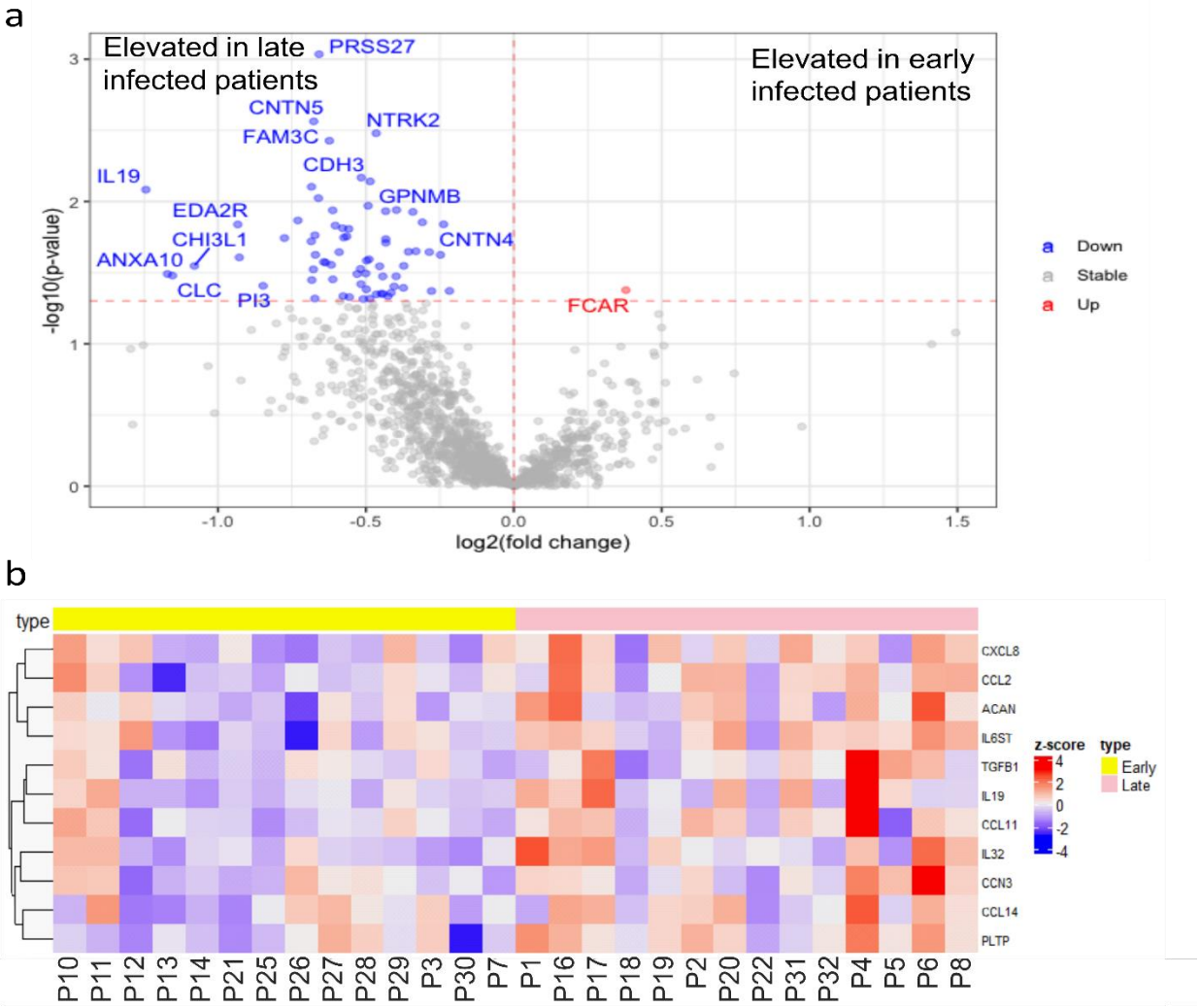

Figure 4.3. Characterization of long-term plasma protein responses associated with SARS-CoV-2 infection. (a) Volcano plot of statistical significance against fold-change of proteins between patients with plasma samples collected less than 3 months after infection (early) and collected more than 3 months after infection (late). Red (upregulated in patients with early collected) and blue (Upregulated in patients

with late collected plasma) dots indicate statistically DEPs. (b) Heatmap of selected statistically DEPs between patients with plasma collected less than 3 months after infection (early) and collected more than 3 months after infection (late) with z-score by row normalization and distributed by hierarchical clustering.

4.3.4. Plasma protein changes associated with SARS-CoV-2 antibody generation

To investigate whether the plasma protein changes are associated with antibody generation, we detected SARS-CoV-2 IgM and IgG antibodies against different antigens, including spike receptor-binding domain (RBD), and Nucleocapsid protein (NP). For IgG, Spike combined with NP was also measured. As a result, 12 out of 28 (42.8%) of the SARS-CoV-2 infected patients did not generate any detectable antibodies against any of the antigens (Table 4.1). Among the SARS-CoV-2 infected patients, no association of antibody generation with comorbidities was found. From those IgM antibody positive patients, 80% of RBD positive patients had a SARS-CoV-2 infection less than 3 months before the antibody measurement, whereas 67% of RBD IgG positive patients had an infection more than 3 months (Figure 4.4a). The detected IgM and IgG antibodies against SARS-CoV-2 RBD follow the regular pattern of antibody generation after a viral infection. However, the pattern in antibody detection against NP antigen is less clear (Figure 4.4a), suggesting that RBD is more suitable than NP as an antigen for the specific testing of a SARS-CoV-2 infection.

Table 4.1. Description of the clinical samples.

COVID-19 patients	Sex	Age (years)	Disease control (DC)		Healthy control	Sampling <3 months of infection	Antibodies				
			DC1	DC2			IgG RBD	IgG NP	IgG (spike+NP)	IgM RBD	IgM NP
Case 1	F	55			+		++	+	++	-	+
Case 2	M	78			+		++	+	+++	-	-
Case 3	F	41			+	+	-	+	+	+	-
Case 4	F	71	+	+	+		-	-	-	-	-
Case 5	F	36			+		-	-	-	-	-
Case 6	F	71	+		+		-	-	-	-	-
Case 7	F	54	+		+	+	-	-	-	-	-
Case 8	F	58	+		+		+++	++	+++	-	-
Case 10	F	53	+		+	+	-	-	-	-	++
Case 11	M	64			+	+	+++	+++	+++	+	+++
Case 12	F	61	+		+	+	-	-	+	-	-
Case 13	F	27	+		+	+	-	+	+	-	-
Case 14	F	45	+		+	+	+	+	+	+	-
Case 16	M	52	+	+	+		+	-	+	-	-
Case 17	M	64	+		+		-	-	-	-	-
Case 18	F	35			+		-	+	+	-	+
Case 19	F	31			+		-	-	-	-	-
Case 20	F	23	+		+		-	-	-	+	-
Case 21	F	53	+		+	+	+++	+++	+++	+	-
Case 22	F	30			+		-	-	+	-	-
Case 25	F	27			+	+	-	-	-	-	-
Case 26	M	62	+		+	+	-	-	-	-	-
Case 27	F	79	+		+	+	-	-	-	-	-
Case 28	F	36	+		+	+	-	-	-	-	-
Case 29	M	74	+		+	+	-	-	-	-	-
Case 30	F	19	+		+	+	-	-	-	-	-
Case 31	M	72	+		+		+	++	+++	-	+
Case 32	F	53			+		+	+	+++	-	-

* The SARS-CoV-2 antibody generation for different antigens in which the quantity of '+' indicates the relative quantification for each antibody. Cases colored in blue indicate no comorbidities and highlighted in yellow indicates no detectable antibodies. F, Female; IgG, Immunoglobulin G; IgM, Immunoglobulin M; M, Male; NP; Nucleocapsid Protein; RBD, Receptor Binding Domain.

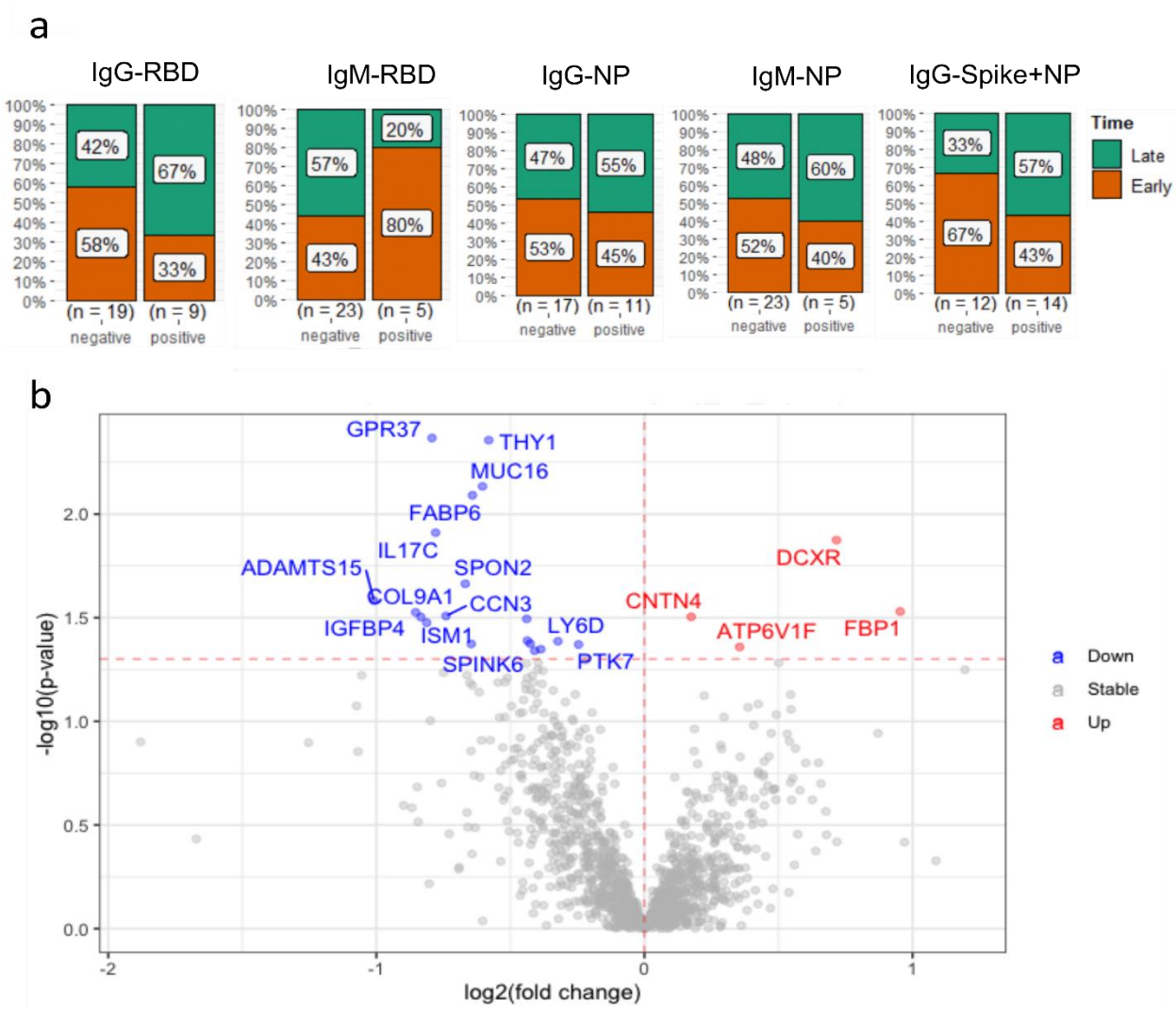

Figure 4.4. Plasma protein changes associated with SARS-CoV-2 antibody generation. (a) Bar plots represent the percentage of patients with positive and negative antibody generation for different SARS-CoV2 antigens from patients with early collected plasma (< 3 months) and late collected plasma (> 3 months). (b) Volcano plot of statistical significance against fold-change of proteins between patients with positive antibody generation and patients with negative antibody generation. The red (upregulated in patients with positive antibody generation) and blue (downregulated in patients with positive antibody generation) dots statistically indicate DEPs.

To identify which protein changes may be associated with the generation of antibodies in response to this viral infection, we compared the detected plasma proteins from patients with and without detectable anti-SARS-CoV-2 IgM or IgG antibodies. Among the 12 patients who did not generate detectable antibodies, we observed a significantly elevated level of IL17C, and a protein in the PPAR signaling pathway, namely FABP6, compared to COVID-19 patients with detectable antibodies (Figure 4.4b, Appendix II Table S6). Moreover, SPON2, involved in innate immune responses [373] was increased in these antibody negative patients. ATP6V1 was upregulated in patients with antibody generation. This protein acidifies cellular compartments as well as the extracellular environment which was previously found upregulated in SARS-CoV-2 infected cells [374]. Interestingly, it has been reported that an acidic pH environment facilitates SARS-CoV-2 infection [375]. Taken together, the comparison of the samples from the SARS-CoV-2 antibody negative and positive patients revealed plasma protein changes associated with SARS-CoV-2 antibody generation, suggesting potential roles of these proteins in SARS-CoV-2 immune responses.

4.3.5. Identification of key immune signatures and novel protein changes caused by a SARS-CoV-2 infection

In this study, 220 DEPs were identified when accounting for all of the comparisons. Among which, 59 proteins were identified in the human blood secretome according to HPA (Appendix II Table S2). We revealed the upregulation of several chemokines and cytokines, including CCL26, HGF, TNFSF13B/BAFF, especially in patients with comorbidities, which promote inflammation, innate immune responses as well as T/B-cell-driven immune responses (Figure 4.5a). Interestingly,

pleiotrophin (PTN), a secreted growth factor that induces the stimulation of the expression of inflammatory cytokines was upregulated in patients with comorbidities compared to their DCs (Figure 4.5b). This is consist with the previous findings in which PTN was elevated at the mRNA level in the upper airway of COVID-19 patients as well as COVID-19 patients over the age of 60 [376,377]. Another growth factor upregulated in this group of patients was fibroblast growth factor (FGF)21, which is linked to mitochondrial dysfunction in peripheral blood mononuclear cells (PBMCs) that drives a systemic immune response in COVID-19 pathogenesis [378]. Moreover, the patients with comorbidities showed the upregulation of well-established COVID-19-related inflammatory markers, such as AREG, an IL18 induced cytokine, which is involved in restoring tissue integrity [379–381]. Simultaneously, the upregulation of LRRC25 (Figure 4.5b), which is a potent negative regulator of NF-κB signaling and inflammation [382], may counteract the hyper-inflammation state in COVID-19 patients. Another possible function of LRRC25 after a SARS-CoV-2 infection is the downregulation of RLR-mediated type I interferon (IFN) signaling [383]. The elevated plasma LRRC25 was detected in patients with comorbidity compared with HCs as well as with DCs, suggesting the elevated plasma LRRC25 is caused by SARS-CoV-2 infection but not by the comorbidities. Notably, the elevated plasma level of LRRC25 detected in COVID-19 patients has not previously been reported.

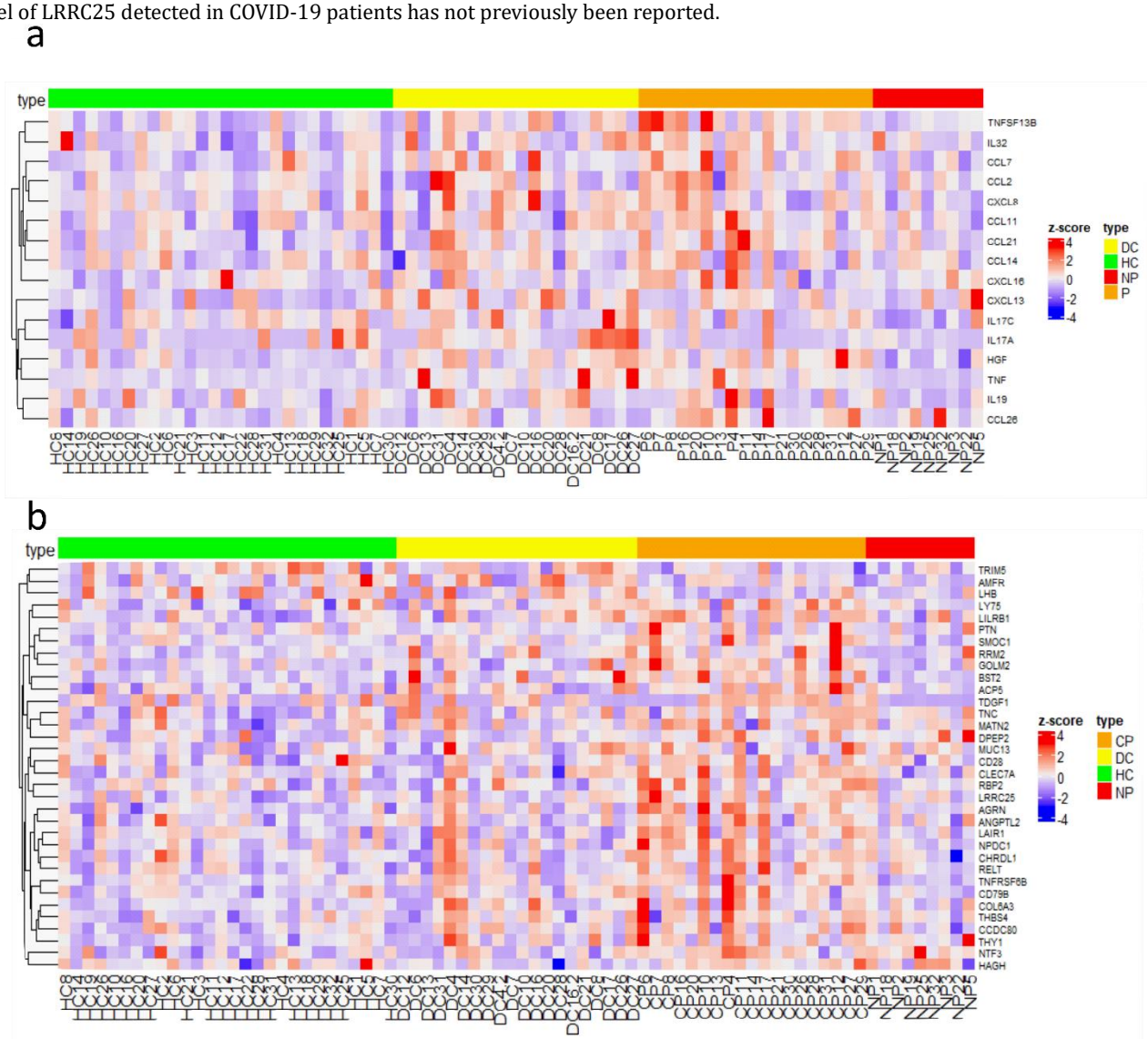

Figure 4.5. Identification of key immune signatures and novel proteins after a SARS-CoV-2 infection. (a) Heatmap of differentially expressed cytokines and (b) novel protein changes among different clinical groups (patients with and without comorbidities, disease controls, and healthy controls) with z-score by row normalization and distributed by hierarchical clustering.

Several proteins related to innate immune cell activation are upregulated in COVID-19 patients, such as FOLR2, anti-inflammatory macrophage marker [384,385], and CCL26. Notably, the elevated levels of these two proteins were detected in COVID-19 patients with and without comorbidities (Appendix II Figure S3). CCL26 is highly expressed in vascular endothelium, fibroblasts, epithelial, and blood endothelial cells [386,387]. It is a chemoattractant to recruit inflammatory cells, especially eosinophils and mast cells in allergic reaction and other immune diseases. It may also block the recruitment of Th1 and monocytes via CCR1, -2, and -5 [388]. Studies have shown that SARS-CoV-2 ORF7a activates the NF-KB pro-

inflammatory release of CCL26, an eosinophil and basophil chemoattractant,among other proinflammatory cytokines [389]. Previously, CCL26 was found upregulated in COVID-19 patients vs. HC in plasma and correlates with disease severity [390,391].

Apart from the elevated plasma proteins, we also identified several downregulated proteins after a SARS-CoV-2 infection. Among which, besides the above-mentioned RNF41, surprisingly, the inflammatory cytokine IL17A was found increased in DCs but reduced in COVID-19 patients with comorbidities compared to their DCs (Figure 4.1c, 4.5a, Appendix II Figure S3). The elevated plasma IL17A detected in DCs, especially in patients with multiple sclerosis, arthropathic psoriasis, and diverticular disease of the large intestine supports the previous findings showing its association with autoimmune inflammation [392]. However, we did not observe a significant change of plasma IL17A in response to a SARS-CoV-2 infection, regardless of with or without comorbidities. This suggests that there is no clear association between IL17A and this viral infection in the COVID-19 patients' samples that we detected.

Furthermore, we identified several plasma proteins, including retinol binding protein (RBP)-2, MATN2, THY1, SMOC1, CHRDL1, NPDC1, GOLM2 and RELT, which were not previously associated with COVID-19 (Figure 4.5b). RBP2 has a function in the absorption of dietary retinoid. Human RBP2 bound all-trans-retinol and all-trans-retinaldehyde but not all trans-retinoic acid. RBP2 protein is highly expressed in the intestine (Appendix II Figure S4) and plays a central role in maintaining intestinal innate immunity: Dendritic cells use all-trans-retinoic acid to promote intestine-specific immune responses, including Foxp3+ Treg conversion, lymphocyte gut homing molecule expression, and IgA production. RBP2 is required for CD103+ DCs with the ability to generate gut tropic T cells [393]. We found elevated plasma RBP2 in COVID-19 patients with comorbidities compared to the HCs as well as to their DCs (Figure 4.5b, Appendix II Figure S3), but no significant changes were observed in patients without comorbidities vs. HCs. This indicates that the elevated RBP2 is the synergistic effect of comorbidities and SARS-CoV-2 infection.

Taken together, plasma proteomics analysis in COVID-19 patients revealed protein changes with immunological signatures. In addition to a SARS-CoV-2 infection, comorbidities cause plasma protein changes. Several novel potential biomarkers associated with a SARS-CoV-2 infection in plasma were identified. Further functional characterization of these proteins in COVID-19 will lead to a better understanding of the immune response in a SARS-CoV-2 infection and may facilitate the development of novel therapeutic targets, diagnosis, and prognosis of COVID-19.

4.3.6. Discussion

In the present study, we applied an antibody-based proteomic technology, Proximity Extension Assay (PEA) to detect 1,463 proteins, including inflammation, oncology, neurology, and cardiometabolic panels from just a few micro-liters of COVID-19 patients' plasma samples. The technology is based on target-specific antibodies conjugated with unique complementary DNA. The antibody pairs targeting one protein bind to the target and a barcoded DNA duplex is formed, which is amplified by next-generation sequencing (NGS) [174]. Due to its high sensitivity and low sample volume requirement, PEA technology has been applied on the discovery and monitoring of biomarkers as well as on the diagnosis and prognosis of several diseases, such as infection, inflammation, cardiovascular diseases, neurological diseases, and cancer [394–398]. However, like all other antibody-based approaches, this technology is limited by the availability and specificity of antibodies and, more importantly, the number of proteins is preselected. Nevertheless, the application of PEA plasma proteomics enabled us to identify 34 novel potential biomarkers for SARS-CoV-2 infection, such as RNF41, FOLR2, RBP2, PTN, LILRA5, and CLEC7A among others.

In the present study, immune signatures including both innate and adaptive immunity were identified in the plasma samples of COVID-19 patients. Several proteins with a function in innate immune responses were found upregulated in COVID-19 patients. In addition to CCL26 and FOLR2, LILRA5, which is an orphan receptor that stimulates cytokine production in monocytes, [399] was also upregulated. To the best of our knowledge, this is the first time that LILRA5 has been identified upregulated in COVID-19 patients' serum as a novel potential biomarker, which was only previously found upregulated in the kidney of COVID-19 patients at the mRNA level [400]. Moreover, patients with comorbidities showed the upregulation of Dectin-1 (CLEC7A), a pattern-recognition receptor (PRR) that stimulates NFAT activation in DCs and macrophages [401] and may participate in cross-communication with TLRs during S protein and DAMP identification and stimulation [402]. Simultaneously, some immune suppressors are upregulated that potentially dampen the immune response, such as VSIG4 inhibiting macrophage and T cell cytotoxicity [403–405], PILRA that was previously found upregulated in monocytes from severe stage COVID-19 patients [406], and SIGLEC9 that potentially inhibits NK cells and neutrophils [407,408]. As expected, SARS-CoV-2 infection promotes the expression of genes directly related with antiviral activity such as CCL26, GRN, and BST2. In fact, tetherin (BST2) is a transmembrane protein with antiviral activity by tethering nascent virions in the plasma membrane, which can be retained or mobilized for endocytic internalization and subsequent ubiquitin-based degradation [409].

SARS-CoV-2 infection promotes T cell activation and cycling [410]. In fact, COVID-19 patients presented the upregulation of T cell markers, such as CD4, CCL21, CD48, and TNFRSF1B/TNFR2. Interestingly, our dataset revealed several upregulated plasma proteins that have not previously been reported to be associated with a SARS-CoV-2 infection, such as RELT that is a receptor capable of stimulating T-cell proliferation in the presence of CD3 signaling [411]. In concordance, several markers of APC activation were also found upregulated such as CD83, and CD74 involved in MHCII antigen presentation [412]. Simultaneously, inhibitory markers were upregulated in COVID-19 patients that potentially counterbalance the hyperactivation of the immune response, such as NT5E/CD73, PD1 and TNFRSF8/CD30. Another novel elevated protein in this group is LAIR1, which downregulates IL2 and IFNγ expression in CD4+ T cells as well as IgG production, IL8, IL10, and

TNF secretion in B cells [413,414]. Moreover, the upregulation of LGALS9 and its receptor HAVCR2 (TIM3) in this group of patients suggests the exhaustion of CD4+ T cells [415].

Apart from immune-related proteins, our dataset revealed the upregulation of proteins involved in metabolic processes, such as leukotrienes synthesis (DPEP2 and LTA4H), histamine degradation (HNMT), retinoic acid synthesis (RBP2), and fatty acid metabolism (FABP1). In addition, proteins related with neuronal damage were found upregulated, such as AGRN, NTF3, and CHRDL1/CHL1. CHRDL1 plays essential roles in many developmental processes, including neurogenesis, vascular development, angiogenesis, and osteogenesis [416]. Although it is secreted by astrocytes, now, for the first time, we reported elevated plasma CHRDL1 in COVID-19 patients with comorbidities. Moreover, extracellular matrix proteins were found upregulated in COVID-19 patients that indicates the potential tissue damage and remodeling due to a SARS-CoV-2 infection such as the pro-inflammatory TNC [417], COL6A3, and THBS4/TSP4, which is involved in tissue regeneration and wound healing [418], not previously found in COVID-19 patients. In agreement with previous findings, the Von Willebrand factor (VWF) was upregulated in COVID-19 patients which can contribute to their hypercoagulable state and increased venous thromboembolism rate [419,420].

Several studies have demonstrated that the underlying medical conditions may increase the risk of infection with SARS-CoV-2 or the severity of COVID-19 [354–357]. Importantly, the majority of the population with SARS-CoV-2 infection does not develop severe symptoms that require admission in ICUs, therefore, we cohort nicely represents the overall situation in population [339]. However, recently, more attention has been drawn to the fact that considerable amount of people infected by SARS-CoV-2 may develop long COVID, a condition currently with very limited knowledge. In the present study, the application of the PEA, an antibody-based proteomic strategy, resulted in the quantification of over 1,000 plasma protein changes in COVID-19 patients mostly with mild symptoms. Further analysis unveiled immunological signatures and several novel protein changes associated with a SARS-CoV-2 infection, their underlying diseases as well as antibody generation. In fact, our data demonstrated that COVID-19 patients with comorbidities shared plasma protein signatures that reflect their underlying immune and physiological responses, despite the heterogeneity of medical conditions. Furthermore, the characterization of long-term plasma protein responses resulted in the identification of novel potential biomarkers for post-COVID-19 condition. Our data provides a valuable resource for the further functional characterization of novel players in a SARS-CoV-2 infection that could also lead to the development of novel biomarkers for the diagnosis and prognosis of COVID-19.

CHAPTER 5. Mass spectrometry proteomics characterization of plasma biomarkers for colorectal cancer associated with inflammation

Once sample preparation protocols and bioinformatics analysis were set up and optimized with previous COVID-19 studies, the second main part of the thesis focused on CRC. CRC biomarkers are urgently needed together with deeper understanding of the systemic response to CRC. In this chapter, LC-MS/MS proteomics was applied to plasma samples from CRC patients and healthy controls to characterize the protein changes caused by CRC development, progression and cancer-associated inflammation. This multi-center study included samples from four different biobanks 3P–Medicine Laboratory, Medical University of Gdansk, Biobank HARC, Medical University of Lodz (Poland), Bank of Biological Material at Masaryk Memorial Cancer Institute (Czech Republic) and Leipzig Medical Biobank (Germany) and it was performed in collaboration with the IFB Laboratory of Mass spectrometry.

This study was originally published in **Biomarker Insights** and is presented with minor modifications.:

***Urbiola-Salvador V*, Jabłońska A, Miroszewska D, Kamysz W, Duzowska K, Drężek-Chyła K, Baber R, Thieme R, Gockel I, Zdrenka M, Śrutek E, Szylberg Ł, Jankowski M, Bała D, Zegarski W, Nowikiewicz T, Makarewicz W, Adamczyk A, Ambicka A, Przewoźnik M, Harazin-Lechowicz A, Ryś J, Macur K, Czaplewska P, Filipowicz N, Piotrowski A, Dumanski JP, Chen Z. Mass spectrometry proteomics characterization of plasma biomarkers for colorectal cancer associated with inflammation*. *Biomark. Insights*. 19, 11772719241257739 (2024, in press).**

5.1. Introduction

Colorectal cancer (CRC) is the third most incident malignancy and the second most deadly cancer worldwide [78]. Despite the great advances in CRC treatment with recently developed immunotherapies, about 20–25% of diagnosed CRC patients present advanced cancer stages and metastasis that is linked to a 5-year survival rate lower than 10% and low therapeutic response [421,422]. In contrast, diagnosis at early stages leads to reduced tumor-related mortality and a 90% 5-year survival rate after radical surgical resection [423]. Apart from the disease stage at diagnosis, CRC prognosis depends on multiple factors such as location, genetic factors, molecular expression profiles, tumor immune infiltration, and inflammation [422]. The low therapeutic response to immunotherapies such as immune checkpoint inhibitors may be caused by the influence of other non-targeted inflammatory and immunosuppressive mechanisms [52]. Notably, cancer-associated inflammation is considered a well-established hallmark of cancer, especially in CRC [424]. Inflammatory modulators including chemokines, cytokines, and growth factors influence the interactions between cancer cells and the tumor microenvironment driving tumor progression and the immune response [425]. Moreover, CRC progression can promote systemic inflammation impacting other organs and facilitating metastasis [424].

Currently, the gold standard for CRC prevention is colonoscopy complemented with fecal occult blood tests [426]. However, colonoscopy is expensive and has poor patient compliance, due to its invasiveness and risks, while stool-based tests have low sensitivity and specificity [114,427]. Therefore, alternative, non-invasive, cost-effective, and easily measurable CRC screening strategies are urgently needed. Mass spectrometry (MS)-based proteomics approaches have been successfully applied to determine blood-based biomarkers of CRC development and progression [114]. MS-based proteomics characterization of low-abundance proteins in serum/plasma is limited by the high dynamic range of protein concentrations over nine orders of magnitude with 99% of the total protein content from only 20 abundant proteins [428]. However, the technological evolution of high-resolution MS instruments such as time-of-flight (TOF) or Orbitrap provides the possibility to discover blood-based biomarkers with high sensitivity and specificity [429].

Nowadays, the most common blood protein biomarker used in clinical CRC diagnosis is carcinoembryonic antigen (CEA), but its accuracy requires improvement [430]. Interestingly, untargeted tandem MS coupled with liquid chromatography (LC-MS/MS) proteomics strategies could discover novel potential CRC biomarkers that can be validated by using targeted MS techniques as well as antibody-based assays [114]. For instance, proteomics analysis discovered that several SERPIN family members are altered in patients with CRC and adenomatous polyps which were validated as potential diagnostic biomarkers by ELISA [431]. Moreover, plasma proteomics analysis combined with neural network classification identified five candidate biomarkers to distinguish between CRC stages [432]. Another glycoproteomics study detected novel diagnostic biomarkers includingElevated levels of complement C9 and fibronectin improved the diagnostic performance of a commercial CEA CRC biomarker [433]. In addition, targeted proteomics analysis in a non-metastatic CRC cohort determined a five protein signature with efficient discrimination of CRC cases from healthy subjects [434]. However, despite advances in CRC biomarker discovery and validation by proteomics, further studies are needed in larger cohorts to implement reliable biomarkers in clinical practice.

The aim of this study was to discover novel plasma protein signatures involved in CRC development and progression by untargeted LC-MS/MS proteomics analysis. Importantly, significant changes in plasma protein levels were identified that were associated with cholesterol metabolism, members of the SERPIN family as well as increased levels of complement cascade proteins in CRC patients versus healthy subjects. Furthermore, high complement C5 levels were confirmed in the validation cohort, being a potential diagnostic CRC biomarker. Plasma protein levels of 11 proteins, including complement C8A and serpin family A member 4 (SERPINA4) were linked to cancer-associated inflammation, while 4 proteins, including C8A and C4B, distinguished early from advanced CRC stages.

5.2. Materials and Methods

5.2.1. Study cohorts and design

This multi-center retrospective study included 36 patients with CRC surgery (age mean: 66.1 ± 11.6 years; 44.4% male) from June 2019 to April 2021 and 26 healthy subjects (age mean: 61.1 ± 10.5 years; 42.3% male) in the discovery cohort. Included patients were with positive colonoscopy and pathologist-confirmed malignant neoplasm. Patients with prior neoadjuvant therapy administration were excluded from the analysis. 69.4% (25 of 36) of diagnosed patients were with advanced CRC stages (III-IV) according to the Union for International Control of Cancer TNM classification and 30.5% (11 of 36) presented cancer-associated inflammation post-operatively assessed by pathologists. Blood samples of healthy subjects and CRC patients were obtained from Biobank HARC, Medical University of Łódź and the 3P–Medicine Laboratory, Medical University of Gdańsk. The independent validation cohort included 60 CRC patients (age mean: 61.8 ± 11.4 years; 51.7% male) without neoadjuvant therapy and 44 sex-and-age-matched healthy subjects. Serum samples were obtained from the Leipzig Medical Biobank, Germany and the Bank of Biological Material at Masaryk Memorial Cancer Institute, Czech Republic. The collection of whole blood samples was with sterile BD Vacutainer® K2EDTA tubes or Sarstedt S-Monovette® 2.7 mL, K3 EDTA (LMB) before the CRC resection followed by centrifugation, aliquoting, and storage at -80°C until use.

5.2.2. Sample preparation for mass spectrometry

Proteins were extracted from plasma samples with lysis buffer (1% SDS, 50 mM DTT, 100 mM Tris-HCl pH 8.0) (Merck KGaA, Darmstadt, Germany) containing phosphatase and protease inhibitors (Thermo Fisher Scientific, Waltham, MA, USA) followed by an incubation at 95°C for 10 min. Protein concentrations were determined at 280 nm in a µDrop plate with a Multiskan Thermo Nanodrop. Then, 100 µg of proteins were transferred to Microcon 10 kDa filters (Merck KGaA) and were processed based on the Filter Aided Sample Preparation (FASP) protocol [110]. Briefly, three washes with 200 µl of urea buffer (8 M urea,100 mM Tris-HCl pH 8.5) at 10,000 rcf for 20 min at room temperature (RT) were applied to the protein mixtures. Free cysteines were alkylated by incubation in the darkness for 20 min at RT with 55 mM iodoacetamide (100 µl) in urea buffer (Merck KGaA). Samples were centrifuged at 10,000 rcf for 15 min and washed three times with urea (100 µl) and two times with digestion buffer (50 mM Tris-HCl pH 8.0). Afterward, the filters were transferred into new tubes and proteins were digested by incubation at 37°C with 1 µg of Sequencing Grade Modified Trypsin (Promega, Madison, WI, USA) in 60 µl of digestion buffer overnight. Then, the elution of peptides was performed with the same centrifugation conditions and washed two times with 125 and 100 µl digestion buffer. Next, 0.1% trifluoroacetic acid quenched trypsin activity. Peptide concentrations were measured as previously and 20 µg of peptides were desalted with STop And Go Extraction (STAGE) Tips

[294] in Empore C18 extraction disks (3M, Neuss, Germany). Peptides were eluted with 60% acetonitrile and 1% acetic acid. Desalted peptides were dried in a SpeedVac at 45°C and samples were in storage at -20°C until analysis.

5.2.3. LC-MS/MS analysis

LC-MS/MS analysis of prepared samples was performed with a TripleTOF 5600+ mass spectrometer (SCIEX, Framingham, MA, USA) and with an EkspertMicroLC 200 Plus System (Eksigent, Redwood City, CA, USA). AB SCIEX Analyst TF 1.6 software was used to control the LC-MS/MS system. Samples were run in triplicates with 1.5 µg injected peptides in each technical replicate. Analyses were in a ChromXP C18CL column (3 µm, 120 Å, 150 × 0.3 mm) at 5 µl/min and 35°C, for 60 min with an 11–35%. acetonitrile gradient in 0.1% formic acid. TripleTOF 5600+ was set in data-dependent acquisition mode and the m/z range of the TOF MS survey scan was at 400-1200 Da with an accumulation time of 250 ms. The selection for collision-induced dissociation (CID) fragmentation was set to a maximum of top 20 precursor ions with +2 to +5 charges. The exclusion of precursor ions from reselection was for 5 s after two occurrences. Product ions spectra were acquired between 100 and 1800 Da with 50 ms accumulation time.

5.2.4. MS data analysis

Acquired raw SCIEX files were converted to mzML format with MSConvertGUI 3.0 and analyzed using PeaksStudio Xpro 10.6 software (Bioinformatics Solutions, Waterloo, ON, Canada). Peptide sequence search was against the *Homo sapiens* UniProtKB/Swiss-Prot database (release 2022_03) for trypsin digested peptides with maximum 3 missed cleavages per peptide. Carbamidomethylation was as fixed post-translational modification (PTM), whereas N-terminal acetylation and methionine oxidation as variable PTMs. Peptide and protein identification was with a < 1% false discovery rate (FDR). Label-free quantification was performed based on the integration of the peptide areas under the curve (AUC).

5.2.5. Complement C5 validation

Complement C5 serum concentrations were quantified in the validation cohort by an ELISA kit with a coated antibody to human C5 (Abcam ab125963, Cambridge, UK) commercially available, following manufacturer's instructions.

5.2.6. Proteomics data and statistical analysis

Statistical analysis was performed with R (version 4.0.3) (R Foundation for Statistical Computing, Vienna, Austria) in RStudio (version 1.3.1093) (RStudio, PBC, Boston, MA, USA). Data preprocessing was performed by summarization of technical replicates with medians and logarithmic transformation of relative abundances. Proteins with missing values in over 50% of patients and 50% of healthy controls were filtered. Random forest imputation was applied to the remaining missing values with the "missForest" R package (version 1.5) followed by quantile normalization. Differences in protein levels between groups were analyzed by the general linear model regression approach with contrast analysis with the "emmeans" R package (version 1.6.2.1). First, for each protein, a general linear model was generated to fit its expression to determine significant changes in CRC patients compared to healthy volunteers including age as a confounding factor. Then, for each protein expression, a general linear model was generated including only CRC patients with the independent variables inflammation and tumor stage while sex was considered a confounding factor. FDR control was applied with the Benjamini & Hochberg correction. Significant changes were considered with FDR-adjusted p-value < 0.05. Point-biserial correlation of protein abundance with inflammation status or tumor stage was calculated with the built-in R function cor.test and correlation was significant with a p-value < 0.05. Principal Component Analysis (PCA) was performed using prcomp built-in R function and PCA visualization using "factoextra" R package (version 1.0.7). Functional annotation of biological process and cellular component GO terms was performed by a two-sided hypergeometric test with FDR correction using the Cytoscape cluGO plugin (version 2.5.7). Pathway enrichment analysis of KEGG terms supported by active subnetworks was applied with the R package "pathfindR" (version 1.6.3) using the STRING database and FDR correction. The generation of graphics was with the R package "ggplot2" (version 3.3.5), with the exception of heatmaps generation by the R package "ComplexHeatmap" (version 2.6.2). The construction of the protein network was with Cytoscape (version 3.8.2) using the STRING database and a 0.7 confidence cut-off.

5.3. Results

5.3.1. Identification and quantification of the plasma proteome of CRC patients using LC-MS/MS

To study the protein profile changes in blood involved in CRC development, we applied LC-MS/MS proteomics analysis to plasma samples of 36 CRC patients and 26 healthy controls. As a result, 322 proteins were identified with at least 1 unique peptide with FDR < 0.01, from which the majority of proteins were identified in both groups (Figure 5.1a). Interestingly, IgGFc-binding protein (FCGBP), which is a mucin responsible for innate immune defense in the intestine and is associated with CRC metastasis by promoting cell adhesion, was only identified in CRC patients [63].

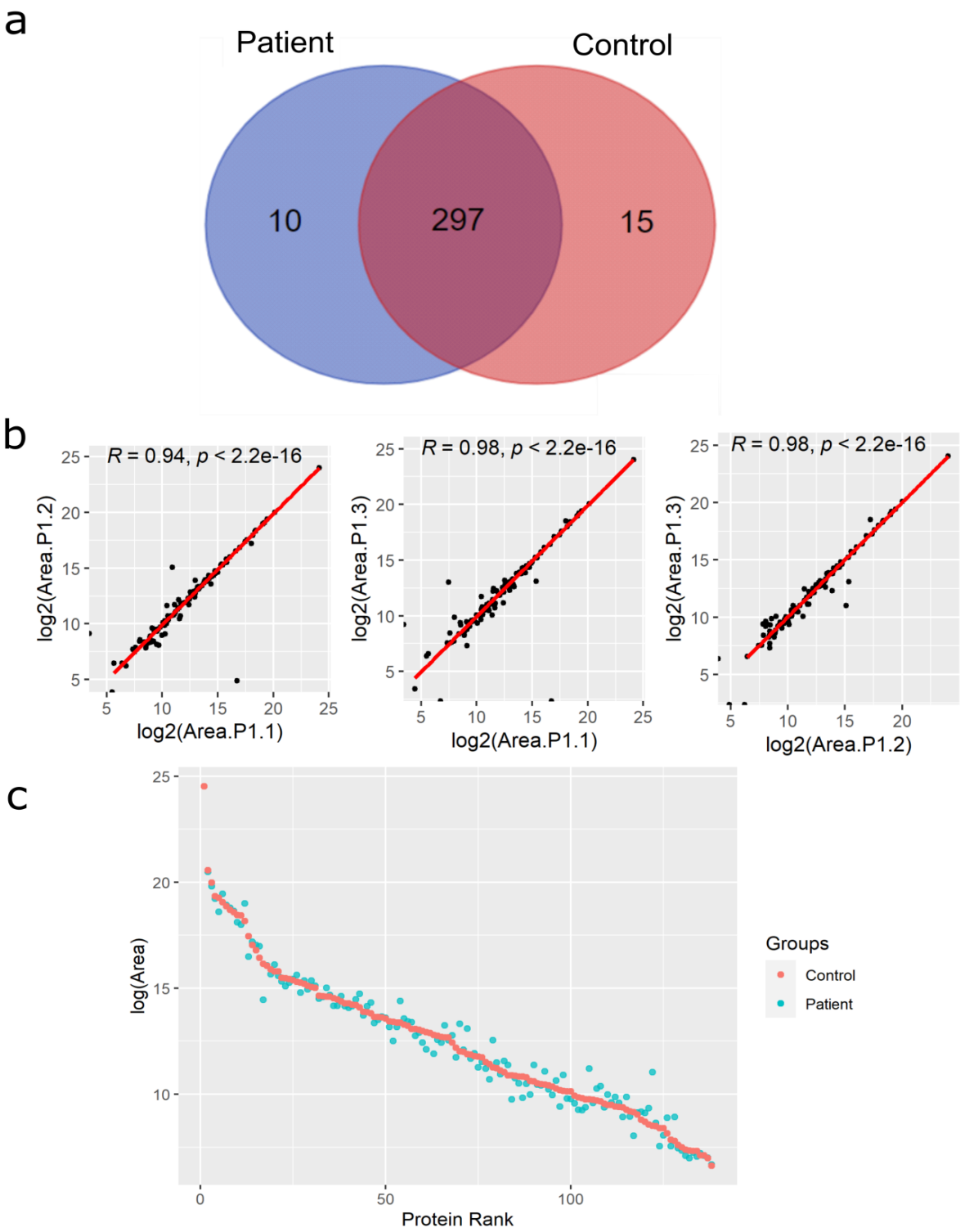

Figure 5.1. LC-MS/MS analysis of plasma proteome from CRC patients and healthy controls. (a) Venn diagram of identified proteins in CRC patients and healthy individuals. (b) Representative scatter plots of log-transformed areas for the three technical replicates from a CRC patient (P1) with their corresponding Pearson correlation coefficients and p-values. (c) Abundance protein ranking plot with the mean of log-transformed areas from healthy subjects (red) and CRC patients (blue).

After filtering proteins with a high % of missing values, 138 protein groups were analyzed with reliable quantification across the samples. The relative protein abundance was reproducible along technical replicates with high Pearson's correlation coefficients (Figure 5.1b). LC-MS/MS analysis quantified proteins in a high dynamic range of concentrations

from high-abundance albumin in the range of mg/mL to chemokines such as C-X-C motif chemokine ligand (CXCL)-7 in the range of ng/mL (Figure 5.1c).

Functional annotation of the identified proteins determined that the majority were from the extracellular organelles, blood, and lipoprotein microparticles, as well as the vesicle/vacuolar lumen (Figure 5.2a). However, proteins from the plasma membrane, cytoplasm, and nucleus, such as histone H4, were also detected that may circulate in the peripheral blood due to tissue damage and cell turnover. (Appendix III Table S2). Identified proteins were included in several biological processes such as blood coagulation, homeostasis, proteolysis, and several metabolic processes including cholesterol and fatty acid metabolism, vesicle-mediated transport, cell death as well as humoral immune and inflammatory responses (Figure 5.2b). Interestingly, over-represented biological process GO terms were associated with different humoral immune and inflammatory responses due to the presence of immunoglobulins, complement proteins, and some chemokines such as CXCL7 (Figure 5.2c, Appendix III Table S2). Overall, our proteomics analysis identified plasma proteins associated with different biological processes including immune responses and quantified 138 proteins in a high dynamic range of concentrations with high reproducibility.

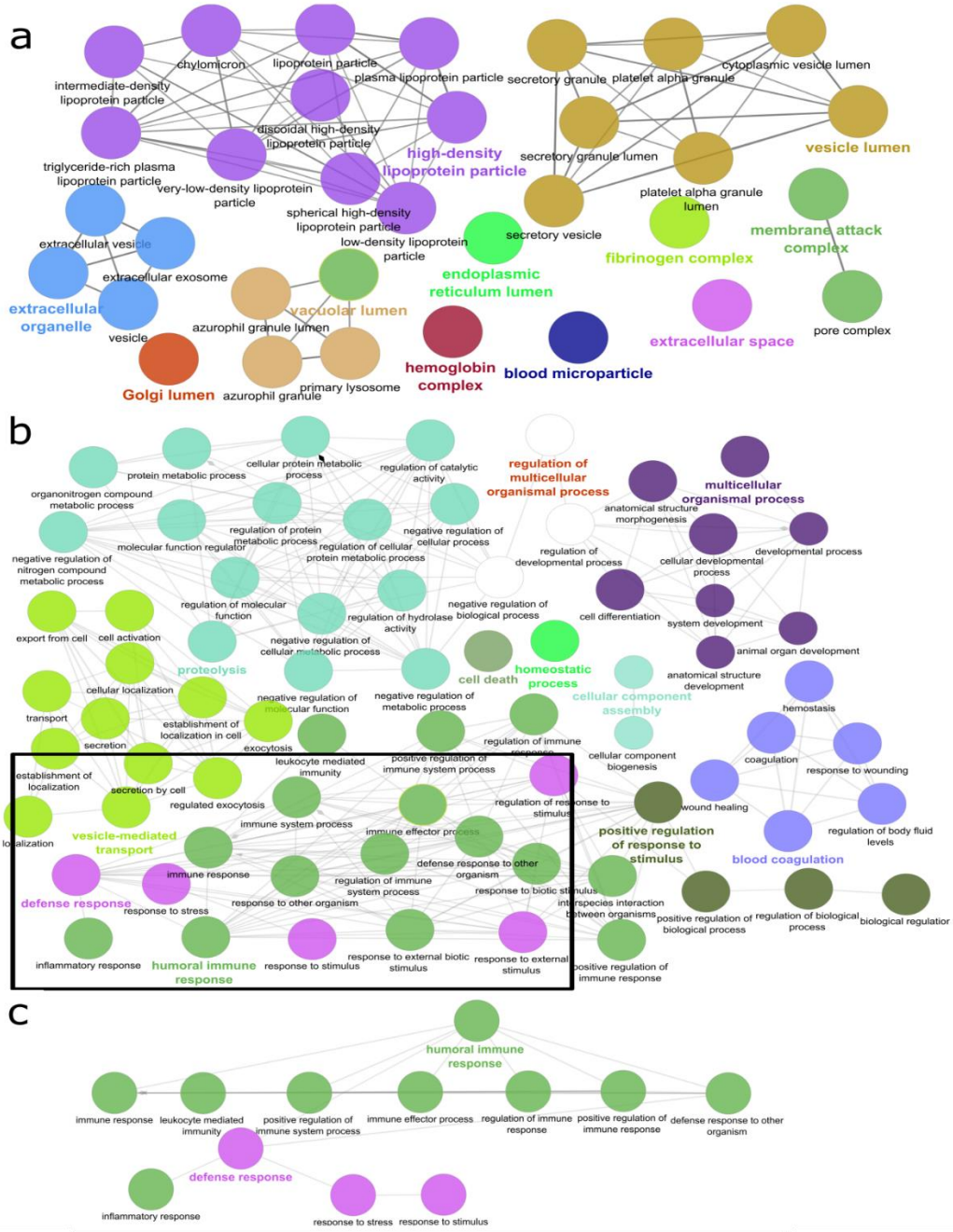

Figure 5.2. Functional annotation of the identified plasma proteins. (a) Interaction network of over-represented cellular component Gene Ontology (GO) terms with an organic l1ayout. (b) Interaction network of over-represented GO terms of biological processes with an organic layout. (c) Amplification of the subnetwork of GO terms from immune and defense responses with a tree layout.

5.3.2. CRC development causes protein plasma changes associated with the complement cascade and cholesterol metabolism

To determine whether the plasma levels of quantified proteins differ in CRC patients versus healthy volunteers, PCA was performed. PCA showed a clear separation of plasma from CRC patients and healthy subjects, indicating that CRC development affects the protein plasma profiles in examined patients (Figure 2.3a). To unveil these protein changes, differential protein expression analysis was applied by linear regression modeling with FDR correction, resulting in 17 proteins with enhanced levels and 20 decreased proteins in CRC patients versus healthy volunteers (Figure 5.3b, Appendix III Table S3). Among the differentially expressed proteins (DEPs), inter- ITIH3, leucine-rich alpha-2-glycoprotein (A2GL), C9, and LBP showed the highest levels in CRC patients, while APOA4, acid labile subunit (ALS), and kallikrein B1 (KLKB1) showed the lowest levels compared to healthy controls. ITIH3, a hyaluronan essential for multiple cellular processes, which transports and regulates hyaluronan turnover in the blood circulation, was found with the highest fold change. Unsupervised hierarchical clustering showed that these 37 DEPs separated CRC from control samples (Appendix III Figure S1).

Pathway enrichment analysis of KEGG terms by active subnetworks revealed that complement and coagulation pathways were activated with elevated protein levels (C4B, C5, C1QB, and C9) in CRC patients (Figure 5.3c, Appendix III Table S4). Moreover, cholesterol metabolism, vitamin digestion, and adsorption were down-regulated in CRC patients, involving two apolipoproteins, APOA2 and APOA4 (Figure 5.3b-c). Both APOA2 and APOA4 are associated with obesity and hypercholesterolemia that are independent risk factors for CRC development [436,437]. Similarly, the STRING protein-protein interaction network showed the interaction between the complement proteins with elevated levels (Figure 5.3d). In addition, SERPINC1 was the most interconnected node linking complement proteins to other DEPs in the network. SERPINC1, also called antithrombin III, is the main inhibitor of blood coagulation which can attenuate inflammatory responses [438]. Collectively, our analysis indicates that development of CRC causes plasma protein changes which are associated with complement cascade and cholesterol metabolism.

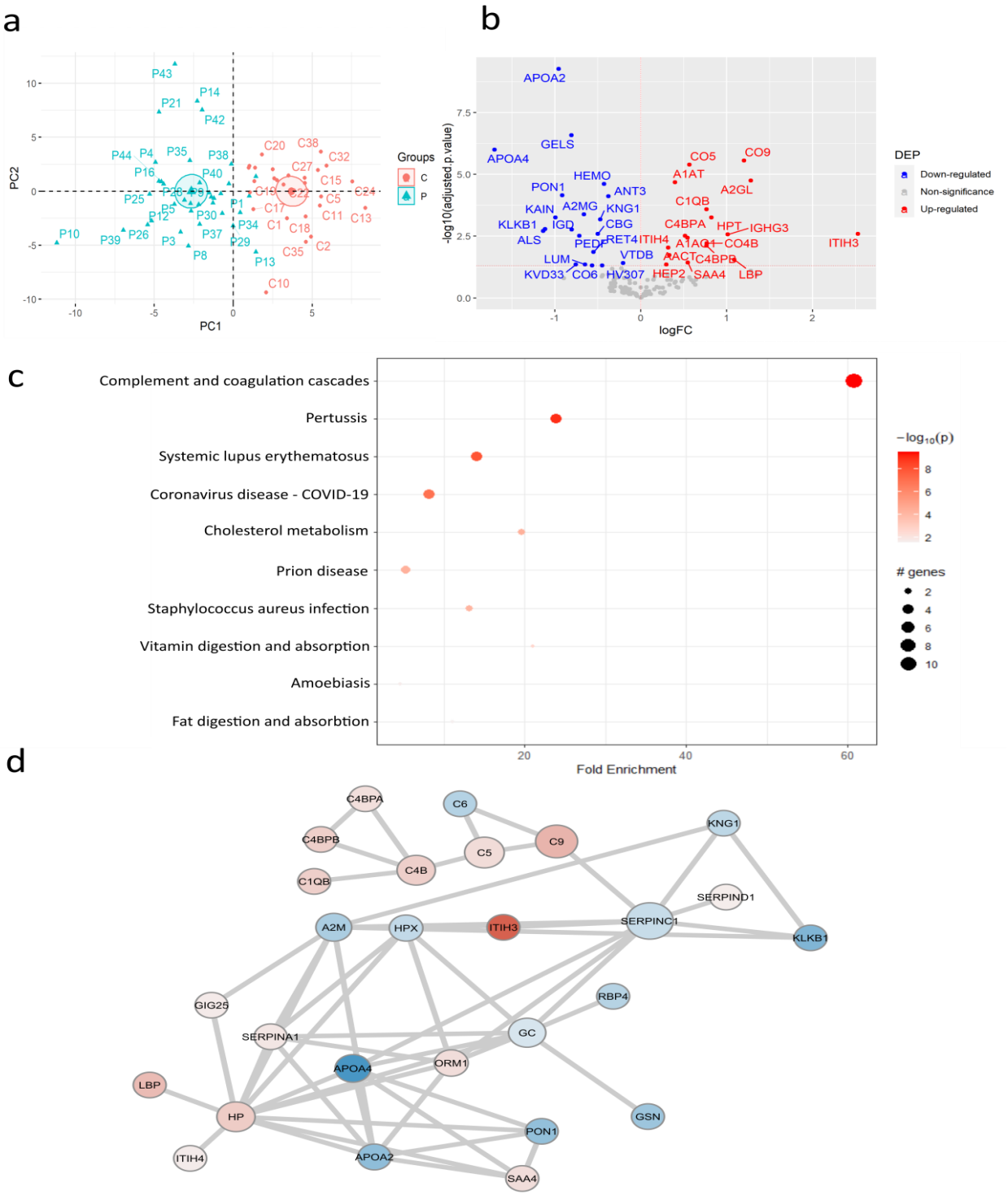

Figure 5.3. Colorectal cancer (CRC) development causes plasma protein changes involved in complement cascades and cholesterol metabolism. (a) Principal Component Analysis of CRC patients and healthy subjects using the relative abundances of all quantified proteins. (b) Volcano plot of statistical significance against fold-change of proteins between CRC patients and healthy individuals. Colored dots indicate statistically differentially expressed proteins (DEPs) calculated by the general linear model approach. (c) Dot plot of KEGG pathway enrichment combined with STRING protein-protein interaction network analysis from DEPs between CRC patients and healthy subjects. (d) Protein-protein interaction network of DEPs between CRC patients and healthy individuals from STRING database query with a 0.7 confidence cut-off. The size of nodes indicates the degree of connectivity of the nodes. The red and blue dots/nodes represent up-regulation and down-regulation in CRC patients, respectively. FC, Fold Change; p, p-value; PC, Principal Component.

5.3.3. Plasma protein changes linked to cancer-associated inflammation in CRC patients

Inflammation is a well-established hallmark of cancer that influences CRC progression. To analyze protein changes in plasma associated with inflammatory status, the protein levels were compared between CRC patients with cancer-associated inflammation (11 of 36 cases) and without. First, correlation analysis determined significant correlation of 18 proteins with cancer-associated inflammation, including 9 proteins correlated positively such as C8A, A2GL, and CERU, while another 9 proteins including retinol-binding protein 4 (RET4) were correlated negatively (Figure 5.4a, Appendix III Table S5).

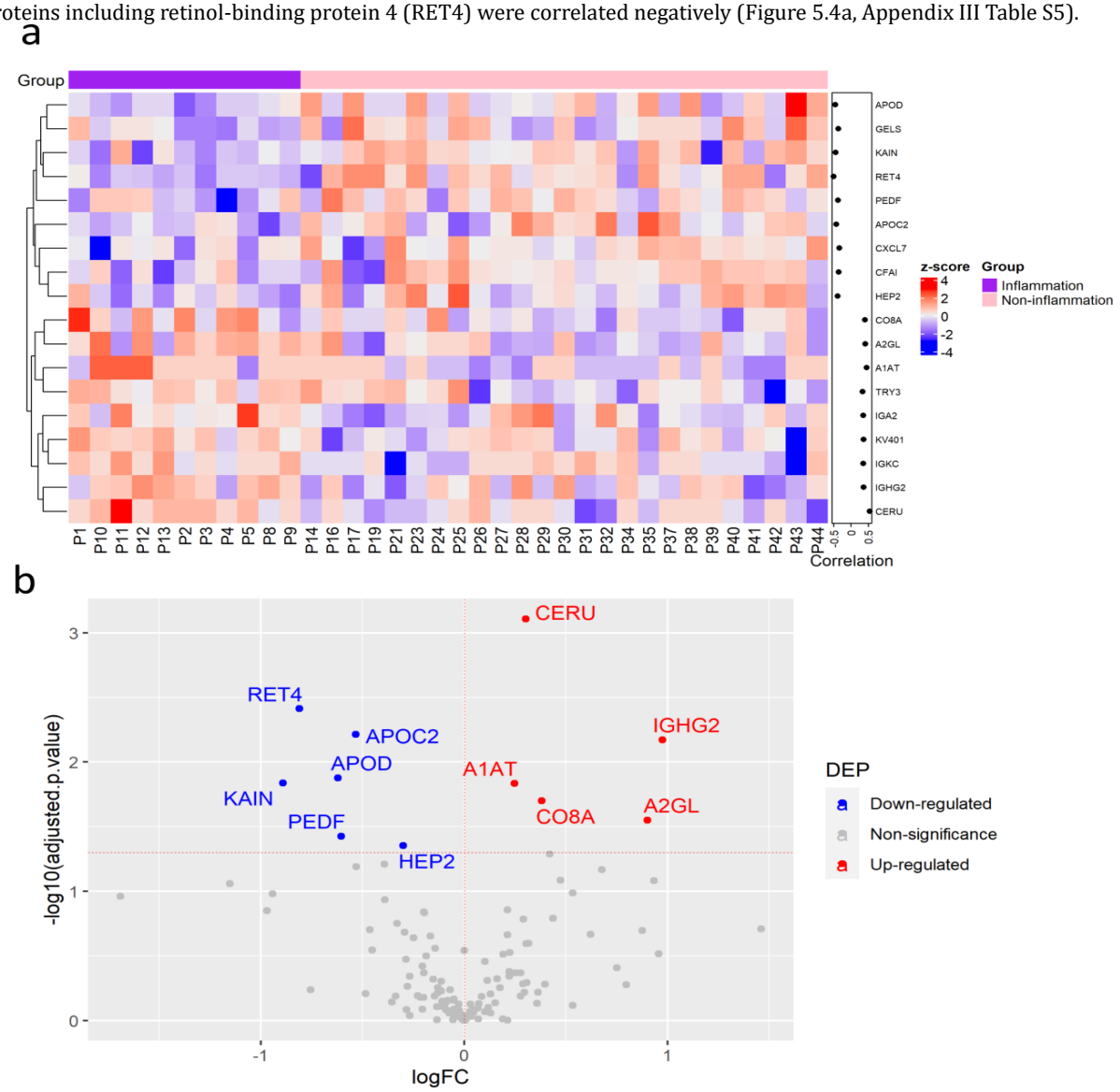

Figure 5.4. Plasma protein changes induced by cancer-associated inflammation in CRC patients. (a) Heatmap of proteins with significant correlation with inflammatory status. Protein expression is transformed with a z-score by row normalization and distributed by hierarchical clustering. The correlation coefficients (right) indicate a positive/negative correlation for each protein. (b) Volcano plot of statistical significance against fold-change of proteins between CRC patients with inflammation and without inflammation. Dots indicate individual proteins and the red and blue dots represent significant up-regulation and down-regulation in CRC patients with inflammation, respectively.

To determine the link between protein abundance and cancer-associated inflammation, the differential protein expression was evaluated by linear regression analysis. This analysis resulted in 11 DEPs that were previously identified with significant correlation (Figure 5.4b, Appendix III Table S6). Some downregulated proteins were SERPIN family members, e.g., SERPINA4 (KAIN) and SERPIND1 (HEP2). Noteworthy, SERPINA4 is an anti-angiogenic and anti-inflammatory agent that was decreased in CRC patients versus healthy volunteers and its downregulation was common in inflammatory processes as well as in cancer [439].Additionally, C8A and IGHG2 may be related to cancer-associated inflammation thus promoting an exacerbated immune response in these patients. Collectively, this analysis determined plasma protein signatures in CRC patients linked to cancer-associated inflammation.

5.3.4. Evaluation of plasma protein signatures linked to CRC stages

The main complication of CRC development is tumor progression and metastasis, resulting in increased CRC mortality. Therefore, CRC prognostic biomarkers are urgently needed. Plasma protein changes linked to CRC progression were determined by comparing protein levels in early-stage patients (I and II) versus late-patients (III and IV). Correlation analysis indicated that 5 proteins were correlated positively, while 6 proteins were correlated negatively (Figure 5.5a, Appendix III Table S7). Among them, enhanced fibrinogen alpha chain (FIBA) levels in late CRC stages and their association with distant metastasis were previously reported [440]. Also, increased alpha-1-acid glycoprotein 2 (A1AG2) was linked to shorter survival rates in a CRC cohort [441]. Similar to the previous comparison, the regression analysis showed that only were 4 DEPs (Figure 5.5b). Among them, C8A and C4B may play a relevant role in CRC progression, while the immunoglobulin IGHG2 may be associated with the immune response in CRC early stages by promoting inflammation as enhanced levels were linked to cancer-associated inflammation.

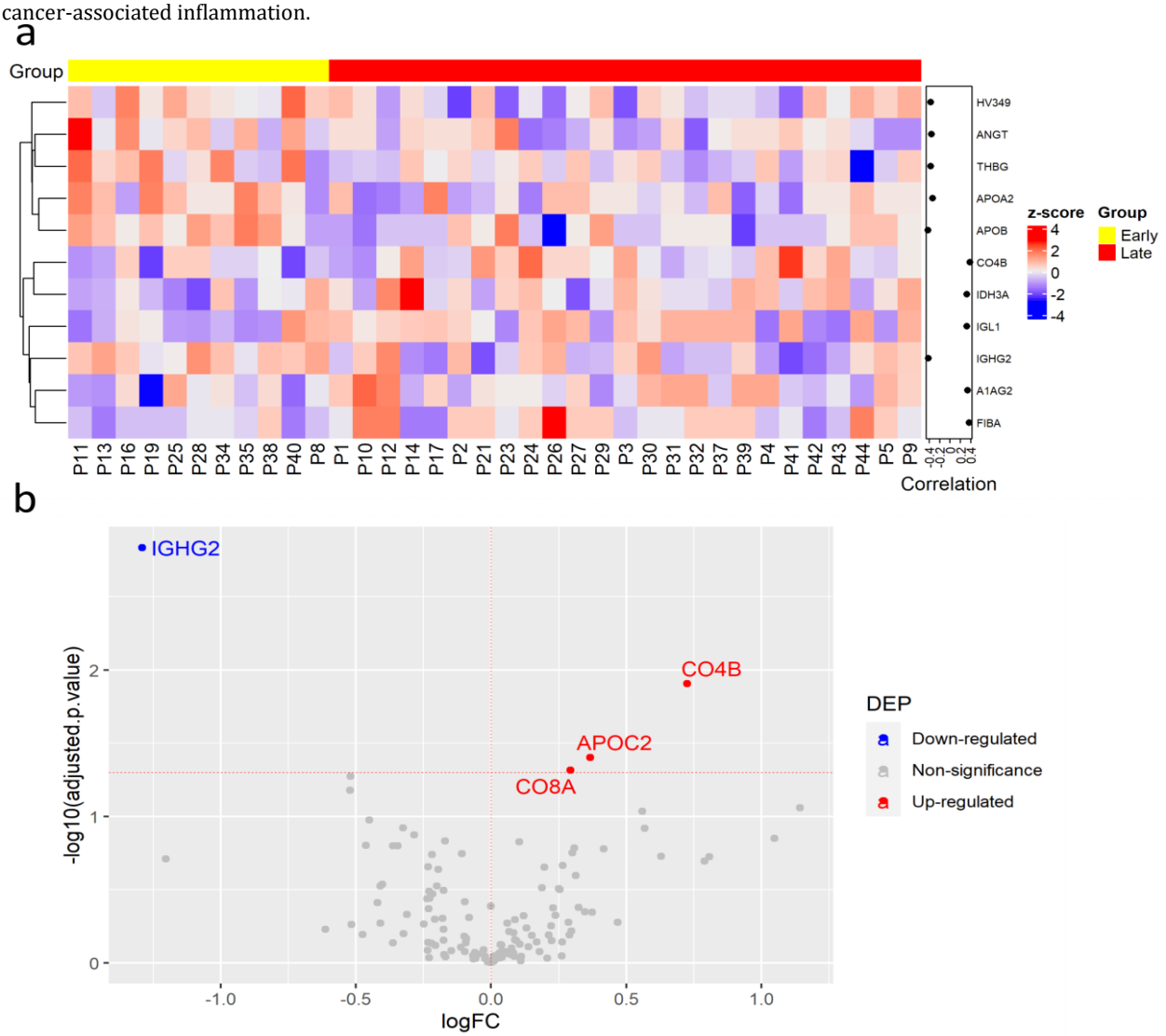

Figure 5.5. Plasma protein expression differences between early and late stages of CRC. (a) Heatmap of proteins with significant correlation with tumor stage. Protein expression is transformed with a z-score by row normalization and distributed by hierarchical clustering. The correlation coefficients (right) indicate a positive/negative correlation for each protein. (b) Volcano plot of statistical significance against fold-change of proteins between CRC patients with early tumor stage and with late tumor stage. Dots indicate individual proteins and the red and blue dots represent significant up-regulation and down-regulation in CRC patients with late tumor stage, respectively.

5.3.5. Complement protein C5 plasma levels are enhanced in CRC patients

Among the complement proteins, we found elevated C5 levels in plasma of CRC patients versus healthy volunteers by LC-MS/MS analysis (Figure 5.2b, 5.6a). To validate this finding, C5 concentrations were measured by ELISA in an independent validation cohort, including 60 CRC patients and 44 healthy subjects (Figure 5.6b). ELISA results confirmed LC-MS/MS findings. Noteworthy, a peptide from C5a was also enhanced in CRC patient's plasma (Figure 5.6c). In fact, C5

proteolytic degradation promotes the release of the anaphylatoxin C5a that is an inflammatory mediator (Figure 5.6d) [442]. Collectively, the enhanced plasma level of complement C5 is a novel promising biomarker for CRC diagnosis and may promote the release of the pro-inflammatory C5a.

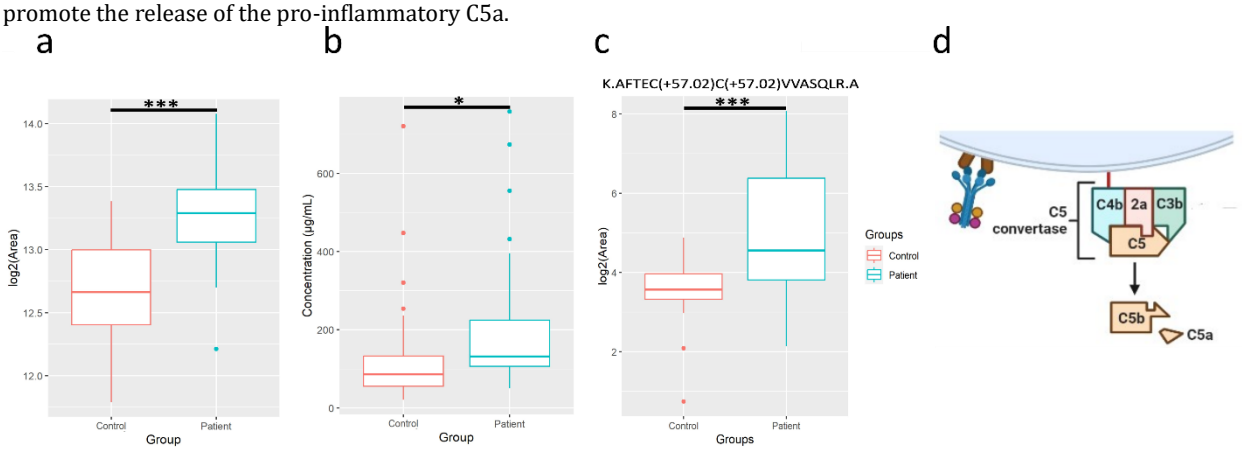

Figure 5.6. Complement protein C5 is a potential diagnostic biomarker for CRC. Box and whisker plots of (a) log-transformed areas of C5 in the discovery cohort calculated the significance by general linear model approach, (b) C5 concentrations measured by ELISA in the validation cohort calculated by Student t-test, and (c) log-transformed areas of a quantified peptide from C5a with the sequence AFTECCVVASQLR in the discovery cohort for CRC patients and healthy subjects calculated by Student t-test. * indicates statistical significance with a p-value < 0.05, and *** indicates a p-value < 0.001.

5.4. Discussion

In this study, we performed LC-MS/MS analysis to characterize the protein changes in plasma involved in CRC development by unbiased proteomics characterization of CRC patients and healthy individuals. Not only secreted proteins were detected but also released intracellular proteins from damaged tissues and cell turnover. Moreover, we quantified 138 proteins with high reproducibility and a high dynamic range of concentrations from ng/mL to mg/mL. Deep plasma proteomics characterization is challenging because high-abundance proteins, such as albumin and immunoglobulins hinder low-abundance protein identification . Immunodepletion of high-abundance proteins is a common strategy to reduce high-abundance protein levels [443]. However, immunodepletion could lead to the removal of non-targeted proteins associated to albumin and other depleted proteins that have been previously implicated as potential biomarkers [443,444]. In our pilot study, FASP combined with STAGE tips method prioritized the quantification reproducibility over the potential increase in the number of identifications by immunodepletion. Therefore, FASP combined with STAGE tips method was selected to perform proteomics analysis in this study.

Several plasma proteins were identified with significant changes in CRC patients compared to healthy individuals. These findings were consistent with previously published data performed with LC-MS/MS as well as antibody-based techniques including ELISA and Western blot [431–434,445–447]. For instance, ITIH3, the DEP with the highest fold change, was reported as increased in CRC patients' serum and serum of a CRC mice model [432,434,448], while another study showed opposite results [445]. Despite the role of ITIH3 in CRC development has not been determined yet, ITIH4 was found upregulated in CRC tissue versus normal-matched tissue and seems to be involved in the extracellular matrix remodeling and the systemic inflammatory response during CRC development [445]. Moreover, an increased level of several SERPIN family members was observed in the examined CRC cohort, which is consistent with previously reported data [431,446]. Among them, SERPINC1 might play a central role in the systemic response to CRC as it is the most interconnected node in the protein-protein interaction network. Moreover, SERPINC1 downregulation may avoid its suppressive tumor activity and inhibit tumor angiogenesis and proliferation [431]. Interestingly, another family member, SERPINF1 also revealed a link to cancer-associated inflammation. It was reported that this antiangiogenic protein was downregulated in CRC tissue and sera and its low levels were associated with a poor survival prognosis [447].

Importantly, in this study, the increased level of the complement cascade and its components were found in CRC patients. This indicates that these proteins might play a relevant role in CRC development. Enhanced level of the complement proteins such as C9 [433], complement component 4 binding protein alpha and beta (C4BPA and C4BPB) [431,449] was previously reported in CRC patients while increased C1QB is novel. C1QB was found upregulated in tumor tissue versus normal-matched tissue but not in CRC patients' plasma [450]. Another novel complement protein with enhanced plasma level is C4B, which is a non-enzymatic component of C3/C5 convertases and was reported as upregulated in the serum of *Apc*$^{Min/+}$ CRC mice versus wild-type mice [448]. In our study, increased C4B was found in advanced-stage CRC patients, suggesting that this complement protein might play a key role in the disease progression. In addition to C4B, another member of the complement cascade, C8A, was also enhanced in the advanced stages of CRC patients. C8A is a key constituent of the membrane attack complex that regulates the pore formation in target cells and regulates the underlying innate and adaptive immune responses [442]. The high *C8a* expression was previously reported in CRC metastasis compared to the primary tumor which supports its potential role in CRC progression [451]. Moreover, the C8A level was also enhanced in patients with cancer-associated inflammation,

suggesting that this complement protein is linked to the systemic inflammation promoted by CRC to facilitate metastasis from the primary tumor. More importantly, enhanced C5 was found in CRC patients' plasma, which was confirmed in the validation cohort. Increased C5 expression in colon tissue versus normal-matched tissue and its association with metastasis was recently reported in another study [451]. Proteomics analysis also revealed an enhanced level of a peptide corresponding to the C5A anaphylatoxin in examined CRC patients. Although there were no previous reports associating C5A with CRC, another complement anaphylatoxin, C3A, was proposed as a potential CRC diagnostic biomarker [452]. Moreover, several studies suggest that C5A may promote CRC tumorigenesis, metastasis, and immunosuppressive microenvironment within the tumor [452–454]. However, further validation studies are needed to confirm the association between C5A plasma levels and CRC. Another enriched pathway in CRC patients was cholesterol metabolism, with two downregulated apolipoproteins APOA2 and APOA4, that were previously reported [455]. It was found that APOA2 polymorphisms were associated with CRC prognosis and might play a relevant role in disease development and progression [456]. These proteins were also related to metabolic syndrome which is a well-established CRC risk factor [457].

Interestingly, our analysis reported novel plasma protein changes associated with CRC development. For instance, serum amyloid A4 (SAA4), one of the major acute-phase reactants, was enhanced in CRC patients versus healthy individuals. The increased circulating levels of SAA have been linked to several inflammatory conditions including neoplasia [458]. *SAA4* was only detected in CRC tissue but not in normal tissue, suggesting a potential role in tumorigenesis [459]. Another enhanced acute-phase response protein was LBP, which promotes cytokine release in response to bacterial lipopolysaccharide [460]. Noteworthy, our recently published study demonstrated the increased level of several pro-inflammatory cytokines in the same CRC cohort by proximity extension assay [461]. It was previously found that LBP polymorphisms were associated with CRC susceptibility [462] and high serum levels were associated with obesity [463].

Our analysis identified novel links between plasma protein levels in CRC patients and cancer-associated inflammation. The secreted glycoprotein A2GL, also called LRG1, was upregulated in CRC patients with positive inflammatory status and overall CRC patients versus healthy individuals [431]. LRG1 was also overexpressed in CRC tissue where it induced cancer proliferation [464]. Hence, it has been suggested that LRG1 plays an important role in CRC progression and may have an exacerbated pro-inflammatory effect in patients with cancer-associated inflammation due to its link to the acute-phase response [465]. Another enhanced protein in positive-inflammation CRC patients was CERU while higher levels in CRC patients versus healthy individuals were revealed in another study [466]. The metalloprotein CERU binds copper in plasma and is associated with inflammatory responses by promoting nitric oxide synthase activity and cytokine secretion [467]. On the contrary, this study found low levels of the retinol-binding protein (RBP)-4, which is related to cancer-associated inflammation. Downregulation of RBP4 in CRC patients versus healthy individuals in serum and tumor tissue was previously reported [468]. Other adipokines with antitumorigenic effects such as adiponectin (APOD) was also reduced in cancer patients and RBP4 may play a role in the reduction of inflammation [469]. A lower level of APOD, a protein associated with cancer-associated inflammation, was also observed in our cohort. This blood transporter was inversely correlated with CRC tumorigenesis and was associated with early stages of CRC, however, further functional studies are needed to elucidate its role in CRC development [470].

A comparison early-stage and late-stage CRC patients revealed four potential biomarkers associated with cancer progression, including C4B, C8A, APOC2, and IGHG2. The lipoprotein metabolism regulator, APOC2, was found elevated in advanced stages of cancer for the first time, while it was previously described as a potential biomarker of CRC development [432]. On the contrary, IGHG2 plasma levels were increased in CRC early stages and in patients with cancer-associated inflammation. The *IGHG2* expression was previously detected enhanced in cancer tissues of CRC patients but not in plasma [471]. Further analysis in larger cohorts will validate our findings to determine the suitability of these potential biomarkers to predict the cancer stage and the association with inflammation.

By using LC-MS/MS proteomics analysis, we quantified 138 plasma proteins in CRC patients and healthy subjects. However, the high dynamic range of proteins limited the quantification of proteins with low abundance and statistical analysis was only applied to those proteins with reliable quantification. Moreover, due to the relatively low number of patients in the discovery CRC cohort, further validation of the novel potential biomarkers in a larger cohort by targeted MS techniques or other quantitative methods such as antibody-based strategies is required. To evaluate the application of these biomarkers in early CRC detection, further validation with a cohort including a higher percentage of early-stage patients is in the plan. Further studies using more advanced LC-MS/MS instrumentation that combines nanoparticle protein coronas with high-resolution Orbitrap mass spectrometer DIA analysis (DIA) mode followed by MRM analysis will improve the low-abundance protein detection [472]. The discovery cohort was also limited by the higher percentage of women, while CRC incidence is higher in men. Finally, CRC family history information and molecular expression profiles of the tumor were missing, which are relevant factors in CRC development and progression.

In this study, LC-MS/MS plasma proteomics application in CRC patients identified novel protein signatures compared to healthy subjects including complement proteins as well as proteins such as SAA4 and LBP associated with pro-inflammatory conditions. Importantly, we confirmed the enhanced levels of C5 in patients of a validation cohort as a potential diagnostic biomarker of CRC. Moreover, several proteins linked to cancer-associated inflammation, such as LRG1, CERU, RBP4, and APOD were identified. Plasma level of several proteins, including C4B, C8A, APOC2, and IGHG2 that may be served as potential early detection biomarkers in clinics to improve patient care after further validation in larger cohorts.

CHAPTER 6. Plasma protein changes reflect colorectal cancer development and associated inflammation

In this chapter, similarly to COVID-19 studies, PEA was applied to plasma samples from CRC patients and healthy controls from the same biobanks as in the previous chapter to characterize the protein changes caused by CRC development, progression, and cancer-associated inflammation. This study resulted in a published article in **Frontiers of Oncology** and is presented with minor modifications:

***Urbiola-Salvador V*, Jabłońska A, Miroszewska D, Huang Q, Duzowska K, Drężek-Chyła K, Zdrenka M, Śrutek E, Szylberg Ł, Jankowski M, Bała D, Zegarski W, Nowikiewicz T, Makarewicz W, Adamczyk A, Ambicka A, Przewoźnik M, Harazin-Lechowicz A, Ryś J, Filipowicz N, Piotrowski A, Dumanski JP, Li B, Chen Z. Plasma protein changes reflect colorectal cancer development and associated inflammation. *Front. Oncol.* 13, 1158261 (2023).**

6.1. Introduction

Colorectal cancer (CRC) is the third most common malignancy and the second most lethal cancer, causing 935,000 cancer-related deaths in 2020 [78]. CRC prognosis depends mainly on the tumor stage, location, and time of detection. However, despite the huge progress in cancer research, a large number of CRC cases are diagnosed at the advanced stage where cancers are aggressive, malignant, and metastatic [422].

Currently, the most commonly-used diagnostic tools for CRC screening and prevention include colonoscopy and flexible sigmoidoscopy, as well as the guaiac-based fecal occult blood test (gFOBT) or the immunochemical fecal occult blood test, also known as the fecal immune test (FIT) [426]. The traditional stool-based tests, such as gFOBT and FIT, have low sensitivity and specificity [473], while colonoscopy and sigmoidoscopy, despite the high sensitivity, have relatively low compliance, high cost, and are invasive which limits their efficacy in a population screening programs [427]. Therefore, alternative, non-invasive, and efficient screening strategies to improve the early detection of cancers are urgently needed. Until now, several potential blood-based protein biomarkers for CRC screening and cancer prevention have been reported, including methylated Septin9 [474], extracellular vesicle microRNAs [475], and cell-free circulating DNA [476], but all lack the sensitivity and/or specificity for use as a stand-alone marker.

Advances in proteomic-based technologies in the last decade have expanded the number of candidate biomarkers and led to a better comprehension of the CRC progression as well as the identification and characterization of related molecular signatures. The most recent advancement of Proximity Extension Assay (PEA) allows the quantification of over 3,000 proteins from low amounts of a sample by the combination of DNA-conjugated antibodies and next-generation sequencing (NGS) [174]. Application of the PEA technology has led to the identification of carcinoembryonic antigen (CEA) as one of the best-studied blood-based prognostic biomarkers used in clinical practice [477–479]. CEA is expressed in the embryonic endodermal epithelium, colorectal cancer, and other malignancies, such as inflammatory bowel disease (IBD), peptic ulcer, and pancreatitis [480]. CEA is a promising plasma biomarker for the detection of CRC with high specificity and sensitivity [479,481], however, due to the limited organ specificity [482], it is not the best sole biomarker for population-based screening, yet it might be useful in CRC recurrence monitoring [483] and metastasis [484]. Currently, the trend in biomarkers discovery is to focus on the biomarker panels rather than on a single-target protein as the broader spectrum of the analysis may help to address the cancer prognosis and detection more precisely.

It was recently reported that two various multimarker panels consisting of five circulating proteins might be used as an efficient tool for the early and late-stage detection of CRC, including advanced adenomas, or in the prediction of overall survival in Germany and Chinese cohorts [48,485]. In a recent study, Harlid et al. (2021) showed that fibroblast growth factor (FGF)-21 was associated with early, but not late stages of colon cancer, while pancreatic prohormone (PPY) was a promising biomarker for rectal cancer detection [486]. However, neither FGF21 nor PPY could be used as stand-alone biomarkers for colon or rectal cancer but might be used as an efficient tool to discriminate between different subtypes of CRC. Therefore, there is an urgent need for the identification of a reliable blood-based biomarker panel that would detect the early stages of CRC as well as assesses the prognosis at the population-based screening.

Both chronic inflammation, such as IBD, and sporadic, cancer-associated inflammation are well-known as key factors in CRC progression and development. Inflammation alters the communication between a variety of cell types, including innate and adaptive immune cells, epithelial cells, and stem cells. These intricate networks of cytokines, growth factors, receptors, and other molecules interaction result in either tumor-promoting or inhibiting environment [425]. Thus, in the development of plasma biomarkers for CRC diagnosis, prognosis, and immunotherapy, the inflammatory status is essential.

The purpose of our study was to identify the protein expression changes in the plasma of CRC patients compared to healthy controls as well as between the early and late stages of CRC and inflammatory status. Therefore, an inflammation panel including 368 proteins was selected to be detected in this study. We hypothesized that CRC development, tumor stage, and inflammation-caused changes in protein level will be reflected in the circulating blood and as such, we would be able to obtain a panel of biomarkers with potential translation into clinics to improve patient care. In this study, we quantified the plasma protein profiles derived from 38 CRC patients and their age- and sex-matched 38 healthy subjects using the PEA technology and protein panels consisting of 368 oncology- and 368 inflammation-related protein biomarker candidates. We quantified 690 proteins, among which 78 differentially expressed proteins (DEPs), were elevated and 124 DEPs were reduced in patients with CRC. We found protein signatures associated with cytokine interactions, oncogenic signaling

pathways, exacerbated apoptosis, as well as metabolism reprogramming. Additionally, we determined protein changes linked to cancer-associated inflammation and novel potential prognostic biomarkers associated with tumor stages. Linear regression model analysis revealed that carbonic anhydrase (CA11), CD276, colony-stimulating factor 3 (CSF3), and interleukin 12 receptor subunit beta 1 (IL12RB1), were positively associated with inflammatory status, whilst amyloid beta precursor protein binding family B member 1 interacting protein (APBB1IP) and CXCL6 were negatively associated. Moreover, linear regression model analysis of tumor stage indicated high plasma levels of Fms-related tyrosine kinase 4 (FLT4), MANSC domain-containing protein 1 (MANSC1), and lysophosphatidic acid (LPA) phosphatase type 6 (ACP6), that could be used as potential prognostic biomarkers for advanced CRC. In contrast, high levels of interferon γ (IFNG), interleukin (IL)32, and IL17C in early CRC stages indicate that these proteins can discriminate between early and late stages patients. Additionally, IFNG is proposed as a potential biomarker for the early detection of CRC.

6.2. Materials and Methods

6.2.1. Study cohort

The study was retrospective and consisted of 38 patients who underwent CRC surgery (mean age: 66.7 ± 12.3; 42.1% male) between June 2019 and April 2021 and 38 age- and sex-matched healthy subjects. All CRC patients had a positive colonoscopy and pathology-confirmed malignant neoplasm of the rectum or colon. Among them, 63.2% (24/38) were diagnosed with late-stage CRC (III-IV) according to the Union for International Control (UICC) TNM classification and 28.9% (11/38) had inflammation according to the pathologist assessment. Samples collected from CRC patients and healthy subjects were obtained from the 3P–Medicine Laboratory, Medical University of Gdansk [487] and Biobank HARC, Medical University of Lodz, respectively. Whole blood samples were collected into sterile BD Vacutainer® K2EDTA tubes during the day of the planned CRC resection, centrifuged, aliquoted plasma and serum, and stored at -80°C until use.

6.2.2. Protein profiling

Plasma proteins were analyzed using the multiplex PEA technology (Olink® Explore 384-Oncology and -Immunology panel, Olink Proteomics, Uppsala, Sweden). Briefly, the PEA technology is a dual recognition approach based on matched pairs of oligonucleotide-labeled antibodies that bind to their target proteins. Once the target proteins are bound, the oligonucleotides brought into proximity, hybridize and are detected and quantified by using NGS [174]. PEA quantifies a large number of proteins (> 3,000) with good precision, using a minimal volume of plasma or serum samples, and without loss of specificity and sensitivity. The protein levels are presented in the normalized protein expression (NPX) values on a log2 scale. A high protein concentration corresponds to a high NPX value. For quality assessment and validation of the PEA technology, the protein level of ACP6 was measured by ELISA, while for CSF3, IFNG, IL6, CXCL9, and CCL23 were determined by using Luminex MAGPIX technology.

6.2.3. Statistical analyses

All statistical analyses were performed in RStudio (version 1.3.1093) using R (version 4.0.3). First, proteins were filtered when the quality control was negative or the calculated NPX values were below the respective protein limit of detection (LOD) in at least 50% of samples from one of the study groups. The remaining NPX values below the LOD were imputed with the respective LOD/$\sqrt{2}$. Moderated t-test from the R package "limma" (version 3.46.0) was used to test differential protein abundance between CRC patients and healthy subjects. Additional analysis was performed using the general linear model regression approach with analysis of contrasts using the R package "emmeans" (version 1.6.2.1). A general linear model was fitted to the expression of each protein in all CRC patients using tumor stage and inflammation as independent variables, and sex as a confounding factor. The false discovery rate (FDR) was determined using the Benjamini & Hochberg correction. Proteins were considered differentially expressed when FDR adjusted p-value < 0.05. The built-in R function cor.test was used to calculate the point-biserial correlation between protein expression and tumor stage or inflammation status, the p-value < 0.05 was considered significant. gene set enrichment analysis (GSEA) with Gene Ontology (GO) terms was performed using ClusterProfiler (version 4.6.0), while KEGG pathway enrichment analysis *via* active subnetworks from STRING database was conducted using "pathfindR" (version 1.6.3), with FDR < 0.05. "ggplot2" (version 3.3.5) was used for graphics generation, excluding heatmaps that were generated using "ComplexHeatmap" (version 2.6.2). The hierarchical clustering (Euclidean distance) was implemented to visualize the patterns of DEPs among samples after the z-score transformation of NPX values; DEPs were split by k-means clustering.

6.3. Results

6.3.1. CRC development causes cytokine and oncogenic signaling pathway changes in plasma

To determine the changes in the protein profiles in peripheral blood caused by CRC development, we performed plasma protein analysis by using PEA technology. Out of the total 736 proteins from the Inflammation and Oncology Explore panels, after removing repetitions in the panels and after removal of proteins with low detection rates among the samples, 690 proteins were quantified. Among them, 78 proteins were elevated and 124 were reduced in the 38 CRC patients compared with their age- and sex-matched healthy controls (Figure 6.1a, Appendix IV Figure S1a and Table S1). Of the elevated DEPs,

dipeptidase 2 (DPEP2), hydroxyacylglutathione hydrolase (HAGH), and agouti-related neuropeptidase (AGRP) as well as downregulated DEPs as neutrophil cytosolic factor 2 (NCF2), epidermal growth factor-like protein 7 (EGFL7), and ectonucleotide pyrophosphatase/phosphodiesterase family member 5 (ENPP5) were the DEPs with the most statistical difference. In line with previous studies which were carried out with different technologies for protein detection and quantification [484,488–493],high plasma levels of AGRP, FGF21, midkine (MDK), C-C motif chemokine ligand (CCL)-20, IL6, and CSF3 as well as reduced ribonucleotide reductase regulatory TP53 inducible subunit M2B (RRM2B) on plasma level of CRC patients were also identified in our study. Importantly, we found novel protein changes including high levels of oncogenic proteins such as R-Spondin 3 (RSPO3) and secernin 1 (SCRN1) as well as low levels of tumor suppressors such as Ret proto-oncogen (RET) and Rho guanine nucleotide exchange factor 12 (ARHGEF12) in CRC patients. These results suggest the association between plasma protein levels and protein expression within the tumor microenvironment (TME).

To investigate the involved pathways and the complex protein-protein interactions among these DEPs, KEGG enrichment analysis *via* active subnetworks was performed. Plasma protein changes were mainly associated with the cytokine-cytokine receptor interaction, including high plasma level of T-cell chemoattracting chemokine CXCL9 and the immune cell chemoattractant CCL23, as well as several signaling pathways including mitogen-activated protein kinase (MAPK), Ras, tumor necrosis factor (TNF), nuclear factor-κB (NF-κB), and IL17 signaling pathways (Figure 6.1b-c, Appendix IV Table S2). Notably, proteins involved in Th17 cell differentiation were upregulated in CRC patients, suggesting an active role of this T-cell helper subtype in CRC development. Moreover, proteins related to non-fatty liver disease (NAFLD), a disease previously associated with CRC risk [494], were enriched (Figure 6.1b, Appendix IV Table S1). At the same time, high levels of apoptosis-associated proteins, CASP8 and BH3 interacting domain death agonist (BID) were discovered, with BID having the second highest fold change in the comparison (Figure 6.1a-c). To reveal possible mechanisms of cancer development, DEPs were further evaluated by using GSEA. This analysis revealed that the GO terms including oxidative phosphorylation, aerobic respiration, respiratory electron transport chain, and ATP synthesis coupled electron transport in mitochondria were enriched in CRC patients (Appendix IV Figure S1b and Table S3). Moreover, other metabolic proteins were highly elevated in CRC patients including HAGH and DPEP2 (Figure 6.1c) which may reflect the metabolism reprogramming due to CRC tumorigenesis, a well-known hallmark of cancer [495].

Next, to distinguish which of the protein changes were a consequence of an altered secretion from a certain type of cells and which were a result of destructed tissues or cells released during CRC tumorigenesis, from the 202 DEPs, 50 proteins were identified in the human blood secretome from Human Protein Atlas, including cytokines that modulate the immune responses within the TME, such as IFNG, IL6, IL15, CCL20, CXCL9, and CCL23 (Appendix IV Table S4). Some of these cytokines were previously found with high plasma levels in CRC such as pro-inflammatory cytokine IL6 which is also required for Th17 differentiation [491], the pro-inflammatory MDK involved in multiple biological processes [484], and the chemoattractant of B- and T-cells CCL20 [490], whereas the detected IFNG is a well-recognized pro-inflammatory and antitumorigenic protein [496]. Interestingly, the elevated plasma levels of the chemoattractants CXCL9 and CCL23 in CRC patients were reported for the first time in our study. These results suggest that plasma protein changes can reflect the variety of altered processes involved in tumorigenesis. Collectively, CRC development causes protein changes in plasma that are linked to several signaling pathways, cytokine interactions of underlying immune responses, and altered metabolism.

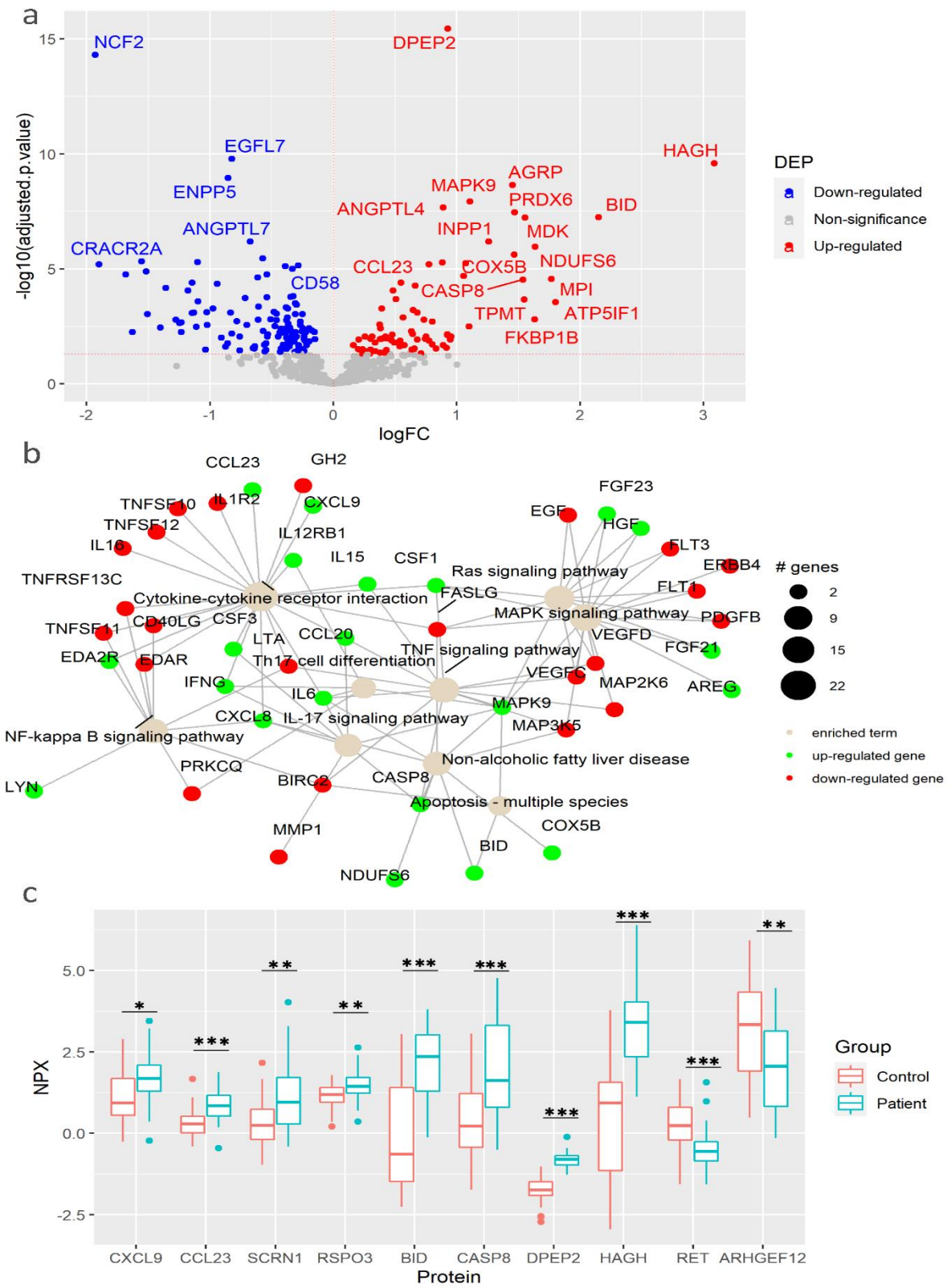

Figure 6.1. Colorectal cancer (CRC) development causes cytokine and oncogenic signaling pathway changes in plasma. (a) Volcano plot of statistical significance against fold-change of proteins between CRC patients and healthy controls. Dots indicate individual proteins and the red and blue dots represent significant up-regulation and down-regulation in patients, respectively. (b) Network of KEGG pathway enrichment analysis combined with STRING protein-protein interaction network analysis. Green and red proteins indicate significant up-regulation and down-regulation, respectively. (c) Box and whisker plots of selected DEPs not previously reported associated with CRC. *

indicates statistically significant with an adjusted p-value < 0.05, ** indicates an adjusted p-value < 0.01, and *** indicates an adjusted p-value < 0.001,DEP, differentially expressed protein; FC, fold change, NPX; normalized protein expression.

6.3.2. Cancer-associated inflammation alters the plasma protein expression in CRC patients

It is well-known that chronic inflammation may contribute to cancer development. To determine plasma protein changes related to inflammatory status in CRC patients, we analyzed DEPs among patients with and without inflammation (11 and 27 cases, respectively). Correlation analysis revealed 56 proteins significantly correlated with inflammation, among which 7 proteins, CA11, CD276, CSF3, IL3RA, IL12RB1, MILR1, and SEMA4C were positively correlated, while 46 proteins including ACP6, APBB1IP, CXCL6, and DCXR were correlated negatively (Figure 6.2a, Appendix IV Table S5). Among them, elevated IL12RB1 and reduced DCXR showed the highest correlation with inflammatory status (Figure 6.2a, Appendix IV Table S5). To confirm the association between protein expression and inflammation, a linear regression analysis was used to determine the differential expression of these proteins. As a result, 26 DEPs were identified which were significantly correlated in the previous analysis (Figure 6.2b, Appendix IV Table S6).

KEGG pathway enrichment analysis demonstrated that the DEPs were mainly assigned to cytokine-cytokine interaction, IL17 and Th17 cell differentiation, and Janus kinase (JAK)-signal transducer and activator of transcription (STAT) signaling pathways, as well as pentose and glucuronate conversion (Appendix IV Table S7). It has been well documented that Th17 cells play an essential role in inflammation *via* the production of pro-inflammatory cytokines IL17A, IL17F, IL22, and IL21. Th17 cell activity is also associated with an increased risk of CRC tumorigenesis [497]. Among the DEPs involved in the IL17, Th17 cell differentiation and JAK-STAT signaling pathways, elevated levels of CSF3 and reduced CXCL6 were previously found in the serum of CRC patients [492,498], while our study also demonstrates their association with cancer-associated inflammation (Figure 6.2a-c). Interestingly, CSF3 is involved in inflammation by inducing bone-marrow neutrophil differentiation and its high levels are related to CRC tumorigenesis [492]. Moreover, we report, for the first time, the association of IL12RB1, CA11, CD276, and APBB1IP with cancer-associated inflammation (Figure 6.2c). Accordingly, IL12RB1 and CSF3 were detected with high plasma levels in the whole CRC patients compared with healthy controls (Figure 3.1a, Appendix IV Table S1). It is worth noting that IL12RB1, CD276, and APBB1IP are involved in cancer surveillance, inhibition of T-cell mediated responses, and T-cell recruitment, respectively [499–501], whereas CA11 may induce proliferation and invasion of gastrointestinal tumors [502]. (Figure 6.2c). In summary, these results suggest that inflammation in CRC patients can influence plasmatic protein levels. Furthermore, these proteins may be useful indicators of cancer-associated inflammation that may complicate the outcome of CRC patients.

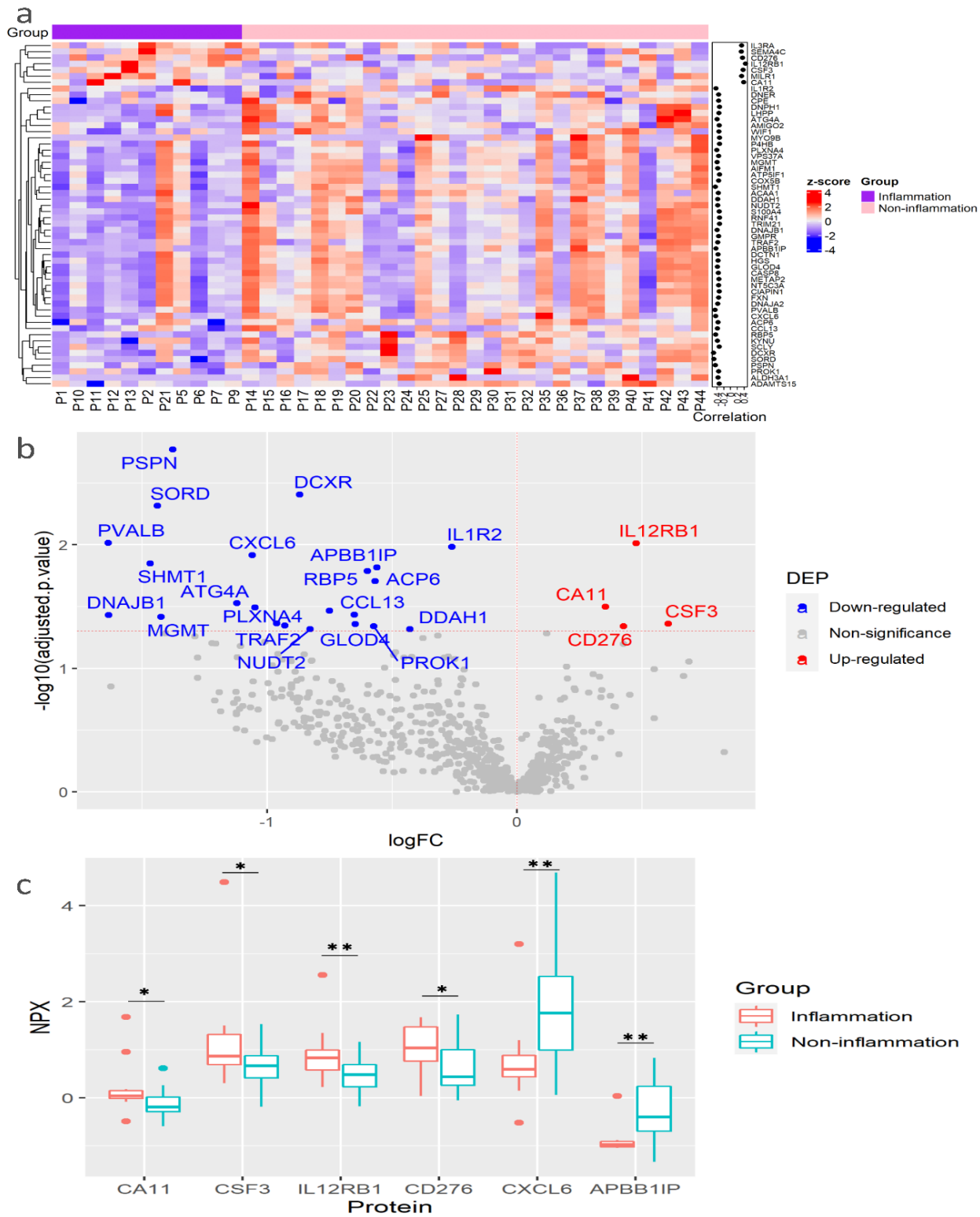

Figure 6.2. Plasma protein changes induced by cancer-associated inflammation in CRC patients. (a) Heatmap of proteins with significant correlation with inflammation status. Protein expression is transformed with a z-score by row normalization and distributed by hierarchical clustering. The correlation coefficients (right) indicate a positive/negative correlation for each protein. (b) Volcano plot of statistical significance against fold-change of proteins between CRC patients with inflammation and without inflammation. Dots indicate individual proteins and the red and blue dots represent significant up-regulation and down-regulation in CRC patients with inflammation, respectively. (c) Box and whisker plots of selected DEPs not previously associated with cancer-related inflammation in CRC patients. *: adjusted p-value < 0.05, **: adjusted p-value < 0.01.

6.3.3. Determination of potential plasma biomarkers associated with CRC stages

The main cause of a patient's death due to CRC is tumor growth and its increased invasiveness, resulting in metastasis. Therefore, it is crucial to find prognostic biomarkers for CRC progression. We determined the plasma protein changes associated with CRC advance by the comparison of patients with early (I and II) and late (III and IV) stages of CRC. The correlation analysis showed that 13 proteins, ACP6, CCL23, C-type lectin domain family 4 member G (CLEC4G), FLT4, IL1R2, IL6, MANSC1, marginal zone B and B1 cell-specific protein (MZB1), S100A12, SCGB1A1, SMOC2, TXNDC15, and WFIKKN2 were positively correlated with tumor stage, whereas 7 proteins, including IFNG, IL32, integrin subunit alpha 11 (ITGA11), ITGAV, selectin P ligand (SELPLG), trefoil factor (TFF)-2, and TMPRSS15 were correlated negatively (Figure 6.3a). Among them, FLT4 showed the best prognostic performance for late-stage CRC with the highest correlation coefficient (Figure 6.3a, Appendix IV Table S8). The elevated plasma FLT4, also named Vascular Endothelial Growth Factor Receptor 3 (VEGFR3) in the late stage of CRC may be associated with VEGF-mediated lymphangiogenesis and angiogenesis.

Similarly to the analysis with inflammatory status, the regression analysis resulted in fewer DEPs than correlated proteins. This analysis revealed that ACP6, FLT4, and MANSC1 were elevated in the late stages of CRC, while IL17C, IL32, and IFNG were elevated in the early stages (Figure 6.3b-c, Appendix IV Table S9). Notably, the enzyme ACP6 which is involved in phospholipid metabolism by hydrolysis of LPA was negatively associated with inflammatory status, suggesting that ACP6 may play a role in both inflammation and CRC progression (Figure 6.3d). Taken together, these results indicate that ACP6, FLT4, and MANSC1 might be potential prognostic markers for advanced CRC. Notably, MANSC1 and ACP6 have not been previously reported to be associated with CRC development. Moreover, the pro-inflammatory cytokines IL17C, IL32, and IFNG were elevated in the early stages of CRC tumorigenesis. Importantly, IFNG levels were elevated in CRC patients compared to healthy subjects as well as in the early stages (Figure 6.3d). Therefore, IFNG can be a novel potential biomarker for the early detection of CRC that may indicate an enhanced anti-tumor activity in the early stages.

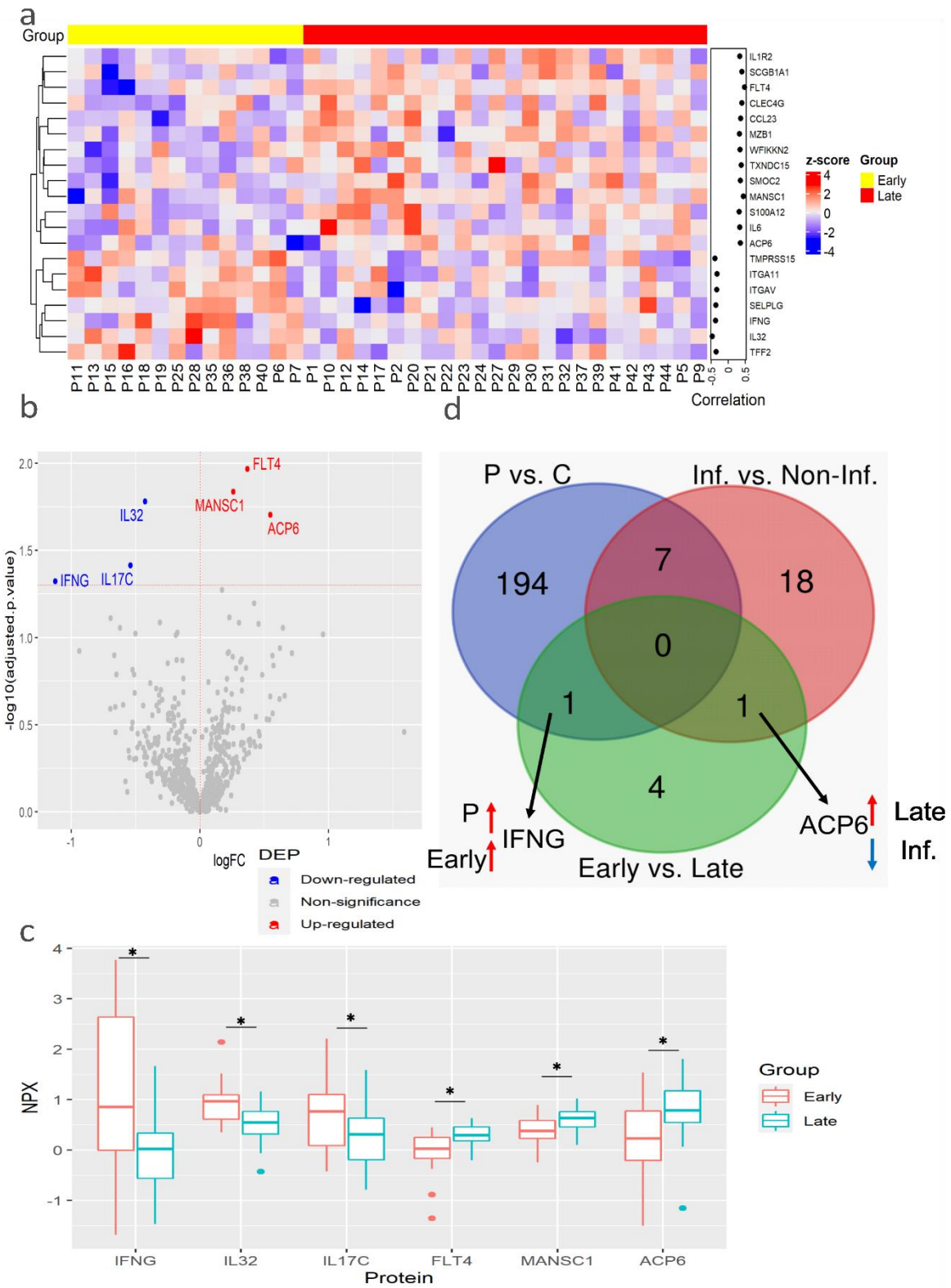

Figure 6.3. Plasma protein expression differences between early and late stages of CRC. (a) Heatmap of proteins with significant correlation with tumor stage. Protein expression is transformed with a z-score by row normalization and distributed by hierarchical clustering. The correlation coefficients (right) indicate a positive/negative correlation for each protein. (b) Volcano plot of statistical significance against fold-change of proteins between CRC patients with early tumor stage and with late tumor stage. Dots indicate individual proteins and the red and blue dots represent significant up-regulation and down-regulation in CRC patients with late tumor stage, respectively. (c) Box and

whisker plots of DEPs that are novel potential prognostic biomarkers associated with cancer stages in CRC patients. *: adjusted p-value < 0.05. D Venn diagram with the differentially expressed proteins for each comparison: CRC patients vs. control, Inflammation vs. Non-inflammation, and Early vs. Late. Black arrows indicate the proteins of interest that are in common between comparisons. Red and blue arrows indicate up-regulation and down-regulation for the specified group, respectively. C, control; Inf., inflammation; Non-Inf., non-inflammation; P, patient.

6.3.4. Validation of identified plasma protein changes with a different cohort

To validate some of the newly identified plasma protein changes in CRC patients, an independent cohort including 41 patients who underwent CRC surgery obtained from the Bank of Biological Material at Masaryk Memorial Cancer Institute, Czech Republic was used. Higher concentrations of IL6 and CSF3 among CRC patients than in healthy volunteers were confirmed in the validation stage of the study (Figure 6.4a). Importantly, increased secretion of IFNG, CXCL9 and CCL23 in the plasma of CRC patients compared to healthy subjects was detected in the validation cohort by Luminex (Figure 6.4a), suggesting that elevated plasma level of IFNG, CXCL9 and CCL23 might be served as a biomarker of CRC. Importantly, similar as detected by PEA (Figure 6.3b), the elevated level of ACP6 in late stage compared with early stage of CRC was confirmed in this cohort as well (Figure 6.4b). Taken together, these results indicate that ACP6 might be a potential prognostic marker for advanced CRC. Notably, MANSC1 and ACP6 have not been previously reported to be associated with CRC development. However, these findings need to be confirmed by using bigger validation cohort.

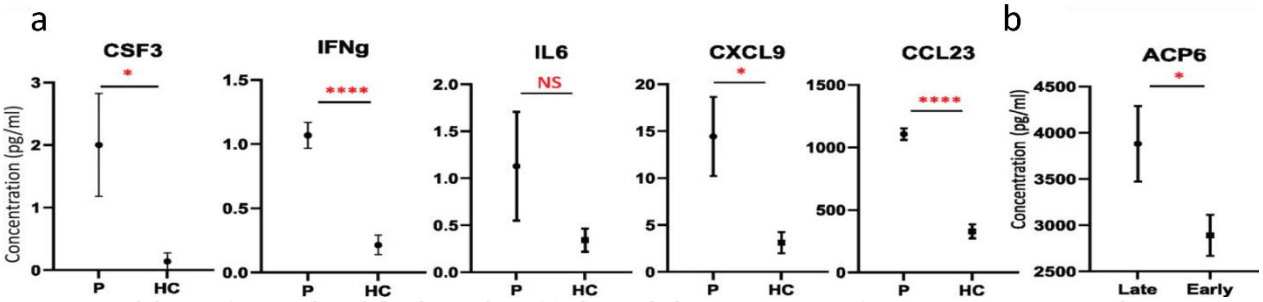

Figure 6.4. Validation of potential candidate biomarkers. (a) Plots with the concentrations of CSF3, IFNG, IL6, CXCL9, and CCL23 in CRC patients (P) and healthy controls (HC) (mean ± SEM) detected by Luminex. (b) Plot with the concentrations of ACP6 detected by ELISA in CRC patients with early and late stages, respectively (mean ± SEM). T test was used for statistical analysis. *: p-value < 0.05, ****: p-value < 0.0001, NS: non-significance.

6.4. Discussion

Cancer, including colorectal cancer, is the leading cause of death worldwide and the most devastating disease as the 21st century begins. Thus, there is an urgent need for the discovery and validation of reliable and efficient non-invasive biomarkers for early CRC detection and prognosis prediction, including biomarkers to detect cancer-associated inflammation. To determine plasma protein changes in CRC patients, by using PEA technology, we quantified 690 proteins, among which 202 were changed compared to healthy subjects.

Among the elevated cytokines in CRC patients, CXCL9 and CCL23 have been identified as novel potential biomarkers. The T-cell chemoattractant *CXCL9* was previously found elevated in CRC tissues compared to normal colon tissues and it was associated with tumor differentiation and invasion, lymph node and distant metastasis, as well as with vascular invasion [503]. An enhanced expression of *CXCL9* in cancer tissue than healthy tissue was also observed in the second Chinese study, where *CXCL9* expression levels were associated with tumor stage and survival [504]. Importantly, CXCL9 may also recruit T-cells to the TME and exerts antitumor activity[505]. The chemokine, CCL23 has been found as a cytokine with both, pro- and anti-cancer properties. It can induce angiogenesis by activating C-C Motif Chemokine Receptor 1 (CCR1) on vascular endothelial cells and increase the proliferation of cancer cells, but also, it can promote immune infiltration [506]. However, what type of immune cells and T-cells are attracted to the TME by CCL23 and CXCL9, respectively, requires further studies. A strong elevation of CCL23 protein was noticed in rectal cancer compared to non-rectal cancer consisting of ascending, transverse, and sigmoid colon [507], while *CCL23* expression was not detected in colon adenocarcinoma cells in a second study [508]. Interestingly, none of the previous studies reported high CXCL9 and CCL23 levels in the plasma of CRC patients.

Apart from cytokines, plasma levels of other immune-related proteins were changed in CRC patients compared to the healthy controls, such as DPEP2 and Peroxiredoxin 6 (PRDX6), which have not been previously reported as plasma diagnostic biomarkers. The protein expression of DPEP2, a dipeptidase involved in leukotriene metabolism, was recently found as a modulator of macrophage inflammatory responses, protecting mice against Coxsackievirus B3-induced viral myocarditis [509]. Interestingly, DPEP1, the paralog of DPEP2 was up-regulated in CRC tissue at mRNA and protein levels and high DPEP1 expression was significantly correlated with cancer stage, location,and poorer prognosis [510], while no association of DPEP2 with CRC has been detected. Similarly, elevated PRDX6, a metabolic enzyme, may modulate inflammation and immune responses through the regulation of antioxidants and reactive oxygen species (ROS) [511]. It was suggested that PRDX6 may promote CRC invasiveness and aggressiveness by inducing an oxidizing TME [512]. Importantly, we found two mediators of apoptosis, CASP8 and BID, which presented high plasma levels in CRC patients compared to healthy subjects, with BID having the second highest fold change. Recently, circulating CASP8 was identified with high expression in

pre-operative serum samples of prostate cancer [513]. BID, belonging to the B-cell lymphoma 2 (BCL-2) family, is a key regulator of apoptosis and a factor associated with CRC initiation and progression [514]. It was found that high expression of proapoptotic BID was a predictor of overall survival in patients with CRC, whereas combined expression of BAD and BID was associated with disease-free survival rates and overall survival [515]. However, further studies are needed to investigate whether the elevated plasma CASP8 and BID are associated with an exacerbated apoptosis of peripheral blood mononuclear cells (PBMC) among these patients, similarly as in the case of melanoma patients [516]. Collectively, the altered cytokines and immune-related proteins suggest an active modulation of the immune system in CRC patients at the systemic level as well as a systemic inflammatory status.

It is well-known that several signaling pathways, such as Ras, NF-κB, and MAPK are altered in CRC patients leading to oncogenesis [517], which was also confirmed in our study at a systemic level. Interestingly, several oncogenic proteins were elevated in plasma, such as SCRN1 and RSPO3, whereas previous studies determine their overexpression in CRC tumor tissue [518,519]. SCRN1 accelerates tumor progression by the regulation of exocytosis of matrix metalloproteinase (MMP)-2/9 [520], while RSPO3 is an oncogenic driver that causes CRC and extensive crypt hyperplasia, concomitantly stimulating stem cells and supportive niche cells [521]. It was found that overexpression of *RSPO2* and *RSPO3* was presented by 4-10% of colon subjects [519] and recurrent R-spondin fusions in colon cancer activate the Wnt signaling and increase the tumorigenesis [522]. Additionally, lower plasma levels of potential tumor suppressor proteins, such as RET and ARHGEF12 were detected in CRC patients. RET, is a transmembrane receptor tyrosine kinase and a receptor for the GDNF-family ligands, which downregulation in CRC tissue compared to healthy tissue was noticed [523]. CRC patients with somatic RET mutations exhibited a lower incidence of liver metastasis but a higher incidence of peritoneal metastasis and more frequently exhibited mucinous histology [524]. On the other hand, a germ-line or somatic RET mutation was linked with more intense and complete angiogenesis in patients with advanced medullary thyroid cancers [525]. ARHGEF12, also known as leukemia-associated Rho guanine-nucleotide exchange factor (LARG), is underexpressed in CRC tissue and is associated with reduced cell proliferation and a slower migration rate in cancer cells [526]. Moreover, it was found that ARHGEF12 regulates cell adhesion and structure morphogenesis in esophageal squamous cell carcinoma tissues [527] and plays a key role in erythroid regeneration after chemotherapy in acute lymphoblastic leukemia patients [528]. These proteins can be potentially used as an oncogenic protein signature for CRC diagnosis in plasma. Apart from oncogenic pathways, NAFLD was also enriched in this cohort. Meta-analyses revealed that NAFLD was associated with an increased risk of gastrointestinal cancers [529] and colon cancers, especially in the right-sided colon[494].

More importantly, our data demonstrated the upregulation of Th17 cell differentiation in CRC patients. Th17 activity has been linked to CRC tumorigenesis and poor prognosis [530]. It is well-known that chronic inflammation contributes to cancer development. We identified upregulation of IL12RB1 and CSF3 in Th17 differentiation and IL17 signaling, indicating their participation in CRC-related inflammation. It is worth noticing that CSF3 expression was previously found elevated in the serum of CRC patients [492]. An increased gene expression of CSF3 was also observed in CRC tissue from two Consensus Molecular Subtypes (CMS) (microsatellite instable immune and mesenchymal), where it was associated with regulators (e.g., CXCL5) of invasion [531]. IL12RB1, a subunit of the interleukin 12 receptors is associated with tyrosine kinase 2 (TYK2), which plays a pivotal role in immunity to viral infection and cancer surveillance [499]. It was found that elevated expression of tumor tissue IL12RB1 was associated with lung cancer progression [532], whereas its correlation with CRC development has not been reported. Moreover, IL12RB1 contributes to both the IL12- and IL23-signaling pathways and is involved in both Th1 and Th17 cell differentiation [533]. A carbonic anhydrase, CA11, was also associated with inflammation which overexpression promotes the proliferation and invasion of gastrointestinal tumors without any previous association with CRC in plasma [502]. Importantly, the immune checkpoint inhibitor CD276, also called B7-H3 was also linked to inflammation. CD276 was previously reported with high expression in CRC tissue and may contribute to the tumor evasion of T-cell mediated responses [500,534]and has been already proposed as a target for immunotherapy [535]. The overexpression of this immune checkpoint molecule in our study further indicates the importance of this protein in the personalized medicine and immune-checkpoint therapy aspect.

In contrast, the reduced plasma level of CXCL6 and APBB1IP in CRC patients with inflammation was observed in our study. It was recently found that low serum CXCL6 levels were associated with an increased risk of CRC development [498], while CXCL6 expression is not altered in CRC tissue [536]. The APBB1IP is a Rap1-binding protein that acts as a regulator of leukocyte recruitment and pathogen clearance through complement-mediated phagocytosis[501]. It was shown that expression of APBB1IP was correlated with the prognosis of various cancer types and its upregulation has been demonstrated as associated with increased immune cell infiltration, especially $CD8^+$ T cells, natural killer (NK) cells, and immune regulators [501]. Bioinformatics analyses revealed that *APBB1IP* may be used as a potential biomarker for osteosarcoma metastasis [537] and suggested its potential role in the evolutionary mechanisms of head and neck squamous cell carcinoma related to inflammation and TME [538]. Moreover, cancer-related inflammation may cause the downregulation of APBB1IP decreasing the recruitment of leukocytes to the TME. In this study, for the first time, we reported the association of reduced plasma APBB1IP level with CRC and inflammation, suggesting that APBB1IP could be a potential biomarker for inflammation-associated CRC.

The next two elevated plasma proteins, MANSC1 and ACP6, identified in our study have never been suggested as associated with CRC risk. Expression of bone marrow *MANSC1*was detected in patients with different hematologic malignancies such as acute myeloid leukemia, myelodysplastic syndromes, and primary myelofibrosis, but no significant correlations between the expression of the gene and survival were observed [539]. In contrast, an association between high

expression of *MANSC1* and a positive prognosis for overall survival was found in patients with non-small cell lung cancer [540]. *MANSC1* Single Nucleotide Polymorphism (SNP) has been also identified as a functional SNP in patients with overall prostate cancer and non-advanced prostate cancer in the genome-wide association study (GWAS) study [541]. The metabolic enzyme ACP6 hydrolyzes LPA to monoacylglycerol and plays a role in regulating lipid metabolism in the mitochondria [542,543]. It has been recently demonstrated that overexpression of ACP6 in hepatocellular carcinoma tissue was positively correlated with clinical progression and worse overall survival of examined patients [542]. On the other hand, decreased expression of ACP6 was found to contribute to increased cell mortality and disease progression in high-grade serous ovarian cancer and esophageal squamous cell carcinoma[543,544]. It was found that CRC cells have abnormal LPA receptor expression that may be associated with enhanced proliferation, survival, and invasion of CRC cells [545]. These results suggest that ACP6 may play a key role in oncogenesis. A positive correlation of plasma ACP6 with the advanced stage of CRC has been revealed for the first time in our study. Moreover, ACP6 was reduced in CRC patients with cancer-related inflammation. The function of ACP6 in cancer-related inflammation and CRC tumorigenesis needs to be further investigated.

More interestingly, three pro-inflammatory cytokines, IL32, IL17C, and IFNG, were increased in the early stages of CRC compared to late-stage patients. IL32 is an intracellular pluripotent cytokine, expressed in various cell types, which affects many cellular and physiological functions such as cell death and survival, angiogenesis, inflammation, and response to pathogens [546]. Increased levels of IL32 were found in cancer tissue [547,548], and primary CRC lymph nodes metastasis [549]. Moreover, IL32 can stimulate NK and T-cell cytotoxicity against primary solid tumors, as well as increase T-cell infiltration [550]. In our study, we observed increased circulating IL32 associated with the early tumor stage, indicating that IL32 may serve as a biomarker for the early stage of CRC. The second pro-inflammatory cytokine, IL17C, a member of the IL17 family, plays an essential role in immunopathology, autoimmune diseases, and cancer progression [551]. It was found that IL17C is higher expressed in CRC tissue and induces tumor angiogenesis of intestinal endothelial cells *via* VEGFR2 production, subsequently enhancing cell invasion and migration of CRC cells [552,553]. Moreover, elevated levels of serum and tissue IL17C were observed in patients with active IBD, which can result in cancer progression [552]. Among these patients, the production of IL17C is induced by the synergic effect of IL17A and TNF-α [554]. Therefore, high circulating IL17C may be associated with tumorigenesis from IBD to early stages of CRC. Lasts of these cytokines, IFNG, is critical to both innate and adaptive immunity [555]. IFNG was reduced in PBMC of patients with recurrent CRC, with the most significantly reduced expression in stage IV tumors [556]. On contrary, the upregulation of *IFNG* mRNA in late-stage CRC tissue and peripheral blood of patients with CRC was observed in another study [557]. IFNG is a well-established anti-tumor factor with controversial findings in CRC at mRNA and protein levels. Several studies did not find a significant association between circulating IFNG and CRC development[558–560]. In contrast, our analysis showed high levels of IFNG in CRC patients supporting previous findings [496]. Moreover, we found high levels of IFNG in the early stages of CRC, suggesting a higher anti-tumor activity of lymphocytes than in the late stages. Taken as a whole, these findings indicate that ACP6, FLT4, MANSC1, IFNG, IL17C, and IL32 may be used as promising prognostic biomarkers that distinguish early-stage from advanced CRC. Moreover, IFNG can be a potential biomarker for early detection of CRC due to its discrimination between early-stage patients with advanced CRC patients as well as healthy controls, which has not been reported before.

In this study, the application of PEA technology enabled us to detect 690 proteins from a low amount of plasma of CRC patients and healthy subjects. Despite the sensitivity and accuracy of PEA, this technology is limited by the availability and specificity of antibodies, and more importantly, the number of preselected proteins. Women are dominant in both study groups, which is different concerning the known population with CRC. We lacked information on family history, which is known as one of the best predictors of CRC risk. Future studies should be conducted to verify our results on a larger number of samples and by using PEA or other quantitative methods.

In conclusion, we identified plasma protein changes in CRC patients related to cytokine interactions, oncogenic pathways, Th17 activity, metabolism reprogramming, as well as cancer-related inflammation with potential usage in CRC diagnosis. We also showed that six proteins, including ACP6, FLT4, IFNG, IL17C, IL32, and MANSC1 may be used as potential prognostic biomarkers to discriminate early-stage and advanced CRC. Moreover, IFNG is a new candidate biomarker for the early detection of CRC.

CHAPTER 7. Deep proteomics characterization of colorectal cancer tumor microenvironment enriched in CD4+ T cells

In the last part of the thesis, deep proteomics analysis was applied to CRC and normal matched tissue samples enriched in CD4+ T cells and other immune cells. This study aimed to characterize protein changes associated with CRC development, progression and immune infiltration that could become novel potential regulators of immune responses within CRC TME.

This study resulted in an unpublished manuscript:

Urbiola-Salvador V, **Miroszewska D, Jabłońska A, Duzowska K, Drężek-Chyła K, Zdrenka M, Śrutek E, Szylberg Ł, Jankowski M, Bała D, Zegarski W, Nowikiewicz T, Makarewicz W, Adamczyk A, Ambicka A, Przewoźnik M, Harazin-Lechowicz A, Ryś J, Filipowicz N, Piotrowski A, Dumanski JP, Chen Z. Deep proteomics characterization of colorectal cancer tumor microenvironment enriched in CD4+ T cells.** ***Unpublished manuscript***

7.1. Introduction

Despite great advances in colorectal cancer (CRC) diagnosis and treatment, CRC remains the second most deadly and the third most common cancer, worldwide [561]. Recently developed immunotherapies ,such as immune checkpoint blockade, have revolutionized CRC treatment, however, CRC can acquire resistance through alternative immunosuppressive mechanisms. As a consequence, only a minor portion of CRC patients exhibit complete responses to therapy [562]. Remarkably, the tumor microenvironment (TME) comprised of CRC cells intermixed with immune and stromal cells plays a central role in CRC development, progression, and immune evasion [563]. Therefore, deeper understanding of the CRC TME immune composition and the underlying immune evasion mechanisms is urgently needed.

Within the CRC TME, Cancer Associated Fibroblasts (CAFs) can support tumor growthand metastasis and they can interact with immune cells through pro-inflammatory and immunosuppressive mediators [564]. Among myeloid cells, M1 macrophages are mainly involved in anti-tumor activity while M2 macrophage phenotypes are related to immunosuppression and tissue remodeling that can recruit Tregs via CCL20, Th2 cells via CCL17, CCL18, and CCL22 to the TME [565]. CD4+ T helper cell subsets are essential regulators of the immune responses within CRC TME from which Th2 and Treg induce multiple immunosuppressive mediators such as IL10, and immune checkpoint inhibitors (PD1, TIM3, and CTLA4), favoring CRC immune evasion [77]. Meanwhile Th1 can recruit cytotoxic CD8+ T cells to enhance anti-tumor activity and Th1 infiltration is linked to better CRC prognosis [566]. Importantly, metabolic reprogramming of CRC TME by cancer and immune cells directly affect the TME cell composition including immunosuppressive mechanisms via exhaustion of effector anti-tumor cells through metabolic deprivation of amino acids such and induced high adenosine levels via CD39/CD73 among others [567].

Recent advances in mass spectrometry (MS)-based proteomics allows to quantify thousands of proteins with high accuracy and sensitivity whilst its clinical application can provide novel insights into CRC research [102]. In 2011, Wiśniewski et al. applied laser capture microdissection (LCM) followed by high-throughput MS proteomics analysis to CRC and normal-matched formalin-fixed paraffin embedded (FFPE) tissue from 3 patients. As a proof of the technology potential, around 4000 proteins were quantified per samples using a peptide pre-fractionation strategy detecting increased CRC marker, CEA, metastasis-associated in colon cancer protein 1 (MACC1), and CD55 [568]. In their follow-up study, the same strategy was applied to cancer, adenoma and normal samples from 16 CRC patients. Similarly, CEA and other tumor related proteins were increasing from normal to adenoma and CRC as well as metabolic reprogramming from oxidative phosphorylation to glycolysis and fatty acid metabolism. Moreover, multiple transporters were elevated to acquire nutrients from the TME [569]. Next, proteogenomics analyses of fresh frozen colon cancer and normal tissues from 110 patients resulted in the discovery of apoptosis dysregulation and increased proliferation, leading to novel potential therapeutic targets [246]. Interestingly, microsatellite unstable cancer sustained increased glycolysis linked to a reduction in CD8+ T cells within colon cancer TME [246]. Recently, another proteogenomics study including 145 CRC patients determined that patients could be stratified in three clusters resembling Consensus Molecular Subtypes (CMS) together with phosphoproteomics data to infer druggable targets. With this approach, minipatient-derived xenograft mouse models were used to evaluate potential druggable targets *in vivo*, to establish a personalized therapy screening platform [570]. Another recent proteogenomics study with 135 primary and 123 metastatic CRC patients' samples unveiled that proteomic hypoxic signatures are linked to metabolic reprogramming and Epithelial-Mesenchymal Transition (EMT) together with Transforming Growth Factor Beta 1 (TGFB1) signaling [571].Also, CRC stemness was associated with altered alternative telomere lengthening (ALT) pathway. Importantly, inferred immune score was associated with active MHCII antigen presentation, proteasome processing, FOXP3 and CD68 expression while immune "cold" tumors were characterized with poor survival [571]. Moreover, proteomics analysis using 12 CRC samples demonstrated that metabolic Phospholipase A2 Group IVA (PLA2G4A) pathway is associated with immunosuppressive CD39+γδ Tregs and prognosis of right-sided CRC [572]. Previous proteogenomics studies used bulk samples of fresh frozen tissue while microdissection allows for Region of Interest (ROI) isolation to characterize their specific proteomes within cancer tissue [571]. For instance, LCM followed by Data Independent Acquisition (DIA) MS-based proteomics was applied to evaluate proteomics changes in epithelial and stromal regions from 22 patients in the transformation from normal to adenoma to CRC . Similar protein features were shared between stromal adenoma and stromal CRC characterized by active antigen presentation and higher proportions of CD4+ and CD8+ T cells [573]. Another recently applied approach consisted of fresh frozen tissue disaggregation and cell sorting of CD4+ and CD8+ T cells from 13 CRC patients. As a result, lipocalin-2 (LCN2) was increased in CRC promoting T cell apoptosis via deregulation of iron efflux [574]. Interestingly, microdissection can be also combined with IHC staining to isolate ROIs enriched with specific cell markers. Huang et al. [575] applied IHC-LCM followed by DIA MS-proteomics to isolate CAFs and hepatocellular carcinoma cells within the cancer tissue demonstrating the main specific cell type isolation by representative CAF and cancer markers.

In this study, CD4 IHC followed by macrodissection was applied to isolate ROIs enriched with high CD4+ T cells and immune infiltration from CRC and normal-matched FFPE tissue samples. Deep DIA MS-based proteomics analysis aimed to determine protein changes involved in CRC development, progression, and associated immune infiltration within enriched ROIs. Several tumorigenic processes were altered such as cell cycle associated pathways and key epigenetic and transcriptional regulators, altered apoptosis with high levels of anti-apoptotic proteins. Importantly, protein profiles revealed a complex immune network of cancer-associated inflammation and adaptive immune processes composed of pro-inflammatory and immunosuppressive mediators such as CD276 and PVR within the CRC TME. Moreover, CRC TME proteome reflected heterogeneous cell compositions with co-existence of CAFs with high FGF2 and monocytes and

immunosuppressive M2 macrophages with high ICOSL linked to CRC progression. Moreover, Treg signatures were associated to high MHCII antigen presentation including tolerogenic-inducer GILT, inflammatory S100A8 and S100A9, and immunosuppressive mechanisms. Noteworthy, CRC TME showed metabolic reprogramming with several immunosuppressive mechanism including NT5E-derived adenosine signaling as well as tryptophan, taurine, and arginine deprivation ongoing simultaneously. Noteworthy, the novel immune-regulatory receptor Mast Cell Expressed Membrane Protein 1 (MCEMP1) was associated with CRC and may be involved in adhesion and migration of CRC infiltrating CD4+ T cells, especially Tregs.

7.2. Materials and methods

7.2.1. Study cohort and sample collection

In this study, 23 CRC patients (age mean 59.2, 42-75 years and 52% males (12 out of 23)) were included who underwent CRC surgery (Appendix V Table S1). 15 CRC patients had tumors with advanced stages according to the Union for International Cancer Control (UICC) classification. Malignant neoplasm was confirmed by a pathologist. Samples were obtained from 3P–Medicine Laboratory, Medical University of Gdansk and Bank of Biological Material at Masaryk Memorial Cancer Institute, Czech Republic. Tissue samples were collected after surgery, rinsed with PBS to remove blood and were formalin-fixed paraffin embedded. Then, sections of 5 µm thickness were cut with a microtome and sections were placed on glass slides and stored at room temperature.

7.2.2. IHC staining and ROI selection

For IHC staining, tissue sections of 5 µm thickness were deparaffinized for 2h at 60 °C, incubated in xylene for 20 minutes, and rehydrated in descending concentrations of ethanol (100%, 95%, 80%, 70%) for 5 minutes each. Slides were washed in Milli-Q water and placed in the preheated target retrieval solution pH=9 (DAKO, S2367) at 96 °C for 30 minutes, washed in Milli-Q water followed by PBS, and blocked with 3% BSA in PBS. Slides were stained with anti-CD4 antibody (Abcam 133616) at dilution 1:100 in 0.5% BSA/PBS with subsequent detection with HRP/DAB Detection IHC kit (Abcam, ab64261) according to manufacturer's instructions with antibody incubation time extended to overnight in humid chamber at 4°C. Slides were counterstained with hematoxylin (Sigma-Aldrich, GH5332) for 1 minute and mounted with Pertex® (Histolab, 00801-EX). Mounted slides were scanned with Axio Scan.Z1 (ZEISS, Oberkochen (Germany), digital images were uploaded to QuPath v.5.2. as Brightfield image (H-DAB) and ROI areas were marked. Color deconvolution was performed with Estimate stain vectors command followed by Positive Cell Detection command to detect CD4+ cells with default parameters, except for Intensity threshold parameters which were optimized for each sample individually. Score compartment parameter was set as 'Nucleus: DAB OD mean' whilst triple-thresholds values were adjusted individually for each of the samples due to variance of the intensity of the staining.

7.2.3. Sample preparation for proteomics analysis

Selected ROIs were scraped from the glass slide with a razor blade and transferred to a Protein LowBind Eppendorf tube with 3 µL of lysis buffer (4% SDS, 100 mM Tris-HCl pH 7.6 supplemented with protease and phosphoprotease inhibitors) per 10 nL of tissue. Samples were sonicated for 5 cycles of 30 s and 30 s pause at RT in an ultrasonicator Qsonica Q700 coupled with a cooler system (Qsonica, Newtown, CT, USA). Protein decrosslinking was performed by incubation in a thermoshaker for 1 h at 99 °C and 600 rpm. Protein concentrations were measured with the Pierce BCA Protein Assay Kit (Thermo Fisher) following manufacturer's instructions. Disulfide bonds were reduced with 50 mM DTT and an incubation at 95 °C for 5 min. Samples were prepared following the Filter Aided Sample Preparation (FASP) protocol [110] in 10 kDa cut-off Microcon filters (Merck, Rahway, NJ, USA). Six consecutive washes were performed with 200 µL of urea buffer (8M in 100mM Tris-HCl pH 7.6) at 10000 RCF for 20 min were performed. Free cysteines were blocked with 50 mM iodoacetamide in urea buffer by incubation 20 min in darkness. Then, samples were washed three times with 100 µL of urea buffer followed by two washes with 100 µL of digestion buffer (50 mM Tris-HCl pH 8.0) at 10000 RCF for 15 min. Proteins were digested with Sequencing Grade Modified Trypsin (Promega, Madison, WI, USA) at a protein:trypsin ratio of 1:50 overnights in 60 µL of digestion buffer at 37 °C. Then, peptides were eluted by centrifugation followed by two elutions with 125 µL and 100 µL of digestion buffer. Trypsin activity was quenched by adding trifluoroacetic acid at 0.1% final concentration. Peptide samples were desalted by STop And Go Extraction (STAGE) Tips protocol [294] in Empore C18 extraction disks (CDS Analytical LLC, Oxford, PA, USA) and eluted with 60% acetonitrile (ACN)/1% acetic acid solution. Samples were dried using SpeedVac and stored at −20 °C until analysis.

7.2.4. High-pH reversed phase liquid chromatography (RPLC) fractionation

Peptides were reconstituted (2% ACN, 0.1% formic acid (FA)), a pool sample was mixed from all the samples and diluted in phase A (5% ACN pH 9.8). The pool sample was fractionated in a Gemini high pH C18 column (5µm, 4.6 x 250mm) coupled in a Shimadzu LC-20AB HPLC system by gradient of phase B (95% ACN, pH 9.8): 5% for 10 min, 5%-35% in 40 min, 35%-95% in 1 min, and 95% for 3 min with a at a flow rate of 1mL/min. Elution was monitored at 214 nm and eluates were collected every minute. Eluates were concatenated in 10 fractions that were freeze-dried.

7.2.5. LC-MS/MS analysis and MS data analysis

Equal amounts of peptide samples with equal amounts of spiked iRT peptides (Biognosys Inc, Newton, MA, USA) were injected into a Thermo UltiMate 3000 UHPLC liquid chromatograph with a trap column to enrich peptides coupled to a self-packed C18 column (150µm internal diameter, 1.8µm column size, 35cm column length). Peptides were separated by a gradient of phase B (98% ACN, 0.1% FA): 5% for 5 min, 5%-25% in 85 min, 25%-35% in 10 min, 35%-80% in 5 min, 80% for 10 min, 80%-5% in 5 min at a flow rate of 500 nL/min. Separated peptides were ionized with spray voltage 2kV and injected to a tandem mass spectrometer Fusion Lumos (Thermo Fisher Scientific, San Jose, CA, USA). Pool fractions were analyzed in Data Dependent Acquisition (DDA) detection mode with 60K resolution MS scan (350-1500m/z) and MS AGC target of 3e6 with maximal injection time (MIT) 50ms by orbitrap mass analyzer that triggered the top 30 precursors. For MS/MS, resolution was 15K (200-2000 m/z) and AGC target was set to 1e5 with MIT 50 ms generated by High-Collision Dissociation (HCD) fragmentation with a normalized collision energy (NCE) of 30%. The dynamic exclusion was 30 s and MS/MS m/z start was fixed to 100. Precursors for MS/MS scan were with positive charge 2-6 and intensity over 2e4. Samples were analyzed in DIA detection mode with the same parameters as DDA, except the MS scan was 400-1500 m/z equally divided into 44 continuous windows and MS/MS resolution was 30K.

A hybrid spectral library was built with FragPipe (version 21.0) with the DIA_SpecLib workflow and default parameters including specific trypsin digestion [285]. *Homo sapiens* UniProtKB/Swiss-Prot database (Release 2024_02) was used as reference.Carbamydomethylation (C) was set as fixed modification. Variables modifications included oxidation (M), N-terminal acetylation, phosphorylation (STY), ubiquitination (K), pyroglutamic acid (QC), methylation (K), formylation (K), formaldehyde adduct (WYH), carbamylation (MLV), and dihydroxylation (WMH). DIA data quantification was performed with DIA-NN (version 1.8.2 beta 39) using the generated hybrid library with default parameters except protein inference was deactivated, while match between runs and peptidoform scoring were used [576].

7.2.6. Proteomics data processing, statistical and bioinformatics analysis

Proteomics data and statistical analysis were performed in RStudio (version 1.3.1093) (RStudio, PBC, Boston, MA, USA) with R (version 4.3.3) (R Foundation for Statistical Computing, Vienna, Austria). Proteomics data report from DIA-NN was used as input and spectral features were filtered by data preprocessing with MSstats R package (version 4.8.7) with default parameters except imputation was deactivated. MSstats preprocessing mainly includes data filtering with low detection rates, logarithmic transformation, feature median center normalization, and Tukey Median Polish summarization to protein abundances [577]. MSstats mixed-linear model was applied to test significantly differentially expressed proteins (DEPs) between paired CRC and normal-matched tissue samples. Proteins were considered differentially expressed with a False Discovery Rate (FDR) < 0.05 cut-off that was controlled by Benjamini and Hochberg correction. Spearman correlation analysis was applied to determine significantly correlated proteins with tumor TNM stage and predicted Treg fractions with a p-value < 0.05 cut-off. CIBERSORT deconvolution of immune cell fractions [578] was applied with the default LM22 signature matrix using IOBR R package (version 0.99.8) in absolute mode. Pathway enrichment analysis supported by active subnetworks was applied to determine enriched GO and KEGG terms from DEPs and significantly correlated proteins with the pathfindR R package (version 2.3.0) based on the STRING protein-protein interaction database and FDR correction [299]. Cytoscape (version 3.10.2) was used to generate protein networks from enriched term proteins based on STRING database. Paired t-test was applied to compare protein levels from a public proteomics dataset [574]. All the figure plots were generated with ggplot2 R package (version 3.4.3), except alluvial plot complemented with ggalluvial R package (version 0.12.5), Uniform Manifold Approximation and Projection (UMAP) plot that was created using the default pipeline using 20 Principal Components of Seurat R package (version 4.4.0), and heatmaps were generated with ComplexHeatmap R package (version 2.16.0) with z-score normalization.

7.3. Results

7.3.1. Deep DIA proteomics characterization of FFPE CRC and normal-matched tissues enriched with CD4+ T cells and immune infiltration

In this study, 23 CRC patients were included ,of which 15 were diagnosed with advanced TNM stages, to perform deep proteomics characterization of CRC tissues with high CD4+ T cells infiltration including secondary lymphoid structures and tertiary lymphoid structures (TLSs) in normal and CRC tissues, respectively. FFPE tissue samples were stained with CD4 antibody to determine CD4 infiltration (Figure 7.1a) and selection according to high immune lymphocyte infiltration was made. ROIs with high percentages of CD4 infiltration were isolated by macrodissection from CRC and normal matched tissue slides followed by protein extraction and sample preparation by FASP protocol for DIA LC-MS/MS proteomics analysis. To create a bona fide spectral library, high-pH fractionation was performed in a pool sample was fractionated and ten fractions were analyzed in DDA mode while the samples were analyzed in DIA mode to improve the protein identification and quantification. As a result, 9249 protein groups supported by spectra from 76448 peptides were included in the spectral library. After data preprocessing and peptide summarization, 7983 protein groups were quantified along the cohort samples supported by 51789 peptides with FDR < 0.01. Most of the protein groups were quantified in cancer and normal, however, some of them were selectively expressed (Figure 7.1b, Appendix V Table S2). Several selectively expressed proteins in normal

matched tissue were related to epithelial integrity such as BEST4, a mature absorptive cell marker [579], PLA2G10 that is expressed in Paneth-like secretory cells and suppression is linked to CRC [580], and recently found tumor suppressor ABCA8 [581] as well as proteins related to normal immune response and Peyer's patches integrity including CCR10, involved in IgA-secreting B cell migration to intestinal mucosa, and CCL19 constitutively expressed in secondary lymphoid tissues to attract CCR7 expressing T cells and other immune cells. The lack of these proteins is a reflection of the disruption of normal tissue integrity in CRC.

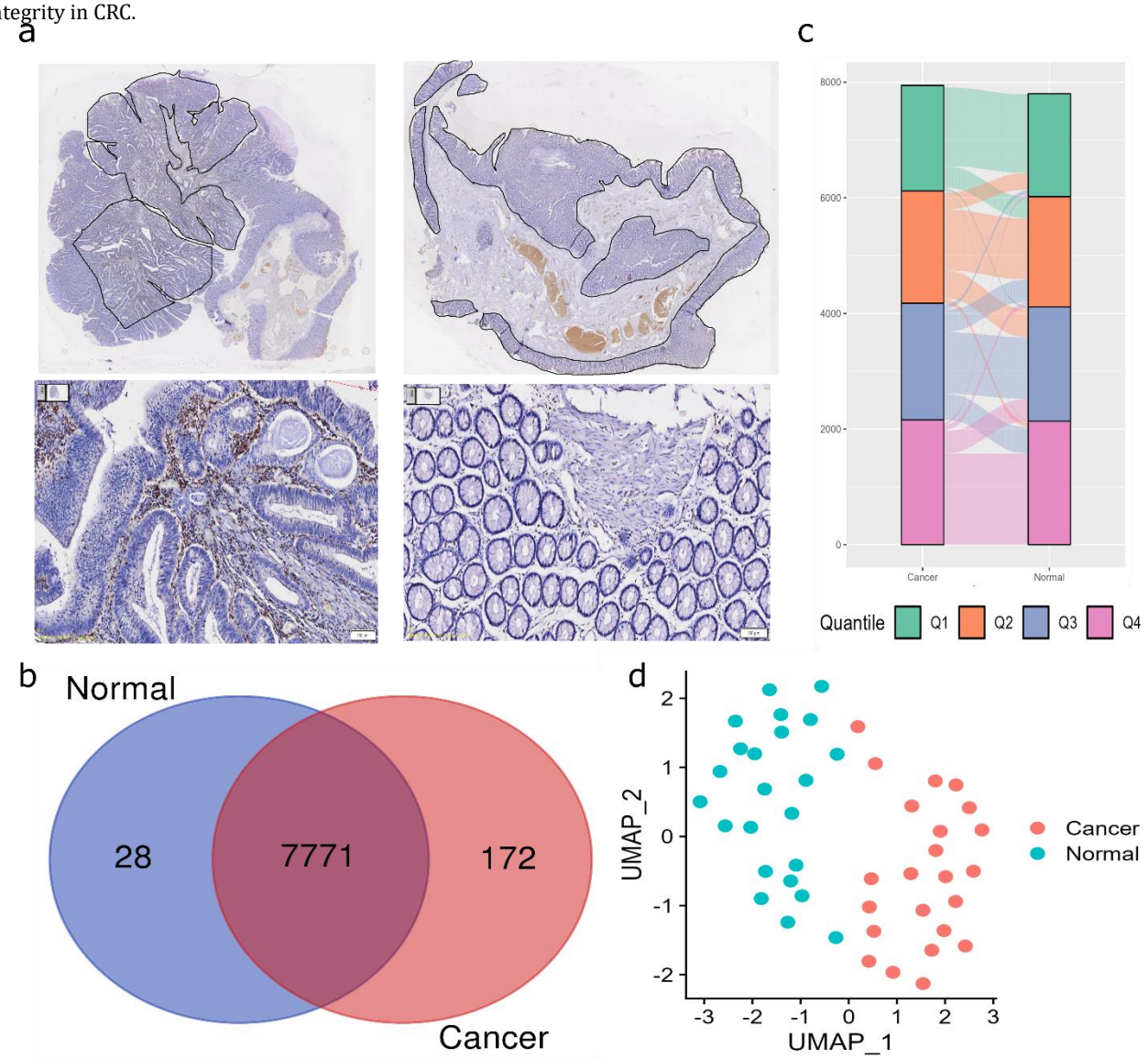

Figure 7.1. Deep proteomics characterization of C4+ T cell enriched CRC tissue and normal matched tissue. (a) CD4 staining of representative CRC and normal matched tissues (left and right) and their corresponding magnifications with high CD4+ T cell infiltration. (b) Venn diagram of quantified proteins between both tissue types. (c) Alluvial plot of proteins commonly identified proteins in CRC and normal tissue divided in four quantiles according to their cumulative distribution of protein abundance mean. (d) UMAP plot of cancer and normal samples.

Among selectively detected proteins in CRC, ASCL2 and LGR5 were associated with CRC stem-like cells with metastatic capacities [582,583], cyclin CDKN2A, epigenetic regulators such as CTCF, and transcriptional factors such as MACC1 that may be involved in CRC epithelial mesenchymal transition (EMT) [584–586], JUN, JUNB, DACH1 or DACH2 and the cell cycle regulator AURKA together with its inducer ARID3A [587] (Figure 7.2b). Also, extracellular cellular matrix (ECM) remodelers were only found in CRC such as cathepsin K (CTSK), metalloproteinases MMP1, MMP11 and MMP12 [588] as well as sulfatases SULF1 and SULF2 involved in CRC progression [589,590]. There were proteins linked to apoptosis such as death receptor 5 (TNFRSF10B) but at the same time the decoy receptor TNFRSF6B that protects against apoptosis was identified. Interestingly, proteins involved in CRC metabolic rewiring were also identified only in CRC including sulfotransferase SULT2B1 that is involved in CRC progression and metastasis by metabolic lipid activation [574], the taurine and amino acid transported SLC6A6, folate receptor FOLR3 related to one-carbon metabolism and KYNU involved in tryptophan metabolism-mediated

immunosuppression, a mechanism exploited in CRC via mainly tumor-associated macrophages and Treg [591]. Although stromal tissue was excluded from ROIs, CRC stromal infiltration andcancer-associated fibroblasts (CAFs) association is reflected by COL10A1 and COL11A1 selective detection in CRC samples [592]. Importantly, several inflammatory players were also selectively detected in CRC such as monocyte and granulocyte chemoattractant S100A12, that contributes to innate immunity by interaction with toll-like receptor 4 (TLR4), receptor for advanced glycation end products (RAGE), or CD36 that is implicated in IBD and CRC [593]. Additionally, other immune related proteins were selectively detected in tumor samples such as C5AR1, a complement C5a receptor, cytolytic perforin PRF1, Interferon Induced Transmembrane Protein 3 (IFITM3), or chemoattractant cytokines CXCL10 and CCL20. Moreover, recently characterized proteins were also detected in CRC samples including C19orf53 that plays a role in metabolic imbalance and excessive cell proliferation [594] and C19orf59 defined as Mast Cell Expressed Membrane Protein 1 (MCEMP1).

To investigate the distribution of the cumulative abundance between cancer and normal tissue, proteins were divided in four quantiles according to the cumulative abundances. The alluvial plot showed similar patterns of cumulative distribution, especially the most abundant proteins (Q1) were shared between cancer and normal including histones, keratins, tubulins, and other structural proteins as well as most of low-abundance proteins (Q3 and Q4). However, several proteins are in different quantiles between cancer and normal, suggesting different protein composition due to tumorigenesis and TME alterations (Figure 7.1c). In addition, UMAP analysis showed that cancer and normal samples were grouped according to their protein abundance (Figure 7.1d). Collectively, DIA proteomics analysis of CRC and normal matched tissues enriched in CD4+ T cells and immune infiltration consistently quantified over 7900 protein groups with several selectively proteins that reflects CRC TME with tissue integrity disruption, oncogenic TFs and cycle proteins from uncontrolled proliferation of CRC cells, ECM remodeling, metabolic rewiring, and prominent presence of pro-inflammatory proteins and cytokines as well as immunosuppressive mechanisms.

7.3.2. Protein changes in CRC TME reflects a complex network of immune processes with cell heterogeneity

To determine protein changes involved in CRC development within the TME enriched in CD4+ T cell infiltration and other immune cells, a mixed-general linear model was applied to protein abundances between paired tumor and norma-matched. As a result, there were 1954 protein groups with higher levels and 607 with lower levels in CRC (Figure 7.2a, Appendix V Table S3). The most elevated protein in CRC was IGF2 that is a mitogen involved in cancer invasion secreted by CRC cells and CAFs, that together with high levels of COL12A1, tenascin C (TNC), Latent Transforming Growth Factor Beta Binding Protein 2 (LTBP2), SPARC and CD90 supports the presence of CAFs promoting cancer-associated inflammation within highly immune infiltrated CRC regions [595–599]. Similarly to selectively detected proteins, the most elevated proteins in CRC included CRC stem cell markers such as OLFM4 and PROM1 as well as adhesion proteins CEA Cell Adhesion Molecule (CEACAM) 6, an anoikis inhibitor, and CEACAM5 [600–603]. In addition, other oncogenes and proteins involved in CRC proliferation were elevated such as cyclins CDK1-2,5-7, MKI67, thymidine kinase 2 (TK2), G protein subunit gamma 4 (GNG4) [604], and Gprotein-coupled receptor, class C, group 5, member A (GPRC5A) [605] among others. Moreover, S100A12 is supported by other family proteins including S100A8 and S100A9 that are also involved in cancer-associated inflammation. Also, elevated S100P contributes to angiogenesis and CRC progression whose regulator MACC1 was selectively detected in CRC samples [606,607]. Also, MMP10, MMP2, and MMP9 together with Neutrophil Gelatinase-Associated Lipocalin (NGAL) are involved in ECM remodeling, cancer progression and metastasis while elastin (ELN) supports CRC growth within TME [608,609]. Interestingly, multiple proteins involved in inflammation and neutrophile/monocyte/macrophage infiltration were elevated in CRC such as defensins DEFA1 or DEFA3 [610], Azurocidin 1 (AZU1) [610], myeloperoxidase (MPO) [611],Macrophage Migration Inhibitory Factor (MIF),myeloid nuclear differentiation antigen (MNDA) involved in inflammosome activation [612], and guanylate-binding protein 1 (GBP1). Interestingly, lipocalin 2(LCN2) was elevated in CRC which recently was associated to T-cell apoptosis by iron efflux deregulation in CRC [574]. Meanwhile, not previously reported, high levels of nucleophosmin 3 (NPM3) may promote mediated immune scape in CRC [613]. Importantly, we also found elevated levels in CRC tissue of CSK binding protein (CBP) that is a negative TCR signaling regulator via CSK-mediated LCK inhibition and a PD-1 signaling mediator [614,615]. In addition the master immunosuppressor TGFB1 was elevated in CRC together with Syndecan Binding Protein (SDBP) that contributes to TGFB1-induced epithelial-to-mesenchymal transition (EMT) as well as elevated levels of immune checkpoint CD276 and PVR [616,617]were identified. Moreover, Ly1 Antibody Reactive (LYAR), which is a negative regulator of NF-kappa-B-mediated pro-inflammatory cytokines expression that can be expressed in CRC, T cells, B cells and myeloid cells [618], was increased in CRC. At the same time, other innate immune related proteins, such as LBP and anti-bacterial Bactericidal Permeability Increasing Protein (BPI) were elevated (Figure 7.2a).

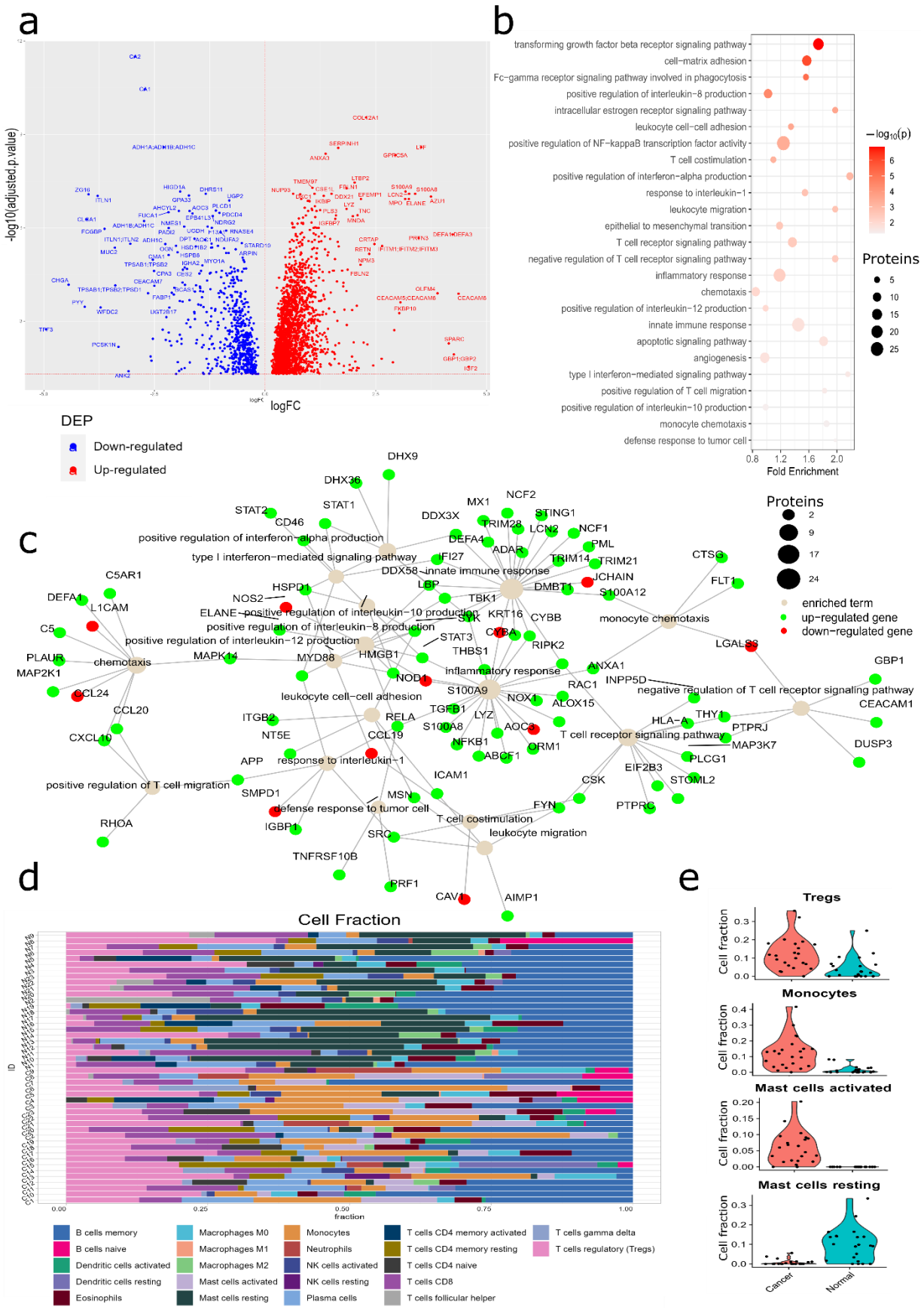
a
-log10(adjusted.p.value)
logFC
DEP
Down-regulated
Up-regulated
b
transforming growth factor beta receptor signaling pathway
cell−matrix adhesion
Fc−gamma receptor signaling pathway involved in phagocytosis
positive regulation of interleukin−8 production
intracellular estrogen receptor signaling pathway
leukocyte cell−cell adhesion
positive regulation of NF−kappaB transcription factor activity
T cell costimulation
positive regulation of interferon−alpha production
response to interleukin−1
leukocyte migration
epithelial to mesenchymal transition
T cell receptor signaling pathway
negative regulation of T cell receptor signaling pathway
inflammatory response
chemotaxis
positive regulation of interleukin−12 production
innate immune response
apoptotic signaling pathway
angiogenesis
type I interferon−mediated signaling pathway
positive regulation of T cell migration
positive regulation of interleukin−10 production
monocyte chemotaxis
defense response to tumor cell
Fold Enrichment
−log10(p)
Proteins
c
Proteins
enriched term
up−regulated gene
down−regulated gene
positive regulation of interferon−alpha production
type I interferon−mediated signaling pathway
DDX58−innate immune response
ELANE−positive regulation of interleukin−10 production
positive regulation of interleukin−8 production
positive regulation of interleukin−12 production
chemotaxis
inflammatory response
leukocyte cell−cell adhesion
response to interleukin−1
defense response to tumor cell
T cell costimulation
leukocyte migration
monocyte chemotaxis
negative regulation of T cell receptor signaling pathway
T cell receptor signaling pathway
positive regulation of T cell migration
d
Cell Fraction
ID
fraction
B cells memory
B cells naive
Dendritic cells activated
Dendritic cells resting
Eosinophils
Macrophages M0
Macrophages M1
Macrophages M2
Mast cells activated
Mast cells resting
Monocytes
Neutrophils
NK cells activated
NK cells resting
Plasma cells
T cells CD4 memory activated
T cells CD4 memory resting
T cells CD4 naive
T cells CD8
T cells follicular helper
T cells gamma delta
T cells regulatory (Tregs)
e
Tregs
Monocytes
Mast cells activated
Mast cells resting
Cell fraction
Cancer
Normal

Figure 2. CRC TME enriched in CD4+ T cells and immune infiltration reflects a complex immune network. (a) Volcano plot of differentially expressed proteins (DEPs) between CRC and normal matched tissue with corresponding logarithmic fold changes and adjusted p-values. (b) Bubble plot of selected GO-BP from DEPs between CRC and normal tissue. (c) Network of innate and adaptive immune processes within CRC TME inferred from DEPs present in the GO-BP terms. (d) Bar plot of CIBERSORT deconvolution results from cancer and normal tissues. Cell fractions are normalized to 1. (e) Violin plot of significantly changed cell fractions between cancer and normal tissues (all of them significant with adjusted p-value < 0.05.

Among reduced proteins in CRC, crucial proteins to maintain intestine mucosa integrity and nutrient intake such as trefoil factor (TFF)-3, peptide-YY (PYY) , and glucagon together with multiple keratins and mucins were identified, reflecting the mucosa disruption in CRC CRC (Figure 7.2a, Appendix V Table S3). Supporting CRC tissue disruption, the lack of tissue integrity is evident by reduced levels of desmin and dermatopontin. Also, the reduction of EPCAM is related to EMT and reduced levels of apoptotic protein GRAMD4, also involved in TLR9-mediated inflammation [619,620],may provide apoptosis resistance to CRC cells.. Moreover, tumor suppressors such as Chromogranin-A [621] and Intelectin 1 (ITLN1), that can inhibit suppressive myeloid cells [622], are also reduced in CRC . Importantly, other anti-tumor macrophage derived proteins were reduced including FOLR2 and MARCKS [623,624], supporting the M2 macrophage infiltration within CRC TME. Regarding B cell proteins, we found reduced levels of IgA, and JCHAIN in CRC from IgA plasma secreting cells required for intestinal immunity as well as protective proteins such as Fc Gamma Binding Protein (FCGBP) and zymogen 16 (ZG16) while IgGs, the BCR signaling transducer CD81 which facilitates clonal expansion and antibody production [625], and the IgM signal transducer Immunoglobulin Binding Protein 1 (IGBP1) are elevated in CRC. Another relevant reduced protein was Neural Cell Adhesion Molecule 1 (NCAM1), CD56, involved mainly in NK cell activation but also in regulation of T cells and B cells.

Next, pathway enrichment analysis of GO and KEGG terms with DEPs and selectively expressed proteins was performed to infer the biological processes involved in CRC TME. Both analyses showed that the top terms were associated with splicing, RNA synthesis and translation, cell cycle, DNA repair, proteasome, and protein folding (Appendix V Figure S1a-b, Table S4 and S5). Noteworthy, metabolic rewiring was reflected in CRC protein changes with reduced oxidative phosphorylation and increased alternative pathways including amino acids metabolism, central carbon cancer metabolism, and purine metabolism (Appendix V Figure S1c). As DEP analysis demonstrated specific immune protein patterns in CRC, we focused on enriched immune and inflammatory processes, however, there were several hallmarks of cancer enriched in CRC including angiogenesis, apoptotic deregulation with high levels of cell death inhibitor BCL2L1, cell-matrix adhesion, and EMT among others (Figure 7.2b). Among the innate immune responses, the complement cascade was activated in CRC with elevated levels of multiple components and regulators (Appendix V Figure S1c). These results are in agreement with our previous LC-MS/MS proteomics analysis of CRC plasma samples in which several of these complement proteins were also elevated in plasma of CRC patients compared to healthy controls including C4B, C9 and C5 as well as other proteins such as LBP and ITIH4 [626]

Interestingly, a complex network of innate and adaptive immune processes was identified with most of the related proteins elevated in CRC (Figure 7.2b). Among them, innate immune and inflammatory responses including type I IFN responses, IL1 response,IL12 production, monocyte and T cell chemotaxis, Fc gamma receptor, TGFB1 and NF-KB signaling pathways were elevated in CRC.. Interestingly, enriched T cell migration and TCR signaling together with anti-tumor cell response proteins, such as PRF1 and the death receptor 5 (TNFRSF10B), may indicate an active anti-tumor response. However, as immunosuppressor IL10 production pathway was activated and several negative T cell regulators and immune checkpoints (PVR and CD276)were elevated in CRC, increased PRF1 may indicate Treg cytolytic action in effector T cells. In fact, other immunosuppressive proteins were elevated in CRC such as 5'-Nucleotidase Ecto (NT5E) that convert ATP to immunosuppressive adenosine, inhibiting T cell activation (Figure 7.2b) [627]. Proteasome and antigen presentation were also enriched in CRC including classical MHCI molecules (HLA-A,-B,-C) and antigen processing proteins together with elevated levels of inhibitory T cell signaling such as CEACAM1, PTPRJ, and GBP1, suggesting defective T cell responses upon MHCI antigen presentation.

Considering the protein signatures of immune infiltration, CIBERSORT cell fraction deconvolution was applied to infer the immune cellular composition within CRC TME and normal tissues. Cell fraction deconvolution resulted in a complex mixture of immune cell types in both tissue types with presence of diverse T cell subsets, mast cells, myeloid cells, and B cells (Figure 7.2d). Importantly, CRC samples were with significantly elevated levels of inferred Treg, monocyte, and activated mast cells but reduced resting mast cells, suggesting an immunosuppressive TME in CRC compared to normal tissue (Figure 7.2e). Taken together, CRC tissue regions with high CD4+ T cell infiltration is characterized by multiple cancer transformation pathways and hallmarks such as metabolic rewiring, cell stemness, and apoptosis dysregulation. At the same time, an intricate network of innate and adaptive immune processes with remarkable cancer-associated inflammation and underlying immunosuppressive mechanisms from Treg and monocytes/macrophages supported CRC immune evasion.

7.3.3. CRC progression is associated EMT, CAFs infiltration and ICOSL mediated immunosuppression

CD4+ T cells and other immune cells are fundamental players within the TME in CRC progression and metastasis [628]. To determine protein changes associated with CRC progression, correlation analysis was performed between cancer TNM stage and protein abundances. As a result, 319 proteins were positively correlated (rho>0.4, p-value<0.05) proteins with CRC stage and 198 with negative correlation (rho<0.4, p-value<0.05) (Figure 7.3a, Appendix V Table S6). Altered proteins along CRC

progression were associated with multiple biological processes such as RNA synthesis and processing, apoptosis with higher levels of CASP1 in early CRC stages, positive regulation of I−kappaB kinase/NF−kappaB signaling, and T cell costimulation among others (Figure 7.3b, Appendix Table S7). The protein with highest correlation was receptor protein tyrosine phosphatase δ (PTPRD) that aberrant expression can affectSTAT3 and SRC signaling contributing to cell metastasis and PD-L1 signaling [629,630] (Figure 7.3c). Several proteins involved in cell cycle and DNA repair were also elevated in advanced CRC such as POLE4 and CGGBP1 as well as key transcriptional factors AHR, FBH1, MBD1, and TCF20 or TMF1 involved in androgen receptor signaling. Metabolic alterations were also observed in CRC progression with elevated levels of glycosidases MAN2A2 and GANC that release glucose from N-glycans as well as PANK1 involved in coenzyme A biosynthesis that is essential for multiple metabolic processes.

Interestingly, several positively correlated proteins were likely derived from cancer-TME interactions and CRC heterogeneity. For instance, the neuroendocrine marker INSL5 was previously found elevated in CRC neuroendocrine tumors and may indicate CRC tumor heterogeneity within adenocarcinoma CRC tumors [631]. In addition, increasing vimentin may indicate an increasing EMT phenotype within CRC progression [632] together with increasing ECM remodeling by MMP2 and MMP14 as well as CRC stemness via ASCL2. Another positively correlated protein was IGFBP5 that is involved CRC progression and metastasis via ECM remodeling and associated with CAFs [633–635]. Importantly, FGF2 was positively correlated with CRC stage which may be involved in cancer invasion and secreted mainly by CAFs [636] (Figure 7.3c). In addition, increasing platelet factor 4 (PF4) levels were only detected in advanced CRC stages while high PF4 levels in circulating platelets of CRC patients were previously reported, supporting cancer invasiveness by TGFB1 platelet production [637,638]. In fact, increasing TGFB1 with CRC progression was also detected, supporting an immunosuppressive TME. Immune related proteins were also altered through CRC progression such as positively correlated cathepsin CSTE which may be involved in CRC tumorigenesis, previously proposed as a marker of colon cancer and sessile serrated adenomas [639,640]. Moreover, increasing levels of Leukotriene B4 receptor (LTB4R), which mRNA expression was previously reported as CRC prognostic predictor, is associated with CRC progression, intestinal inflammation, and is required for leukocyte migration [641,642]. At the same time, increasing TCR transducer LCK and TRAF6 may indicate T cell stimulation although cancer and other immune cells can express LCK and TRAF6 with tumorigenic functions [643,644]. In fact, TRAF6 was previously correlated with lymphangiogenesis and lymph node metastasis in CRC as well as pro-inflammatory cytokine secretion in innate immune responses [645]. Interestingly, SUSD2 was positively correlated with CRC progression that is involved in tumorigenesis and inhibit CD8+ T cell anti-tumor activity [646]. Noteworthy, increasing immune checkpoint ICOSL with CRC progression may indicate an immunosuppressive TME (Figure 7.3c).

Several immune related proteins were also negatively correlated with CRC progression such as ITGB6 involved in CD8+ T cell suppression or CCL24, eotaxin-2, that was downregulated compared to CRC tissue (Figure 2a, 3c). CCL24 is produced in secondary lymphoid structures in normal tissue and attracts granulocytes mainly eosinophils as well as resting T cells, but also, CCL24 is secreted in CRC were is associated with eosinophil infiltration [647,648]. Noteworthy, decreasing levels of MHC-I proteins (HLA-B,-C) and MHC-I adaptor B2M according to CRC progression, confirm a well-established mechanism of immune evasion [649].Moreover, IRF5 reduction with CRC progression may be associated with M2 macrophage polarization [650]. Similarly, the immune checkpoint LGALS9, that can promote the immunosuppressive TME through TIM3 interactions with CD8+ T cells and Tregs [651], was decreasing with CRC progression. Taken together, protein changes associated with CRC progression suggested the transcriptional reprogramming within TME with metabolic adaptions, EMT signatures, and ECM remodeling. Moreover, these protein changes may reflect TME heterogeneity with increased neuroendocrine INSL5, CAF-related proteins including FGF2 and IGFBP5, activated platelets, and relevant immunosuppressive mechanisms including SUSD2, ICOSL, IRF5, LGALS9, and reduction of MHC-I presentation among others.

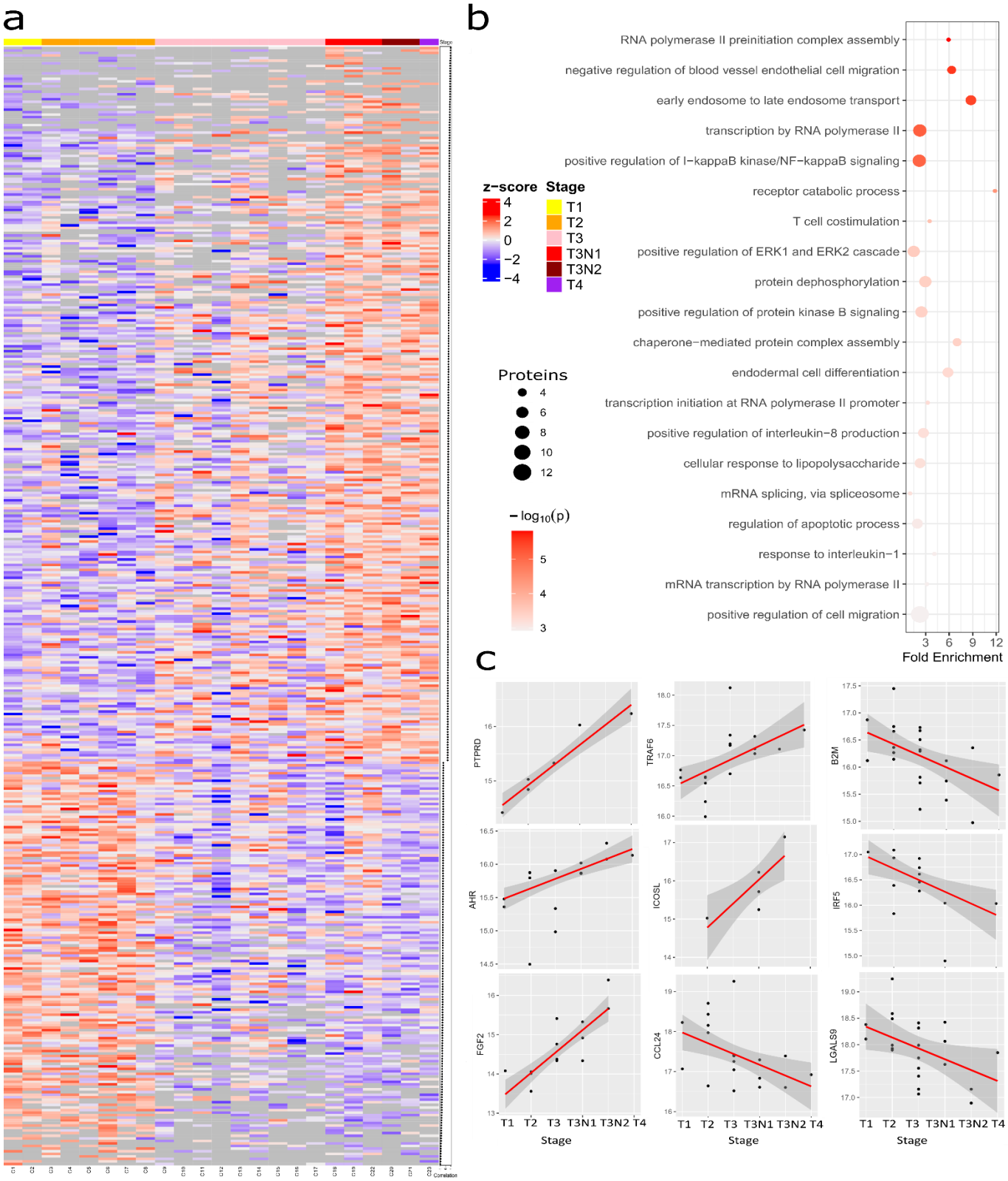

Figure 7.3. CRC progression and protein changes associated with EMT, CAFs, and immunosuppression. (a) Heatmap of significantly correlated proteins with CRC progression order by decreasing rho value and samples order in increasing CRC stage with z-score normalization (correlation rho vales on the right). (b) Bubble plot of selected GO-BP from significantly correlated proteins with TNM stage. (c) Scatter plots of selected significant correlated proteins with cancer TNM stage and relative protein abundance.

7.3.4. Proteomic changes linked to Treg infiltration in CD4+ enriched CRC tissues

Previous comparison between CRC and normal tissue resulted in higher Treg infiltration. Considering the fundamental role of Treg infiltration in CRC immune evasion, correlation analysis was performed between CIBERSORT inferred Treg fractions and protein abundances along CRC tissues to determine potential protein association with Tregs. As a result, 380 proteins were significantly positively correlated and 77 were negatively correlated with inferred Treg fractions (Figure 7.4a, Appendix V Table S8). Pathway enrichment analysis via active subnetworks demonstrated elevated oxidative phosphorylation that supports Treg differentiation [652],proteasome processing and high MHC-II antigen presentation, TCR signaling and T cell cytotoxicity as well as other processes such as Arp2/3 complex–mediated actin nucleation, apoptosis regulation, cell-cell adhesion, Fc–epsilon receptor signaling, and histone deacetylation among others (Figure 7.4b, Appendix V Table S9).

High MHC-II presentation, increasing MHC-I proteins HLA-A and HLA-F, immunoproteasome subunits PSMB8 and PSMB9, and MHC-I processing TAPBP and GILT may indicate co-existence of APCs with increasing Treg fractions (Figure 7.4c). Meanwhile, inflammatory-associated proteins S100A8, S100A9, and pro-tumorigenic S100P were positively correlated with Treg content. Increasing levels of integrins ITGAM and ITGB2 were associated with Treg fractions and both are involved in innate immune responses to complement-opsonized pathogens as well as T cell migration [653,654]. Similarly, FCER1G was linked to Treg content that could indicate the presence of mature regulatory DCs, a subset of highly immunosuppressive DCs with high indoleamine 2,3-dioxygenase (IDO)-1 production and CXCL9 deregulation [655]. In fact, IDO1 and arginase ARG1 were positively correlated with Treg fraction, supporting an immunosuppressive metabolic rewiring within enriched CD4+ T cell CRC tissues with high Treg fractions [656,657]. Increasing deacetylases SIRT1 and SIRT2 with Treg content may be associated with this immunosuppressive metabolic TME via deacetylation of key metabolic enzymes [131]. In contrast, ITGA4 and IFITM3 were negatively correlated with Treg fractions. Low integrin ITGA4 was previously associated with poor CRC prognosis and positively correlated with Th17 and immature DCs in CRC [658]. Importantly, IFITM3 is required to promote anti-tumor activity via perturbation of FOXP3 in tumor-infiltrating Tregs [659] (Figure 7.4c). Taken together, Treg-linked protein changes within CRC TME are associated with APCs with immunosuppressive phenotypes via metabolic alterations including tryptophan and arginine T cell depravation.

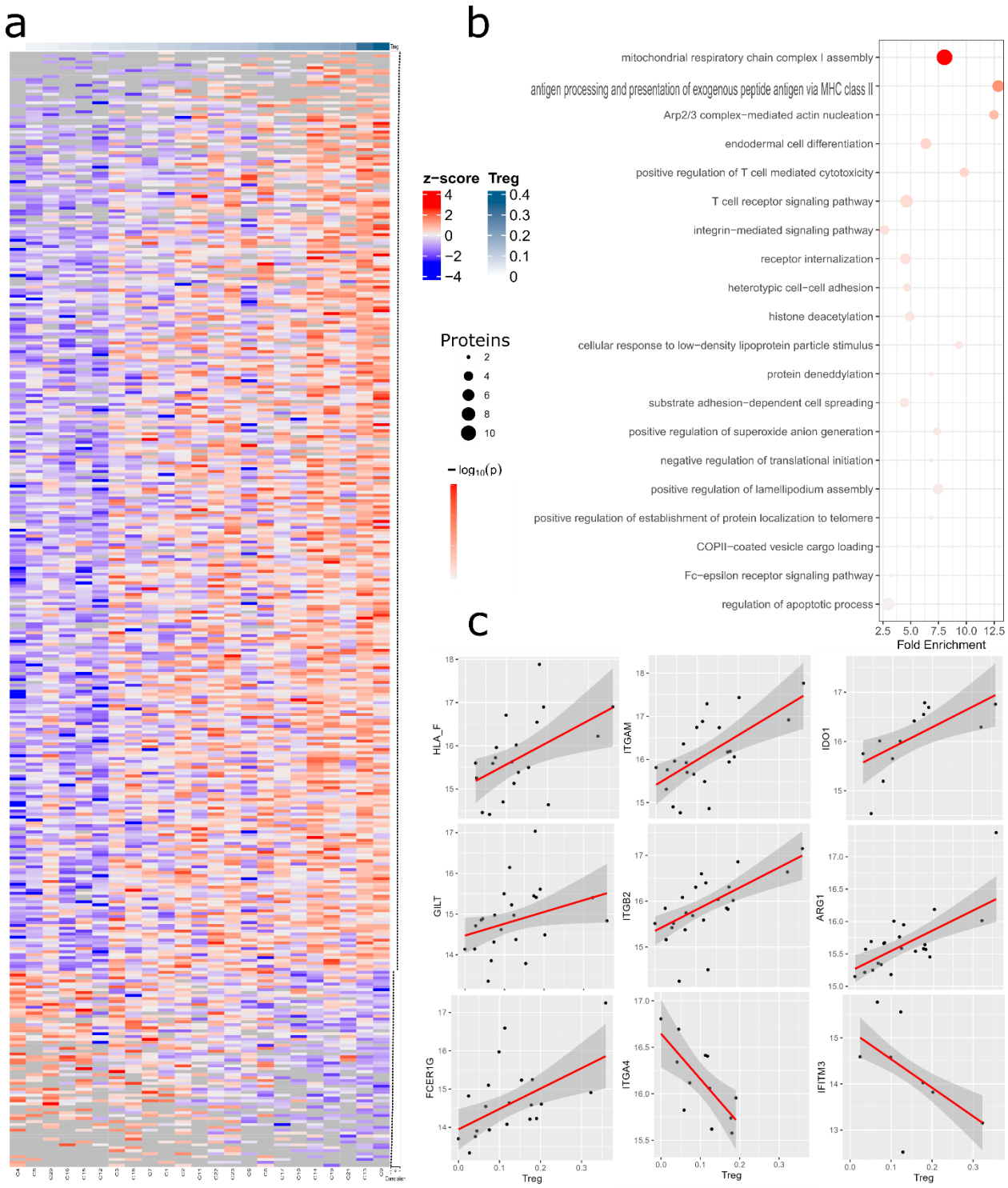

Figure 7.4. Treg fractions and associated protein changes. (a) Heatmap of significantly correlated proteins with Treg fractions order by decreasing rho value and samples order in increasing inferred Treg fraction with z-score normalization (correlation rho values on the right) (b) Bubble plot of selected GO-BP from significantly correlated proteins with Treg fractions. (c) Scatter plots of selected significantly correlated proteins with inferred Treg fractions and relative protein abundance.

7.3.4. CRC TME protein changes are associated with CD4+ T cell pro-inflammatory factors, CEACAMs, ARG1, and the novel chemotactic receptor MCEMP1

Our analysis revealed immune heterogeneity and the inflammatory/immunosuppressive processes within CRC TME enriched in CD4+ T cells and other immune cells. However, bulk proteomics analysis of complex tissue samples is limited in determining the cells expressing identified proteins. Therefore, a complementary analysis of public datasets to infer specific cell expression was performed. The first dataset comprised proteomics data from sorted CD4+ T cells and CD8+ T cells of CRC and normal-matched tissue [574] to infer protein changes in CRC TME derived from T cells. The main findings from Che et al. [608] were that elevated LCN2 induces T-cell apoptosis via iron efflux deregulation which our analysis supported with elevated LCN2 levels within immune CRC TME (Figure 7.2a). According to their reported results, cancer-associated inflammation proteins S100A8, S100A9, and S100P were higher in both tumor-infiltrating CD4+ and CD8+ T cells compared to normal counterparts while S100A12 and MMP2 were only elevated in tumor infiltrating CD8+ T cells. Moreover, adhesion proteins, CEACAM5 and CEACAM6, as well as pro-inflammatory MIF were elevated in CRC CD4+ T cells but not in CD8+ T cells. In fact, CEACAM6 was associated with FOXP3 and T cells in CRC progression [660]. These common results suggest that protein changes within TME are partly originated from infiltrated T cells.

Next, scRNA-seq data from CRC and normal matched tissue of 72 CRC patients [661] was used to infer specific cell expression of detected protein changes. For instance, genes from commonly increased proteins between our analysis and Che et al. [574] were with higher expression in tumor-infiltrating T cells than normal counterparts in scRNA-seq dataset. However, gene expression stems from other cell types such as LCN2, CEACAM5, and CEACAM6 highly expressed in tumor cells or S100A8 and S100A12 in myeloid cells (Figure 7.5a). Interestingly, a tendency to higher fractions of LCN2+ Treg and anti-tumor CXCL13+CD8+ T cells [662] was observed within CRC T cell subsets, suggesting a potential role in Treg-mediated immunosuppression of cytotoxic T cells. Another increased protein in CRC tissue was CEACAM1 which was confirmed in scRNA-seq data to be highly expressed by CRC infiltrating Tregs. Interestingly, the Gram-negative bactericidal BPI was only found in CRC tissues, mainly expressed by stem-like CRC cells and some macrophages/monocytes. Meanwhile, increasing proteins with CRC progression including FGF2 and PTPRD were mainly expressed by CAFs and CRC cells. From correlated proteins with inferred Treg fractions, ARG1 was only expressed from a minor portion of cells in this scRNA-seq dataset, mainly macrophage/granulocytes, low fraction of epithelial cells, and CRC infiltrating Treg and follicular helper T cells (Tfol) (Figure 7.5a). Noteworthy, the recently characterized immune receptor MCEMP1 was consistently quantified in CD4+ and CD8+ T cells with a tendency to be expressed at higher levels in tumor-infiltrating CD4+ T cells (paired t-test: logFC = 0.53, p-value = 0.06), but not in CD8+ T cells, suggesting a relevant role in CD4+ T cells within the CRC TME (Figure 7.5b). Similarly, scRNA-seq confirmed MCEMP1 expression mainly in CRC monocyte/macrophage, granulocytes, CRC stem-like cells, and importantly, several T cell subsets with a high fraction of MCEMP1+Tregs within them (Figure 7.5a, 7.5c, Appendix V Figure S2). Taken together, our proteomics analysis reflected protein changes that may be associated with specific cell subsets within the TME, especially from myeloid and T cells, several of them playing a relevant role in tumor immunity and immunosuppression.

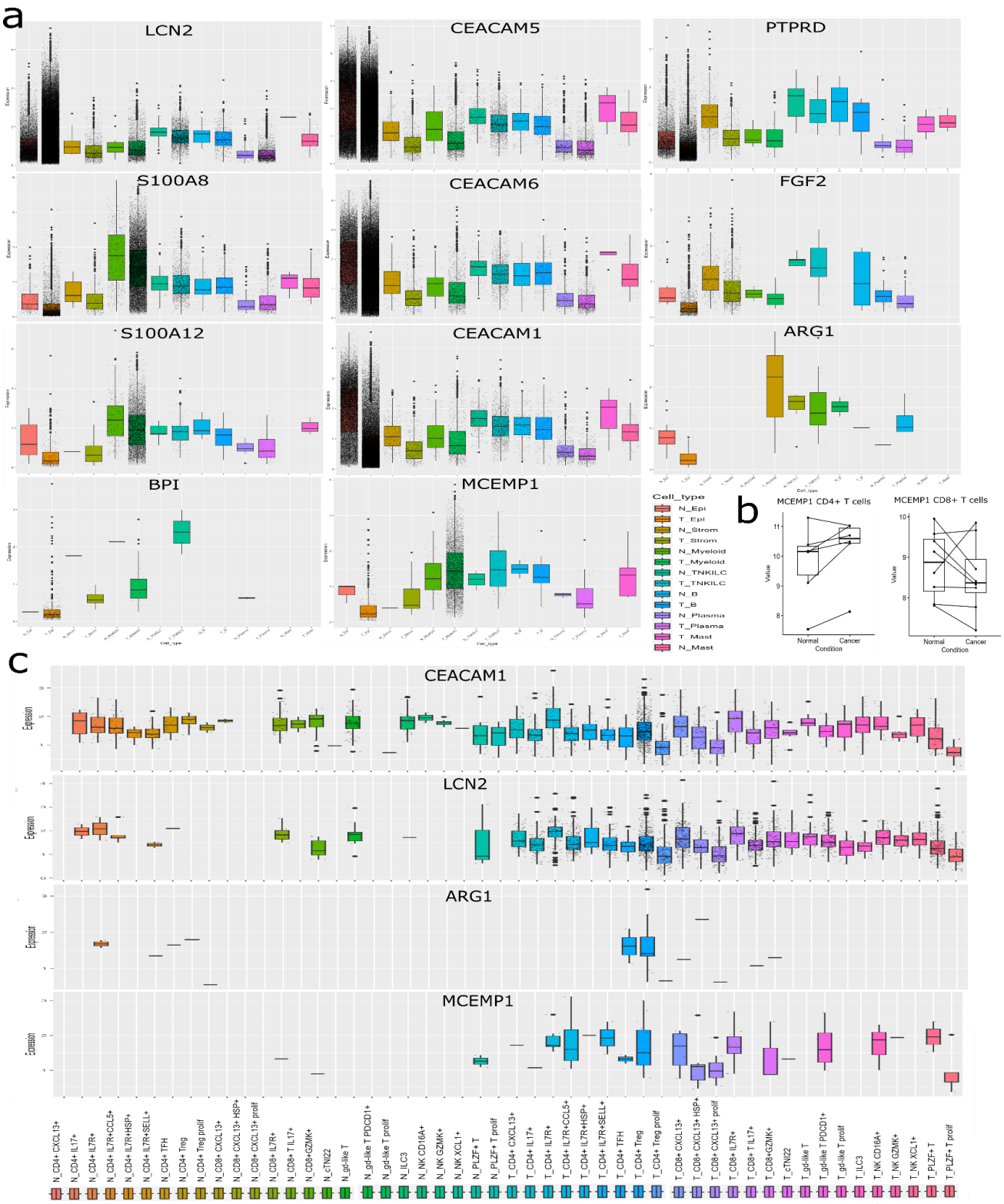

Figure 7.5. Protein changes are associated with specific cell components of CRC TME. (a) Box and whisker plots of normalized mRNA expression in scRNA-seq dataset from Pelka et al. [661] in main cell types separated by tumor and normal tissues. Each small black dot represents the gene expression of a single cell. (b) Paired box plots of normalized abundances of MCEMP1 between CRC-infiltrating CD4+ and CD8+ T cells and normal counterparts, respectively. (c) Box and whisker plots of normalized mRNA expression along T cell subtypes separated from CRC and normal tissue.

7.4. Discussion

Immune cell infiltration plays a relevant role within CRC TME via supporting tumor growth, invasion, and drug resistance among others [565]. In this study, deep DIA proteomics characterization of CRC tissue and normal matched tissue enriched in CD4+ T cells and other immune cells was performed to determine protein changes within CRC TME involved in CRC development, progression, and immune infiltration. CRC and normal-matched tissue samples were consistently different in their protein composition. In fact, selectively detected proteins reflected CRC tumorigenic processes with key epigenetic and transcriptional regulators. At the same time selective expression in only normal tissue of BEST4, a marker of absorptive cells, and CCL19, specific for secondary lymph nodes and previously found associated with anti-tumor CD8+ T cells in TCGA datasets and breast cancer [663], confirmed disrupted tissue integrity and reduction of anti-tumor CD8+T cells within studied CRC tissue, respectively. Mainly, protein changes within TME CRC reflected an active cell cycle with associated biological processes such as DNA repair, RNA synthesis, spliceosome, and protein refolding under high proliferation. Moreover, key CRC stem-cell markers such as ASCL2 involved in CRC progression, PROM1, and LGR5, for which CAR-T cell therapy is currently under development [664] were identified in CRC samples. Other proteins related to hallmarks of cancer were altered such as angiogenesis with elevated FLT1 associated with increasing microvessels in CRC [665], apoptosis deregulation with anti-apoptotic BCL2L1 and TNFRSF6B associated with poor CRC prognosis [666,667] or EMT with EPCAM reduction and transcriptional factor MACC1 via HGF/MET signaling [586].

Importantly, protein expression patterns in CRC TME reflected changes in their cell composition compared to normal-matched tissue. Although several key cell markers are missing due to their low abundance within CRC TME, CIBERSORT can deconvolute and estimate cell fractions in bulk samples. Inferred immune cell fractions suggested an increased immunosuppressive environment compared to normal tissue with higher fractions of Treg, monocytes, and activated mast cells. Meanwhile, protein changes revealed further cell composition changes linked to alterations in innate and adaptive immune responses. For instance, innate immune proteins dedicated to intestinal defense against pathogens such as ZG16 and FGCBP together with plasma-derived IgA were reduced while B cells may be increased with high IgGs, CD81, and IGBP1 in CRC TME. Also, NK cell marker CD56 was reduced similarly to previous studies based on IHC staining [668]. Among innate immune proteins, BPI was elevated in CRC tissues and in CRC stem-like and myeloid cells from the scRNA-seq dataset. BPI was previously reported to be associated with IBD, but not in CRC, and was proposed as an anti-angiogenic factor [669]. BPI may play a role within TME-infiltrated microbiota and can attenuate inflammation by competing with LBP [670]. Surprisingly, complement cascade members are involved in tolerogenic cell death and immunosuppressive TME with the recruitment of Treg, M2 macrophages, and myeloid-derived suppressor cells (MDSCs). Among the complement-related proteins, we identified elevated soluble Factor H, responsible for creating anti-inflammatory responses, and elevated negative complement regulators CD46 and CD55, responsible for dampening the complement cascade [671].

Several CAF-related proteins were increased in CRC tissue including SPARC, associated with CRC invasion [596],CD90, supported by a previous study which demonstrated CRC CD90+ stromal cells are main IL6 producers, promoting cancer-associated inflammation [597], and also can produce immunosuppressive NT5E [672]. Active ECM remodeling with high levels MMPs in CRC tissue may indicate the co-existence of CAFs and myeloid cells that are involved in CRC progression and metastasis. In fact, multiple myeloid related proteins were elevated including GBP1, together with GBP2, that is required for autophagosome maturation and activated by interferon type I and II stimulations. In fact, GBP1 was associated with immunosuppressive M2 macrophage phenotype in previous studies [673]. Moreover, GBP1 can inhibit T cell activation by reducing IL2 production via IFNG [674]. Elevated levels of LCN2 together with lactotransferrin (LTF) may also counteract exacerbated inflammation and promote T-cell death via ferroptosis [675,676]. Interestingly, high LCN2 expression in infiltrating Treg populations found in scRNA-seq data is according to a previous study that demonstrated Treg differentiation via LCN2 characterized by non-classical HLA-G expression *in vitro* [677], suggesting additional immunoregulatory functions of LCN2 within CRC TME. Although antigen presentation and TCR signaling related proteins were found to be elevated in CRC TME, we also identified several immunosuppressive signals present in CRC tissue, such as immune checkpoints (CD276 and PVR) and the well-established immunosuppressor TGFB1 which also correlated with CRC progression. At the same time, PRPRJ was also increased in CRC while a recent proteomics study performed in circulating immune cells of CRC patients demonstrated that increasing PTPRJ levels are responsible for effector CD4+ T cell suppression along CRC progression [678].Moreover, increased non-classical HLA-F can negatively regulate NK cells via this alternative antigen presentation of a limited variety of peptides which is a potential reason for the observed reduced CD56 marker in CRC tissues [679].

Among identified proteins, we identified novel associations of FGF2 and ICOSLwith CRC progression. Previous stitudies found that FGF2 contributes to CRC invasiveness *in vitro* and angiogenesis. FGF2 is secreted mainly by CAFs butcan also be produced in an autocrine manner by CRC cells as observed in the CRC scRNA-seq dataset [680,681]. Noteworthy, the immune checkpoint ICOSL positively correlated with CRC progression may increase M2 macrophages and plasmacytoid DCs that induce Treg phenotypes characterized by TGFB1 and IL10 secretion [682]. Meanwhile, inferred Treg fractions were associated with oxidative phosphorylation and active antigen presentation from APCs including GILT. GILT was previously linked to tolerogenic responses to breast cancer and melanoma tumor antigens via induction of Treg differentiation. Recently, GILT was also associated with PD-L1 signaling in breast cancer [683],what highlights the potential role of GILT in immune tolerance to CRC [684,685]. Interestingly, ITGB2 was previously correlated with Treg and MDSC infiltration in non-small-cell lung cancer (NSCLC) and Treg-DC interaction via ITGB2 resulting in impaired antigen presentation and T cell activation [686,687].

Importantly, protein changes reflected metabolic rewiring within CRC TME showed by reduction of oxidative phosphorylation, alterations in glycolysis compensated with one carbon, amino acid, nucleotide, and lipid metabolism activation. Fatty acid metabolism inducers SULT2B1 was recently reported linked to CRC development and LYAR may contribute to cancer invasion via activation of lipid metabolism [688,689]. Accumulated evidence demonstrated that metabolic alterations are responsible of immune evasion and immunosuppressive mechanisms [690]. For instance, in this study, multiple proteins involved in metabolic immunosuppression were reported. Selective expression of KYNU was only detected in CRC and IDO1 was correlated with inferred Treg fractions while AHR was elevated in CRC and associated with CRC progression. Taken together, these results suggested that active tryptophan deprivation based immunosuppression was increased in CRC TME, previously associated with IBD and CRC [591]. Interestingly, SIRT1 and SIRT2 were correlated with Treg fractions and previous proteomics studies in CRC, melanoma, and lung cancer mice models determined that SIRT5 enhances Tregs [131] while SIRT2 can promote T cell exhaustion via deacetylation [132]. Apart from tryptophan metabolism, increased levels of the transporter SLC6A6 in CRC TME are involved in a novel immunosuppressive mechanism through taurine deprivation. Previously, SCL6A6 overexpression was found in CRC tissues and associated with chemotherapy resistance *in vitro* and *in vivo* [691], while a recent study in gastric cancer demonstrated that cancer SLC6A6 overexpression cause taurine deprivation resulting in CD8+ T cell dysfunction and exhaustion [692]. Another metabolic enzyme, ARG1, was found associated with CD15+ bone-marrow derived cells in CRC and was associated with poor prognosis [693,694]. Recently, ARG1+ granulocytes and IDO1 monocytes were analyzed in CRC and their spatial pattern distribution may be associated with CRC prognosis, however, immunosuppression was not addressed in this tissue microarray multiplex IHC study [695]. In our study, ARG1 was positively correlated with Treg fractions and scRNA-seq dataset analysis confirmed its expression in myeloid cells and in CRC-infiltrating Treg and Tfol, suggesting T cell effector suppression via arginine deprivation.

For the first time, the novel chemotactic regulator MCEMP1 was only detected within the CRC TME and confirmed in CD4+ T cells isolated from CRC compared to normal tissue [574].Public CRC single-cell expression showed MCEMP1 expression mainly in CRC monocyte/macrophage, granulocytes, CRC stem-like cells, and importantly, several T cells subsets with a high fraction of Tregs. MCEMP1 was primarily found in mast cells in which can regulate proliferation within lungs [696,697]. Recently, TGFB1-mediated activation of MCEMP1 was found in classical monocytes and alveolar macrophages in which MCEMP1 regulates migration and adhesion [698]. Another study in a mice sepsis model showed that MCEMP1 is upregulated in sepsis and promotes T cell apoptosis and inhibits their viability [699]. Interestingly, an *in silico* study in CRC mRNA data showed that FOXP3 and MCEMP1 were correlated with liver metastasis while in gastric cancer it was included in a gene prognostic signature associated with Treg [700,701] Further studies are needed to unveil MCEMP1 function in T cells, especially in infiltrating Treg within CRC TME as well as other immune cells.

This study was limited by the number of involved patients and limited tissue material, although robust proteomics analysis ensured deep proteome characterization of the CRC TME. Although bulk proteomics analysis of composed tissue samples limits to determine cell specific protein expression and their full spatial distribution, the usage of public proteomics and scRNA dataset could facilitate cell composition of the CRC TME. However, new combinational multi-omics approaches with emerging single-cell proteomics and spatial proteomics would provide deeper understanding of CRC underlying immune responses.

This study characterized the immune regulatory network that included co-stimulatory and inhibitory signals reflecting the complexity of immune responses within CRC TME. Moreover, protein signatures linked to CRC progression and Treg content within CRC TME were found. Cancer-associated inflammation and immunosuppressive mechanisms are disbalanced with co-existence of multiple processes from exacerbated inflammation to immune checkpoints and metabolic deprivation immunosuppressive mechanisms. Moreover, proteomics changes within CRC TME enriched in CD4+ T cells and other immune cells reflected immune TME heterogeneity with higher inferred fractions of Treg and monocytes. Our studied unveiled novel immune regulators involved in CRC that may facilitate the functional validation of immune-regulatory proteins for therapeutical application.

CHAPTER 8. CONCLUSIONS AND FUTURE DIRECTIONS

This thesis focused on the application of proteomics strategies to characterize protein changes related to immune responses in inflammation and cancer in two different disease contexts, COVID-19 and CRC. We hypothesized that proteomics characterization could allow to discover novel immune related proteins and potential biomarkers.

In SARS-CoV-2 studies, the combination of orthogonal proteomics LC-MS/MS and proximity extension assay technologies determined protein signatures common for these COVID-19 patients with comorbidities. Several novel innate immune proteins were altered as expected in viral infections with activated complement and coagulation cascade, acute-phase proteins such as α-2-antiplasmin possibly associated with post-COVID19, and anti-viral BST2. Complementary adaptive immune responses were also altered in COVID-19 patients with comorbidities, such asCD4, CD28as well as orphan receptor LILRA5, which is involved in monocyte activation. Moreover, several markers of tissue remodeling and damage were elevated in COVID-19 plasma, such as K22E, MATN2, COL6A3, and ECM1. Not previously reported,RNF41 was reduced in COVID-19, being a potential biomarker. At the same time, several protein changes were associated with antibody generation and time of infection that provided novel insights within plasma proteome changes in SARS-CoV-2 infection.

The optimized protocols were applied with the same proteomics approaches in plasma samples collected from CRC patients. This led to the characterization of altered protein levels caused by CRC development. Mainly, pro-inflammatory proteins LBP and SAA4 and cytokines such as MDK, IL6, CSF3, and CCL20, and SERPIN family members were increased in CRC. Importantly, complement cascade components were increased in CRC plasma samples. Increased C5 levels in CRC patients were validated in an additional cohort while C4B and C8A were linked to CRC progression and cancer-associated inflammation. Other proteins linked to cancer-associated inflammation including pro-inflammatory CSF3, acute-phase protein LGR1, metabolic regulator ACP6 also associated with CRC progression, and immune checkpoint CD276, suggesting different immune checkpoint patterns under cancer-associated inflammation conditions. For the first time, elevated CXCL9 and CCL23 plasma levels together with IFNG were identified and validated in an additional cohort, both chemoattractants of T cells and other immune cells but not well-established their effect in immune responses to CRC. Noteworthy, altered proteins were enriched in Th17 differentiation, oncogenic pathways, increased apoptotic markers, and altered metabolic proteins. Taken together, CRC development causes plasma protein changes that may reflect altered cancer pathways and promote systemic inflammation that can facilitate CRC progression and metastasis [702].

Proteomics characterization of CD4+ T cell enriched CRC tissues enriched ROIs identified tumor related proteins involved in multiple hallmarks of cancer. Moreover, protein changes reflected a complex interaction between fundamental proteins of innate and adaptive immune responses with most of them elevated within CRC TME. In the TME, co-existence of pro-inflammatory signals such as S100A12, S100A8, S100A9 and immunosuppressive signals such as TGFB1, PVR, NT5E, and CD276 orchestrate CRC immune evasion. Moreover, key markers showed the cell heterogeneity within CRC TME in which CAFs secreted FGF12 and myeloid cells derived ICOSL were linked to CRC progression. Multiple immune cell types were inferred from proteomics data with reductions of NK cell CD56 marker and IgA while increased Treg, monocytes, and activated mast cells supporting CRC TME compared to normal tissue. Potential effects on T cell dysfunction by SIRT1/2 deacetylation within CRC TME emphasized the need of further studies to evaluate PTMs profiles. Metabolic reprogramming and immunosuppression by metabolic deprivation occurred simultaneously in CRC. Among related proteins, Treg/Tfol ARG1 production may be a novel immunosuppressive mechanism in CRC. Moreover, our findings suggest that the novel chemotactic MCEMP1 may be involved in migration and adhesion of CD4+ T cells in tumor tissue, especially Tregs. Further validation studies are required to confirm specific cell expression and location of these novel markers. Then, functional studies *in vitro* and *in vivo* with orthogenic mice models may lead to uncover their role in CRC.

As the observed multiple orchestrated immunosuppression in CRC TME, one of the main causes of immunotherapy failure is acquired resistances by alternatives immunosuppressive mechanisms [562]. Therefore, further studies must focus on the characterization of the TME to design successful therapies that can counteract this plethora of pro-tumorigenic processes. As complex tissue bulk proteomics cannot determine cell type origin, public scRNA data and sorted CRC T cells proteomics data were used to infer specific sample expression. However, tissue disruption losses the spatial location which is fundamental within the TME[209]. Further studies must consider the isolation of specific cells within the tissue but keeping the tissue structure and location information. Current advances in single-cell proteomics allow for single-cell isolation from tissue slides by LCM followed by high-resolution mass spectrometry analysis. Therefore, further studies must apply single-cell proteomics with spatial information. An attractive approach can be combined with targeted Imaging Mass Cytometry in consecutive slides to gain insights into the 3D spatial distribution of the proteomic landscape of the TME. This strategy would be scalable to multi-omics analysis from consecutive slides for a more comprehensive molecular characterization of the immune responses underlying in the CRC TME.

Both plasma proteomics studies determined novel protein changes and immune regulators that are associated with the corresponding pathophysiology of SARS-CoV-2 infection and CRC, respectively. Further studies with larger and more diverse cohorts will facilitate the implementation of these potential biomarkers. From the technical point of view, untargeted MS-based proteomics analysis is limited to identify low abundance proteins due to the high dynamic range of blood protein content is fully untargeted without prior selection. Meanwhile, antibody-based approaches are pre-selected and rely on the availability of the specific antibodies. Therefore, the combination of both orthogonal proteomics strategies maximized protein identification and quantification by circumventing the limitations of each other. Furthermore, recent advances in Proximity Extension Assay with over 5000 targeted proteins while MS-based proteomics combined with nanoparticle protein coronas improve the limited low-abundance protein detection being promising tools for future experiments [472].

ACKNOWLEDGEMENTS

This PhD project has been an tempestuous and exciting journey that would not have been possible without the continuous support of an amazing crew of excellent colleagues.

At first, I would like to thank my supervisor Dr. Zhi Jane Chen for giving me this opportunity to develop my scientific career and transmit me her knowledge in the intricate word of immunology. Thanks for the provided experience, for introducing me into the world of science and how to build collaborations and working interactions. I fully appreciate her fabulous supervision of this project allowing for the development of my research skills and for always being willing to discuss new ideas.

Then, I would like to thank to my colleague and friend, Dominika Miroszewska, for her continued help and support since even before we started the project together from the academic and the personal side. Dominika has been a fundamental part

of this PhD thesis with her unbreakable energy. I finish this project keeping great memories and laughs in the lab, just the perfect buffer to keep ourselves in stressful and extensive protocols. Thanks for her patience with me, my "Spanglish", and nonending bureaucracy that she helped me to go through.

From the IFB mass spectrometry facility, I would like to thanks to Dr. Paulina Czaplewska for introducing to the exciting filed of proteomics and mass spectrometry. From the first week of my thesis, Paulina has been always there to answer my multiple questions and doubts. She allowed me to accommodate in her laboratory as if I was one of her students and taught all the aspects of the technology from the sample preparation to the informatics analysis. Thanks for giving me access to this revolutionizing technology and all the tools including software and instrumentation to carry my experiments out. I would like to thanks to Dr. Katarzyna Macur also for sharing with me her expertise and her patience to answer all my questions and solve multiple technical issues along the way. From the colleagues of proteomics lab, I would like to thanks to Natalia, Inez, Marc, and Michał. All of them have been always welcome and helpful when something went wrong with the mass spectrometer, that unfortunately, it is more frequent than we want. Specially, I would like to thanks to Natalia as becoming the major responsible of the lab and the organization within us to share instruments and software. Also, thanks to Marc for the inspiring conversations and great moments in the lab and outside.

Then, I would like to thanks to Dr. Agnieszka Jabłońska, although she left before the end of the project, she has been a fundamental part thanks to her expertise and collaboration skills, she achieved to establish collaborations with clinicians and biobanks to have the required patient samples that would make not possible to bring the project ahead. Also, thanks for her support in the lab always willing to order anything needed even much before I could run out. Thanks for her advices for my ongoing career and potential courses to develop my skills.

Also, I would like to thanks to Agnieszka Sankiewicz, she was always available to help me to find the correct bureaucratic procedures and our nice chats in the office to chill a bit between experiments.

I would like to thanks to Dr. Peter Grešner for helping me to understand statistical models when I was struggling, He provided me the basis to build my current bioinformatics skills.

Also, I would like to thanks to all the colleagues from Oulu, they made me feel like in my lab, since the first time I went for collaboration visits.

I would like to thanks the irreplaceable group of clinicians and their hard work to provide the required patient samples for this project. Of course, I would like to thanks to all the participants and patients involved in this project. This thesis is from them and for them, hopefully, this thesis together with further projects will finally contribute to improving their wellbeing.

Finally, I would like to thank to my family and girlfriend, they are my soul and love, the reason that I continue this way and my will to develop as a person and as a scientist.

REFERENCES

1. Abul Abbas, Andrew Lichtman, & Shiv Pillai. *Cellular and Molecular Immunology*. (Elsevier, 2021).
2. da Gama Duarte, J., Woods, K., Andrews, M. C. & Behren, A. The good, the (not so) bad and the ugly of immune homeostasis in melanoma. *Immunology & Cell Biology* **96**, 497–506 (2018).
3. Huntington, N. D. & Gray, D. H. Immune homeostasis in health and disease. *Immunology & Cell Biology* **96**, 451–452 (2018).
4. Hannoodee, S. & Nasuruddin, D. N. Acute Inflammatory Response. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2024).
5. Tang, D., Kang, R., Coyne, C. B., Zeh, H. J. & Lotze, M. T. PAMPs and DAMPs: signal 0s that spur autophagy and immunity. *Immunol Rev* **249**, 158–175 (2012).
6. Noris, M. & Remuzzi, G. Overview of Complement Activation and Regulation. *Seminars in Nephrology* **33**, 479–492 (2013).
7. Alberts, B. *et al.* Innate Immunity. in *Molecular Biology of the Cell. 4th edition* (Garland Science, 2002).
8. Acute Inflammation - Creative Diagnostics. https://www.creative-diagnostics.com/acute-inflammation.htm.
9. Girardi, G., Lingo, J. J., Fleming, S. D. & Regal, J. F. Essential Role of Complement in Pregnancy: From Implantation to Parturition and Beyond. *Front. Immunol.* **11**, (2020).
10. Stone, K. D., Prussin, C. & Metcalfe, D. D. IgE, mast cells, basophils, and eosinophils. *J Allergy Clin Immunol* **125**, S73-80 (2010).
11. Aristizábal, B. & González, Á. Innate immune system. in *Autoimmunity: From Bench to Bedside [Internet]* (El Rosario University Press, 2013).
12. Zhang, N. & Bevan, M. J. CD8(+) T cells: foot soldiers of the immune system. *Immunity* **35**, 161–168 (2011).
13. Luckheeram, R. V., Zhou, R., Verma, A. D. & Xia, B. CD4+T Cells: Differentiation and Functions. *Clin Dev Immunol* **2012**, 925135 (2012).
14. Wyatt, R. C., Lanzoni, G., Russell, M. A., Gerling, I. & Richardson, S. J. What the HLA-I!—Classical and Non-classical HLA Class I and Their Potential Roles in Type 1 Diabetes. *Curr Diab Rep* **19**, 159 (2019).
15. Cruz-Tapias, P., Castiblanco, J. & Anaya, J.-M. Major histocompatibility complex: Antigen processing and presentation. in *Autoimmunity: From Bench to Bedside [Internet]* (El Rosario University Press, 2013).
16. Jr, C. A. J. *et al. Immunobiology*. (Garland Science, 2001).
17. Thome, M., Charton, J. E., Pelzer, C. & Hailfinger, S. Antigen Receptor Signaling to NF-κB via CARMA1, BCL10, and MALT1. *Cold Spring Harb Perspect Biol* **2**, a003004 (2010).
18. Shah, K., Al-Haidari, A., Sun, J. & Kazi, J. U. T cell receptor (TCR) signaling in health and disease. *Sig Transduct Target Ther* **6**, 1–26 (2021).
19. Veillette, A. SLAM-Family Receptors: Immune Regulators with or without SAP-Family Adaptors. *Cold Spring Harb Perspect Biol* **2**, a002469 (2010).
20. Pearce, E. L. Metabolism in T cell activation and differentiation. *Curr Opin Immunol* **22**, 314–320 (2010).
21. Peng, S. & Bao, Y. A narrative review of immune checkpoint mechanisms and current immune checkpoint therapy. *Annals of Blood* **7**, (2022).
22. Lauterbach, N., Wieten, L., Popeijus, H. E., Voorter, C. E. M. & Tilanus, M. G. J. HLA-E regulates NKG2C+ natural killer cell function through presentation of a restricted peptide repertoire. *Hum Immunol* **76**, 578–586 (2015).
23. Dieckmann, N. M. G., Frazer, G. L., Asano, Y., Stinchcombe, J. C. & Griffiths, G. M. The cytotoxic T lymphocyte immune synapse at a glance. *J Cell Sci* **129**, 2881–2886 (2016).
24. Geginat, J. *et al.* The CD4-centered universe of human T cell subsets. *Seminars in Immunology* **25**, 252–262 (2013).
25. Crotty, S. T follicular helper cell differentiation, function, and roles in disease. *Immunity* **41**, 529–542 (2014).
26. Shevyrev, D. & Tereshchenko, V. Treg Heterogeneity, Function, and Homeostasis. *Front. Immunol.* **10**, (2020).
27. Lu, Y. & Craft, J. T Follicular Regulatory Cells: Choreographers of Productive Germinal Center Responses. *Front. Immunol.* **12**, (2021).
28. Geginat, J. *et al.* Plasticity of Human CD4 T Cell Subsets. *Front. Immunol.* **5**, (2014).
29. Kleinewietfeld, M. & Hafler, D. A. The plasticity of human Treg and Th17 cells and its role in autoimmunity. *Semin Immunol* **25**, 305–312 (2013).
30. Muroyama, Y. & Wherry, E. J. Memory T-Cell Heterogeneity and Terminology. *Cold Spring Harb Perspect Biol* **13**, a037929 (2021).
31. Nielsen, M. M., Witherden, D. A. & Havran, W. L. γδ T cells in homeostasis and host defence of epithelial barrier tissues. *Nat Rev Immunol* **17**, 733–745 (2017).
32. Wu, L. & Van Kaer, L. Natural killer T cells in health and disease. *Front Biosci (Schol Ed)* **3**, 236–251 (2011).
33. Godfrey, D. I., Uldrich, A. P., McCluskey, J., Rossjohn, J. & Moody, D. B. The burgeoning family of unconventional T cells. *Nat Immunol* **16**, 1114–1123 (2015).
34. Cyster, J. G. & Allen, C. D. C. B Cell Responses: Cell Interaction Dynamics and Decisions. *Cell* **177**, 524–540 (2019).
35. Harwood, N. E. & Batista, F. D. Early events in B cell activation. *Annu Rev Immunol* **28**, 185–210 (2010).
36. Bannard, O. & Cyster, J. G. Germinal centers: programmed for affinity maturation and antibody diversification. *Curr Opin Immunol* **45**, 21–30 (2017).

37. Allman, D. & Pillai, S. Peripheral B cell subsets. *Curr Opin Immunol* **20**, 149–157 (2008).
38. Heyman, B. Antibody Feedback Inhibition – a Biological Principle of Immune Regulation. *Transfusion Medicine and Hemotherapy* **32**, 348–354 (2005).
39. Justiz Vaillant, A. A. & Qurie, A. Immunodeficiency. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2024).
40. Broide, D. H. Molecular and cellular mechanisms of allergic disease. *Journal of Allergy and Clinical Immunology* **108**, S65–S71 (2001).
41. Al-Ma'amouri, M. Y. A Review Article: Hypersensitivity and its Disorders. *Journal for Research in Applied Sciences and Biotechnology* **2**, 168–172 (2023).
42. Manzo, G. COVID-19 as an Immune Complex Hypersensitivity in Antigen Excess Conditions: Theoretical Pathogenetic Process and Suggestions for Potential Therapeutic Interventions. *Front. Immunol.* **11**, (2020).
43. Pahwa, R., Goyal, A. & Jialal, I. Chronic Inflammation. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2024).
44. Singh, N. *et al.* Inflammation and Cancer. *Ann Afr Med* **18**, 121–126 (2019).
45. Ejaz, H. *et al.* COVID-19 and comorbidities: Deleterious impact on infected patients. *Journal of Infection and Public Health* **13**, 1833–1839 (2020).
46. Upadhyay, A. Cancer: An unknown territory; rethinking before going ahead. *Genes Dis* **8**, 655–661 (2020).
47. Zarour, H. M., DeLeo, A., Finn, O. J. & Storkus, W. J. Categories of Tumor Antigens. in *Holland-Frei Cancer Medicine. 6th edition* (BC Decker, 2003).
48. Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer–immune set point. *Nature 2017 541:7637* **541**, 321–330 (2017).
49. Marabelle, A. & Gray, J. Tumor-targeted and immune-targeted monoclonal antibodies: Going from passive to active immunotherapy. *Pediatric Blood & Cancer* **62**, 1317–1325 (2015).
50. Nicolò, E. *et al.* Combining antibody-drug conjugates with immunotherapy in solid tumors: current landscape and future perspectives. *Cancer Treatment Reviews* **106**, 102395 (2022).
51. Dahlén, E., Veitonmäki, N. & Norlén, P. Bispecific antibodies in cancer immunotherapy. *Therapeutic Advances in Vaccines and Immunotherapy* **6**, 3–17 (2018).
52. Jenkins, R. W., Barbie, D. A. & Flaherty, K. T. Mechanisms of resistance to immune checkpoint inhibitors. *British Journal of Cancer 2018 118:1* **118**, 9–16 (2018).
53. Waldman, A. D., Fritz, J. M. & Lenardo, M. J. A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nat Rev Immunol* **20**, 651–668 (2020).
54. Kuipers, E. J. *et al.* Colorectal cancer. *Nature Reviews Disease Primers* **1**, 15065 (2015).
55. Pan, J., Xin, L., Ma, Y.-F., Hu, L.-H. & Li, Z.-S. Colonoscopy Reduces Colorectal Cancer Incidence and Mortality in Patients With Non-Malignant Findings: A Meta-Analysis. *Am J Gastroenterol* **111**, 355–365 (2016).
56. Ciombor, K. K. & Bekaii-Saab, T. A Comprehensive Review of Sequencing and Combination Strategies of Targeted Agents in Metastatic Colorectal Cancer. *The Oncologist* **23**, 25–34 (2018).
57. Leslie, A., Carey, F. A., Pratt, N. R. & Steele, R. J. C. The colorectal adenoma–carcinoma sequence. *BJS (British Journal of Surgery)* **89**, 845–860 (2002).
58. Fleming, M., Ravula, S., Tatishchev, S. F. & Wang, H. L. Colorectal carcinoma: Pathologic aspects. *J Gastrointest Oncol* **3**, 153–173 (2012).
59. Bogaert, J. & Prenen, H. Molecular genetics of colorectal cancer. *Annals of Gastroenterology* vol. 27 9–14 Preprint at https://doi.org/10.1146/annurev-pathol-011110-130235 (2014).
60. Colussi, D., Brandi, G., Bazzoli, F. & Ricciardiello, L. Molecular pathways involved in colorectal cancer: implications for disease behavior and prevention. *International journal of molecular sciences* vol. 14 16365–16385 Preprint at https://doi.org/10.3390/ijms140816365 (2013).
61. Vasen, H. F. A., Tomlinson, I. & Castells, A. Clinical management of hereditary colorectal cancer syndromes. *Nature Reviews Gastroenterology and Hepatology* vol. 12 88–97 Preprint at https://doi.org/10.1038/nrgastro.2014.229 (2015).
62. Rejali, L. *et al.* Principles of Molecular Utility for CMS Classification in Colorectal Cancer Management. *Cancers* **15**, (2023).
63. Guinney, J. *et al.* The Consensus Molecular Subtypes of Colorectal Cancer. *Nat Med* **21**, 1350–1356 (2015).
64. Nadeem, M. S., Kumar, V., Al-Abbasi, F. A., Kamal, M. A. & Anwar, F. Risk of colorectal cancer in inflammatory bowel diseases. *Seminars in Cancer Biology* **64**, 51–60 (2020).
65. Zhang, Y.-Z. & Li, Y.-Y. Inflammatory bowel disease: Pathogenesis. *World J Gastroenterol* **20**, 91–99 (2014).
66. Lee, S. H., Kwon, J. eun & Cho, M.-L. Immunological pathogenesis of inflammatory bowel disease. *Intest Res* **16**, 26–42 (2018).
67. Du, L. & Ha, C. Epidemiology and Pathogenesis of Ulcerative Colitis. *Gastroenterology Clinics* **49**, 643–654 (2020).
68. Kurokawa, K., Hayakawa, Y. & Koike, K. Plasticity of Intestinal Epithelium: Stem Cell Niches and Regulatory Signals. *International Journal of Molecular Sciences* **22**, 357 (2021).
69. Fantini, M. C. & Guadagni, I. From inflammation to colitis-associated colorectal cancer in inflammatory bowel disease: Pathogenesis and impact of current therapies. *Digestive and Liver Disease* **53**, 558–565 (2021).
70. Lucafò, M., Curci, D., Franzin, M., Decorti, G. & Stocco, G. Inflammatory Bowel Disease and Risk of Colorectal Cancer: An Overview From Pathophysiology to Pharmacological Prevention. *Front. Pharmacol.* **12**, (2021).

71. Razi, S., Baradaran Noveiry, B., Keshavarz-Fathi, M. & Rezaei, N. IL17 and colorectal cancer: From carcinogenesis to treatment. *Cytokine* **116**, 7–12 (2019).

72. Pagès, F. *et al.* International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet* **391**, 2128–2139 (2018).

73. Lin, C.-C., Liao, T.-T. & Yang, M.-H. Immune Adaptation of Colorectal Cancer Stem Cells and Their Interaction With the Tumor Microenvironment. *Front. Oncol.* **10**, (2020).

74. Wozniakova, M., Skarda, J. & Raska, M. The Role of Tumor Microenvironment and Immune Response in Colorectal Cancer Development and Prognosis. *Pathol. Oncol. Res.* **28**, 1610502 (2022).

75. Zheng, Z. *et al.* T Cells in Colorectal Cancer: Unravelling the Function of Different T Cell Subsets in the Tumor Microenvironment. *International Journal of Molecular Sciences* **24**, 11673 (2023).

76. Xia, C., Yin, S., To, K. K. W. & Fu, L. CD39/CD73/A2AR pathway and cancer immunotherapy. *Molecular Cancer* **22**, 44 (2023).

77. Aristin Revilla, S., Kranenburg, O. & Coffer, P. J. Colorectal Cancer-Infiltrating Regulatory T Cells: Functional Heterogeneity, Metabolic Adaptation, and Therapeutic Targeting. *Front. Immunol.* **13**, (2022).

78. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians* **71**, 209–249 (2021).

79. Wu, S., Zhu, W., Thompson, P. & Hannun, Y. A. Evaluating intrinsic and non-intrinsic cancer risk factors. *Nature Communications 2018 9:1* **9**, 1–12 (2018).

80. Colotta, F., Allavena, P., Sica, A., Garlanda, C. & Mantovani, A. Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability. *Carcinogenesis* **30**, 1073–1081 (2009).

81. Mantovani, A., Allavena, P., Sica, A. & Balkwill, F. Cancer-related inflammation. *Nature* **454**, 436–444 (2008).

82. Sun, B. & Zhang, Y. Overview of orchestration of CD4+ T cell subsets in immune responses. in *Advances in Experimental Medicine and Biology* vol. 841 1–13 (Springer New York LLC, 2014).

83. Castro, F., Cardoso, A. P., Gonçalves, R. M., Serre, K. & Oliveira, M. J. Interferon-Gamma at the Crossroads of Tumor Immune Surveillance or Evasion. *Frontiers in Immunology* **0**, 847 (2018).

84. Walker, J. A. & McKenzie, A. N. J. TH2 cell development and function. *Nature Reviews Immunology* **18**, 121–133 (2018).

85. Ellyard, J. I., Simson, L. & Parish, C. R. Th2-mediated anti-tumour immunity: friend or foe? *Tissue Antigens* **70**, 1–11 (2007).

86. Knochelmann, H. M. *et al.* When worlds collide: Th17 and Treg cells in cancer and autoimmunity. *Cellular & Molecular Immunology 2018 15:5* **15**, 458–469 (2018).

87. Zhao, J., Chen, X., Herjan, T. & Li, X. The role of interleukin-17 in tumor development and progression. *Journal of Experimental Medicine* **217**, (2020).

88. Bailey, S. R. *et al.* Th17 Cells in Cancer: The Ultimate Identity Crisis. *Frontiers in Immunology* **0**, 276 (2014).

89. Ohue, Y. & Nishikawa, H. Regulatory T (Treg) cells in cancer: Can Treg cells be a new therapeutic target? *Cancer Science* **110**, 2080 (2019).

90. Ward-Hartstonge, K. A. & Kemp, R. A. Regulatory T-cell heterogeneity and the cancer immune response. *Clinical & Translational Immunology* **6**, e154 (2017).

91. Cui, G. TH9, TH17, and TH22 Cell Subsets and Their Main Cytokine Products in the Pathogenesis of Colorectal Cancer. *Frontiers in Oncology* **9**, 1(2019).

92. Hetta, H. F. *et al.* T follicular helper and T follicular regulatory cells in colorectal cancer: A complex interplay. *Journal of Immunological Methods* 480 112753 (2020).

93. Zou, W., Wolchok, J. D. & Chen, L. PD-L1 (B7-H1) and PD-1 pathway blockade for cancer therapy: Mechanisms, response biomarkers, and combinations. *Science Translational Medicine* **8**, 328rv4-328rv4 (2016).

94. Binnewies, M. *et al.* Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nature Medicine* **24**, 541–550 (2018).

95. Chang, K. *et al.* Immune profiling of premalignant lesions in patients with Lynch syndrome. in *JAMA Oncology* vol. 4 1085–1092 (2018).

96. Yang, X. *et al.* Single-Cell Analysis Reveals Characterization of Infiltrating T Cells in Moderately Differentiated Colorectal Cancer. *Frontiers in Immunology* **11**, 1 (2021).

97. Hopkins, A. C. *et al.* T cell receptor repertoire features associated with survival in immunotherapy-treated pancreatic ductal adenocarcinoma. *JCI Insight* **3**, (2018).

98. Sharma, A. *et al.* Non-Genetic Intra-Tumor Heterogeneity Is a Major Predictor of Phenotypic Heterogeneity and Ongoing Evolutionary Dynamics in Lung Tumors. *Cell Reports* **29**, 2164-2174.e5 (2019).

99. Mohammad, I. *et al.* Quantitative proteomic characterization and comparison of T helper 17 and induced regulatory T cells. *PLoS Biology* **16**, e2004194 (2018).

100. Cenik, C. *et al.* Integrative analysis of RNA, translation, and protein levels reveals distinct regulatory variation across humans. *Genome Research* **25**, 1610–1621 (2015).

101. Ludvigsen, M. *et al.* Proteomic characterization of colorectal cancer cells versus normal-derived colon mucosa cells: Approaching identification of novel diagnostic protein biomarkers in colorectal cancer. *International Journal of Molecular Sciences* **21**, (2020).

102. Macklin, A., Khan, S. & Kislinger, T. Recent advances in mass spectrometry based clinical proteomics: Applications to cancer research. *Clinical Proteomics* **17**, 1–25 (2020).

103. Carbonara, K., Andonovski, M. & Coorssen, J. R. Proteomes Are of Proteoforms: Embracing the Complexity. *Proteomes 2021, Vol. 9, Page 38* **9**, 38 (2021).

104. Shah, T. R. & Misra, A. Proteomics. *Challenges in Delivery of Therapeutic Genomics and Proteomics* 387–427 (2011).

105. Verrills, N. M. Clinical Proteomics: Present and Future Prospects. *Clinical Biochemist Reviews* **27**, 99 (2006).

106. Scurrah, C. R., Simmons, A. J. & Lau, K. S. Single-cell mass cytometry of archived human epithelial tissue for decoding cancer signaling pathways. *Methods in molecular biology (Clifton, N.J.)* **1884**, 215 (2019).

107. Simmons, A. J. *et al.* Impaired coordination between signaling pathways is revealed in human colorectal cancer using single-cell mass cytometry of archival tissue blocks. *Science signaling* **9**, rs11 (2016).

108. Gerdes, M. J. *et al.* Highly multiplexed single-cell analysis of formalinfixed, paraffin-embedded cancer tissue. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 11982–11987 (2013).

109. Gundry, R. L. *et al.* Preparation of Proteins and Peptides for Mass Spectrometry Analysis in a Bottom-Up Proteomics Workflow. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]* **CHAPTER**, Unit10.25 (2009).

110. Wiśniewski, J. R. Filter Aided Sample Preparation – A tutorial. *Analytica Chimica Acta* **1090**, 23–30 (2019).

111. Ly, L. & Wasinger, V. C. Protein and peptide fractionation, enrichment and depletion: Tools for the complex proteome. *PROTEOMICS* **11**, 513–534 (2011).

112. Wichmann, C. *et al.* MaxQuant.Live Enables Global Targeting of More Than 25,000 Peptides. *Molecular & Cellular Proteomics : MCP* **18**, 982 (2019).

113. Hao, J.-J. *et al.* Comprehensive Proteomic Characterization of the Human Colorectal Carcinoma Reveals Signature Proteins and Perturbed Pathways. *Scientific Reports 2017 7:1* **7**, 1–15 (2017).

114. Martins, B. A. A. *et al.* Biomarkers in Colorectal Cancer: The Role of Translational Proteomics Research. *Frontiers in Oncology* **9**, 1284 (2019).

115. MacMullan, M. A., Dunn, Z. S., Graham, N. A., Yang, L. & Wang, P. Quantitative Proteomics and Metabolomics Reveal Biomarkers of Disease as Potential Immunotherapy Targets and Indicators of Therapeutic Efficacy. *Theranostics* **9**, 7872 (2019).

116. Stoeckli, M., Chaurand, P., Hallahan, D. E. & Caprioli, R. M. Imaging mass spectrometry: A new technology for the analysis of protein expression in mammalian tissues. *Nature Medicine 2001 7:4* **7**, 493–496 (2001).

117. Aichler, M. & Walch, A. MALDI Imaging mass spectrometry: current frontiers and perspectives in pathology research and practice. *Laboratory Investigation 2015 95:4* **95**, 422–431 (2015).

118. Kazdal, D. *et al.* Digital PCR After MALDI–Mass Spectrometry Imaging to Combine Proteomic Mapping and Identification of Activating Mutations in Pulmonary Adenocarcinoma. *PROTEOMICS – Clinical Applications* **13**, 1800034 (2019).

119. Erich, K. *et al.* Spatial distribution of endogenous tissue protease activity in gastric carcinoma mapped by MALDI mass spectrometry imaging. *Molecular and Cellular Proteomics* **18**, 151–161 (2019).

120. Powers, T. W. *et al.* MALDI Imaging Mass Spectrometry Profiling of N-Glycans in Formalin-Fixed Paraffin Embedded Clinical Tissue Blocks and Tissue Microarrays. *PLOS ONE* **9**, e106255 (2014).

121. Paine, M. R. L. *et al.* Three-Dimensional Mass Spectrometry Imaging Identifies Lipid Markers of Medulloblastoma Metastasis. *Scientific Reports 2019 9:1* **9**, 1–10 (2019).

122. Lou, S., Balluff, B., Cleven, A. H. G., Bovée, J. V. M. G. & McDonnell, L. A. Prognostic Metabolite Biomarkers for Soft Tissue Sarcomas Discovered by Mass Spectrometry Imaging. *Journal of the American Society for Mass Spectrometry* **28**, 376–383 (2017).

123. Taylor, M. J., Lukowski, J. K. & Anderton, C. R. Spatially Resolved Mass Spectrometry at the Single Cell: Recent Innovations in Proteomics and Metabolomics. *Journal of the American Society for Mass Spectrometry* **32**, 872–894 (2021).

124. Ahmed, M. *et al.* Next-generation protein analysis in the pathology department. *Journal of Clinical Pathology* **73**, 1–6 (2020).

125. Timp, W. & Timp, G. Beyond mass spectrometry, the next step in proteomics. *Science Advances* **6**, (2020).

126. Steen, H. & Mann, M. The abc's (and xyz's) of peptide sequencing. *Nature Reviews Molecular Cell Biology 2004 5:9* **5**, 699–711 (2004).

127. Baumeister, S. H., Freeman, G. J., Dranoff, G. & Sharpe, A. H. Coinhibitory Pathways in Immunotherapy for Cancer. *Annurev Immunol.***34**, 539–573 (2016).

128. Allard, B. *et al.* Immuno-oncology-101: overview of major concepts and translational perspectives. *Seminars in Cancer Biology* **52**, 1–11 (2018).

129. Munn, D. H. & Bronte, V. Immune suppressive mechanisms in the tumor microenvironment. *Current Opinion in Immunology* **39**, 1–6 (2016).

130. Chang, C. H. *et al.* Metabolic Competition in the Tumor Microenvironment Is a Driver of Cancer Progression. *Cell* **162**, 1229–1241 (2015).

131. Wang, K. *et al.* SIRT5 Contributes to Colorectal Cancer Growth by Regulating T Cell Activity. *Journal of Immunology Research* **2020**, 3792409 (2020).

132. Hamaidi, I. *et al.* Sirt2 Inhibition Enhances Metabolic Fitness and Effector Functions of Tumor-Reactive T Cells. *Cell Metabolism* **32**, 420-436.e12 (2020).

133. Cheng, F. *et al.* SIRT1 promotes epithelial–mesenchymal transition and metastasis in colorectal cancer by regulating Fra-1 expression. *Cancer Letters* **375**, 274–283 (2016).

134. Shi, L. *et al.* SIRT5-mediated deacetylation of LDHB promotes autophagy and tumorigenesis in colorectal cancer. *Molecular Oncology* **13**, 358–375 (2019).

135. Wang, B. *et al.* SIRT2-dependent IDH1 deacetylation inhibits colorectal cancer and liver metastases. *EMBO reports* **21**, e48183 (2020).

136. Martí i Líndez, A. A. *et al.* Mitochondrial arginase-2 is a cell-autonomous regulator of CD8+ T cell function and antitumor efficacy. *JCI Insight* **4**, (2019).

137. Xue, T. *et al.* Interleukin-6 induced 'Acute' phenotypic microenvironment promotes Th1 anti-tumor immunity in cryo-thermal therapy revealed by shotgun and parallel reaction monitoring proteomics. *Theranostics* **6**, 773–794 (2016).

138. Servais, L. *et al.* Platelets contribute to the initiation of colitis-associated cancer by promoting immunosuppression. *Journal of Thrombosis and Haemostasis* **16**, 762–777 (2018).

139. Cianciaruso, C. *et al.* Molecular Profiling and Functional Analysis of Macrophage-Derived Tumor Extracellular Vesicles. *Cell Reports* **27**, 3062-3080.e11 (2019).

140. Xu, R. *et al.* Spatial-Resolution Cell Type Proteome Profiling of Cancer Tissue by Fully Integrated Proteomics Technology. *Analytical Chemistry* **90**, 5879–5886 (2018).

141. Maria Bosisio, F. *et al.* Functional heterogeneity of lymphocytic patterns in primary melanoma dissected through single-cell multiplexing. *eLife* **9**, (2020).

142. Stacker, S. A. *et al.* Lymphangiogenesis and lymphatic vessel remodelling in cancer. *Nature Reviews Cancer 2014 14:3* **14**, 159–172 (2014).

143. Dieterich, L. C. *et al.* Tumor-associated lymphatic vessels upregulate PDL1 to inhibit T-cell activation. *Frontiers in Immunology* **8**, 66 (2017).

144. Stevenson, J. *et al.* Proteomics of REPLICANT perfusate detects changes in the metastatic lymph node microenvironment. *npj Breast Cancer* **7**, 1–13 (2021).

145. Müller, A. K. *et al.* Proteomic Characterization of Prostate Cancer to Distinguish Nonmetastasizing and Metastasizing Primary Tumors and Lymph Node Metastases. *Neoplasia (United States)* **20**, 140–151 (2018).

146. Naidoo, K. *et al.* Proteome of formalin-fixed paraffin-embedded pancreatic ductal adenocarcinoma and lymph node metastases. *Journal of Pathology* **226**, 756–763 (2012).

147. Krantz, D. *et al.* IL16 processing in sentinel node regulatory T cells is a factor in bladder cancer immunity. *Scandinavian Journal of Immunology* **92**, e12926 (2020).

148. Kriegsmann, M. *et al.* Reliable entity subtyping in non-small cell lung cancer by matrix-assisted laser desorption/ionization imaging mass spectrometry on formalin-fixed paraffinembedded tissue specimens. *Molecular and Cellular Proteomics* **15**, 3081–3089 (2016).

149. Casadonte, R. *et al.* Development of a Class Prediction Model to Discriminate Pancreatic Ductal Adenocarcinoma from Pancreatic Neuroendocrine Tumor by MALDI Mass Spectrometry Imaging. *PROTEOMICS – Clinical Applications* **13**, 1800046 (2019).

150. Phillips, L., Gill, A. J. & Baxter, R. C. Novel prognostic markers in triple-negative breast cancer discovered by MALDI-mass spectrometry imaging. *Frontiers in Oncology* **9**, 379 (2019).

151. Meding, S. *et al.* Tumor classification of six common cancer types based on proteomic profiling by MALDI imaging. *Journal of Proteome Research* **11**, 1996–2003 (2012).

152. Gawin, M. *et al.* Intra-tumor heterogeneity revealed by mass spectrometry imaging is associated with the prognosis of breast cancer. *Cancers* **13**, 4349 (2021).

153. Berghmans, E. *et al.* Mass Spectrometry Imaging Reveals Neutrophil Defensins as Additional Biomarkers for Anti-PD-(L)1 Immunotherapy Response in NSCLC Patients. *Cancers 2020, Vol. 12, Page 863* **12**, 863 (2020).

154. Alberts, D. *et al.* MALDI Imaging-Guided Microproteomic Analyses of Heterogeneous Breast Tumors—A Pilot Study. *PROTEOMICS – Clinical Applications* **12**, 1700062 (2018).

155. Berghmans, E. *et al.* MALDI Mass Spectrometry Imaging Linked with Top-Down Proteomics as a Tool to Study the Non-Small-Cell Lung Cancer Tumor Microenvironment. *Methods and Protocols 2019, Vol. 2, Page 44* **2**, 44 (2019).

156. Davalieva, K. *et al.* Comparative Proteomics Analysis of Urine Reveals Down-Regulation of Acute Phase Response Signaling and LXR/RXR Activation Pathways in Prostate Cancer. *Proteomes 2018, Vol. 6, Page 1* **6**, 1 (2017).

157. Massimi, L., Martelli, C., Caldarelli, M., Castagnola, M. & Desiderio, C. Proteomics in pediatric cystic craniopharyngioma. *Brain Pathology* **27**, 370–376 (2017).

158. Martelli, C. *et al.* Investigating the protein signature of adamantinomatous craniopharyngioma pediatric brain tumor tissue: Towards the comprehension of its aggressive behavior. *Disease Markers* **2019**, (2019).

159. Rossetti, D. V. *et al.* Ependymoma Pediatric Brain Tumor Protein Fingerprinting by Integrated Mass Spectrometry Platforms: A Pilot Investigation. *Cancers 2020, Vol. 12, Page 674* **12**, 674 (2020).

160. Ntai, I. *et al.* Integrated Bottom-Up and Top-Down Proteomics of Patient-Derived Breast Tumor Xenografts. *Molecular & Cellular Proteomics : MCP* **15**, 45 (2016).

161. Skinner, O. S. *et al.* Top-down characterization of endogenous protein complexes with native proteomics. *Nature Chemical Biology 2017 14:1* **14**, 36–41 (2017).

162. Deighan, W. I. *et al.* Development of novel methods for non-canonical myeloma protein analysis with an innovative adaptation of immunofixation electrophoresis, native top-down mass spectrometry, and middle-down de novo sequencing. *Clinical Chemistry and Laboratory Medicine* **59**, 653–661 (2021).

163. Ntai, I. *et al.* Precise characterization of KRAS4b proteoforms in human colorectal cells and tumors reveals mutation/modification cross-talk. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 4140–4145 (2018).

164. Resing, K. A. & Ahn, N. G. Proteomics strategies for protein identification. *FEBS Letters* **579**, 885–889 (2005).

165. Griss, J. *et al.* Recognizing millions of consistently unidentified spectra across hundreds of shotgun proteomics datasets. *Nature Methods 2016 13:8* **13**, 651–656 (2016).

166. Geyer, P. E., Holdt, L. M., Teupser, D. & Mann, M. Revisiting biomarker discovery by plasma proteomics. *Molecular Systems Biology* **13**, 942 (2017).

167. Jungblut, P. R., Thiede, B. & Schlüter, H. Towards deciphering proteomes via the proteoform, protein speciation, moonlighting and protein code concepts. *Journal of Proteomics* **134**, 1–4 (2016).

168. Zhang, Y., Fonslow, B. R., Shan, B., Baek, M. C. & Yates, J. R. Protein Analysis by Shotgun/Bottom-up Proteomics. *Chemical Reviews* **113**, 2343–2394 (2013).

169. Kim, M. S., Zhong, J. & Pandey, A. Common errors in mass spectrometry-based analysis of post-translational modifications. *Proteomics* **16**, 700 (2016).

170. Cesano, A. & Warren, S. Bringing the Next Generation of Immuno-Oncology Biomarkers to the Clinic. *Biomedicines 2018, Vol. 6, Page 14* **6**, 14 (2018).

171. Zhao, C. *et al.* Identification of immune checkpoint and cytokine signatures associated with the response to immune checkpoint blockade in gastrointestinal cancers. *Cancer Immunology, Immunotherapy 2021 70:9* **70**, 2669–2679 (2021).

172. Calu, V. *et al.* Key biomarkers within the colorectal cancer related inflammatory microenvironment. *Scientific Reports 2021 11:1* **11**, 1–14 (2021).

173. Lundberg, M., Eriksson, A., Tran, B., Assarsson, E. & Fredriksson, S. Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood. *Nucleic Acids Research* **39**, e102 (2011).

174. Zhong, W. *et al.* Next generation plasma proteome profiling to monitor health and disease. *Nature Communications 2021 12:1* **12**, 1–12 (2021).

175. Erkan, E. P. *et al.* Circulating Tumor Biomarkers in Meningiomas Reveal a Signature of Equilibrium Between Tumor Growth and Immune Modulation. *Frontiers in Oncology* **0**, 1031 (2019).

176. Ali, A. S. *et al.* Candidate protein biomarkers in pancreatic neuroendocrine neoplasms grade 3. *Scientific Reports 2020 10:1* **10**, 1–9 (2020).

177. Kasanen, H. *et al.* Age-associated changes in the immune system may influence the response to anti-PD1 therapy in metastatic melanoma patients. *Cancer Immunology, Immunotherapy 2020 69:5* **69**, 717–730 (2020).

178. Eltahir, M. *et al.* Profiling of donor-specific immune effector signatures in response to rituximab in a human whole blood loop assay using blood from CLL patients. *International Immunopharmacology* **90**, 107226 (2021).

179. Árnadóttir, S. S. *et al.* Transcriptomic and proteomic intra-tumor heterogeneity of colorectal cancer varies depending on tumor location within the colorectum. *PLOS ONE* **15**, e0241148 (2020).

180. Wik, L. *et al.* Proximity Extension Assay in Combination with Next-Generation Sequencing for High-throughput Proteome-wide Analysis. *Molecular and Cellular Proteomics* **20**, 100168 (2021).

181. Coarfa, C. *et al.* Reverse-Phase Protein Array: Technology, Application, Data Processing, and Integration. *Journal of Biomolecular Techniques : JBT* **32**, jbt.2021-3202-001 (2021).

182. Díez, P. *et al.* NAPPA as a Real New Method for Protein Microarray Generation. *Microarrays* **4**, 214 (2015).

183. Lin, Z. *et al.* Intratumor Heterogeneity Correlates With Reduced Immune Activity and Worse Survival in Melanoma Patients. *Frontiers in Oncology* **10**, 2687 (2020).

184. González-González, M. *et al.* Tracking the antibody immunome in sporadic colorectal cancer by using antigen self-assembled protein arrays. *Cancers* **13**, 1–23 (2021).

185. Bradbury, A. & Plückthun, A. Reproducibility: Standardize antibodies used in research. *Nature* vol. 518 27–29 Preprint at https://doi.org/10.1038/518027a (2015).

186. Runa, F. *et al.* Tumor Microenvironment Heterogeneity: Challenges and Opportunities. *Current Molecular Biology Reports* **3**, 218–229 (2017).

187. Vickman, R. E. *et al.* Deconstructing tumor heterogeneity: The stromal perspective. *Oncotarget* vol. 11 3621–3632 27736 (2020).

188. Sanders, C. K. & Mourant, J. R. Advantages of full spectrum flow cytometry. *Journal of Biomedical Optics***18**, 037004 (2013).

189. Schmutz, S., Valente, M., Cumano, A. & Novault, S. Spectral Cytometry Has Unique Properties Allowing Multicolor Analysis of Cell Suspensions Isolated from Solid Tissues. *PLOS ONE* **11**, e0159961 (2016).

190. Bonilla, D. L., Reinin, G. & Chua, E. Full Spectrum Flow Cytometry as a Powerful Technology for Cancer Immunotherapy Research. *Frontiers in Molecular Biosciences* **0**, 495 (2021).

191. van de Velde, L. A. *et al.* Neuroblastoma Formation Requires Unconventional CD4 T Cells and Arginase-1–Dependent Myeloid Cells. *Cancer Research* **81**, 5047–5059 (2021).

192. Yang, M. *et al.* Checkpoint molecules coordinately restrain hyperactivated effector T cells in the tumor microenvironment. *Oncoimmunology* **9**, (2020).

193. Di, J. *et al.* Phenotype molding of T cells in colorectal cancer by single-cell analysis. *International Journal of Cancer* **146**, 2281–2295 (2020).

194. Norton, S. E. *et al.* High-Dimensional Mass Cytometric Analysis Reveals an Increase in Effector Regulatory T Cells as a Distinguishing Feature of Colorectal Tumors. *The Journal of Immunology* **202**, 1871–1884 (2019).

195. Fu, W. *et al.* Single-Cell Atlas Reveals Complexity of the Immunosuppressive Microenvironment of Initial and Recurrent Glioblastoma. *Frontiers in Immunology* **0**, 835 (2020).

196. Khalsa, J. K. *et al.* Immune phenotyping of diverse syngeneic murine brain tumors identifies immunologically distinct types. *Nature Communications 2020 11:1* **11**, 1–14 (2020).

197. Thrash, E. M. *et al.* High-Throughput Mass Cytometry Staining for Immunophenotyping Clinical Samples. *STAR Protocols* **1**, 100055 (2020).

198. Friebel, E. *et al.* Single-Cell Mapping of Human Brain Cancer Reveals Tumor-Specific Instruction of Tissue-Invading Leukocytes. *Cell* **181**, 1626-1642.e20 (2020).

199. Wierz, M., Janji, B., Berchem, G., Moussay, E. & Paggetti, J. High-dimensional mass cytometry analysis revealed microenvironment complexity in chronic lymphocytic leukemia. *Oncoimmunoogy* **7**, (2018).

200. Chevrier, S. *et al.* An Immune Atlas of Clear Cell Renal Cell Carcinoma. *Cell* **169**, 736-749.e18 (2017).

201. Wagner, J. *et al.* A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. *Cell* **177**, 1330-1345.e18 (2019).

202. Ma, C. *et al.* A clinical microchip for evaluation of single immune cells reveals high functional heterogeneity in phenotypically similar T cells. *Nature Medicine* **17**, 738–743 (2011).

203. Lu, Y. *et al.* Highly multiplexed profiling of single-cell effector functions reveals deep functional heterogeneity in response to pathogenic ligands. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E607–E615 (2015).

204. Diab, A. *et al.* Progression-free survival and biomarker correlates of response with BEMPEG plus NIVO in previously untreated patients with metastatic melanoma: results from the PIVOT-02 study. *Journal for ImmunoTherapy of Cancer* **8**, A446–A446 (2020).

205. Zhao, P., Bhowmick, S., Yu, J. & Wang, J. Highly Multiplexed Single-Cell Protein Profiling with Large-Scale Convertible DNA-Antibody Barcoded Arrays. *Advanced Science* **5**, 1800672 (2018).

206. Pham, T., Tyagi, A., Wang, Y. & Guo, J. Single-cell proteomic analysis. *WIREs Mechanisms of Disease* **13**, e1503 (2021).

207. Agasti, S. S., Liong, M., Peterson, V. M., Lee, H. & Weissleder, R. Photocleavable DNA barcode-antibody conjugates allow sensitive and multiplexed protein analysis in single cells. *Journal of the American Chemical Society* **134**, 18499–18502 (2012).

208. Ullal, A. V. *et al.* Cancer cell profiling by barcoding allows multiplexed protein analysis in fine-needle aspirates. *Science Translational Medicine* **6**, 219ra9-219ra9 (2014).

209. Yuan, Y. Spatial heterogeneity in the tumor microenvironment. *Cold Spring Harbor Perspectives in Medicine* **6**, (2016).

210. Parra, E. R. *et al.* Identification of distinct immune landscapes using an automated nine-color multiplex immunofluorescence staining panel and image analysis in paraffin tumor tissues. *Scientific Reports 2021 11:1* **11**, 1–14 (2021).

211. Lundberg, E. & Borner, G. H. H. Spatial proteomics: a powerful discovery tool for cell biology. *Nature Reviews Molecular Cell Biology* vol. 20 285–302 Preprint at https://doi.org/10.1038/s41580-018-0094-y (2019).

212. Paul, I., White, C., Turcinovic, I. & Emili, A. Imaging the future: the emerging era of single-cell spatial proteomics. *The FEBS Journal* febs.15685 (2021) doi:10.1111/febs.15685.

213. Berndt, U. *et al.* Systematic High-Content Proteomic Analysis Reveals Substantial Immunologic Changes in Colorectal Cancer. *Cancer Research* **68**, 880–888 (2008).

214. Bhattacharya, S. *et al.* Toponome imaging system: In situ protein network mapping in normal and cancerous colon from the same patient reveals more than five-thousand cancer specific protein clusters and their subcellular annotation by using a three symbol code. *Journal of Proteome Research* **9**, 6112–6125 (2010).

215. Duose, D. Y., Schweller, R. M., Hittelman, W. N. & Diehl, M. R. Multiplexed and reiterative fluorescence labeling via DNA circuitry. *Bioconjugate Chemistry* **21**, 2327–2331 (2010).

216. Giedt, R. J. *et al.* Single-cell barcode analysis provides a rapid readout of cellular signaling pathways in clinical specimens. *Nature Communications* **9**, 1–10 (2018).

217. Zrazhevskiy, P. & Gao, X. Quantum dot imaging platform for single-cell molecular profiling. *Nature Communications* **4**, 1–12 (2013).

218. Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968-981.e15 (2018).

219. Schürch, C. M. *et al.* Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. *Cell* **182**, 1341-1359.e19 (2020).

220. Mondal, M., Liao, R., Xiao, L., Eno, T. & Guo, J. Highly Multiplexed Single-Cell In Situ Protein Analysis with Cleavable Fluorescent Antibodies. *Angewandte Chemie - International Edition* **56**, 2636–2639 (2017).

221. Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nature Methods* **11**, 417–422 (2014).

222. Jackson, H. W. *et al.* The single-cell pathology landscape of breast cancer. *Nature* **578**, 615–620 (2020).

223. Singh, M. *et al.* Highly Multiplexed Imaging Mass Cytometry Allows Visualization of Tumor and Immune Cell Interactions of the Tumor Microenvironment in FFPE Tissue Sections. *Blood* **130**, 2751–2751 (2017).

224. Angelo, M. *et al.* Multiplexed ion beam imaging of human breast tumors. *Nature Medicine* **20**, 436–442 (2014).

225. Hartmann, F. J. *et al.* Single-cell metabolic profiling of human cytotoxic T cells. *Nature Biotechnology* **39**, 186 (2020).

226. Baharlou, H., Canete, N. P., Cunningham, A. L., Harman, A. N. & Patrick, E. Mass Cytometry Imaging for the Study of Human Diseases—Applications and Data Analysis Strategies. *Frontiers in Immunology* **10**, 2657 (2019).

227. Maibach, F., Sadozai, H., Seyed Jafari, S. M., Hunger, R. E. & Schenk, M. Tumor-Infiltrating Lymphocytes and Their Prognostic Value in Cutaneous Melanoma. *Frontiers in Immunology* vol. 11 Preprint at https://doi.org/10.3389/fimmu.2020.02105 (2020).

228. Woolley, C. F., Hayes, M. A., Mahanti, P., Douglass Gilman, S. & Taylor, T. Theoretical limitations of quantification for noncompetitive sandwich immunoassays. *Analytical and Bioanalytical Chemistry 2015 407:28* **407**, 8605–8615 (2015).

229. Slavov, N. Single-cell protein analysis by mass spectrometry. *Current Opinion in Chemical Biology* vol. 60 1–9 Preprint at https://doi.org/10.1016/j.cbpa.2020.04.018 (2021).

230. Kelly, R. T. Single-cell Proteomics: Progress and Prospects. *Molecular and Cellular Proteomics* vol. 19 1739–1748 (2020).

231. Zhu, Y. *et al.* Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells. *Nature Communications* **9**, (2018).

232. Li, Z. Y. *et al.* Nanoliter-Scale OILAir-Droplet Chip-Based Single Cell Proteomic Analysis. *Analytical Chemistry* **90**, 5430–5438 (2018).

233. Specht, H. *et al.* Automated sample preparation for high-throughput single-cell proteomics. *bioRxiv* 399774 (2018).

234. Cong, Y. *et al.* Improved Single-Cell Proteome Coverage Using Narrow-Bore Packed NanoLC Columns and Ultrasensitive Mass Spectrometry. *Analytical Chemistry* **92**, 2665–2671 (2020).

235. Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biology* **19**, (2018).

236. Thompson, A. *et al.* Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical Chemistry* **75**, 1895–1904 (2003).

237. Specht, H. *et al.* Single-cell mass-spectrometry quantifies the emergence of macrophage heterogeneity. *bioRxiv* 665307 (2019).

238. Tsai, C. F. *et al.* An Improved Boosting to Amplify Signal with Isobaric Labeling (iBASIL) Strategy for Precise Quantitative Single-cell Proteomics. *Molecular and Cellular Proteomics* **19**, 828–838 (2020).

239. Williams, S. M. *et al.* Automated Coupling of Nanodroplet Sample Preparation with Liquid Chromatography-Mass Spectrometry for High-Throughput Single-Cell Proteomics. *Analytical Chemistry* **92**, 10588–10596 (2020).

240. Schoof, E. M. *et al.* Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nature Communications 2021 12:1* **12**, 1–15 (2021).

241. Slavov, N. Driving Single Cell Proteomics Forward with Innovation. *Journal of Proteome Research* 639 (2021).

242. de Vries, N. L., Mahfouz, A., Koning, F. & de Miranda, N. F. C. C. Unraveling the Complexity of the Cancer Microenvironment With Multidimensional Genomic and Cytometric Technologies. *Frontiers in Oncology* vol. 10 1254 Preprint at https://doi.org/10.3389/fonc.2020.01254 (2020).

243. Efremova, M., Vento-Tormo, R., Park, J.-E., Teichmann, S. A. & James, K. R. Immunology in the Era of Single-Cell Technologies. *Annual Review of Immunology* **38**, 727–757 (2020).

244. Mertins, P. *et al.* Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature 2016 534:7605* **534**, 55–62 (2016).

245. Zhu, Y. *et al.* Discovery of coding regions in the human genome by integrated proteogenomics analysis workflow. *Nature Communications 2018 9:1* **9**, 1–14 (2018).

246. Vasaikar, S. *et al.* Proteogenomic Analysis of Human Colon Cancer Reveals New Therapeutic Opportunities. *Cell* **177**, 1035-1049.e19 (2019).

247. Zhang, H. *et al.* Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. *Cell* **166**, 755–765 (2016).

248. Demaree, B. *et al.* Joint profiling of DNA and proteins in single cells to dissect genotype-phenotype associations in leukemia. *Nature Communications 2021 12:1* **12**, 1–10 (2021).

249. Frei, A. P. *et al.* Highly multiplexed simultaneous detection of RNAs and proteins in single cells. *Nature Methods* **13**, 269–275 (2016).

250. Duckworth, A. D. *et al.* Multiplexed profiling of RNA and protein expression signatures in individual cells using flow or mass cytometry. *Nature Protocols* **14**, 901–920 (2019).

251. Peterson, V. M. *et al.* Multiplexed quantification of proteins and transcripts in single cells. *Nature Biotechnology* **35**, 936–939 (2017).

252. Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nature Methods* **14**, 865–868 (2017).

253. Wu, S. Z. *et al.* A single-cell and spatially resolved atlas of human breast cancers. *Nature Genetics 2021 53:9* **53**, 1334–1347 (2021).

254. Granja, J. M. *et al.* Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia. *Nature Biotechnology* **37**, 1458–1465 (2019).

255. Chung, H. *et al.* Simultaneous single cell measurements of intranuclear proteins and gene expression. *bioRxiv* 2021.01.18.427139 (2021) doi:10.1101/2021.01.18.427139.

256. Katzenelenbogen, Y. *et al.* Coupled scRNA-Seq and Intracellular Protein Activity Reveal an Immunosuppressive Role of TREM2 in Cancer. *Cell* **182**, 872-885.e19 (2020).

257. Kearney, C. J. *et al.* SUGAR-seq enables simultaneous detection of glycans, epitopes, and the transcriptome in single cells. *Science Advances* **7**, 3610 (2021).

258. Zhang, L. *et al.* Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature* **564**, 268–272 (2018).

259. De Vries, N. L. *et al.* High-dimensional cytometric analysis of colorectal cancer reveals novel mediators of antitumour immunity. *Gut* **69**, 691–703 (2020).

260. Blundell, M. P., Sanderson, S. L. & Long, T. A. Flow Cytometry as an Important Tool in Proteomic Profiling. *Methods in Molecular Biology* **2261**, 213–227 (2021).

261. Holzner, G. *et al.* High-throughput multiparametric imaging flow cytometry: toward diffraction-limited sub-cellular detection and monitoring of sub-cellular processes. *Cell Reports* **34**, (2021).

262. Saka, S. K. *et al.* Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues. *Nature Biotechnology* **37**, 1080–1090 (2019).

263. Swaminathan, J. *et al.* Highly parallel single-molecule identification of proteins in zeptomole-scale mixtures. *Nature Biotechnology* **36**, 1076–1091 (2018).

264. Emili A, McLaughlin M, Zagorovsky K, Olsen JB, C. W. & S. S. Protein sequencing method and reagents. (2017).

265. Ouldali, H. *et al.* Electrical recognition of the twenty proteinogenic amino acids using an aerolysin nanopore. *Nature Biotechnology* vol. 38 176–181 (2020).

266. Drachman, N. *et al.* A nanopore ion source delivers single amino acid and peptide ions directly into the gas phase. *bioRxiv* 2021.08.15.456243 (2021) doi:10.1101/2021.08.15.456243.

267. Lenčo, J. *et al.* Reversed-Phase Liquid Chromatography of Peptides for Bottom-Up Proteomics: A Tutorial. *J. Proteome Res.* **21**, 2846–2892 (2022).

268. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198–207 (2003).

269. Ho, C. *et al.* Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications. *Clin Biochem Rev* **24**, 3–12 (2003).

270. Glish, G. L. & Burinsky, D. J. Hybrid mass spectrometers for tandem mass spectrometry. *J. Am. Soc. Spectrom.* **19**, 161–172 (2008).

271. Lesur, A. & Domon, B. Advances in high-resolution accurate mass spectrometry application to targeted proteomics. *PROTEOMICS* **15**, 880–890 (2015).

272. Mellon, F. A. MASS SPECTROMETRY | Principles and Instrumentation. in *Encyclopedia of Food Sciences and Nutrition (Second Edition)* (ed. Caballero, B.) 3739–3749 (Academic Press, Oxford, 2003).

273. Niessen, W. M. A. MS–MS and MS*n*. in *Encyclopedia of Spectroscopy and Spectrometry (Second Edition)* (ed. Lindon, J. C.) 1675–1681 (Academic Press, Oxford, 1999).

274. Michalski, A. *et al.* Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer *. *Molecular & Cellular Proteomics* **10**, (2011).

275. Zubarev, R. A. & Makarov, A. Orbitrap Mass Spectrometry. *Anal. Chem.* **85**, 5288–5296 (2013).

276. Senko, M. W. *et al.* Novel Parallelized Quadrupole/Linear Ion Trap/Orbitrap Tribrid Mass Spectrometer Improving Proteome Coverage and Peptide Identification Rates. *Anal. Chem.* **85**, 11710–11714 (2013).

277. Lin, L. *et al.* Types, principle, and characteristics of tandem high-resolution mass spectrometry and its applications. *RSC Adv.* **5**, 107623–107636 (2015).

278. Li, J., Smith, L. & Zhu, H.-J. Data-independent acquisition (DIA): an emerging proteomics technology for analysis of drug-metabolizing enzymes and transporters. *Drug Discov Today Technol* **39**, 49–56 (2021).

279. Ankney, J. A., Muneer, A. & Chen, X. Relative and Absolute Quantitation in Mass Spectrometry–Based Proteomics. *Annual Review of Analytical Chemistry* **11**, 49–77 (2018).

280. Ting, Y. S. *et al.* Peptide-Centric Proteome Analysis: An Alternative Strategy for the Analysis of Tandem Mass Spectrometry Data *. *Molecular & Cellular Proteomics* **14**, 2301–2307 (2015).

281. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367–1372 (2008).

282. Eng, J. K., Fischer, B., Grossmann, J. & Maccoss, M. J. A fast SEQUEST cross correlation algorithm. *J Proteome Res* **7**, 4598–4602 (2008).

283. Moosa, J. M., Guan, S., Moran, M. F. & Ma, B. Repeat-Preserving Decoy Database for False Discovery Rate Estimation in Peptide Identification. *J. Proteome Res.* **19**, 1029–1036 (2020).

284. Gillet, L. C., Leitner, A. & Aebersold, R. Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. *Annual Review of Analytical Chemistry* **9**, 449–472 (2016).

285. Yu, F. *et al.* Analysis of DIA proteomics data using MSFragger-DIA and FragPipe computational platform. *Nat Commun* **14**, 4154 (2023).

286. Willems, P., Fels, U., Staes, A., Gevaert, K. & Van Damme, P. Use of Hybrid Data-Dependent and -Independent Acquisition Spectral Libraries Empowers Dual-Proteome Profiling. *J. Proteome Res.* **20**, 1165–1177 (2021).

287. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* **17**, 41–44 (2020).

288. Qu, X., Tang, Y. & Hua, S. Immunological Approaches Towards Cancer and Inflammation: A Cross Talk. *Front Immunol* **9**, 563 (2018).

289. Husi, H. & Albalat, A. Chapter 9 - Proteomics. in *Handbook of Pharmacogenomics and Stratified Medicine* (ed. Padmanabhan, S.) 147–179 (Academic Press, San Diego, 2014). doi:10.1016/B978-0-12-386882-4.00009-8.

290. Parasher, A. COVID-19: Current understanding of its Pathophysiology, Clinical presentation and Treatment. *Postgraduate Medical Journal* **97**, 312–320 (2021).

291. Merad, M., Blish, C. A., Sallusto, F. & Iwasaki, A. The immunology and immunopathology of COVID-19. *Science* **375**, 1122–1127 (2022).

292. Belchior-Bezerra, M., Lima, R. S., Medeiros, N. I. & Gomes, J. A. S. COVID-19, obesity, and immune response 2 years after the pandemic: A timeline of scientific advances. *Obesity Reviews* **23**, e13496 (2022).

293. Yu, S., Li, X., Xin, Z., Sun, L. & Shi, J. Proteomic insights into SARS-CoV-2 infection mechanisms, diagnosis, therapies and prognostic monitoring methods. *Frontiers in Immunology* **13**, 923387 (2022).

294. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols 2007 2:8* **2**, 1896–1906 (2007).

295. R Core Team. A Language and Environment for Statistical Computing. Preprint at (2020).

296. Team, Rs. RStudio: Integrated Development Environment for R. Preprint at (2019).

297. Sticker, A., Goeminne, L., Martens, L. & Clement, L. Robust Summarization and Inference in Proteome-wide Label-free Quantification. *Molecular & cellular proteomics : MCP* **19**, 1209–1219 (2020).

298. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society* **57**, 289–300 (1995).

299. Ulgen, E., Ozisik, O. & Sezerman, O. U. PathfindR: An R package for comprehensive identification of enriched pathways in omics data through active subnetworks. *Frontiers in Genetics* **10**, 858 (2019).

300. Overmyer, K. A. *et al.* Large-Scale Multi-omic Analysis of COVID-19 Severity. *Cell Systems* **12**, 23-40.e7 (2021).

301. Huang, J. J. *et al.* Role of the extracellular matrix in COVID-19. *World journal of clinical cases* **11**, 73–83 (2023).

302. Yeung, M. L. *et al.* Soluble ACE2-mediated cell entry of SARS-CoV-2 via interaction with proteins related to the renin-angiotensin system. *Cell* **184**, 2212-2228.e12 (2021).

303. Hilser, J. R. *et al.* Association of serum HDL-cholesterol and apolipoprotein A1 levels with risk of severe SARS-CoV-2 infection. *Journal of Lipid Research* **62**, 100061 (2021).

304. Wu, J. *et al.* APOL1 risk variants in individuals of African genetic ancestry drive endothelial cell defects that exacerbate sepsis. *Immunity* **54**, 2632-2649.e6 (2021).

305. Oczypok, E. A., Perkins, T. N. & Oury, T. D. All the "RAGE" in lung disease: The receptor for advanced glycation endproducts (RAGE) is a major mediator of pulmonary inflammatory responses. *Paediatric respiratory reviews* **23**, 40 (2017).

306. Bierhaus, A. *et al.* Diabetes-associated sustained activation of the transcription factor nuclear factor-kappaB. *Diabetes* **50**, 2792–2808 (2001).

307. Cavezzi, A., Troiani, E. & Corrao, S. COVID-19: hemoglobin, iron, and hypoxia beyond inflammation. A narrative review. *Clinics and Practice* **10**, (2020).

308. Camps, J., Iftimie, S., García-Heredia, A., Castro, A. & Joven, J. Paraoxonases and infectious diseases. *Clinical Biochemistry* **50**, 804–811 (2017).

309. Schaller, J. & Gerber, S. S. The plasmin-antiplasmin system: structural and functional aspects. *Cellular and molecular life sciences : CMLS* **68**, 785–801 (2011).

310. Java, A. *et al.* The complement system in COVID-19: friend and foe? *JCI Insight* **5**, (2020).

311. Henry, B. M. *et al.* Complement levels at admission as a reflection of coronavirus disease 2019 (COVID-19) severity state. *Journal of Medical Virology* **93**, 5515–5522 (2021).

312. Marfia, G. *et al.* Decreased serum level of sphingosine-1-phosphate: a novel predictor of clinical severity in COVID-19. *EMBO Molecular Medicine* **13**, e13424 (2021).

313. Kurt, N. *et al.* Evaluation of fetuin-A, CRP, and CRP/fetuin-A values in COVID-19 patients. *International Journal of Medical Biochemistry* **5**, 125–131 (2022).

314. Corneo, E. *et al.* Coagulation biomarkers and coronavirus disease 2019 phenotyping: a prospective cohort study. *Thrombosis Journal 2023 21:1* **21**, 1–10 (2023).

315. Lee, J. S. *et al.* Longitudinal proteomic profiling provides insights into host response and proteome dynamics in COVID-19 progression. *PROTEOMICS* **21**, 2000278 (2021).

316. Duke-Cohan, J. S. *et al.* Attractin (DPPT-L), a member of the CUB family of cell adhesion and guidance proteins, is secreted by activated human T lymphocytes and modulates immune cell interactions. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 11336–11341 (1998).

317. He, L. *et al.* Extracellular matrix protein 1 promotes follicular helper T cell differentiation and antibody production. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 8621–8626 (2018).

318. Rostami, M., Mansouritorghabeh, H. & Parsa-Kondelaji, M. High levels of Von Willebrand factor markers in COVID-19: a systematic review and meta-analysis. *Clinical and Experimental Medicine* **22**, 347–357 (2022).

319. Rahnavard, A., Mann, B., Giri, A., Chatterjee, R. & Crandall, K. A. Metabolite, protein, and tissue dysfunction associated with COVID-19 disease severity. *Scientific Reports 2022 12:1* **12**, 1–16 (2022).

320. Filbin, M. R. *et al.* Longitudinal proteomic analysis of severe COVID-19 reveals survival-associated signatures, tissue-specific cell death, and cell-cell interactions. *Cell Reports Medicine* **2**, 100287 (2021).

321. Zhang, N., Wang, S. & Wong, C. C. L. Proteomics research of SARS-CoV-2 and COVID-19 disease. *Medical Review* **2**, 427–445 (2022).

322. Alosaimi, B. *et al.* Complement Anaphylatoxins and Inflammatory Cytokines as Prognostic Markers for COVID-19 Severity and In-Hospital Mortality. *Frontiers in Immunology* **12**, 668725 (2021).

323. Risitano, A. M. *et al.* Complement as a target in COVID-19? *Nature Reviews Immunology 2020 20:6* **20**, 343–344 (2020).

324. Pretorius, E. *et al.* Persistent clotting protein pathology in Long COVID/Post-Acute Sequelae of COVID-19 (PASC) is accompanied by increased levels of antiplasmin. *Cardiovascular Diabetology* **20**, 1–18 (2021).

325. Razanakolona, M. *et al.* Anti-inflammatory Activity of the Protein Z-Dependent Protease Inhibitor. *TH Open: Companion Journal to Thrombosis and Haemostasis* **5**, e220 (2021).

326. Hill, L. A., Bodnar, T. S., Weinberg, J. & Hammond, G. L. Corticosteroid-Binding Globulin is a biomarker of inflammation onset and severity in female rats. *The Journal of endocrinology* **230**, 215 (2016).

327. van Lier, D., Kox, M. & Pickkers, P. Commentary: Plasma angiotensin II is increased in critical coronavirus disease 2019. *Frontiers in Cardiovascular Medicine* **9**, 1012452 (2022).

328. van Lier, D. *et al.* Increased blood angiotensin converting enzyme 2 activity in critically ill COVID-19 patients. *ERJ Open Research* **7**, (2021).

329. Urbiola-Salvador, V., Lima de Souza, S., Grešner, P., Qureshi, T. & Chen, Z. Plasma Proteomics Unveil Novel Immune Signatures and Biomarkers upon SARS-CoV-2 Infection. *International Journal of Molecular Sciences* **24**, 6276 (2023).

330. Hoch, N. E. *et al.* Regulation of T-cell function by endogenously produced angiotensin II. *American journal of physiology. Regulatory, integrative and comparative physiology* **296**, (2009).

331. Gabaldó, X. *et al.* Usefulness of the Measurement of Serum Paraoxonase-1 Arylesterase Activity in the Diagnoses of COVID-19. *Biomolecules* **12**, (2022).

332. Geyer, P. E. *et al.* High-resolution serum proteome trajectories in COVID-19 reveal patient-specific seroconversion. *EMBO Molecular Medicine* **13**, e14167 (2021).

333. Muhammad, Y. *et al.* Deficiency of antioxidants and increased oxidative stress in COVID-19 patients: A cross-sectional comparative study in Jigawa, Northwestern Nigeria. *SAGE Open Medicine* **9**, (2021).

334. Dworzański, J. *et al.* Glutathione peroxidase (GPx) and superoxide dismutase (SOD) activity in patients with diabetes mellitus type 2 infected with Epstein-Barr virus. *PLoS ONE* **15**, (2020).

335. Vollbracht, C. & Kraft, K. Oxidative Stress and Hyper-Inflammation as Major Drivers of Severe COVID-19 and Long COVID: Implications for the Benefit of High-Dose Intravenous Vitamin C. *Frontiers in Pharmacology* **13**, 1 (2022).

336. Yang, Q. *et al.* Human Hemoglobin Subunit Beta Functions as a Pleiotropic Regulator of RIG-I/MDA5-Mediated Antiviral Innate Immune Responses. *Journal of Virology* **93**, (2019).

337. Zhang, Z. *et al.* Proteomic profiling reveals a distinctive molecular signature for critically ill COVID-19 patients compared with asthma and chronic obstructive pulmonary disease. *International Journal of Infectious Diseases* **116**, 258–267 (2022).

338. D'Alessandro, A. *et al.* Serum Proteomics in COVID-19 Patients: Altered Coagulation and Complement Status as a Function of IL6 Level. *Journal of Proteome Research* **19**, 4417–4427 (2020).

339. WHO. COVID-19 weekly epidemiological update. *World Health Organization* 1–23 (2022).

340. Pietro, R. Di, Calcagno, S., Biondi-Zoccai, G. & Versaci, F. Is COVID-19 the deadliest event of the last century? *European Heart Journal* **42**, 2876–2879 (2021).

341. Bergamaschi, L. *et al.* Longitudinal analysis reveals that delayed bystander CD8+ T cell activation and early immune pathology distinguish severe COVID-19 from mild disease. *Immunity* **54**, 1257-1275.e8 (2021).

342. Jamilloux, Y. *et al.* Should we stimulate or suppress immune responses in COVID-19? Cytokine and anti-cytokine interventions. *Autoimmunity Reviews* **19**, 102567 (2020).

343. Cao, X. COVID-19: immunopathology and its implications for therapy. *Nature Reviews Immunology* **20**, 269–270 (2020).

344. Selva, K. J. & Chung, A. W. Insights into how SARS-CoV2 infection induces cytokine storms. *Trends in Immunology* **43**, 417–419 (2022).

345. Geng, J. *et al.* CD147 antibody specifically and effectively inhibits infection and cytokine storm of SARS-CoV-2 and its variants delta, alpha, beta, and gamma. *Signal Transduction and Targeted Therapy* **6**, 347 (2021).

346. Fathizadeh, H. *et al.* SARS-CoV-2 (Covid-19) vaccines structure, mechanisms and effectiveness: A review. *International Journal of Biological Macromolecules* **188**, 740–750 (2021).

347. Park, G. & Hwang, B. H. SARS-CoV-2 Variants: Mutations and Effective Changes. *Biotechnology and Bioprocess Engineering* **26**, 859–870 (2021).

348. Fan, Y. *et al.* SARS-CoV-2 Omicron variant: recent progress and future perspectives. *Signal Transduction and Targeted Therapy* **7**, 141 (2022).

349. Yong, S. J. Long COVID or post-COVID-19 syndrome: putative pathophysiology, risk factors, and treatments. *Infectious Diseases* **53**, 737–754 (2021).

350. Zhang, J., Dong, X., Liu, G. & Gao, Y. Risk and Protective Factors for COVID-19 Morbidity, Severity, and Mortality. *Clinical Reviews in Allergy & Immunology* (2022) doi:10.1007/s12016-022-08921-5.

351. Pennington, A. F. *et al.* Risk of Clinical Severity by Age and Race/Ethnicity Among Adults Hospitalized for COVID-19—United States, March–September 2020. *Open Forum Infectious Diseases* **8**, 1–8 (2021).

352. Liu, H., Chen, S., Liu, M., Nie, H. & Lu, H. Comorbid Chronic Diseases are Strongly Correlated with Disease Severity among COVID-19 Patients: A Systematic Review and Meta-Analysis. *Aging and disease* **11**, 668 (2020).

353. Downes, D. J. *et al.* Identification of LZTFL1 as a candidate effector gene at a COVID-19 risk locus. *Nature Genetics* **53**, 1606–1615 (2021).

354. Pathania, A. S. *et al.* COVID-19 and Cancer Comorbidity: Therapeutic Opportunities and Challenges. *Theranostics* **11**, 731 (2021).

355. Li, G. *et al.* Diabetes Mellitus and COVID-19: Associations and Possible Mechanisms. *International Journal of Endocrinology* **2021**, (2021).

356. Nishiga, M., Wang, D. W., Han, Y., Lewis, D. B. & Wu, J. C. COVID-19 and cardiovascular disease: from basic mechanisms to clinical perspectives. *Nature Reviews Cardiology 2020 17:9* **17**, 543–558 (2020).

357. Romagnolo, A. *et al.* Neurological comorbidity and severity of COVID-19. *Journal of Neurology* **268**, 762–769 (2021).

358. van Kessel, S. A. M., Olde Hartman, T. C., Lucassen, P. L. B. J. & van Jaarsveld, C. H. M. Post-acute and long-COVID-19 symptoms in patients with mild diseases: a systematic review. *Family Practice* **39**, 159–167 (2022).

359. Perrotta, F., Matera, M. G., Cazzola, M. & Bianco, A. Severe respiratory SARS-CoV2 infection: Does ACE2 receptor matter? *Respiratory Medicine* **168**, 105996 (2020).

360. Jefferies, W. A., Green, J. R. & Williams, A. F. Authentic T helper CD4 (W3/25) antigen on rat peritoneal macrophages. *Journal of Experimental Medicine* **162**, 117–127 (1985).

361. Peakman, M., Senaldi, G., Foote, N., McManus, T. J. & Vergani, D. Naturally Occurring Soluble CD4 in Patients with Human Immunodeficiency Virus Infection. *Journal of Infectious Diseases* **165**, 799–804 (1992).

362. Milani, D. *et al.* p53/NF-kB Balance in SARS-CoV-2 Infection: From OMICs, Genomics and Pharmacogenomics Insights to Tailored Therapeutic Perspectives (COVIDomics). *Frontiers in Pharmacology* **13**, 1–13 (2022).

363. Tilborghs, S. *et al.* The role of Nuclear Factor-kappa B signaling in human cervical cancer. *Critical Reviews in Oncology/Hematology* **120**, 141–150 (2017).

364. Wauman, J., De Ceuninck, L., Vanderroost, N., Lievens, S. & Tavernier, J. RNF41 (Nrdp1) controls type 1 cytokine receptor degradation and ectodomain shedding. *Journal of Cell Science* **124**, 921 (2011).

365. Zhang, Y., Li, L. F., Munir, M. & Qiu, H. J. RING-domain E3 ligase-mediated host-virus interactions: Orchestrating immune responses by the host and antagonizing immune defense by viruses. *Frontiers in Immunology* **9**, 1083 (2018).

366. Gordon, D. E. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature 2020 583:7816* **583**, 459–468 (2020).

367. Ribero, M. S., Jouvenet, N., Dreux, M. & Nisole, S. Interplay between SARS-CoV-2 and the type I interferon response. *PLOS Pathogens* **16**, e1008737 (2020).

368. Tullett, K. M. *et al.* RNF41 regulates the damage recognition receptor Clec9a and antigen cross-presentation in mouse dendritic cells. *eLife* **9**, 1–31 (2020).

369. Kao, J. & Frankland, P. W. COVID fog demystified. *Cell* **185**, 2391–2393 (2022).

370. Lin, Z. *et al.* A novel role of CCN3 in regulating endothelial inflammation. *Journal of cell communication and signaling* **4**, 141–153 (2010).

371. Gautier, T. *et al.* New therapeutic horizons for plasma phospholipid transfer protein (PLTP): Targeting endotoxemia, infection and sepsis. *Pharmacology & Therapeutics* **236**, 108105 (2022).

372. Lauwers, M., Au, M., Yuan, S. & Wen, C. COVID-19 in Joint Ageing and Osteoarthritis: Current Status and Perspectives. *International Journal of Molecular Sciences 2022, Vol. 23, Page 720* **23**, 720 (2022).

373. Jia, W., Li, H. & He, Y.-W. Pattern Recognition Molecule Mindin Promotes Intranasal Clearance of Influenza Viruses. *The Journal of Immunology* **180**, 6255–6261 (2008).

374. Hou, Y. *et al.* Expression profiles of respiratory V-ATPase and calprotectin in SARS-CoV-2 infection. *Cell Death Discovery 2022 8:1* **8**, 1–7 (2022).

375. Jimenez, L. *et al.* Acid pH Increases SARS-CoV-2 Infection and the Risk of Death by COVID-19. *Frontiers in Medicine* **8**, 1358 (2021).

376. Qi, M., Liu, B., Li, S., Ni, Z. & Li, F. Construction and Investigation of Competing Endogenous RNA Networks and Candidate Genes Involved in SARS-CoV-2 Infection. *International Journal of General Medicine* **14**, 6647–6659 (2021).

377. Laudanski, K. *et al.* Unbiased Analysis of Temporal Changes in Immune Serum Markers in Acute COVID-19 Infection With Emphasis on Organ Failure, Anti-Viral Treatment, and Demographic Characteristics. *Frontiers in Immunology* **12**, 1966 (2021).

378. Ajaz, S. *et al.* Mitochondrial metabolic manipulation by SARS-CoV-2 in peripheral blood mononuclear cells of patients with COVID-19. *American Journal of Physiology - Cell Physiology* **320**, C57–C65 (2021).

379. Satış, H. *et al.* Prognostic value of interleukin-18 and its association with other inflammatory markers and disease severity in COVID-19. *Cytokine* **137**, 155302 (2021).

380. Zaiss, D. M. W., Gause, W. C., Osborne, L. C. & Artis, D. Emerging functions of amphiregulin in orchestrating immunity, inflammation and tissue repair. *Immunity* **42**, 216 (2015).

381. Harb, H. *et al.* Notch4 signaling limits regulatory T-cell-mediated tissue repair and promotes severe lung inflammation in viral infections. *Immunity* **54**, 1186 (2021).

382. Feng, Y. *et al.* LRRC25 Functions as an Inhibitor of NF-κB Signaling Pathway by Promoting p65/RelA for Autophagic Degradation. *Scientific Reports 2017 7:1* **7**, 1–12 (2017).

383. Du, Y. *et al.* LRRC25 inhibits type I IFN signaling by targeting ISG15-associated RIG-I for autophagic degradation. *The EMBO Journal* **37**, 351–366 (2018).

384. Puig-Kröger, A. *et al.* Folate Receptor β Is Expressed by Tumor-Associated Macrophages and Constitutes a Marker for M2 Anti-inflammatory/Regulatory Macrophages. *Cancer Research* **69**, 9395–9403 (2009).

385. Warmink, K. *et al.* Folate Receptor Expression by Human Monocyte–Derived Macrophage Subtypes and Effects of Corticosteroids. *CARTILAGE* **13**, 194760352210814 (2022).

386. Ogilvie, P., Paoletti, S., Clark-Lewis, I. & Uguccioni, M. Eotaxin-3 is a natural antagonist for CCR2 and exerts a repulsive effect on human monocytes. *Blood* **102**, 789–794 (2003).

387. Sugaya, M. Chemokines and Skin Diseases. *Archivum Immunologiae et Therapiae Experimentalis* **63**, 109–115 (2015).

388. Petkovic, V., Moghini, C., Paoletti, S., Uguccioni, M. & Gerber, B. Eotaxin-3/CCL26 is a natural antagonist for CC chemokine receptors 1 and 5. A human chemokine with a regulatory role. *Journal of Biological Chemistry* **279**, 23357–23363 (2004).

389. Su, C. M., Wang, L. & Yoo, D. Activation of NF-κB and induction of proinflammatory cytokine expressions mediated by ORF7a protein of SARS-CoV-2. *Scientific Reports* **11**, 1–12 (2021).

390. Vanderbeke, L. *et al.* Monocyte-driven atypical cytokine storm and aberrant neutrophil activation as key mediators of COVID-19 disease severity. *Nature Communications 2021 12:1* **12**, 1–15 (2021).

391. Abers, M. S. *et al.* An immune-based biomarker signature is associated with mortality in COVID-19 patients. *JCI Insight* **6**, (2021).

392. Kikly, K., Liu, L., Na, S. & Sedgwick, J. D. The IL23/Th17 axis: therapeutic targets for autoimmune inflammation. *Current Opinion in Immunology* **18**, 670–675 (2006).

393. Blaner, W. S., Brun, P.-J., Calderon, R. M. & Golczak, M. Retinol-binding protein 2 (RBP2): biology and pathobiology. *Critical Reviews in Biochemistry and Molecular Biology* **55**, (2020).

394. da Silva Antunes, R. *et al.* A system-view of Bordetella pertussis booster vaccine responses in adults primed with whole-cell versus acellular vaccine in infancy. *JCI Insight* **6**, (2021).

395. Gorreja, F. *et al.* MEFV and NLRP3 Inflammasome Expression Is Attributed to Immature Macrophages and Correlates with Serum Inflammatory Proteins in Crohn´s Disease Patients. *Inflammation* **45**, 1631–1650 (2022).

396. Henry, N. *et al.* Short Chain Fatty Acids Taken at Time of Thrombectomy in Acute Ischemic Stroke Patients Are Independent of Stroke Severity But Associated With Inflammatory Markers and Worse Symptoms at Discharge. *Frontiers in Immunology* **12**, 5883 (2022).

397. Heath, L. *et al.* Manifestations of Alzheimer's disease genetic risk in the blood are evident in a multiomic analysis in healthy adults aged 18 to 90. *Scientific Reports 2022 12:1* **12**, 1–17 (2022).

398. Dagnino, S. *et al.* Prospective identification of elevated circulating CDCP1 in patients years before onset of lung cancer. *Cancer Research* **81**, 3738–3748 (2021).

399. Borges, L., Kubin, M. & Kuhlman, T. LIR9, an immunoglobulin-superfamily-activating receptor, is expressed as a transmembrane and as a secreted molecule. *Blood* **101**, 1484–1486 (2003).

400. Hammoudeh, S. M. *et al.* Systems Immunology Analysis Reveals the Contribution of Pulmonary and Extrapulmonary Tissues to the Immunopathogenesis of Severe COVID-19 Patients. *Frontiers in Immunology* **12**, 1–16 (2021).

401. Kalia, N., Singh, J. & Kaur, M. The role of dectin-1 in health and disease. *Immunobiology* **226**, 152071 (2021).

402. Gadanec, L. K. *et al.* Can SARS-CoV-2 Virus Use Multiple Receptors to Enter Host Cells? *International Journal of Molecular Sciences 2021, Vol. 22, Page 992* **22**, 992 (2021).

403. Li, J. *et al.* VSIG4 inhibits proinflammatory macrophage activation by reprogramming mitochondrial pyruvate metabolism. *Nature Communications* **8**, (2017).

404. Jung, K., Seo, S. K. & Choi, I. Endogenous VSIG4 negatively regulates the helper T cell-mediated antibody response. *Immunology letters* **165**, 78–83 (2015).

405. Widyagarini, A., Nishii, N., Kawano, Y., Zhang, C. & Azuma, M. VSIG4/CRIg directly regulates early CD8+ T cell activation through its counter-receptor in a narrow window. *Biochemical and Biophysical Research Communications* **614**, 100–106 (2022).

406. Saheb Sharif-Askari, N. *et al.* Enhanced expression of immune checkpoint receptors during SARS-CoV-2 viral infection. *Molecular therapy. Methods & clinical development* **20**, 109–121 (2021).

407. Zheng, Y. *et al.* The Roles of Siglec7 and Siglec9 on Natural Killer Cells in Virus Infection and Tumour Progression. *Journal of Immunology Research* **2020**, (2020).

408. Delaveris, C. S. *et al.* Synthetic Siglec-9 Agonists Inhibit Neutrophil Activation Associated with COVID-19. *ACS Central Science* **7**, 650–657 (2021).

409. Martin-Sancho, L. *et al.* Functional landscape of SARS-CoV-2 cellular restriction. *Molecular Cell* **81**, 2656-2668.e8 (2021).

410. Altmann, D. M. & Boyton, R. J. SARS-CoV-2 T cell immunity: Specificity, function, durability, and role in protection. *Science Immunology* **5**, 6160 (2020).

411. Sica, G. L. *et al.* RELT, a new member of the tumor necrosis factor receptor superfamily, is selectively expressed in hematopoietic tissues and activates transcription factor NF-κB. *Blood* **97**, 2702–2707 (2001).

412. Schröder, B. The multifaceted roles of the invariant chain CD74 - More than just a chaperone. *Biochimica et Biophysica Acta - Molecular Cell Research* **1863**, 1269–1281 (2016).

413. Saverino, D. *et al.* Surface density expression of the leukocyte-associated Ig-like receptor-1 is directly related to inhibition of human T-cell functions. *Human immunology* **63**, 534–546 (2002).

414. Merlo, A. *et al.* Inhibitory receptors CD85j, LAIR-1, and CD152 down-regulate immunoglobulin and cytokine production by human B lymphocytes. *Clinical and Diagnostic Laboratory Immunology* **12**, 705–712 (2005).

415. Modabber, Z. *et al.* TIM-3 as a potential exhaustion marker in CD4+ T cells of COVID-19 patients. *Immunity, Inflammation and Disease* **9**, 1707–1715 (2021).

416. Rezzola, S. *et al.* VEGFR2 activation mediates the pro-angiogenic activity of BMP4. *Angiogenesis* **22**, 521–533 (2019).

417. Sur, S. *et al.* Exosomes from COVID-19 Patients Carry Tenascin-C and Fibrinogen-β in Triggering Inflammatory Signals in Cells of Distant Organ. *International journal of molecular sciences* **22**, 1–11 (2021).

418. Chi, Y., Chai, J., Xu, C., Luo, H. & Zhang, Q. The extracellular matrix protein matrilin-2 induces post-burn inflammatory responses as an endogenous danger signal. *Inflammation research : official journal of the European Histamine Research Society ... [et al.]* **64**, 833–839 (2015).

419. Ladikou, E. E. *et al.* Von Willebrand factor (vWF): marker of endothelial damage and thrombotic risk in COVID-19? *Clinical Medicine* **20**, e178–e182 (2020).

420. Babkina, A. S. *et al.* The Role of Von Willebrand Factor in the Pathogenesis of Pulmonary Vascular Thrombosis in COVID-19. *Viruses* **14**, 211 (2022).

421. Cervantes, A. *et al.* Metastatic colorectal cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Annals of oncology : official journal of the European Society for Medical Oncology* **34**, 10–32 (2023).

422. Xi, Y. & Xu, P. Global colorectal cancer burden in 2020 and projections to 2040. *Translational Oncology* **14**, 101174 (2021).

423. Alves Martins, B. A. *et al.* Biomarkers in Colorectal Cancer: The Role of Translational Proteomics Research. *Frontiers in Oncology* **9**, 1284 (2019).

424. Sirniö, P. *et al.* Systemic inflammation is associated with circulating cell death released keratin 18 fragments in colorectal cancer. *OncoImmunology* **9**, (2020).

425. Terzić, J., Grivennikov, S., Karin, E. & Karin, M. Inflammation and Colon Cancer. *Gastroenterology* **138**, 2101-2114.e5 (2010).

426. Ladabaum, U., Dominitz, J. A., Kahi, C. & Schoen, R. E. Strategies for Colorectal Cancer Screening. *Gastroenterology* **158**, 418–432 (2020).

427. Young, P. E. & Womeldorph, C. M. Colonoscopy for Colorectal Cancer Screening. *Journal of Cancer* **4**, 217 (2013).

428. Anderson, N. L. & Anderson, N. G. The human plasma proteome: history, character, and diagnostic prospects. *Molecular & cellular proteomics : MCP* **1**, 845–867 (2002).

429. Zhao, Y. *et al.* Evolution of Mass Spectrometry Instruments and Techniques for Blood Proteomics. *Journal of Proteome Research* **22**, 1009–1023 (2023).

430. Kankanala, V. L. & Mukkamalla, S. K. R. Carcinoembryonic Antigen. *StatPearls* (2022).

431. Peltier, J., Roperch, J. P., Audebert, S., Borg, J. P. & Camoin, L. Quantitative proteomic analysis exploring progression of colorectal cancer: Modulation of the serpin family. *Journal of Proteomics* **148**, 139–148 (2016).

432. Ahn, S. B. *et al.* Potential early clinical stage colorectal cancer diagnosis using a proteomics blood test panel. *Clinical Proteomics* **16**, 1–20 (2019).

433. Chantaraamporn, J. *et al.* Glycoproteomic Analysis Reveals Aberrant Expression of Complement C9 and Fibronectin in the Plasma of Patients with Colorectal Cancer. *Proteomes 2020, Vol. 8, Page 26* **8**, 26 (2020).

434. Ivancic, M. M. *et al.* Noninvasive Detection of Colorectal Carcinomas Using Serum Protein Biomarkers. *Journal of Surgical Research* **246**, 160–169 (2020).

435. Liu, Q. *et al.* Role of the mucin-like glycoprotein FCGBP in mucosal immunity and cancer. *Frontiers in Immunology* **13**, (2022).

436. Domínguez-Reyes, T. *et al.* Interaction of dietary fat intake with APOA2, APOA5 and LEPR polymorphisms and its relationship with obesity and dyslipidemia in young subjects. *Lipids in Health and Disease* **14**, 1–10 (2015).

437. Carmena-Ramon, R., Ascaso, J. F., Real, J. T., Ordovas, J. M. & Carmena, R. Genetic Variation at the ApoA-IV Gene Locus and Response to Diet in Familial Hypercholesterolemia. *Arteriosclerosis, Thrombosis, and Vascular Biology* **18**, 1266–1274 (1998).

438. Hepner, M. & Karlaftis, V. Antithrombin. *Methods in Molecular Biology* **992**, 355–364 (2013).

439. Sun, H. M. *et al.* SERPINA4 is a novel independent prognostic indicator and a potential therapeutic target for colorectal cancer. *American Journal of Cancer Research* **6**, 1636 (2016).

440. Tang, L. *et al.* High preoperative plasma fibrinogen levels are associated with distant metastases and impaired prognosis after curative resection in patients with colorectal cancer. *Journal of Surgical Oncology* **102**, 428–432 (2010).

441. Gao, F., Zhang, X., Wang, S. & Zheng, C. Prognostic impact of plasma ORM2 levels in patients with stage II colorectal cancer - PubMed. *Ann Clin Lab Sci* . **44**, 388–393 (2014).

442. Merle, N. S., Church, S. E., Fremeaux-Bacchi, V. & Roumenina, L. T. Complement system part I - molecular mechanisms of activation and regulation. *Frontiers in Immunology* **6**, 262 (2015).

443. Depletion of Abundant Plasma Proteins and Limitations of Plasma Proteomics | Journal of Proteome Research. https://pubs.acs.org/doi/10.1021/pr100646w.

444. Plasma/Serum Proteomics: Depletion Strategies for Reducing High-Abundance Proteins for Biomarker Discovery: *Bioanalysis* **11**, 19 (2019).

445. Jiang, X. *et al.* Plasma Inter-Alpha-Trypsin Inhibitor Heavy Chains H3 and H4 Serve as Novel Diagnostic Biomarkers in Human Colorectal Cancer. *Disease markers* **2019**, 5069614 (2019).

446. Wågsäter, D., Löfgren, S., Zar, N., Hugander, A. & Dimberg, J. Pigment Epithelium-Derived Factor Expression in Colorectal Cancer Patients. *Cancer Investigation* **28**, 872–877 (2010).

447. Ji, D. *et al.* Prognostic role of serum AZGP1, PEDF and PRDX2 in colorectal cancer patients. *Carcinogenesis* **34**, 1265–1272 (2013).

448. Ivancic, M. M. *et al.* Candidate serum biomarkers for early intestinal cancer using 15N metabolic labeling and quantitative proteomics in the ApcMin/+ mouse. *Journal of Proteome Research* **12**, 4152–4166 (2013).

449. Battistelli, S. *et al.* Coagulation Factor Levels in Non-Metastatic Colorectal Cancer Patients. *The International Journal of Biological Markers* **23**, 36–41 (2018).

450. Yang, Q., Roehrl, M. H. & Wang, J. Y. Proteomic profiling of antibody-inducing immunogens in tumor tissue identifies PSMA1, LAP3, ANXA3, and maspin as colon cancer markers. *Oncotarget* **9**, 3996 (2018).

451. Chang, H. *et al.* Complement C5 is a novel biomarker for liver metastasis of colorectal cancer. *Journal of Gastrointestinal Oncology* **13**, 2351–2365 (2022).

452. Ajona, D., Ortiz-Espinosa, S. & Pio, R. Complement anaphylatoxins C3a and C5a: Emerging roles in cancer progression and treatment. *Seminars in Cell & Developmental Biology* **85**, 153–163 (2019).

453. Ding, P. *et al.* Intracellular complement C5a/C5aR1 stabilizes β-catenin to promote colorectal tumorigenesis. *Cell Reports* **39**, 110851 (2022).

454. Piao, C. *et al.* Complement 5a stimulates macrophage polarization and contributes to tumor metastases of colon cancer. *Experimental Cell Research* **366**, 127–138 (2018).

455. Voronova, V. *et al.* Diagnostic Value of Combinatorial Markers in Colorectal Carcinoma. *Frontiers in Oncology* **10**, 832 (2020).

456. Vargas, T. *et al.* Genes associated with metabolic syndrome predict disease-free survival in stage II colorectal cancer patients. A novel link between metabolic dysregulation and colorectal cancer. *Molecular Oncology* **8**, 1469–1481 (2014).

457. Esposito, K. *et al.* Colorectal cancer association with metabolic syndrome and its components: A systematic review with meta-analysis. *Endocrine* **44**, 634–647 (2013).

458. Buck, M. *et al.* Structure and Expression of Different Serum Amyloid A (SAA) Variants and their Concentration-Dependent Functions During Host Insults. *Current Medicinal Chemistry* **23**, 1725–1755 (2016).

459. Michaeli, A. *et al.* Serum amyloid A enhances plasminogen activation: Implication for a role in colon cancer. *Biochemical and Biophysical Research Communications* **368**, 368–373 (2008).

460. Heumann, D. *et al.* Competition between Bactericidal/Permeability-Increasing Protein and Lipopolysaccharide-Binding Protein for Lipopolysaccharide Binding to Monocytes. *Journal of Infectious Diseases* **167**, 1351–1357 (1993).

461. Urbiola-Salvador, V. *et al.* Plasma protein changes reflect colorectal cancer development and associated inflammation. *Frontiers in Oncology* **13**, 1635 (2023).

462. Chen, R., Luo, F. K., Wang, Y. L., Tang, J. L. & Liu, Y. S. LBP and CD14 polymorphisms correlate with increased colorectal carcinoma risk in Han Chinese. *World journal of gastroenterology* **17**, 2326–2331 (2011).

463. González-Sarrías, A. *et al.* Consumption of pomegranate decreases plasma lipopolysaccharide-binding protein levels, a marker of metabolic endotoxemia, in patients with newly diagnosed colorectal cancer: a randomized controlled clinical trial. *Food & Function* **9**, 2617–2622 (2018).

464. Zhou, Y. *et al.* LRG1 promotes proliferation and inhibits apoptosis in colorectal cancer cells via RUNX1 activation. *PLOS ONE* **12**, e0175122 (2017).

465. Shirai, R., Hirano, F., Ohkura, N., Ikeda, K. & Inoue, S. Up-regulation of the expression of leucine-rich α2-glycoprotein in hepatocytes by the mediators of acute-phase response. *Biochemical and Biophysical Research Communications* **382**, 776–779 (2009).

466. Nayak, S. B., Bhat, V. R., Upadhyay, D. & Udupa, S. L. Copper and ceruloplasmin status in serum of prostate and colon cancer patients - PubMed. *Indian J Physiol Pharmacol* .**47**, 108–110 (2003).

467. Linder, M. C. Ceruloplasmin and other copper binding components of blood plasma and their functions: an update. *Metallomics* **8**, 887–905 (2016).

468. Fei, W. *et al.* RBP4 and THBS2 are serum biomarkers for diagnosis of colorectal cancer. *Oncotarget* **8**, 92254 (2017).

469. Raucci, R. *et al.* Functional and structural features of adipokine family. *Cytokine* **61**, 1–14 (2013).

470. Kopylov, A. T. *et al.* Revelation of Proteomic Indicators for Colorectal Cancer in Initial Stages of Development. *Molecules 2020, Vol. 25, Page 619* **25**, 619 (2020).

471. Boncheva, V. B. *et al.* Identification of the Antigens Recognised by Colorectal Cancer Patients Using Sera from Patients Who Exhibit a Crohn's-like Lymphoid Reaction. *Biomolecules* **12**, 1058 (2022).

472. Ma, C. *et al.* Comprehensive and deep profiling of the plasma proteome with protein corona on zeolite NaY. *Journal of Pharmaceutical Analysis* **13**, 503–513 (2023).

473. Nicholson, B. D. *et al.* Experience of adopting faecal immunochemical testing to meet the NICE colorectal cancer referral criteria for low-risk symptomatic primary care patients in Oxfordshire, UK. *Frontline gastroenterology* **10**, 347–355 (2019).

474. Lu, D. *et al.* Methylated Septin9 has moderate diagnostic value in colorectal cancer detection in Chinese population: a multicenter study. *BMC Gastroenterology* **22**, 232 (2022).

475. Yang, C. K. *et al.* EV-miRome-wide profiling uncovers miR-320c for detecting metastatic colorectal cancer and monitoring the therapeutic response. *Cellular Oncology* **45**, 621–638 (2022).

476. Agah, S. *et al.* Quantification of Plasma Cell-Free Circulating DNA at Different Stages of Colorectal Cancer. *Cancer Invest.* **35**, 625–632 (2017).

477. Bhardwaj, M. *et al.* Multiplex screening of 275 plasma protein biomarkers to identify a signature for early detection of colorectal cancer. *Molecular Oncology* **14**, 8–21 (2020).

478. Mahboob, S. *et al.* A novel multiplexed immunoassay identifies CEA, IL8 and prolactin as prospective markers for Dukes' stages A-D colorectal cancers. *Clinical Proteomics* **12**, 1–12 (2015).

479. Thorsen, S. B. *et al.* Detection of serological biomarkers by proximity extension assay for detection of colorectal neoplasias in symptomatic individuals. *Journal of Translational Medicine* **11**, 1–13 (2013).

480. Kankanala, V. L. & Mukkamalla, S. K. R. *Carcinoembryonic Antigen*. *StatPearls* (StatPearls Publishing, 2022).

481. Chen, H., Zucknick, M., Werner, S., Knebel, P. & Brenner, H. Head-to-Head Comparison and Evaluation of 92 Plasma Protein Biomarkers for Early Detection of Colorectal Cancer in a True Screening Setting. *Clinical cancer research : an official journal of the American Association for Cancer Research* **21**, 3318–3326 (2015).

482. Sekiguchi, M. & Matsuda, T. Limited usefulness of serum carcinoembryonic antigen and carbohydrate antigen 19-9 levels for gastrointestinal and whole-body cancer screening. *Scientific Reports* **10**, 1–10 (2020).

483. Ushigome, M. *et al.* Changing pattern of tumor markers in recurrent colorectal cancer patients before surgery to recurrence: serum p53 antibodies, CA19-9 and CEA. *International journal of clinical oncology* **25**, 622–632 (2020).

484. Ji, L., Gu, L., Zhang, X. & Wang, J. Evaluation of serum midkine and carcinoembryonic antigen as diagnostic biomarkers for rectal cancer and synchronous metastasis. *Translational cancer research* **9**, 2814–2823 (2020).

485. Sun, X. *et al.* Prospective Proteomic Study Identifies Potential Circulating Protein Biomarkers for Colorectal Cancer Risk. *Cancers* **14**, 3261 (2022).

486. Harlid, S. *et al.* A two-tiered targeted proteomics approach to identify pre-diagnostic biomarkers of colorectal cancer risk. *Scientific Reports* **11**, 1–12 (2021).

487. Filipowicz, N. *et al.* Comprehensive cancer-oriented biobanking resource of human samples for studies of post-zygotic genetic variation involved in cancer predisposition. *PLOS ONE* **17**, e0266111 (2022).

488. Cui, M. *et al.* IL8, MSPa, MIF, FGF-9, ANG-2 and AgRP collection were identified for the diagnosis of colorectal cancer based on the support vector machine model. *Cell cycle* **20**, 781–791 (2021).

489. Qian, J., Tikk, K., Weigl, K., Balavarca, Y. & Brenner, H. Fibroblast growth factor 21 as a circulating biomarker at various stages of colorectal carcinogenesis. *British Journal of Cancer* **119**, 1374–1382 (2018).

490. Wang, D. *et al.* Serum CCL20 combined with IL17A as early diagnostic and prognostic biomarkers for human colorectal cancer. *Journal of Translational Medicine* **17**, 1–11 (2019).

491. Xu, J. *et al.* Diagnostic and Prognostic Value of Serum Interleukin-6 in Colorectal Cancer. *Medicine* **95**, (2016).

492. Mroczko, B. *et al.* The diagnostic value of G-CSF measurement in the sera of colorectal cancer and adenoma patients. *Clinica chimica acta; international journal of clinical chemistry* **371**, 143–147 (2006).

493. Mobarra, N. *et al.* Serum level and tumor tissue expression of Ribonucleotide-diphosphate Reductase subunit M2 B: a potential biomarker for colorectal cancer. *Molecular Biology Reports* **49**, 3657–3663 (2022).

494. Lin, X. L. *et al.* Site-specific risk of colorectal neoplasms in patients with non-alcoholic fatty liver disease: A systematic review and meta-analysis. *PLOS ONE* **16**, e0245921 (2021).

495. Nenkov, M., Ma, Y., Gaßler, N. & Chen, Y. Metabolic Reprogramming of Colorectal Cancer Cells and the Microenvironment: Implication for Therapy. *International Journal of Molecular Sciences* **22**, 6262 (2021).

496. Crucitti, A. *et al.* Laparoscopic surgery for colorectal cancer is not associated with an increase in the circulating levels of several inflammation-related factors. *Cancer Biology & Therapy* **16**, 671–677 (2015).

497. De Simone, V., Pallone, F., Monteleone, G. & Stolfi, C. Role of TH17 cytokines in the control of colorectal cancer. *Oncoimmunology* **2**, e26617 (2013).

498. Song, M. *et al.* Circulating inflammatory markers and colorectal cancer risk: A prospective case-cohort study in Japan. *International Journal of Cancer* **143**, 2767–2776 (2018).

499. Vielnascher, R. M. *et al.* Conditional ablation of TYK2 in immunity to viral infection and tumor surveillance. *Transgenic Research* **23**, 519–529 (2014).

500. Sun, J. *et al.* Clinical significance and regulation of the costimulatory molecule B7-H3 in human colorectal carcinoma. *Cancer immunology, immunotherapy : CII* **59**, 1163–1171 (2010).

501. Ge, Q. *et al.* Immunological Role and Prognostic Value of APBB1IP in Pan-Cancer Analysis. *Journal of Cancer* **12**, 595–610 (2021).

502. Morimoto, K. *et al.* Overexpression of carbonic anhydrase-related protein XI promotes proliferation and invasion of gastrointestinal stromal tumors. *Virchows Archiv* **447**, 66–73 (2005).

503. Wu, Z. *et al.* The chemokine CXCL9 expression is associated with better prognosis for colorectal carcinoma patients. *Biomedicine & Pharmacotherapy* **78**, 8–13 (2016).

504. Yu, L. *et al.* Comprehensive analysis of the expression and prognostic value of CXC chemokines in colorectal cancer. *International immunopharmacology* **89**, (2020).

505. Tian, Y. *et al.* CXCL9-modified CAR T cells improve immune cell infiltration and antitumor efficacy. *Cancer immunology, immunotherapy : CII* **71**, 2663–2675 (2022).

506. Korbecki, J. *et al.* CC Chemokines in a Tumor: A Review of Pro-Cancer and Anti-Cancer Properties of the Ligands of Receptors CCR1, CCR2, CCR3, and CCR4. *International Journal of Molecular Sciences* **21**, 1–29 (2020).

507. Miyoshi, H. *et al.* Expression profiles of 507 proteins from a biotin label-based antibody array in human colorectal cancer. *Oncology reports* **31**, 1277–1281 (2014).

508. Farquharson, A. J., Steele, R. J., Carey, F. A. & Drew, J. E. Novel multiplex method to assess insulin, leptin and adiponectin regulation of inflammatory cytokines associated with colon cancer. *Molecular biology reports* **39**, 5727–5736 (2012).

509. Yang, X., Yue, Y. & Xiong, S. Dpep2 emerging as a modulator of macrophage inflammation confers protection against CVB3-induced viral myocarditis. *Frontiers in Cellular and Infection Microbiology* **9**, 57 (2019).

510. Eisenach, P. A. *et al.* Dipeptidase 1 (DPEP1) is a marker for the transition from low-grade to high-grade intraepithelial neoplasia and an adverse prognostic factor in colorectal cancer. *British journal of cancer* **109**, 694–703 (2013).

511. Patel, P. & Chatterjee, S. Peroxiredoxin6 in Endothelial Signaling. *Antioxidants* **8**, 63 (2019).

512. Huang, W. S. *et al.* Expression of PRDX6 Correlates with Migration and Invasiveness of Colorectal Cancer Cells. *Cellular Physiology and Biochemistry* **51**, 2616–2630 (2018).

513. Liu, S. *et al.* Discovery of CASP8 as a potential biomarker for high-risk prostate cancer through a high-multiplex immunoassay. *Scientific Reports* **11**, 1–10 (2021).

514. Ramesh, P. & Medema, J. P. BCL-2 family deregulation in colorectal cancer: potential for BH3 mimetics in therapy. *Apoptosis : an international journal on programmed cell death* **25**, 305–320 (2020).

515. Sinicrope, F. A. *et al.* Proapoptotic Bad and Bid protein expression predict survival in stages II and III colon cancers. *Clinical cancer research : an official journal of the American Association for Cancer Research* **14**, 4128–4133 (2008).

516. Alecu, M., Coman, G. & Dănăilă, L. High levels of sFas and PBMC apoptosis before and after excision of malignant melanoma--case report - PubMed. *Roum Arch Microbiol Immunol* **61**, 267–273 (2002).

517. Mosnier, J. F. *et al.* Mapping of colorectal carcinoma diseases with activation of Wnt/beta-catenin signalling pathway using hierarchical clustering approach. *Journal of clinical pathology* **75**, 168–175 (2022).

518. Miyoshi, N. *et al.* SCRN1 is a novel marker for prognosis in colorectal cancer. *Journal of Surgical Oncology* **101**, 156–159 (2010).

519. Conboy, C. B. *et al.* R-Spondins 2 and 3 Are Overexpressed in a Subset of Human Colon and Breast Cancers. *DNA and Cell Biology* **40**, 70–79 (2021).

520. Lin, S. *et al.* Secernin-1 contributes to colon cancer progression through enhancing matrix metalloproteinase-2/9 exocytosis. *Disease Markers* **2015**, 230703 (2015).

521. Hilkens, J. *et al.* RSPO3 expands intestinal stem cell and niche compartments and drives tumorigenesis. *Gut* **66**, 1095–1105 (2017).

522. Seshagiri, S. *et al.* Recurrent R-spondin fusions in colon cancer. *Nature* **488**, 660–664 (2012).

523. Ashkboos, M., Nikbakht, M., Zarinfard, G. & Soleimani, M. RET Protein Expression in Colorectal Cancer ; An Immunohistochemical Assessment. **22**, 1019–1023 (2021).

524. Yang, Y. Z., Hu, W. M., Xia, L. P. & He, W. Z. Association between somatic RET mutations and clinical and genetic characteristics in patients with metastatic colorectal cancer. *Cancer medicine* **10**, 8876–8882 (2021).

525. Verrienti, A. *et al.* RET mutation and increased angiogenesis in medullary thyroid carcinomas. *Endocrine-related cancer* **23**, 665–676 (2016).

526. Ong, D. C. T. *et al.* LARG at chromosome 11q23 has functional characteristics of a tumor suppressor in human breast and colorectal cancer. **28**, 4189–4200 (2009).

527. Sha, Y., Hong, H., Cai, W. & Sun, T. Single-Cell Transcriptomics of Endothelial Cells in Upper and Lower Human Esophageal Squamous Cell Carcinoma. *Current oncology (Toronto, Ont.)* **29**, 7680–7694 (2022).

528. Xie, Y. *et al.* ARHGEF12 regulates erythropoiesis and is involved in erythroid regeneration after chemotherapy in acute lymphoblastic leukemia patients. *Haematologica* **105**, 925–936 (2020).

529. Mantovani, A. *et al.* Non-alcoholic fatty liver disease and increased risk of incident extrahepatic cancers: a meta-analysis of observational cohort studies. *Gut* **71**, 778–788 (2022).

530. Razi, S., Baradaran Noveiry, B., Keshavarz-Fathi, M. & Rezaei, N. IL17 and colorectal cancer: From carcinogenesis to treatment. *Cytokine* vol. 116 7–12 Preprint at https://doi.org/10.1016/j.cyto.2018.12.021 (2019).

531. Saunders, A. S., Bender, D. E., Ray, A. L., Wu, X. & Morris, K. T. Colony-stimulating factor 3 signaling in colon and rectal cancers: Immune response and CMS classification in TCGA data. *PLOS ONE* **16**, e0247233 (2021).

532. Li, J. *et al.* Interleukin 23 regulates proliferation of lung cancer cells in a concentration-dependent way in association with the interleukin-23 receptor. *Carcinogenesis* **34**, 658–666 (2013).

533. Robinson, R. T. IL12Rβ1: The cytokine receptor that we used to know. *Cytokine* **71**, 348–359 (2015).

534. Bin, Z. *et al.* Overexpression of B7-H3 in CD133+ colorectal cancer cells is associated with cancer progression and survival in human patients. *The Journal of surgical research* **188**, 396–403 (2014).

535. Picarda, E., Ohaegbulam, K. C. & Zang, X. Molecular Pathways: Targeting B7-H3 (CD276) for Human Cancer Immunotherapy. *Clinical cancer research : an official journal of the American Association for Cancer Research* **22**, 3425–3431 (2016).

536. Rubie, C. *et al.* ELR+ CXC chemokine expression in benign and malignant colorectal conditions. *BMC cancer* **8**, (2008).

537. Liang, J. *et al.* Bioinformatics analysis of the key genes in osteosarcoma metastasis and immune invasion. *Translational Pediatrics* **11**, 1656–1670 (2022).

538. Sanati, N., Iancu, O. D., Wu, G., Jacobs, J. E. & McWeeney, S. K. Network-based predictors of progression in head and neck squamous cell carcinoma. *Frontiers in Genetics* **9**, 183 (2018).

539. Haferlach, C. *et al.* CDKN1B, encoding the cyclin-dependent kinase inhibitor 1B (p27), is located in the minimally deleted region of 12p abnormalities in myeloid malignancies and its low expression is a favorable prognostic marker in acute myeloid leukemia. *Haematologica* **96**, 829–836 (2011).

540. Vastrad, C. & Vastrad, B. Investigation into the underlying molecular mechanisms of non-small cell lung cancer using bioinformatics analysis. *Gene Reports* **15**, 100394 (2019).

541. Haiman, C. A. *et al.* Genome-Wide Testing of Putative Functional Exonic Variants in Relationship with Breast and Prostate Cancer Risk in a Multiethnic Population. *PLOS Genetics* **9**, e1003419 (2013).

542. Gao, L. *et al.* The expression characteristics and clinical significance of ACP6, a potential target of nitidine chloride, in hepatocellular carcinoma. *BMC Cancer* **22**, 1–14 (2022).

543. Chryplewicz, A. *et al.* Mutant p53 regulates LPA signaling through lysophosphatidic acid phosphatase type 6. *Scientific Reports* **9**, 1–10 (2019).

544. Ando, T. *et al.* Expression of ACP6 is an independent prognostic factor for poor survival in patients with esophageal squamous cell carcinoma. *Oncology Reports* **15**, 1551–1555 (2006).

545. Lee, S. J. & Yun, C. C. Colorectal cancer cells – Proliferation, survival and invasion by lysophosphatidic acid. *The International Journal of Biochemistry & Cell Biology* **42**, 1907–1910 (2010).

546. Gautam, A. & Pandit, B. IL32: The multifaceted and unconventional cytokine. *Human Immunology* **82**, 659–667 (2021).

547. Shamoun, L., Kolodziej, B., Andersson, R. E. & Dimberg, J. Protein Expression and Genetic Variation of IL32 and Association with Colorectal Cancer in Swedish Patients. *Anticancer research* **38**, 321–328 (2018).

548. Ozmen, F. Interleukin-21 and Interleukin-32 gene expression levels and their relationship with clinicopathological parameters in colorectal cancer. *Annali italiani di chirurgia* **92**, 78–87 (2022).

549. Yang, Y. *et al.* Dysregulation of over-expressed IL32 in colorectal cancer induces metastasis. *World Journal of Surgical Oncology* **13**, 1–5 (2015).

550. Sloot, Y. J. E., Smit, J. W., Joosten, L. A. B. & Netea-Maier, R. T. Insights into the role of IL32 in cancer. *Seminars in Immunology* **38**, 24–32 (2018).

551. Meehan, E. V. & Wang, K. Interleukin-17 Family Cytokines in Metabolic Disorders and Cancer. *Genes 2022* **13**, 1643 (2022).

552. Swedik, S., Madola, A. & Levine, A. IL17C in human mucosal immunity: More than just a middle child. *Cytokine* **146**, 155641 (2021).

553. Lee, Y. *et al.* miR-23a-3p is a Key Regulator of IL17C-Induced Tumor Angiogenesis in Colorectal Cancer. *Cells* **9**, 1363 (2020).

554. Friedrich, M., Diegelmann, J., Schauber, J., Auernhammer, C. J. & Brand, S. Intestinal neuroendocrine cells and goblet cells are mediators of IL17A-amplified epithelial IL17C production in human inflammatory bowel disease. *Mucosal immunology* **8**, 943–958 (2015).

555. Schroder, K., Hertzog, P. J., Ravasi, T. & Hume, D. A. Interferon-gamma: an overview of signals, mechanisms and functions. *Journal of leukocyte biology* **75**, 163–189 (2004).

556. Ganapathi, S. K., Beggs, A. D., Hodgson, S. V. & Kumar, D. Expression and DNA methylation of TNF, IFNG and FOXP3 in colorectal cancer and their prognostic significance. *British Journal of Cancer* **111**, 1581–1589 (2014).

557. Söylemez, Z. *et al.* Investigation of the expression levels of CPEB4, APC, TRIP13, EIF2S3, EIF4A1, IFNg, PIK3CA and CTNNB1 genes in different stage colorectal tumors. *Turkish journal of medical sciences* **51**, 661–674 (2021).

558. Henry, C. J. *et al.* Lack of significant association between serum inflammatory cytokine profiles and the presence of colorectal adenoma. *BMC Cancer* **15**, 1–9 (2015).

559. Farc, O. *et al.* A role for serum cytokines and cell adhesion molecules in the non-invasive diagnosis of colorectal cancer. *Oncology Letters* **24**, 1–6 (2022).

560. Ma, Y. *et al.* Diagnostic value of carcinoembryonic antigen combined with cytokines in serum of patients with colorectal cancer. *Medicine* **101**, e30787 (2022).

561. Bray, F. *et al.* Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **74**, 229–263 (2024).

562. Weng, J. *et al.* Exploring immunotherapy in colorectal cancer. *J Hematol Oncol* **15**, 95 (2022).

563. Gallo, G., Vescio, G., De Paola, G. & Sammarco, G. Therapeutic Targets and Tumor Microenvironment in Colorectal Cancer. *J Clin Med* **10**, 2295 (2021).

564. Kamali Zonouzi, S., Pezeshki, P. S., Razi, S. & Rezaei, N. Cancer-associated fibroblasts in colorectal cancer. *Clin Transl Oncol* **24**, 757–769 (2022).

565. Zhang, Y., Rajput, A., Jin, N. & Wang, J. Mechanisms of Immunosuppression in Colorectal Cancer. *Cancers* **12**, 3850 (2020).

566. DiToro, D. & Basu, R. Emerging Complexity in CD4+T Lineage Programming and Its Implications in Colorectal Cancer. *Front. Immunol.* **12**, (2021).

567. Chen, X. *et al.* The effects of metabolism on the immune microenvironment in colorectal cancer. *Cell Death Discov.* **10**, 1–11 (2024).

568. Wiśniewski, J. R., Ostasiewicz, P. & Mann, M. High Recovery FASP Applied to the Proteomic Analysis of Microdissected Formalin Fixed Paraffin Embedded Cancer Tissues Retrieves Known Colon Cancer Markers. *J. Proteome Res.* **10**, 3040–3049 (2011).

569. Wiśniewski, J. R. *et al.* Absolute Proteome Analysis of Colorectal Mucosa, Adenoma, and Cancer Reveals Drastic Changes in Fatty Acid Metabolism and Plasma Membrane Transporters. *J. Proteome Res.* **14**, 4005–4018 (2015).

570. Li, C. *et al.* Integrated Omics of Metastatic Colorectal Cancer. *Cancer Cell* **38**, 734-747.e9 (2020).

571. Tanaka, A. *et al.* Proteogenomic characterization of primary colorectal cancer and metastatic progression identifies proteome-based subtypes and signatures. *Cell Reports* **43**, (2024).

572. Zhan, Y. *et al.* PLA2G4A promotes right-sided colorectal cancer progression by inducing CD39+γδ Treg polarization. *JCI Insight* **6**, (2021).

573. Yan, K. *et al.* The Comparable Microenvironment Shared by Colorectal Adenoma and Carcinoma: An Evidence of Stromal Proteomics. *Front Oncol* **12**, 848782 (2022).

574. Che, R., Wang, Q., Li, M., Shen, J. & Ji, J. Quantitative Proteomics of Tissue-Infiltrating T Cells From CRC Patients Identified Lipocalin-2 Induces T-Cell Apoptosis and Promotes Tumor Cell Proliferation by Iron Efflux. *Molecular & Cellular Proteomics* **23**, 100691 (2024).

575. Huang, P. *et al.* Spatial proteome profiling by immunohistochemistry-based laser capture microdissection and data-independent acquisition proteomics. *Anal Chim Acta* **1127**, 140–148 (2020).

576. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput | Nature Methods. **17**, 41–44 (2020).

577. MSstats Version 4.0: Statistical Analyses of Quantitative Mass Spectrometry-Based Proteomic Experiments with Chromatography-Based Quantification at Scale | Journal of Proteome Research. 22, 1466–1482 (2023).

578. Newman, A. M. *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* **37**, 773–782 (2019).

579. Malonga, T., Vialaneix, N. & Beaumont, M. BEST4+ cells in the intestinal epithelium. *American Journal of Physiology-Cell Physiology* **326**, C1345–C1352 (2024).

580. Schewe, M. *et al.* Secreted Phospholipases A2 Are Intestinal Stem Cell Niche Factors with Distinct Roles in Homeostasis, Inflammation, and Cancer. *Cell Stem Cell* **19**, 38–51 (2016).

581. Yang, K. *et al.* Tumour suppressor ABCA8 inhibits malignant progression of colorectal cancer via Wnt/β-catenin pathway. *Digestive and Liver Disease* **56**, 880–893 (2024).

582. Li, R. *et al.* Single-cell transcriptomic analysis deciphers heterogenous cancer stem-like cells in colorectal cancer and their organ-specific metastasis. *Gut* **73**, 470–484 (2024).

583. Defining the role of Lgr5+ stem cells in colorectal cancer: from basic research to clinical applications. *Genome Medicine* **9**, 66 (2017).

584. Shi, W.-K., Li, Y.-H., Bai, X.-S. & Lin, G.-L. The Cell Cycle-Associated Protein CDKN2A May Promotes Colorectal Cancer Cell Metastasis by Inducing Epithelial-Mesenchymal Transition. *Front. Oncol.* **12**, 13123 (2022).

585. CTCF promotes colorectal cancer cell proliferation and chemotherapy resistance to 5-FU via the P53-Hedgehog axis. *Aging (Albany NY)* 12, 16270–16293 (2020).

586. Boardman, L. A. Overexpression of MACC1 leads to downstream activation of HGF/METand potentiates metastasis and recurrence of colorectal cancer. *Genome Med* **1**, 36 (2009).

587. Tang, J. *et al.* ARID3A promotes the development of colorectal cancer by upregulating AURKA. *Carcinogenesis* **42**, 578–586 (2021).

588. Yu, J. *et al.* Comprehensive Analysis of the Expression and Prognosis for MMPs in Human Colorectal Cancer. *Front Oncol* **11**, 771099 (2021).

589. Zhu, W. *et al.* SULF1 regulates malignant progression of colorectal cancer by modulating ARSH via FAK/PI3K/AKT/mTOR signaling. *Cancer Cell International* **24**, 201 (2024).

590. Tao, Y., Han, T., Zhang, T. & Sun, C. Sulfatase-2 promotes the growth and metastasis of colorectal cancer by activating Akt and Erk1/2 pathways. *Biomedicine & Pharmacotherapy* **89**, 1370–1377 (2017).

591. Ala, M. Tryptophan metabolites modulate inflammatory bowel disease and colorectal cancer by affecting immune system. *International Reviews of Immunology* **41**, 326–345 (2022).

592. Li, T. *et al.* TGF-β1-SOX9 axis-inducible COL10A1 promotes invasion and metastasis in gastric cancer via epithelial-to-mesenchymal transition. *Cell Death Dis* **9**, 1–18 (2018).

593. Carvalho, A. *et al.* S100A12 in Digestive Diseases and Health: A Scoping Review. *Gastroenterology Research and Practice* **2020**, 2868373 (2020).

594. Bertomeu, T. *et al.* A High-Resolution Genome-Wide CRISPR/Cas9 Viability Screen Reveals Structural Features and Contextual Diversity of the Human Cell-Essential Proteome. *Molecular and Cellular Biology* **38**, e00302-17 (2018).

595. Kasprzak, A. & Adamek, A. Insulin-Like Growth Factor 2 (IGF2) Signaling in Colorectal Cancer—From Basic Research to Potential Clinical Applications. *International Journal of Molecular Sciences* **20**, 4915 (2019).

596. Drev, D. *et al.* Impact of Fibroblast-Derived SPARC on Invasiveness of Colorectal Cancer Cells. *Cancers* **11**, 1421 (2019).

597. Huynh, P. T. *et al.* CD90+ stromal cells are the major source of IL6, which supports cancer stem-like cells and inflammation in colorectal cancer. *International Journal of Cancer* **138**, 1971–1981 (2016).

598. Li, M. *et al.* Proteomic analysis of stromal proteins in different stages of colorectal cancer establishes Tenascin-C as a stromal biomarker for colorectal cancer metastasis. *Oncotarget* **7**, 37226–37237 (2016).

599. Giguelay, A. *et al.* The landscape of cancer-associated fibroblasts in colorectal cancer liver metastases. *Theranostics* **12**, 7624–7639 (2022).

600. van der Flier, L. G., Haegebarth, A., Stange, D. E., van de Wetering, M. & Clevers, H. OLFM4 is a robust marker for stem cells in human intestine and marks a subset of colorectal cancer cells. *Gastroenterology* **137**, 15–17 (2009).

601. Gisina, A. *et al.* CEACAM5 overexpression is a reliable characteristic of CD133-positive colorectal cancer stem cells. *Cancer Biomark* **32**, 85–98 (2021).

602. Zhao, D. *et al.* CEACAM6 expression and function in tumor biology: a comprehensive review. *Discov Onc* **15**, 186 (2024).

603. Karim, B. O., Rhee, K.-J., Liu, G., Yun, K. & Brant, S. R. Prom1 function in development, intestinal inflammation, and intestinal tumorigenesis. *Front Oncol* **4**, 323 (2014).

604. Liang, L. *et al.* GNG4 Promotes Tumor Progression in Colorectal Cancer. *Journal of Oncology* **2021**, 9931984 (2021).

605. Zhou, H. & Rigoutsos, I. The emerging roles of GPRC5A in diseases. *Oncoscience* **1**, 765–776 (2014).

606. Zhou, H., Zhao, C., Shao, R., Xu, Y. & Zhao, W. The functions and regulatory pathways of S100A8/A9 and its receptors in cancers. *Front Pharmacol* **14**, 1187741 (2023).

607. Schmid, F. *et al.* Calcium-binding protein S100P is a new target gene of MACC1, drives colorectal cancer metastasis and serves as a prognostic biomarker. *Br J Cancer* **127**, 675–685 (2022).

608. Bauvois, B. & Susin, S. A. Revisiting Neutrophil Gelatinase-Associated Lipocalin (NGAL) in Cancer: Saint or Sinner? *Cancers* **10**, 336 (2018).

609. Li, J. *et al.* Elastin is a key factor of tumor development in colorectal cancer. *BMC Cancer* **20**, 217 (2020).

610. Sethi, M. K. *et al.* Quantitative proteomic analysis of paired colorectal cancer and non-tumorigenic tissues reveals signature proteins and perturbed pathways involved in CRC progression and metastasis. *J Proteomics* **126**, 54–67 (2015).

611. Roncucci, L. *et al.* Myeloperoxidase-Positive Cell Infiltration in Colorectal Carcinogenesis as Indicator of Colorectal Cancer Risk. *Cancer Epidemiology, Biomarkers & Prevention* **17**, 2291–2297 (2008).

612. Bottardi, S. *et al.* MNDA, a PYHIN factor involved in transcriptional regulation and apoptosis control in leukocytes. *Front. Immunol.* **15**, (2024).

613. Wang, H. *et al.* Pumilio1 regulates NPM3/NPM1 axis to promote PD-L1-mediated immune escape in gastric cancer. *Cancer Letters* **581**, 216498 (2024).

614. Brdicka, T. *et al.* Phosphoprotein associated with glycosphingolipid-enriched microdomains (PAG), a novel ubiquitously expressed transmembrane adaptor protein, binds the protein tyrosine kinase csk and is involved in regulation of T cell activation. *J Exp Med* **191**, 1591–1604 (2000).

615. Strazza, M., Azoulay-Alfaguter, I., Peled, M., Adam, K. & Mor, A. Transmembrane adaptor protein PAG is a mediator of PD-1 inhibitory signaling in human T cells. *Commun Biol* **4**, 672 (2021).

616. Hwangbo, C. *et al.* Syntenin regulates TGF-β1-induced Smad activation and the epithelial-to-mesenchymal transition by inhibiting caveolin-mediated TGF-β type I receptor internalization. *Oncogene* **35**, 389–401 (2016).

617. Qian, C.-J. *et al.* ADAR-mediated RNA editing regulates PVR immune checkpoint in colorectal cancer. *Biochem Biophys Res Commun* **695**, 149373 (2024).

618. Yang, C. *et al.* LYAR Suppresses Beta Interferon Induction by Targeting Phosphorylated Interferon Regulatory Factor 3. *J Virol* **93**, e00769-19 (2019).

619. Stanelle, J., Tu-Rapp, H. & Pützer, B. M. A novel mitochondrial protein DIP mediates E2F1-induced apoptosis independently of p53. *Cell Death Differ* **12**, 347–357 (2005).

620. Rodet, F. *et al.* Deciphering molecular consequences of the proprotein convertase 1/3 inhibition in macrophages for application in anti-tumour immunotherapy. *J Biotechnol* **282**, 80–85 (2018).

621. IJMS | Free Full-Text | Chromogranin-A Expression as a Novel Biomarker for Early Diagnosis of Colon Cancer Patients. *International Journal of Molecular Sciences* **20**, 12 (2019).

622. Chen, L. *et al.* ITLN1 inhibits tumor neovascularization and myeloid derived suppressor cells accumulation in colorectal carcinoma. *Oncogene* **40**, 5925–5937 (2021).

623. Nalio Ramos, R. *et al.* Tissue-resident FOLR2+ macrophages associate with CD8+ T cell infiltration in human breast cancer. *Cell* **185**, 1189-1207.e25 (2022).

624. Bickeböller, M. *et al.* Functional characterization of the tumor-suppressor MARCKS in colorectal cancer and its association with survival. *Oncogene* **34**, 1150–1159 (2015).

625. Cherukuri, A. *et al.* B cell signaling is regulated by induced palmitoylation of CD81. *J Biol Chem* **279**, 31973–31982 (2004).

626. Urbiola-Salvador, V. *et al.* Mass spectrometry proteomics characterization of plasma biomarkers for colorectal cancer associated with inflammation. *Biomarker Insights* **19**, 11772719241257739 (2024).

627. Chen, S. *et al.* CD73: an emerging checkpoint for cancer immunotherapy. *Immunotherapy* **11**, 983–997.

628. Li, J., Chen, D. & Shen, M. Tumor Microenvironment Shapes Colorectal Cancer Progression, Metastasis, and Treatment Responses. *Front Med (Lausanne)* **9**, 869010 (2022).

629. Qian, Z. *et al.* Manipulating PTPRD function with ectodomain antibodies. *Genes Dev.* **37**, 743–759 (2023).

630. Huang, X., Qin, F., Meng, Q. & Dong, M. Protein tyrosine phosphatase receptor type D (PTPRD)—mediated signaling pathways for the potential treatment of hepatocellular carcinoma: a narrative review. *Ann Transl Med* **8**, 1192 (2020).

631. Mashima, H., Ohno, H., Yamada, Y., Sakai, T. & Ohnishi, H. INSL5 may be a unique marker of colorectal endocrine cells and neuroendocrine tumors. *Biochemical and Biophysical Research Communications* **432**, 586–592 (2013).

632. Du, L. *et al.* High Vimentin Expression Predicts a Poor Prognosis and Progression in Colorectal Cancer: A Study with Meta-Analysis and TCGA Database. *BioMed Research International* **2018**, 6387810 (2018).

633. Wang, B. *et al.* PANK1 associates with cancer metabolism and immune infiltration in clear cell renal cell carcinoma: a retrospective prognostic study based on the TCGA database. *Transl Cancer Res* **11**, 2321–2337 (2022).

634. Deng, Y., Yang, X., Hua, H. & Zhang, C. IGFBP5 is Upregulated and Associated with Poor Prognosis in Colorectal Cancer. *Int J Gen Med* **15**, 6485–6497 (2022).

635. Waters, J. A., Urbano, I., Robinson, M. & House, C. D. Insulin-like growth factor binding protein 5: Diverse roles in cancer. *Front. Oncol.* **12**, (2022).

636. Ferguson, H. R., Smith, M. P. & Francavilla, C. Fibroblast Growth Factor Receptors (FGFRs) and Noncanonical Partners in Cancer Signaling. *Cells* **10**, 1201 (2021).

637. Peterson, J. E. *et al.* VEGF, PF4 and PDGF are elevated in platelets of colorectal cancer patients. *Angiogenesis* **15**, 265–273 (2012).

638. Lam, M. *et al.* The potential role of platelets in the consensus molecular subtypes of colorectal cancer. *Cancer Metastasis Rev* **36**, 273–288 (2017).

639. Cathepsin E as a marker of colon cancer. (2016) doi:10.7274/24773283.v1.

640. Caruso, M. *et al.* Over-expression of cathepsin E and trefoil factor 1 in sessile serrated adenomas of the colorectum identified by gene expression analysis. *Virchows Arch* **454**, 291–302 (2009).

641. Ma, X.-B., Xu, Y.-Y., Zhu, M.-X. & Wang, L. Prognostic Signatures Based on Thirteen Immune-Related Genes in Colorectal Cancer. *Front. Oncol.* **10**, (2021).

642. Tang, C. *et al.* Leukotriene B4 receptor knockdown affects PI3K/AKT/mTOR signaling and apoptotic responses in colorectal cancer. *Biomolecules and Biomedicine* (2024) doi:10.17305/bb.2024.10119.

643. Bommhardt, U., Schraven, B. & Simeoni, L. Beyond TCR Signaling: Emerging Functions of Lck in Cancer and Immunotherapy. *Int J Mol Sci* **20**, 3500 (2019).

644. Xie, J.-J., Liang, J.-Q., Diao, L.-H., Altman, A. & Li, Y. TNFR-associated factor 6 regulates TCR signaling via interaction with and modification of LAT adapter. *J Immunol* **190**, 4027–4036 (2013).

645. Guangwei, Z. *et al.* TRAF6 regulates the signaling pathway influencing colorectal cancer function through ubiquitination mechanisms. *Cancer Science* **113**, 1393–1405 (2022).

646. Zhao, B. *et al.* SUSD2 suppresses CD8+ T cell antitumor immunity by targeting IL2 receptor signaling. *Nat Immunol* **23**, 1588–1599 (2022).

647. Buri, C., Gutersohn, A., Hauser, C., Kappeler, A. & Mueller, C. The chemokines CCL11, CCL20, CCL21, and CCL24 are preferentially expressed in polarized human secondary lymphoid follicles. *The Journal of Pathology* **204**, 208–216 (2004).

648. Cho, H. *et al.* Eosinophils in Colorectal Neoplasms Associated with Expression of CCL11 and CCL24. *J Pathol Transl Med* **50**, 45–51 (2015).

649. Taylor, B. C. & Balko, J. M. Mechanisms of MHC-I Downregulation and Role in Immunotherapy Response. *Front. Immunol.* **13**, (2022).

650. Ugai, T. *et al.* Smoking and Incidence of Colorectal Cancer Subclassified by Tumor-Associated Macrophage Infiltrates. *JNCI: Journal of the National Cancer Institute* **114**, 68–77 (2022).

651. Lv, Y., Ma, X., Ma, Y., Du, Y. & Feng, J. A new emerging target in cancer immunotherapy: Galectin-9 (LGALS9). *Genes & Diseases* **10**, 2366–2382 (2023).

652. Chen, X., Li, S., Long, D., Shan, J. & Li, Y. Rapamycin facilitates differentiation of regulatory T cells *via* enhancement of oxidative phosphorylation. *Cellular Immunology* **365**, 104378 (2021).

653. DiScipio, R. G., Daffern, P. J., Schraufstätter, I. U. & Sriramarao, P. Human polymorphonuclear leukocytes adhere to complement factor H through an interaction that involves alphaMbeta2 (CD11b/CD18). *J Immunol* **160**, 4057–4066 (1998).

654. Ostermann, G., Weber, K. S. C., Zernecke, A., Schröder, A. & Weber, C. JAM-1 is a ligand of the beta(2) integrin LFA-1 involved in transendothelial migration of leukocytes. *Nat Immunol* **3**, 151–158 (2002).

655. Li, J. *et al.* Mature dendritic cells enriched in immunoregulatory molecules (mregDCs): A novel population in the tumour microenvironment and immunotherapy target. *Clinical and Translational Medicine* **13**, e1199 (2023).

656. Mondanelli, G., Ugel, S., Grohmann, U. & Bronte, V. The immune regulation in cancer by the amino acid metabolizing enzymes ARG and IDO. *Current Opinion in Pharmacology* **35**, 30–39 (2017).

657. Sieminska, I. & Baran, J. Myeloid-Derived Suppressor Cells in Colorectal Cancer. *Front. Immunol.* **11**, (2020).

658. Mo, J. *et al.* The early predictive effect of low expression of the ITGA4 in colorectal cancer. *J Gastrointest Oncol* **13**, 265–278 (2022).

659. Liu, X. *et al.* FOXP3+ regulatory T cell perturbation mediated by the IFNγ-STAT1-IFITM3 feedback loop is essential for anti-tumor immunity. *Nat Commun* **15**, 122 (2024).

660. Liu, Y. *et al.* FOXP3 and CEACAM6 expression and T cell infiltration in the occurrence and development of colon cancer. *Oncology Letters* **11**, 3693–3701 (2016).

661. Pelka, K. *et al.* Spatially organized multicellular immune hubs in human colorectal cancer. *Cell* **184**, 4734-4752.e20 (2021).

662. Wang, F. *et al.* Single-cell and spatial transcriptome analysis reveals the cellular heterogeneity of liver metastatic colorectal cancer. *Science Advances* **9**, eadf5464 (2023).

663. Gu, Q. *et al.* CCL19: a novel prognostic chemokine modulates the tumor immune microenvironment and outcomes of cancers. *Aging (Albany NY)* **15**, 12369–12387 (2023).

664. Thompson, E. J. *et al.* Abstract 5545: Assessing LGR5 expression levels on colorectal cancer tissue samples for use in LGR5-targeting CAR-T cell therapy clinical trial. *Cancer Research* **83**, 5545 (2023).

665. Zygoń, J., Szajewski, M., Kruszewski, W. J. & Rzepko, R. VEGF, Flt-1, and microvessel density in primary tumors as predictive factors of colorectal cancer prognosis. *Molecular and Clinical Oncology* **6**, 243–248 (2017).

666. LAGOU, S., GRAPSA, D., SYRIGOS, N. & BAMIAS, G. The Role of Decoy Receptor DcR3 in Gastrointestinal Malignancy. *Cancer Diagn Progn* **2**, 411–421 (2022).

667. Thomadaki, H. & Scorilas, A. BCL2 Family of Apoptosis-Related Genes: Functions and Clinical Implications in Cancer. *Critical Reviews in Clinical Laboratory Sciences* **43**, 1–67 (2006).

668. Halama, N. *et al.* Natural Killer Cells are Scarce in Colorectal Carcinoma Tissue Despite High Levels of Chemokines and Cytokines. *Clinical Cancer Research* **17**, 678–689 (2011).

669. Balakrishnan, A., Marathe, S. A., Joglekar, M. & Chakravortty, D. Bactericidal/permeability increasing protein: A multifaceted protein with functions beyond LPS neutralization. *Innate Immun* **19**, 339–347 (2013).

670. Schultz, H. & Weiss, J. P. The bactericidal/permeability-increasing protein (BPI) in infection and inflammatory disease. *Clin Chim Acta* **384**, 12–23 (2007).

671. Pio, R., Ajona, D., Ortiz-Espinosa, S., Mantovani, A. & Lambris, J. D. Complementing the Cancer-Immunity Cycle. *Front. Immunol.* **10**, (2019).

672. Yu, M. *et al.* CD73 on cancer-associated fibroblasts enhanced by the A2B-mediated feedforward circuit enforces an immune checkpoint. *Nat Commun* **11**, 515 (2020).

673. Honkala, A. T., Tailor, D. & Malhotra, S. V. Guanylate-Binding Protein 1: An Emerging Target in Inflammation and Cancer. *Front. Immunol.* **10**, (2020).

674. Forster, F. *et al.* Guanylate binding protein 1-mediated interaction of T cell antigen receptor signaling with the cytoskeleton. *J Immunol* **192**, 771–781 (2014).

675. Zhang, R., Kang, R. & Tang, D. Ferroptosis in gastrointestinal cancer: from mechanisms to implications. *Cancer Letters* **561**, 216147 (2023).

676. Kruzel, M. L., Zimecki, M. & Actor, J. K. Lactoferrin in a Context of Inflammation-Induced Pathology. *Front. Immunol.* **8**, (2017).

677. Manna, G. L. *et al.* Neutrophil Gelatinase-Associated Lipocalin Increases HLA-G+/FoxP3+ T-Regulatory Cell Population in an In Vitro Model of PBMC. *PLOS ONE* **9**, e89497 (2014).

678. Zhu, W., Li, M., Wang, Q., Shen, J. & Ji, J. Quantitative Proteomic Analysis Reveals Functional Alterations of the Peripheral Immune System in Colorectal Cancer. *Molecular & Cellular Proteomics* **23**, (2024).

679. Hazini, A., Fisher, K. & Seymour, L. Deregulation of HLA-I in cancer and its central importance for immunotherapy. *J Immunother Cancer* **9**, e002899 (2021).

680. Knuchel, S., Anderle, P., Werfelli, P., Diamantis, E. & Rüegg, C. Fibroblast surface-associated FGF-2 promotes contact-dependent colorectal cancer cell migration and invasion through FGFR-SRC signaling and integrin αvβ5-mediated adhesion. *Oncotarget* **6**, 14300–14317 (2015).

681. Matsuda, Y., Ueda, J. & Ishiwata, T. Fibroblast Growth Factor Receptor 2: Expression, Roles, and Potential As a Novel Molecular Target for Colorectal Cancer. *Pathology Research International* **2012**, 574768 (2012).

682. Solinas, C., Gu-Trantien, C. & Willard-Gallo, K. The rationale behind targeting the ICOS-ICOS ligand costimulatory pathway in cancer immunotherapy. *ESMO Open* **5**, e000544 (2020).

683. Li, L. *et al.* IFI30 as a key regulator of PDL1 immunotherapy prognosis in breast cancer. *Int Immunopharmacol* **133**, 112093 (2024).

684. Rausch, M. P. *et al.* GILT in Thymic Epithelial Cells Facilitates Central CD4 T Cell Tolerance to a Tissue-Restricted, Melanoma-Associated Self-Antigen. *J Immunol* **204**, 2877–2886 (2020).

685. Rausch, M. P. & Hastings, K. T. GILT modulates CD4+ T-cell tolerance to the melanocyte differentiation antigen tyrosinase-related protein 1. *J Invest Dermatol* **132**, 154–162 (2012).

686. Li, C. *et al.* Identifying ITGB2 as a Potential Prognostic Biomarker in Ovarian Cancer. *Diagnostics* **13**, 1169 (2023).

687. Zu, L. *et al.* The Profile and Clinical Significance of ITGB2 Expression in Non-Small-Cell Lung Cancer. *Journal of Clinical Medicine* **11**, 6421 (2022).

688. Che, G. *et al.* Sulfotransferase SULT2B1 facilitates colon cancer metastasis by promoting SCD1-mediated lipid metabolism. *Clinical and Translational Medicine* **14**, e1587 (2024).

689. Wu, Y. *et al.* LYAR Promotes Colorectal Cancer Progression by Upregulating FSCN1 Expression and Fatty Acid Metabolism. *Oxid Med Cell Longev* **2021**, 9979707 (2021).

690. Arner, E. N. & Rathmell, J. C. Metabolic programming and immune suppression in the tumor microenvironment. *Cancer Cell* **41**, 421–433 (2023).

691. Yasunaga, M. & Matsumura, Y. Role of SLC6A6 in promoting the survival and multidrug resistance of colorectal cancer. *Sci Rep* **4**, 4852 (2014).

692. Cao, T. *et al.* Cancer SLC6A6-mediated taurine uptake transactivates immune checkpoint genes and induces exhaustion in CD8+ T cells. *Cell* **187**, 2288-2304.e27 (2024).

693. Miyoshi, K. *et al.* CD15+ Bone Marrow-derived Cells Are Regulators of Immune Response in ARG1-producing Colorectal Cancer Cells. *Anticancer Research* **42**, 459–470 (2022).

694. Ma, Z. *et al.* Overexpression of Arginase-1 is an indicator of poor prognosis in patients with colorectal cancer. *Pathology - Research and Practice* **215**, 152383 (2019).

695. Elomaa, H. *et al.* Quantitative Multiplexed Analysis of Indoleamine 2,3-Dioxygenase (IDO) and Arginase-1 (ARG1) Expression and Myeloid Cell Infiltration in Colorectal Cancer. *Modern Pathology* **37**, 100450 (2024).

696. Li, K. *et al.* Identification and expression of a new type II transmembrane protein in human mast cells. *Genomics* **86**, 68–75 (2005).

697. Choi, Y. J. *et al.* Lung-specific MCEMP1 functions as an adaptor for KIT to promote SCF-mediated mast cell proliferation. *Nat Commun* **14**, 2045 (2023).

698. Cy, P. *et al.* Mast-cell expressed membrane protein-1 is expressed in classical monocytes and alveolar macrophages in idiopathic pulmonary fibrosis and regulates cell chemotaxis, adhesion, and migration in a TGFβ-dependent manner. *American journal of physiology. Cell physiology* **326**, (2024).

699. Chen, J.-X., Xu, X. & Zhang, S. Silence of long noncoding RNA NEAT1 exerts suppressive effects on immunity during sepsis by promoting microRNA-125-dependent MCEMP1 downregulation. *IUBMB Life* **71**, 956–968 (2019).

700. Wang, Z. & Zhang, J. FOXP3 promotes colorectal carcinoma liver metastases by evaluating MMP9 expression via regulating S-adenosylmethionine metabolism. *Ann Transl Med* **8**, 592 (2020).

701. Hu, G., Sun, N., Jiang, J. & Chen, X. Establishment of a 5-gene risk model related to regulatory T cells for predicting gastric cancer prognosis. *Cancer Cell International* **20**, 433 (2020).

702. Tuomisto, A. E., Mäkinen, M. J. & Väyrynen, J. P. Systemic inflammation in colorectal cancer: Underlying factors, effects, and prognostic significance. *World J Gastroenterol* **25**, 4383–4404 (2019).

APPENDICES

APENDIX I. SUPPLEMENTARY MATERIAL OF CHAPTER 3

Table S1. Clinical details from the COVID-19 patients with comorbidities, corresponding controls, time of sampling, and severity.

Table S2. List of differentially expressed proteins with UNIPROT identifier between patients with comorbidities and their healthy controls with the corresponding fold change as (patients with comorbidities (CP) - healthy controls(HC)) and the adjusted p-value.

Table S3. List of KEGG enriched terms of differentially expressed proteins between patients with comorbidities and their healthy controls.

Table S4. List of differentially expressed proteins with UNIPROT identifier between patients with comorbidities and their disease controls with the corresponding fold change as (patients with comorbidities (CP) - disease controls (DC)) and the adjusted p-value.

Table S5. List of KEGG enriched terms of differentially expressed proteins between patients with comorbidities and their disease controls.

Table S6. List of differentially expressed proteins with UNIPROT identifier between patients without comorbidities and their healthy controls with the corresponding fold change as (patients without comorbidities(NP) - healthy controls(HC)) and the adjusted p-value.

Table S7. List of KEGG enriched terms of differentially expressed proteins between patients without comorbidities and their healthy controls.

Table S8. List of differentially expressed proteins with UNIPROT identifier between patients with early versus late infection with the corresponding fold change as (patients with early infection (E) - patients with late infection (L)) and the adjusted p-value.

Table S9. List of KEGG enriched terms of differentially expressed proteins between early and late SARS-CoV-2 infection.

APENDIX II. SUPPLEMENTARY MATERIAL OF CHAPTER 4

Figure S1. Protein expression in clinical groups and altered pathways in patients. (a) Heatmap of differentially expressed proteins (DEPs) between patients infected with SARS-CoV-2 with comorbidities and healthy controls, with z-score by row normalization and distributed by hierarchical clustering. (b) Network of KEGG-enriched terms from DEPs between patients with comorbidities and healthy controls. Green (upregulated in patients) and red (downregulated in patients) dots indicate statistical DEPs. The node size indicates the number of genes involved in the term.

Figure S2: Protein expression in clinical groups and altered pathways in patients. (a) Heatmap of the DEPs among patients with comorbidities, patients without comorbidities, healthy controls, and disease controls with z-score by row normalization and distributed by hierarchical clustering. The 3 clusters were generated by K-means algorithm. (b) Venn diagram with DEPs of patients with comorbidities versus healthy controls, DEPs of disease controls versus healthy controls, and DEPs of patients with comorbidities versus disease controls. The red and blue arrows represent upregulation and downregulation, respectively.

Figure S3: Clusters of protein expression in clinical groups by mean. Heatmap of the mean of DEPs for each clinical group among patients with comorbidities, patients without comorbidities, healthy controls, and diseases controls with z-score by row normalization and distributed by hierarchical clustering. The 5 clusters were generated by K-means algorithm.

Figure S4: Cell-specific expression of selected proteins from the Human Protein Atlas. Bar plot with relative quantification (score) for each protein in different cell types.

Table S1. List of DEPs between patients with comorbidities and corresponding healthy controls with the fold change expressed as patients with comorbidities—healthy controls (CP-HC).

Table S2: List of DEPs identified as secreted proteins in human blood in the Human Protein Atlas.

Table S3: List of DEPs between patients with comorbidities and disease controls with the fold change expressed as patients with comorbidities—disease controls (CP-DC).

Table S4: List of DEPs between patients without comorbidities and their healthy controls with the fold change expressed as patients without comorbidities—healthy controls (CP-DC).

Table S5: List of DEPs between patients with early collection of sample after SARS-CoV-2 infection and late collection with the fold change expressed as early–late (E-L)

Table S6: List of DEPs between patients with antibody generation against SARS-CoV-2 and without antibody generation with the fold change expressed as antibody generation–non-antibody generation (Ab^+-Ab^-).

Table S7: List of major comorbidities, BMI and medications of the major comorbidities of patients with COVID-19.

APENDIX III. SUPPLEMENTARY MATERIAL OF CHAPTER 5

Figure S1. Heatmap of DEPs between CRC patients and healthy subjects with z-score by row normalization and distributed by hierarchical clustering.

Table S1. List of proteins only identified in CRC patients and healthy subjects, respectively.

Table S2. List of identifed proteins by proteomics analysis with Uniprot entries, names and associated Gene Ontology terms.

Table S3. List of differentially expressed proteins (DEP) between CRC patients and healthy subjects with the corresponding fold change expressed as (Patient-Control) and adjusted p-value.

Table S4. KEGG enriched terms in colorectal cancer patients determined by pathway enrichment analysis via active subnetworks with the DEPs from CRC patients vs healthy subjects.

Table S5. List of significantly correlated proteins with cancer-associated inflammation in CRC patientswith the corresponding coefficients and p-values.

Table S6. List of differentially expressed proteins (DEP) between CRC patients with and without cancer-associated inflammation with the corresponding fold change expressed as (Inf.-Non-Inf.) and adjusted p-value.

Table S7. List of significantly correlated proteins with tumor stages in CRC patientswith the corresponding coefficients and p-values.

Table S8. List of differentially expressed proteins (DEP) between CRC patients with late and early stages with the corresponding fold change expressed as (Late-Early) and adjusted p-value.

APENDIX IV. SUPPLEMENTARY MATERIAL OF CHAPTER 6

Figure S1. Plasma protein changes and biological processes induced by colorectal cancer. (a) Heatmap of the differentially expressed proteins (DEP) among CRC patients and healthy controls with z-score by row normalization and distributed by hierarchical clustering. The 2 clusters are generated by the K-means algorithm. (b) Dot plot with statistically significant Gene Ontology enriched terms from Gene Set Enrichment Analysis (GSEA) after false discovery rate (FDR) correction. p.adjust, adjusted p-value.

Table S1. List of proteins included in the Olink 384-Oncology Explore panel with their respective Uniprot accession and gene name.

Table S2. List of proteins included in the Olink 384-Inflammation Explore panel with their respective Uniprot accession and gene name

Table S3. DEPs between patients and age and sex-matched healthy controls. The Fold Change is defined as Patient-Healthy control (P-C).

Table S4. KEGG enriched terms for the 202 DEPs between CRC patients and healthy controls. Each term contains an associated description, the size of the gene set, the enrichment score, the normal enrichment score (NES), the p-value; adjusted p-value, q-value, the rank, and the core enrichment with the ENTREZ identifiers of the enriched proteins

Table S5. GO enriched terms for the 202 DEPs between CRC patients and healthy controls.

Table S6. DEPs between CRC patients and healthy controls that are identified in the human blood secretome from Human Protein Atlas.

Table S7. Proteins significantly correlated with inflammation status with the corresponding correlation coefficient and *p*-value.

Table S8. DEPs between patients with and without inflammation. The FC is defined as patients with inflammation - patients without inflammation (Inf-NonInf).

Table S9. KEGG enriched terms for the 26 DEPs between CRC patients with and without inflammation. Each term contains an associated description, the fold enrichment, the occurrence, the support, the lowest/highest p-value in the iterations, as well as the up/down-regulated proteins.

Table S10. Proteins significantly correlated with cancer stage with the corresponding correlation coefficient and *p*-value.

APENDIX V. SUPPLEMENTARY MATERIAL OF CHAPTER 7

Figure S1. Main pathway alterations in CRC compared to normal tissue. (a-b) Bubble plots of top 20 enriched terms of Gene Ontology and KEGG. (c) Networks of increased proteins in CRC samples involved in complement and coagulation as well as different metabolic pathways.

Figure S2. Box and whisker plots of normalized expression of ARG1 and MCEMP1 in myeloid cells from Pelka et al. scRNA dataset.

Table S1. Clinical characteristics of included CRC patients.

Table S2. Detection rate percentages of protein groups in cancer and normal samples.

Table S3. List of differentially expressed proteins between cancer and normal tissues with their corresponding protein name (Uniprot), logarithmic fold change (log2FC) and adjusted p-value.

Table S4. List of enriched Gene Ontology (GO) terms using differentially expressed proteins as input for PathfindR analysis with the corresponding fold enrichment, metrics, adjusted p-values and involved up-/down-regulated proteins.

Table S5. List of enriched KEGG terms using differentially expressed proteins as input for PathfindR analysis with the corresponding fold enrichment, metrics, adjusted p-values and involved up-/down-regulated proteins.

Table S6. List of significantly correlated proteins with TNM stage.

Table S7. List of enriched Gene Onotology (GO) terms using correlated proteins with TNM stage as input for PathfindR analysis with the corresponding fold enrichment, metrics, adjusted p-values and involved up-/down-regulated proteins.

Table S8. List of significantly correlated proteins with Treg fractions.

Table S9. List of enriched Gene Onotology (GO) terms using correlated proteins with Treg fractions as input for PathfindR analysis with the corresponding fold enrichment, metrics, adjusted p-values and involved up-/down-regulated proteins.