Dissecting the Stromal Drivers of Gastroesophageal Adenocarcinoma Chemoresistance

Kulsum Tai

Supervisor: Dr. Lorenzo E. Ferri MD, PhD

Faculty of Medicine

Division of Experimental Medicine

McGill University, Montreal, Canada

April 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science in Experimental Medicine

TABLE OF CONTENTS

ABSTRACT

INTRODUCTION: Gastroesophageal adenocarcinoma (GEA) is the fastest-rising cancer in North America. Currently, providing patients with peri-operative systemic docetaxel-based triplet chemotherapy (DCF or FLOT) is the most effective approach to treat locally advanced GEA, the most common presentation stage of this aggressive malignancy. However, most patients recur with metastatic disease due to innate or newly acquired resistance to the therapeutic agents. The tumor microenvironment (TME) has shown to be implicated in modulating the tumor's response to chemotherapy. Distinct populations of cancer-associated fibroblasts (CAFs) have demonstrated a potential role in conferring chemoresistance in other cancer types; however, the data on this is largely gained from animal studies and this paradigm has not been extensively studied in GEA. This project aims to comprehensively characterize the heterogeneity of CAFs in a well annotated cohort of GEA patients, to identify potential biomarkers and targets to overcome therapeutic resistance.

METHODS: Immunofluorescence (IF) and fluorescence-activated cell sorting (FACS) were conducted to validate previously reported CAF markers from the literature in primary patient fibroblasts. To identify specific CAF markers for GEA, an atlas was developed using single-cell RNA sequencing (scRNA-seq) data obtained from 46 GEA patient samples available for this study, including 28 patients with longitudinal samples over the treatment trajectory with docetaxel-based triplet chemotherapy. Differential and gene ontology analyses were performed to characterize distinct CAF subpopulations and dynamic CAF markers across treatment timepoints and correlated with in-patient treatment response to chemotherapy. Alongside, an organoid-CAF co-culture was established for subsequent *in vitro* drug testing with standard of care chemotherapy.

RESULTS: CAF markers, VIM, FAP, PDPN, and S100A4 were found to be differentially expressed in fibroblasts isolated from good and poor pathological response tumours; however, only FAP and PDPN were specific to GEA CAFs. From the scRNA-seq, 204,047 cells were classified according to cell type, of which 3,556 cells were treatment naïve CAFs from tumor biopsy, 2,949 cells were CAFs from mid-treatment tumor biopsy, and 13,729 cells were CAFs from surgical resection tumor. Two main subpopulations of CAFs were identified: myofibroblast CAFs (myCAFs) and inflammatory CAFs (iCAFs). Two markers, CCL20 and CHRDL1, were found to be differentially expressed between good and poor pathological responders in iCAFs, while nine markers, including ISG20 and C20orf27, were differentially expressed between clinical partial responders and poor responders. Dynamic CAF markers, including STAG2, HAT1, and ID3 were differentially expressed across two treatment timepoints and were associated to pathological or clinical response.

CONCLUSIONS: GEA CAFs demonstrate extensive heterogeneity with the identification of two major subpopulations. Several CAF markers were found to be associated with patient chemotherapy response as well as treatment time and will be further investigated to gain a better understanding of their clinical relevance.

RÉSUMÉ

INTRODUCTION : L'adénocarcinome gastro-œsophagien (GEA) est le cancer qui progresse le plus rapidement en Amérique du Nord. Actuellement, l'administration aux patients d'une triple chimiothérapie systémique périopératoire à base de docétaxel (DCF ou FLOT) est l'approche la plus efficace pour traiter le GEA localement avancé, le stade de présentation le plus courant de cette tumeur maligne agressive. Cependant, la plupart des patients récidivent avec une maladie métastatique en raison d'une résistance innée ou nouvellement acquise aux agents thérapeutiques. Il a été démontré que le microenvironnement tumoral (TME) jouait un rôle dans la modulation de la réponse de la tumeur à la chimiothérapie. Des populations distinctes de fibroblastes associés au cancer (CAF) ont démontré un rôle potentiel dans la chimiorésistance dans d'autres types de cancer ; cependant, les données à ce sujet proviennent en grande partie d'études animales et ce paradigme n'a pas été étudié de manière approfondie dans la GEA. Ce projet vise à caractériser de manière exhaustive l'hétérogénéité des CAF dans une cohorte bien annotée de patients atteints de GEA, afin d'identifier des biomarqueurs et des cibles potentielles pour surmonter la résistance thérapeutique.

MÉTHODES : L'immunofluorescence (IF) et le tri cellulaire activé par fluorescence (FACS) ont été effectués pour valider les marqueurs CAF précédemment rapportés dans la littérature dans les fibroblastes primaires des patients. Afin d'identifier les marqueurs CAF spécifiques de la GEA, un atlas a été développé à partir de données de séquençage d'ARN unicellulaire (scRNA-seq) obtenues à partir de 46 échantillons de patients atteints de GEA disponibles pour cette étude, dont 28 patients avec des échantillons longitudinaux au cours de la trajectoire de traitement par trithérapie à base de docétaxel. L'analyse différentielle et l'analyse de l'ontologie des gènes ont été effectuées pour caractériser des sous-populations distinctes de CAF et des marqueurs dynamiques de CAF à travers

des points dans le temps du traitement et corrélés avec la réponse du traitement à la chimiothérapie chez le patient. Parallèlement, une co-culture organoïde-CAF a été établie pour des essais ultérieurs de médicaments *in vitro* avec la chimiothérapie standard.

RÉSULTATS : Les marqueurs CAF, VIM, FAP, PDPN et S100A4 ont été exprimés de manière différentielle dans les fibroblastes isolés à partir de tumeurs répondant bien ou mal à la chimiothérapie ; cependant, seuls FAP et PDPN étaient spécifiques aux CAF de la GEA. À partir du scRNA-seq, 204 047 cellules ont été classées par type cellulaire, dont 3556 cellules étaient des CAF naïves de traitement provenant d'une biopsie de tumeur, 2949 cellules étaient des CAF provenant d'une biopsie de tumeur à mi-traitement, et 13 729 cellules étaient des CAF provenant d'une résection chirurgicale de tumeur. Deux sous-populations principales de CAF ont été identifiées : les CAF myofibroblastes (myCAF) et les CAF inflammatoires (iCAF). Deux marqueurs, CCL20 et CHRDL1, ont été exprimés de manière différentielle entre les bons et les mauvais répondeurs pathologiques dans les iCAF, tandis que neuf marqueurs, dont ISG20 et C20orf27, ont été exprimés de manière différentielle entre les répondeurs partiels cliniques et les mauvais répondeurs. Les marqueurs dynamiques des CAF, dont STAG2, HAT1 et ID3, ont été exprimés de manière différentielle à deux moments du traitement et ont été associés à la réponse pathologique ou clinique.

CONCLUSIONS : Les CAF de la GEA présentent une grande hétérogénéité avec l'identification de deux sous-populations majeures. Plusieurs marqueurs des CAF ont été associés à la réponse des patients à la chimiothérapie ainsi qu'à la durée du traitement et feront l'objet d'études plus approfondies afin de mieux comprendre leur pertinence clinique.

ACKNOWLEDGEMENTS

Completion of this project and thesis would not have been possible without the support of all of those mentioned. I would like to show gratitude to my supervisor, Dr. Lorenzo Ferri, for his support and guidance with this master's project. I truly appreciate his mentorship, leadership, and his constant encouragement to succeed not only in my master's, but also towards my future goals. His work ethic, optimism, and passion have been very inspiring throughout this process, and I hope to implement the skills I've learned from him in my forthcoming endeavours.

I would also like to show my appreciation to all of my lab members; it was truly a pleasure to work alongside each and every single one of them. I learned a great deal from Dr. Veena Sangwan's and Dr. Nick Bertos' research expertise which was essential for helping this project move forward and tackling any obstacles that arose. Our research assistants, Julie and Wotan, helped out immensely with day-to-day tasks within the lab; their kindness and generosity is deeply appreciated, and I couldn't be more grateful to have had their support.

My lab members, Adam, Iris, and Sanjima, made me look forward to coming to the lab every day. They helped to welcome and integrate me into this lab, and not only taught me the foundational concepts and techniques needed to excel, but showed me such an incredible amount of care and support throughout this entire process. They were always there to guide me with this project and shared valuable insight that helped me become a better researcher. I am so thankful for having a team that made this process so enjoyable, and I will dearly cherish the wonderful experiences and memories we've made. I appreciate having Dr. James Tankel work alongside me over the last few months to troubleshoot experiments and for always spreading his positive spirit. I would also like to thank Ruoyu for helping out with optimizing experiments for this project. Betty and the members of Dr. Jonathan Cools-Lartigue's lab were always welcoming and willing

to support me when I would visit the Research Institute. I would like to show my appreciation to Rayan, Justin, Misghana, and Rachad for creating such a fun and light-hearted environment to work in, and for filling our student room with so much laughter and joy. I truly value their constant encouragement to stay motivated towards my goals and for always offering their support.

Finally, my family and friends were my rock throughout this entire journey. I am so deeply grateful for their motivating words, their daily check-ins, and for their overwhelming love and support that pushed me to keep going whenever this process would get tough. I can't put into words how incredibly fortunate I am to have them by my side.

PREFACE & AUTHOR CONTRIBUTIONS

The following thesis was written in compliance with the McGill University Graduate and Postdoctoral Studies Faculty's requirements. The thesis was written in its entirety by the author. Project design was primarily carried out by the author under the supervision of Dr. Lorenzo Ferri. Laboratory experiments and subsequent data analysis was performed by the author with the help of members of the lab: Dr. Veena Sangwan (assistant professor) contributed to the establishment of relevant methods and techniques. Julie Bérubé and Wotan Zeng (research assistants) performed tissue sample dissociation for cell culture and contributed to experiment planning. Betty Giannias (lab manager) oversaw lab activities. Dr. Nick Bertos (biobanking & lab manager) coordinated biobanking of tissue samples and lab activities, and contributed to experiment planning. Dr. Sanjima Pal (postdoctoral fellow), Adam Hoffman (M.Sc. candidate), Iris Kong (Ph.D. candidate), Ruoyu Ma (Ph.D. candidate), and Dr. James Tankel (surgical fellow) contributed to the planning and completion of experiments. Dr. Sui Huang's team, including Michael Strasser and David Gibbs (bioinformaticians at the Institute of Systems Biology), performed single-cell RNA sequencing analysis and creation of the atlas presented in this thesis.

LIST OF ABBREVIATIONS

apCAF: Antigen-presenting CAF

aSMA: Smooth muscle actin

BAX: BCL-2 Associated X protein

BCL-2: B-cell lymphoma-2

BE: Barrett's esophagus

BMP: Bone morphogenetic protein

CAF: Cancer-associated fibroblast

CAV-1: Caveolin-1

CCL: CC motif chemokine ligand

CHRDL1: Chordin-like 1

CIN: Chromosomal instability

CNV: Copy-number variation

CR: Complete response, clinical

CRC: Colorectal cancer

CSC: Cancer stem cell

CT: Computerized tomography

CXCL: CXC motif chemokine ligand

DCF: docetaxel + cisplatin + 5-fluorouracil

DEG: Differentially expressed gene

DFS: Disease-free survival

DPT: Dermatopontin

EAC: Esophageal adenocarcinoma

EBV: Epstein-Barr virus

EC: Esophageal cancer

ECF: Epirubicin + cisplatin + 5-fluorouracil

ECM: Extracellular matrix

EGJ: Esophagogastric junction

EGCDB: Esophageal, Gastric, Colorectal Data- and Bio-bank

EMT: Epithelial-to-mesenchymal transition

EndoMT: Endothelial-to-mesenchymal transition

ESCC: Esophageal squamous cell carcinoma

FAP: Fibroblast activation protein

FACS: Fluorescence-activated cell sorting

FGF2: Fibroblast growth factor 2

FLOT: 5-fluorouracil + leucovorin + oxaliplatin + docetaxel

FMO: Fluorescence-minus-one control

GAC: Gastric adenocarcinoma

GC: Gastric cancer

GEA: Gastroesophageal adenocarcinoma

GEJ: Gastroesophageal junction

GERD: Gastroesophageal reflux disease

GO: Gene ontology

HAT: Histone acetyltransferase

HER2: Human epidermal growth factor receptor 2

iCAF: inflammatory CAF

inCAF: intermediate CAF

IF: Immunofluorescence

IFNγ: Interferon gamma

IL: Interleukin

IRF: Interferon regulatory factor

ISG: Interferon stimulated exonuclease gene

MDSC: Myeloid-derived suppressor cell

MMP: Matrix metalloproteinase

myCAF: myofibroblast CAF

MSI: Microsatellite instability

OGN: Osteoglycin

OS: Overall survival

PD-1: Programmed cell death 1

PDAC: Pancreatic ductal adenocarcinoma

PDO: Patient-derived organoid

PDGF: Platelet-derived growth factor

PDGFR: Platelet-derived growth factor receptor

PDPN: Podoplanin

PR: Partial response, clinical

PSC: Pancreatic stellate cell

POSTN: Periostin

RECIST: Response Evaluation Criteria in Solid Tumors

RGS: regulator of G protein signaling

S100A4: Fibroblast specific protein

scRNA-seq: Single-cell RNA sequencing

SD: Stable disease, clinical

STAG: Stromal antigen

TAM: Tumor-associated macrophage

Tbio: Tumor biopsy

Tmid: Mid-treatment tumor

Tsur: Surgical resection tumor

TGFβ: Transforming growth factor β

TNF: Tumor necrosis factor

TME: Tumor microenvironment

TLR: Toll-like receptor

UMAP: Uniform manifold approximation and projection

VEGF: Vascular endothelial growth factor

VEGFR2: Vascular endothelial growth factor receptor 2

VIM: Vimentin

VST: Variance stabilized transform

5-FU: 5-fluorouracil

LIST OF FIGURES & TABLES

Supplemental Table 5. Gene ontology analysis of CAF subpopulations

Supplemental Table 6. myCAF and iCAF gene signatures

Supplemental Figure 1. Differentially expressed genes in iCAF and myCAF populations according to chemoresponse

Supplementary Figure 2. DEGs in iCAFs across two timepoints, treatment naïve biopsy and mid-treatment biopsy, according to pathological response to neoadjuvant chemotherapy

Supplementary Figure 3. DEGs in myCAFs across two timepoints, treatment naïve biopsy and mid-treatment biopsy, according to (A) pathological response, and (B) clinical response

Supplemental Figure 4. Validation of commonly reported markers in scRNA-seq atlas

CHAPTER 1: REVIEW OF THE LITERATURE

1.1. Overview of gastric and esophageal adenocarcinomas

1.1.1. Etiology and epidemiology

Gastric and esophageal cancers are the fourth and sixth leading causes of cancer-related mortality worldwide, respectively, resulting in over 1.3 million estimated deaths annually[1]. The incidence and mortality of these two cancers is 2-3 times higher in men, with the majority of cases and deaths arising in East Asian countries[1]. Risk factors for gastric cancer (GC) and esophageal cancer (EC) include age over 55 years, substance use (tobacco, alcohol, drugs), poor diet, and obesity[1,2]. In addition, infection by *Helicobacter pylori* or Epstein-Barr virus is prominently associated with GC, while gastroesophageal reflux disease (GERD) and its progression to Barrett's esophagus (BE) are associated with increased incidence of EC[2,3].

The aggressive nature of the disease and diagnosis at advanced stages drive a low 5-year survival rate of <20%[2]. Although there has been a striking decline in distal GC cases, the incidence of esophageal adenocarcinomas (EAC) and GCs of the proximal stomach have been on the rise[2,3]. GC and EC were previously viewed as distinct malignancies, whereby squamous cell carcinoma was said to be the predominant histological type in the esophagus in contrast to adenocarcinoma in the stomach[2,3]. However, adenocarcinomas of the proximal stomach and esophagus have shown to demonstrate shared genomic and pathological features, thereby classifying them as a single condition – gastroesophageal adenocarcinoma (GEA)[4,5].

1.1.2. Classification of disease and clinical biomarkers

A) Histological classification

EC is histologically stratified as squamous cell carcinoma (ESCC) and adenocarcinoma (EAC), where the latter typically affects the lower esophagus and is associated with obesity, GERD, and BE[5]. GC, normally presenting as adenocarcinomas (GAC), has been more extensively categorized into two subtypes by the Lauren classification: intestinal- and diffuse-type. Intestinal-type carcinomas are identified by their glandular structures and is commonly derived from intestinal metaplasia (IM), while diffuse-type carcinomas are characterized by populations of scattered cells with signet-ring cell features[3,6,7]. Both EAC and intestinal-type GC share their origin from IM following chronic inflammation, whereas diffuse-type GC emerges directly from chronic inflammatory stimuli and circumvents the intermediate step of IM[6,8].

B) Anatomical classification

Siewert classification has been used for decades to stage gastroesophageal junction (GEJ) adenocarcinomas anatomically into three types: type I carcinomas (distal esophageal tumors) are located at an epicenter 2.5 cm above the GEJ, type II carcinomas (true GEJ tumors) are within 2 cm of the GEJ, and type III carcinomas (subcardial gastric tumors) have an epicenter 2-5 cm below the GEJ[3,9]. TNM classification is further used to differentiate between EC and GC: (1) esophageal carcinomas are deemed as tumors whose epicenter are located in the distal esophagus or esophagogastric junction (EGJ), while (2) gastric carcinomas are regarded as tumors whose epicentres are >5 cm distal from the esophagus and do not extend into the GEJ[3].

C) Molecular classification

Advances in genomic techniques have facilitated the molecular characterization of GC and EAC to identify disease heterogeneity and overlapping features between the two malignancies[7]. TCGA analysis of 295 treatment naïve GACs, including GEJ tumors, distinguished four distinct

molecular subtypes: (1) chromosomal instability (CIN, 49.8%), (2) microsatellite instability (MSI)-high (21.7%), genomically stable (GS, 19.7%), and Epstein-Barr virus (EBV) positive (8.8%)[4,10]. CIN, the most predominant GC subtype, favors a gradient towards the proximal stomach, which coincides with the increasing cases of malignancies in this region. This subtype is typified by aneuploidy and alterations within the chromosome, along with the presence of TP53 mutations and receptor-tyrosine kinase/RAS/cell-cycle gene amplifications[7,10]. EBV+ tumors demonstrate increased DNA methylation and tend to appear in the proximal stomach as well. GS lacks the presence of the hypermethylation, and aneuploidy seen in the previous two subgroups. Finally, MSI-high tumors emerge in the absence of DNA mismatch repair proteins and localize in the distal stomach[10].

Additionally, the Asian Cancer Research Group characterizing 300 GACs identified an MSI-high group (22.7%) and a TP53-deficient group (35.7%) reminiscent of the CIN subtype, but did not detect the EBV+ subgroup[7,11]. Given that the large majority of tumors were of the diffuse-type, a microsatellite stable group with epithelial-to-mesenchymal transition (EMT) (15.3%) was also observed[11].

Interestingly, EACs display comparable genomic alterations and aneuploidy to CIN GCs, but less frequently have MSI-high profiles and almost never portray EBV+ profiles[5,7]. A comparative analysis between GC and EC demonstrated that out of 72 EACs, 71 tumors identified as CIN, and greater similarity was observed in the genetic aberrations between EACs and CIN GCs than ESCC. Furthermore, when looking specifically at GEAs, CIN became more pervasive proximally towards the GEJ. MSI-high and EBV+ profiles were present in GEJ adenocarcinomas, but not in EACs[5]. This coherence seen between the genomic profile of CIN GCs and EACs

highlights the classification and subsequent treatment of these two malignancies as a single group[5,10].

1.1.3. Standard-of-care & currently available treatments

Surgical resection following systemic neoadjuvant chemotherapy and/or radiotherapy has been considered the gold-standard curative treatment for locally advanced gastric and esophageal adenocarcinoma. Despite these measures, five-year survival outcomes post-resection are still poor with most clinical series reporting a rate of 30-50%[12–15]. In patients with EACs, chemotherapy prior to surgical resection reduced recurrence rates, locally and distantly[13]. However, given that patients must undergo complete resection, along with the toxicity associated with this approach, chemoradiotherapy has not been feasible to include as part of the standard-of-care[16].

Introduction of platinum-based perioperative treatment into the regimen for EACs and GACs has shown promising patient outcomes. The MAGIC trial, which implemented six combined cycles of perioperative ECF (epirubicin, cisplatin, and 5-fluoruouracil (5-FU)), demonstrated a significantly higher five-year survival rate among the perioperative chemotherapy group compared to the surgery alone group (36% vs. 23%, respectively), as well as greater tumor shrinkage at the time of resection[17]. Although this study highlighted the importance of perioperative chemotherapy in the clinical setting to improve survival, 14% of patients receiving ECF presented with local recurrence and 24% of patients presented with metastasis, prior to death within the five-year period[17]. These findings demonstrate the need to find a more efficacious regimen that reduces the probability of recurrence while maintaining the improvement in patient survival seen from the use of perioperative chemotherapy.

Docetaxel, a semi-synthetic taxoid agent, has previously shown cytotoxic activity against metastatic gastric cancer as both a first-line combination therapy and second-line monotherapy[18–

[20]. Ferri et al.[16] established the use of four cycles of DCF (docetaxel, cisplatin, and 5-FU) in the neoadjuvant setting for GEA patients, which resulted in complete resection among all patients of the clinical trial and tumor reduction in 89% of patients. Perioperative FLOT (5-FU, leucovorin, oxaliplatin, and docetaxel) has been shown to be a more tolerable therapeutic option compared to DCF and improved median overall survival (OS) relative to ECF[20].

Given the wide molecular heterogeneity of patients' tumors, there has been an urgent need to develop more targeted therapies that can improve patient prognosis and reduce the exposure to toxicity. Human epidermal growth factor receptor 2 (HER2), also known as *ERBB2*, has become a notable biomarker in GEA, given its amplification among 15-20% of patients and association with aggressive, poor prognostic cancer[10,21,22]. The ToGA phase 3 clinical trial tested trastuzumab, a monoclonal antibody targeted-therapy against HER2 commonly used for breast cancer, in combination with chemotherapy (cisplatin alone or cisplatin plus fluorouracil) among patients with advanced or metastatic HER2-positive gastric or GEJ cancer. Comparing to patients who received chemotherapy alone, the addition of trastuzumab resulted in improved OS by 3 months, and by 4.2 months among patients with high HER2 expression[21]. These findings were considered clinically significant, resulting in the approval of trastuzumab plus chemotherapy as a first-line therapy for HER2-positive GEA patients in the metastatic setting (i.e., stage IV)[10].

Addition of targeted therapies against commonly altered genes in GEA, such as *MET*, *EGFR*, and *FGFR2*, to perioperative chemotherapy have reached phase II and phase III clinical trials, but have reported disappointing results, lacking improvement in efficacy and survival compared to chemotherapy alone[23–26]. In contrast, antiangiogenic therapies such as ramucirumab (monoclonal antibody against VEGFR2) have shown moderate improvement to patient OS when administered as a second-line monotherapy or in combination with chemotherapy, resulting in approval by the

FDA[27–29]. Similarly, immune checkpoint inhibitors such as pembrolizumab and nivolumab have been FDA-approved for advanced, metastatic gastric and GEJ adenocarcinomas, as well as MSI-high tumors; however, there have been mixed results regarding their efficacy[30–33].

1.2. Therapy resistance in GEA

Although the currently established standard-of-care treatments for GEA have shown moderate improvement of OS for patients with advanced disease relative to surgery alone, recurrence and metastasis is still an overarching issue that many patients encounter. The use of docetaxel-based chemotherapies in the neoadjuvant and adjuvant setting for GEA has demonstrated an initial response of >60%, making it the most effective therapy to date. However, approximately 40% of patients are innately resistant to the chemotherapy and 50% of initial responders experience recurrence as a result of acquired genetic or non-genetic chemoresistance [34]. Likewise, treatment with targeted therapies has shown mixed results regarding its efficacy due to the presence of resistance mechanisms against the drug and the lack of predictive biomarkers for the disease[35–37].

1.2.1. Tumor heterogeneity and cancer stem cells

Intratumoral heterogeneity, defined by genomic, phenotypic, and functional variations within the tumor population, is a key factor that challenges the efficacy and use of chemotherapy or targeted therapies, given that various components of the tumor will respond differently to the treatment[38]. The clonal evolution model of cancer states that there are distinct clones present within a tumor that can become mutated independently, resulting in different patterns of cell growth and survival. This suggests that certain clones will be selected due to advantageous driver mutations

that promote growth and expansion[38,39]. As a result of this selection pressure, certain pre-existing resistant clones that survive following cytotoxic treatment will become dominant and expand, resulting in therapy-resistant tumor progression. Furthermore, the presence of multiple initial clones can also drive partial intrinsic chemoresistance[40].

Concordantly, the cancer stem cell (CSC) hypothesis describes that only a small subset of malignant cells has proliferative, self-renewing, tumorigenic potential and the ability to differentiate into various cell types[41,42]. CSCs have the capacity to enter a quiescent or dormant state, whereby they can resist cancer therapies[41]. These two hypotheses, in combination, explain the presence of diverse tumor cells, which respond heterogeneously to cytotoxic drugs. Additionally, non-genetic cancer cell plasticity, explained by either spontaneous or induced cell switching from a differentiated, drug-sensitive state to a more resistant, stem-like state via non-genetic elevation of survival and mitogenic pathways, has been shown to play a role in the development of therapy resistance[43,44].

1.2.2. Molecular mechanisms of therapy resistance

A multitude of mechanisms involved in resistance have been described at the molecular level. These mechanisms will not be explained in extensive depth here, but they include decreased influx and increased efflux by drug transporters, drug inactivation by detoxification systems, promotion of anti-apoptotic pathways, and enhanced DNA repair[45,46]. Furthermore, alteration of key cell functions such as drug metabolism, and epigenetics are all factors influencing therapy response[46]. CSCs confer resistance through the use of these mechanisms in addition to their ability to promote the rise of resistant daughter cells, therefore facilitating the creation of a therapy-resistant heterogeneous population[41].

1.2.3. Tumor microenvironment

The tumor microenvironment (TME) is comprised of tumor cells, endothelial cells, cancer-associated fibroblasts (CAFs), and cells of the innate and adaptive immune system, along with their secreted products (growth factors, chemokines, cytokines), and the extracellular matrix (ECM)[47]. The complex interplay between the components of the TME (tumor-tumor, tumor-stromal, cell-cell, and tumor-ECM interactions), along with environmental conditions, such as hypoxia and low pH, all have been implicated in modulating the response to therapy and promotion of tumor progression and growth, through the release of soluble factors[46]. Intrinsic resistance conferred by the TME includes high interstitial fluid pressure and inefficient blood supply, both of which hinder the distribution of the therapy towards the tumor. Meanwhile, the shifting adaptability of the TME composition in response to external stresses and therapeutic interventions has been described as a form of acquired resistance seen in many cancer types[48].

Kim et al., [49] demonstrated the remodeling of the tumor-immune microenvironment (TIME) among advanced gastric cancer patients following treatment with the standard-of-care 5-FU and platinum chemotherapy, whereby non-responders demonstrated a greater proportions of tumor-associated macrophages (TAMs) in both pre-treatment and on-treatment samples, while responders presented with greater T cell infiltration. Furthermore, a lack of response to therapy is associated with the development of an immunosuppressive environment via TAMs and regulatory T cells. The findings from this study highlighted the extensive heterogeneity of the TME among patients, as well as its ability to remodel itself, either favorably or unfavorably, to modulate the response to chemotherapy. Along with cellular alterations within the TME, the release of certain cytokines and growth factors help the tumor epithelium to resist and evade apoptosis, such as IL-6 and IGF-1[47].

1.3. Cancer-associated fibroblasts (CAFs) in GEA

1.3.1. Origin & differentiation of CAFs

CAFs are the most abundant and heterogenous cell type in the TME and are derived from fibroblasts which play a major role in producing and maintaining the ECM. Fibroblasts function as wound healing cells, such that they enter their activated, myofibroblast state upon injury and release inflammatory factors, such as transforming growth factor β (TGFβ), while simultaneously adopting a highly contractile nature via the expression of smooth muscle actin (αSMA)[50]. Fibroblasts generally can be defined by their lack of expression of epithelial and endothelial lineage markers, such as EPCAM and CD31, respectively, and positive expression of mesenchymal markers such as vimentin (VIM) and fibroblast activation protein (FAP)[50–52]. Activated CAFs express particular markers such as fibroblast specific protein (FSP1; S100A4), periostin (POSTN), Tenascin-C, and caveolin-1 (CAV1)[52,53].

CAFs display a unique characteristic compared to normal fibroblasts and myofibroblasts, such that they remain in the activated state and do not undergo apoptosis or revert back to their normal state[51]. Activation of CAFs can occur through various mechanisms besides TGFβ signalling, such as the presence of inflammatory signals from interleukins (IL-1 and IL-6) and tumor necrosis factor (TNF), contact with tumor cells and other cells, DNA damage, and oxidative stress[50]. Although CAFs are classically believed to originate from activated fibroblasts, several studies have shown their origin from various other cell types via cell-cell contact and release of soluble factors. Interestingly, epithelial cells are capable of differentiating into CAFs via EMT, such that they begin to lose their epithelial features and begin to adopt mesenchymal morphology, function, and gene signature[52,54]. This transition often progresses to a state where the cells retain their epithelial features but also assimilate mesenchymal properties, leading them to express a

combination of markers. The epithelial cells will begin to lose their cell-cell junctions and undergo ECM reorganization[54]. Simultaneously, CAFs can induce epithelial cells to go through EMT through the release of cytokines and chemokines, such as TGFβ, resulting in the acquisition of more aggressive features by tumor cells and increased metastatic potential[52,55]. Similarly, endothelial cells undergo a similar process of endothelial-mesenchymal transition (EndMT) to differentiate into CAFs[52,56]. Furthermore, other cell types, such as mesenchymal cells, adipocytes, and CSCs, are also capable of differentiating into CAFs through various signalling pathways. This wide diversity in how CAFs can be derived provides an explanation for their extensive molecular, morphological, and functional heterogeneity[52].

1.3.2. Heterogeneity of CAFs

A) Marker heterogeneity & CAF subtyping

CAFs demonstrate great heterogeneity regarding the markers they express. Although αSMA is a commonly known marker to identify CAFs, it still falls short in being able to distinguish all CAFs within the TME as well as its expression by other cell types[57,58]. Finding a specific marker to identify CAFs remains a challenge as no single marker has been found to be exclusive to CAFs; alternatively, CAF subtyping based on expression of a combination of markers as well as function can be used to characterize distinct populations[59].

CAF subpopulations have been extensively studied and characterized in various cancer types, such as pancreatic and breast, and have been associated to functions that they play within the TME.

Table 1. Subtype characterization of CAFs in other cancer types

Cancer	Subtype	Gene Signature	Subcluster	Reference
Pancreatic	myCAF	aSMA, TAGLN, MYL9, TPM1, TPM2, POSTN, MMP11, HOPX, TWIST1		60, 61
	iCAF	IL-6, IL-8, CXCL1, CXCL2, CCL2, CXCL12, HAS1, HAS2, PDGFRa, HIF1a, NRF2, SOD2		60, 61
	apCAFs	Human: HLA-DRA, CD74, HLA-DPA1, HLA-DQA1		61
Breast	CAF-S1	FAP, aSMA, CD29, S100A4, PDGFRβ, CAV1	ecm-myCAF: THY1, COL3A1, LRRC17, FBN1, SFRP2, CCDC80	62, 63
			detox-iCAF: C7, ADH1B, CXCL12, DCN, IL6, OGN, CCL2, CXCL3	62, 63
			IL-iCAF: LAMC3, C7, CCL8, VCAM1, IL6, FBLN5	62, 63
			TGFβ-myCAF: TGFB1, TGFB3, ID4, THBS4, LOXL1	62, 63
			wound-myCAF: SMOC2, RARRES2, SFRP4, FBLN2, IGF1, IGF2	62, 63
			INFγ-iCAF: CCL19, RBP5, C7, CXCL13, CXCL9, CCL3, IL34, CCL8	62, 63
	CAF-S2	$CD29^{low}$		62, 63
	CAF-S3	CD29, S100A4, PDGFRβ		62, 63
	CAF-S4	aSMA, CD29, S100A4, PDGFRβ		62, 63
Gastric	myCAF	TPM1, TPM2, MYL9, TAGLN, POSTN		64
	iCAFs	IL6, IL11, IL24, CXCL1, CXCL2, MMP1, MMP3, MMP10		64
	inCAFs	PDGFRa, POSTN, ID1, and ID3		64

Pancreatic cancer. Öhlund et al[60] identified two subtypes of CAFs, myofibroblastic CAFs (myCAFs) and inflammatory CAFs (iCAFs), in pancreatic ductal adenocarcinoma (PDAC) using immunofluorescence (IF) analysis on human tissue samples and a murine co-culture model between pancreatic stellate cells (PSCs) and tumor organoids. Between these two subtypes, myCAFs were described as FAP^+ CAFs expressing high levels of aSMA, while iCAFs expressed lower levels of aSMA with increased levels of cytokine and chemokine release. Interestingly, the location of these CAFs relative to the tumor were found to be distinct: myCAFs were found to be adjacent to the tumor, while iCAFs were found further away from the tumor, around desmoplastic regions[60].

These findings were further validated and elaborated using single-cell RNA sequencing (scRNA-seq) on human PDAC tumor samples from six patients as well as tumor samples from four KPC mice[61]. In concordance with the *in vitro* PDAC models, scRNA-seq identified the myCAF and iCAF clusters, where myCAFs were distinguished by their highly contractile nature, involvement in ECM organization, and increased collagen formation, while inflammatory signalling pathways were predominant among iCAFs[61]. In addition, distinct signalling pathways were found to be implicated among these two CAF subtypes; the IL1/JAK-STAT3 pathway was found to be activated among iCAFs, while the TGFβ/SMAD2/3 pathway was associated strongly with myCAFs[61]. Analysis of CAFs in the KPC mouse tumors identified a novel subtype of CAFs, known as antigen-presenting CAFs (apCAFs), which express genes from the MHC class II family, normally expressed by antigen-presenting cells within the immune system[61].

Breast cancer. Costa et al.[62] identified four major CAF subpopulations in human breast cancer using multicolor flow cytometry with six known CAF markers (FAP, CD29, aSMA, S100A4, PDGFRβ, and CAV1). The four subtypes were defined as: CAF-S1, CAF-S2, CAF-S3, and CAF-

S4. Notably, CAF-S1 was the only subtype demonstrating positive expression of FAP and presented an ECM and inflammatory signature. This subtype of CAFs was also shown to create an immunosuppressive environment by increasing $CD4^+$ $CD25^+$ $FOXP3^+$ regulatory T cells within the TME. On the other hand, CAF-S4 was defined by a gene signature similar to that of pericytes. Given the lack of expression of the chosen markers by CAF-S2 and CAF-S3 and their presence in healthy tissue, the subtypes were likened to normal fibroblasts[62,63].

Further analysis using scRNA-seq on seven breast cancer patient samples allowed for eight distinct CAF subgroups to be identified within the CAF-S1 population. Consistent with PDAC, these subgroups fell into the major myCAFs (cluster 0, ecm-myCAF; cluster 3, TGFβ-myCAF; cluster 4, wound-myCAF, cluster 6 IFNαβ-myCAF, cluster 7, acto-myCAF), iCAFs (cluster 1, detox-iCAF; cluster 2, IL-iCAF; cluster 5, IFNγ-iCAF), and apCAFs populations[63]. It was found that ecm-myCAF and TGFβ-myCAF were the subgroups associated with an immunosuppressive environment, such that their overall content was correlated with the abundance of PD-1^+ and $CTLA4^+$ $CD4^+$ T cells. It was also found that ecm-myCAFs are capable of converting into TGFβ-myCAFs through interaction with $CD4^+$ $CD25^+$ T cells, which may contribute to resistance to immunotherapy[63].

Gastric cancer. Studies characterizing CAF subtypes in GC have been sparse within the literature. Transcriptome mapping of the stroma from nine GC patients using laser capture microdissection and scRNA-seq revealed 129 DEGs between good and poor prognostic outcome groups, such that genes and pathways related to ECM organization (*AEBP1, COL10A1, COL11A1, THBS2, TPM2*), growth factor response (*INHBA, FGFR1, HSPB1*), and mesenchymal cell proliferation (*LMNA, UACA*) were significantly upregulated among poor outcome patients. In contrast, genes associated with the humoral immune response (*LCN2, PGC, REG1A, ITLN1, BIRC3*), tissue homeostasis

(*LYZ, MUC6*), and digestive tract development (*GATA6, ITGA6, CLDN18*) were differentially downregulated among poor outcome patients compared to good[65].

A single-cell atlas comprising of >200,000 cells from 31 GC patients identified four clusters within the stromal metacluster: fibroblasts (STF1 and STF3; *LUM* and *DCN* expression), pericytes (STF2; *NOTCH3* and *RGS5* expression), and a novel cell type expressing both endothelial and fibroblast markers (STF4; *PLVAP* and *RGS5* expression). The identification of STF4 stromal cells was hypothesized to be derived from EndoMT differentiation of endothelial cells to CAFs[66]. Importantly, Kim et al.[64] performed single-cell analysis on pre-cancerous and cancerous GC lesions on 30,888 cells from 24 patients and revealed three CAF subtypes reminiscent of those seen in breast cancer and PDAC: myCAFs, iCAFs, and intermediate CAFs (inCAFs). inCAFs were found to share gene signatures between myCAFs and iCAFs. NF-kappa B signalling, TNF signalling, and cytokine interaction pathways were upregulated among iCAFs. While iCAFs were positively correlated with stemness, myCAFs and inCAFs were negatively correlated.

B) *Functional heterogeneity*

The functions of CAFs within the TME are diverse and multifaceted, ranging from promoting tumor growth, immune modulation, pro-metastatic activity, angiogenesis, and resistance to therapy.

Interactions with tumors. Several studies have shown how CAFs can promote tumor growth and proliferation through the release of growth factors (e.g. HGF, IGF1), cytokines (e.g. TGFβ, IL-6, IL-22), chemokines (e.g. CXCL12, CCL2), and activation of their signalling pathways[58,67,68]. Interaction of CAFs with tumor cells via the release of these factors enhances their survival and

can result in anti-apoptotic activity through the upregulation of proteins like B-cell lymphoma-2 (BCL-2) and downregulation of pro-apoptotic protein BCL-2 Associated X protein (BAX)[58,69].

Binding of CAF-secreted CXCL12, a chemoattractant for lymphocytes and macrophages, to its corresponding receptor, CXCR4, expressed on the surface of cancer cells, has shown to stimulate TGF-β expression by the tumor, leading to increased survival and progression[67,70].

IL-6 upregulation in stromal cells has also been shown to be implicated in GC tumorigenesis, such that knockout of the cytokine in N-methyl-N-nitrosourea- (MNU-) induced GC mice resulted in a lower incidence of GC compared to wildtype mice. Co-culture of IL-6 positive fibroblasts with the GC NUGC4 cell line further induced genes associated with STAT3, a known marker involved in promoting GC progression and inducing vascular endothelial growth factor (VEGF) expression[69,71].

Angiogenesis. Neovascularization is a common occurrence in tumorigenesis as it facilitates invasion and metastasis of cancer cells. CAFs play a role in the development of new blood vessels through the secretion of pro-angiogenic factors, such as VEGFa, which subsequently interact with its receptor found on endothelial cells, such as VEGF receptor 2 (VEGFR2)[57]. Other important pro-angiogenic factors secreted by CAFs include platelet-derived growth factor (PDGF), CXCL12, and fibroblast growth factor 2 (FGF2), which facilitate endothelial cell proliferation, migration, and recruitment to the site of the tumor[57,58,69]. Interestingly, the release of CXCL12 by CAFs has also been shown to recruit endothelial progenitor cells and modulate their migration in breast cancer, promoting angiogenesis and subsequent tumor expansion[67].

ECM remodelling. The process of neovascularization and promotion of tumor progression is further associated with reorganization and degradation of the ECM mediated by CAFs, given that they are the major producers of ECM components, including collagens, glycoproteins,

proteoglycans, and matrix metalloproteinases (MMPs)[57]. Deposition of ECM components by CAFs is a pathological mechanism resulting from TGFβ secretion by tumor and immune cells, as well as sustained wound inflammatory signals[72]. Furthermore, increased stiffness of the ECM is characteristic of malignancy, whereby tumor cell invasion and metastasis are facilitated, along with creating a barrier for T cell migration towards the tumor site[57,59].

Immunosuppressive activity. The release of inflammatory cytokines and chemokines by CAFs modulates the recruitment of immune cells to the TME. CAFs have also shown to participate in the recruitment of myeloid-derived suppressor cells (MDSCs) and Tregs, which are involved in promoting tumor survival. FAP^+ CAFs demonstrated immunosuppressive activity via the release of CCL2, leading to the activation of the STAT3-CCL2 signaling pathway and recruitment of MDSCs, which was further confirmed by the recruitment of Tregs and $CD4^+$ $CD25^+$ T cells CAF-S1 in breast cancer[62,73].

Chemoresistance. Several mechanisms have been reported for CAF-induced chemoresistance, including the release of cytokines, CAF-tumor cross-talk, decreased drug delivery via ECM deposition, and increasing CSCs in the TME[58,59,74]. Interestingly, CAFs are capable of limiting drug bioavailability via drug scavenging mechanisms, as seen with gemcitabine in PDAC[75]. The secretion of IL-11 by CAFs was found to be involved in conferring resistance to cisplatin in GC cells, through the activation of the gp130/JAK/STAT/Bcl signalling pathway[76]. IL-6 released by CAFs was found to induce the expression of the CXCR7 receptor via STAT3/NF-κB signalling in ESCC cells following co-culture between the two cell types, resulting in cisplatin resistance[77]. TGFβ1 signalling by CAFs has also demonstrated involvement in ESCC chemoresistance to several drugs (cisplatin, taxol, 5-FU, docetaxel, and carboplatin)[78]. Despite these findings, the association of distinct CAF subtypes to chemoresistance in GC is still unclear.

1.3.3. Clinical implication of CAFs

CAF gene signatures and proteomic analysis have shown to provide prognostic relevance in several cancer types[57]. Increased levels of stromal-related genes, such as MMP7 and FAP+ CAFs, have shown to be associated with poor outcome in PDAC and colorectal cancer (CRC), respectively[79,80]. In particular, TGFβ signaling being upregulated in the stromal environment has been shown to be related to poor survival outcomes among CRC patients[81]. The association of stromal gene signatures to disease prognosis can therefore serve to detect cancer early among patients. For example, POSTN expression was found to be associated with poor OS and DFS in EAC tumors[82].

Along with the valuable role of CAF genes as biomarkers within the clinic, strategies to target CAFs based on population subtyping has been of interest, despite the challenge of this approach. For example, sibrotuzumab, a humanized antibody against FAP, has been tested in Phase I and Phase II clinical trials among patients with metastatic CRC, and demonstrated a lack of severe toxicity and efficacy[83,84]. Nab-paclitaxel in combination with gemcitabine was found to inhibit collagen-I production and release of inflammatory cytokines CXCL10 and IL-6 in pancreatic cells and CAF co-culture models[85,86]. Furthermore, nab-paclitaxel in combination with atezolizumab, a PD-LI inhibitor, was found to increase OS among metastatic lung cancer patients[87]. Targeted-therapies against CAFs or CAF-induced activity within the TME show promise to complement standard-of-care treatments, but emphasize the need to identify specific CAF subpopulations involved in conferring therapy resistance in GEA given the heterogenous nature of CAFs.

1.4. Project rationale, objectives, & hypothesis

Rationale. The poor prognosis of GEA patients despite docetaxel-based chemotherapy (DCF), the current standard-of-care treatment, demonstrates an urgent need for alternative, targeted approaches and better biomarkers for this disease. Along with genetic and non-genetic alterations, a better understanding of the interactions within the TME is necessary to identify the factors conferring resistance to chemotherapy. Specifically, CAFs present as the most abundant and heterogenous cell type in the TME, and distinct subtypes have shown to interact with tumor cells to induce therapy resistance in other cancer types. However, studies on the identification of distinct CAF populations are scarce and still remain unclear.

Aim & hypothesis. The following research aims to characterize CAF subtypes in GEA and investigate their role in the patients' chemoresponse. It is hypothesized that heterogenous CAF populations modulate the response to chemotherapy in GEA, such that CAFs from chemo-sensitive and chemo-resistant patients have distinct genomic profiles.

Objective 1. Select patient cohort based on pathological response and availability of treatment-naïve (biopsy) fibroblasts and patient-derived tumor organoids (PDOs).

Objective 2. Validate known CAF markers based on a review of literature of other cancer types in GEA patient-derived treatment naïve biopsy fibroblasts.

Objective 3. Characterize CAF populations using an scRNA-seq atlas based on data from >30 GEA patients and identify specific CAF markers associated with chemotherapy response.

Objective 4. Establish a PDO-CAF 3D co-culture model for future *in vitro* drug screening with standard-of-care and alternative therapies.

CHAPTER 2: EXPERIMENTAL DESIGN & METHODOLOGY

2.1. Study design & patient cohort

2.1.1. Primary tissue sample collection

The McGill University Esophageal, Gastric, Colorectal Data- and Bio-bank (EGCDB) consists of tumor samples collected from >350 GEA patients that have been treated at the Montreal General Hospital Thoracic Surgery Department between January 2018 and February 2023.

GEA tumor and tumor-adjacent normal tissue samples were often collected at three timepoints for patients that received neoadjuvant chemotherapy: 1) initial endoscopic biopsy (Tbio), pre-treatment; (2) endoscopic biopsy, mid-treatment (Tmid); (3) surgical resection (Tsur), post-treatment. Patients who did not receive neoadjuvant treatment and went directly to surgical resection had a treatment-naïve Tsur sample collected. Tissue sample collection from patients was performed under the MUHC Research Ethics Board guidelines, and written informed consent was received from all patients prior to use of their samples.

Clinicopathological information (tumor grade, tumor stage, tumor site, HER2 status, mismatch repair protein status, pathological and clinical response to neoadjuvant treatment) was retrieved from the OACIS Clinical Information System. The clinical objective response rate was determined by Dr. James Tankel (Thoracic Surgery fellow) using computerized tomography (CT) scans of the tumor at the time of biopsy and at the time of surgical resection.

2.1.2. Patient cohort

40 patients from EGCDB were selected to be part of this retrospective cohort based on tumor location, tumor type, neoadjuvant treatment, availability of treatment naïve patient-derived organoids (PDOs) and fibroblasts at the time of biopsy, and information regarding clinical and

pathological response to neoadjuvant treatment. Patients whose biopsy tumor samples had scRNA-seq data were prioritized to be included in the cohort (Supplementary Table 1).

2.1.3. Primary tissue dissociation for cell culture of PDOs and CAFs

After identification of tumor and tumor-adjacent normal regions from tissue samples by a pathologist, tissue dissociation was performed by the research technician. Samples were washed with sterile AD-DF+++ media **(Supplemental Table 1)** to eliminate the excess blood, and necrotic parts of the tumor were removed with razor blades. The tissue was divided into pieces for subsequent experiments. One piece of the tissue sample was used for generation of single cells required for fibroblast and organoid cultures.

Tissue dissociation. Viable tissues were minced with razor blades and forceps in 5 mL of tissue dissociation solution: Collagenase III (2 mg/ml), (Worthington, LS004182), Hyaluronidase (100 U/ml) (Sigma, H3884), Primocin (10 g/ml) (Invivogen, ant-pm-1), AD-DF+++ media. Minced tissue mix was collected into a MACS-C tube (Miltenyi Biotec, 130096334). C-tubes were incubated in GentleMACS^{TM} Octo Dissociator with heaters (Miltenyi Biotec) and program A was run, followed by program B (according to company's tissue dissociation protocol). Following tissue dissociation, 10 mL of PBS + 1mM DTT (Sigma Aldrich, D0632), was added to the tube and pipetted up and down to mix. Cells were passed through a 100 µm cell strainers (Fisherbrand, 22363549) into a sterile 50 mL tube. The sample was then transferred to a 15 mL tube and centrifuged at 500g for 5 min at 4°C. The supernatant was removed with a pipette and the pellet was resuspended in 1 mL of 0.25% trypsin, followed by incubation in a 37°C water bath for 5 min. 9 mL of PBS/10% FBS + 1mM DTT was then added to the cell suspension. The sample was then aliquoted into two 15 mL tubes. 1 mL of solution was taken from the 15 mL tube (10ml cells+PBS+10%FBS) and transferred into a new 15 mL conical tube, then centrifuged at 500g for

5 min at 4°C. The supernatant was removed, and the pellet was resuspended in 4 mL AD-DF+++. The solution was left to stand for 10 minutes before transferring 2 mL of the upper part of the supernatant to a new 15 mL conical tube (this portion is used for fibroblast culture). Both fractions were centrifuged at 500g for 5min at 4°C.

Fibroblast propagation. The supernatant was removed, and 1 mL of fibroblast media **(Supplemental Table 1)** was added to the cells, then transferred onto a Purecol (Advanced BioMatrix, #5005) coated 24-well plate.

Organoid propagation. The supernatant was removed with a pipette, leaving the pellet behind. The tube was placed on ice for a few minutes before adding 80 µL of Matrigel (Corning, 356231). The Matrigel+cells were then plated onto a 24-well plate, allowed to rest at room temperature for 1-2 minutes, before transferring to an incubator (37°C, 3% O_2) for 5-10 minutes to allow the domes to solidify. 600 µL of complete growth media **(Supplemental Table 1)** was gently added to each well.

2.2. Single-Cell RNA Sequencing

2.2.1. RNA collection, extraction, and sequencing

Sample processing. An aliquot of the dissociated tissue sample solution previously mentioned, the sample was centrifuged at 500 g for 5 min at 4°C. The supernatant was removed with a 5 mL pipette, and the pellet was resuspended in 0.5 mL of Dispase (Sigma 10269638001) and 10 µL DNase I (1mg/ml), then incubated in a 37°C water bath for 5 min. 9 mL of PBS was then added and the solution was centrifuged at 500 g for 5 min. The supernatant was removed with a pipette and the pellet was resuspended in 1 mL of ACK lysis buffer (Invitrogen, A1049201) for 5 min at room temperature. The solution was then passed through a 40 µm cell strainer (Falcon, 352340) and then centrifuged at 500 g for 5 min. The supernatant was removed, and the cells were washed

with 1 mL of 2% FBS-PBS, then transferred to a 1.5 mL microtube. The solution was centrifuged for 5 min at 500g. The wash process was repeated twice and then the cells were resuspended in 100 µL of 2% FBS-PBS. RNA sequencing, quality check, and downstream analysis was performed by McGill Genome Center.

2.2.2. scRNA-seq data processing

Creation of a scRNA-seq browser was conducted by bioinformaticians at the Institute of Systems Biology, Seattle, US.

Read processing and alignment. After polyA-trimming via cutadapt (v3.2) [Martin2011], reads were pseudoaligned to the GRCh38 reference transcriptome (ENSEMBL release 96) with kallisto (v0.46.2) [Bray2016] using the default kmer size of 31. The pseudoaligned reads were processed into a cell-by-gene count matrix using bustools (0.40.0) [Melsted2021]. Cell barcodes were filtered using the whitelist (v3) provided by 10xGenomics. All further processing was done in scanpy (v.1.7.1) [Wolf2018].

Quality control and normalization. Quality control was performed for each sample independently as followed: Cell barcodes with less than 1000 counts or less than 500 genes expressed were removed. Cell barcodes with more than 10% mitochondrial gene expression were removed. Doublet cells were identified using scrublet [Wolock2019], removing any cell barcode with a scrublet score > 0.2. Only coding genes were retained in the final count matrix. Expression profiles were normalized by total counts, the 4000 most highly variable genes identified [Zheng2017], renormalized, log- transformed and z-scored. The data was projected onto the first 50 principal components. After the above per-sample preprocessing, samples are pooled and integrated using Harmony [Korsunsky2019] on the first 50 principal components with a maximum of 25 iterations. A nearest neighbor graph (k=15) was calculated on the harmony corrected principal components

space. Datasets were visualized in 2D via UMAP [McInnes2018] and initialized with PAGA [Wolf2019] coordinates. The nearest neighbor graph was clustered with the Leiden algorithm [Traag2019]. The processed datasets were visualized interactively using Cellbrowser [Speir2021] allowing for easy access and exploration across teams and labs.

2.2.3. Differential gene expression analysis and pathway analysis

The top 50 significantly upregulated genes in each of the clusters of the treatment-naïve biopsy fibroblast dataset were identified using pairwise differential analysis defined by a z-score. Clustering of CAFs according to their function was performed using gene ontology analysis with Metascape tool (http://metascape.org) using all significantly upregulated genes (log p-value<0.01) in each of the 14 clusters. Variance stabilized transform (VST) analysis was performed to identify DEGs between pre- and mid-treatment timepoints according to pathological and clinical response.

2.3. Validation in primary fibroblasts

2.3.1. Maintenance of fibroblasts

Collagen-coated petri dishes. Fibroblasts were plated on petri dishes coated with Purecol to facilitate cell adherence. A 1:30 Purecol:1X PBS was made in a 50 mL tube. The volume added to each well was dependent on the size of the dish: 400 µL per well for 24-well plates, 1 mL per well for 6-well plates, and 5 mL for 10 cm petri dish. The plates were incubated for at least 30 minutes, ideally 2 hours, at 37^{o}C. The solution was then removed, and the plates were washed twice with 5 mL of 1X PBS, and then all the liquid was aspirated before storing the plates at 4^{o}C. Plates could be stored in the fridge for several weeks.

Passaging fibroblasts. Once the 10 cm petri dish reached 70-80% confluency, the fibroblasts were passaged. Media was carefully removed, and the cells were washed once with 5 mL of 1X PBS

(Gibco, 10010023). Fibroblasts were detached from the plate using 2 mL 0.25% trypsin (Gibco, 15050- 065) and incubated at 37°C for 90 seconds or until completely detached. 4 mL of 5% FBS-PBS was added to the dish to inactivate the trypsin, and the cells were transferred to a 15 mL labeled tube. Cells were centrifuged at 500xg for 5 minutes. The supernatant was removed, and the cells were resuspended in 1 mL of fibroblast media.

Cell viability. Cell counting and viability were performed using an automated cell counter (Invitrogen Countess 3). 10 µL of cells were mixed with 10 µL of trypan blue (Gibco, 15250-061) and pipetted into the counting slides (Invitrogen, C10283) before inserting into the machine.

Plating fibroblasts. Fibroblasts were plated on pre-coated 10 cm petri dishes with 8 mL of fibroblast media, then placed in a HERAcell V10S 160i incubator (Thermo Scientific) at 37°C, 3% O_2, and 5% CO_2.

Freezing fibroblasts. Cells were resuspended in 90% FBS-10% DMSO and then 500 µL were distributed into 1.5 mL cryovials (Sarstedt, 9081911), labelled with the biobank ID, cell type, passage number, and date. The cryovials were placed in CoolCell freezing containers (Corning, 432002) at -80°C for storage.

2.3.2. Immunofluorescence

20,000 cells in 500 µL of fibroblast media were plated on collagen-coated coverslips placed in wells on a 24-well dish. Cells were grown until they reached 50-60% confluency.

Fixation and permeabilization. Media was aspirated from each well and the cells were washed twice with 500 µL of 1X PBS. The cells were immediately fixed with 500 µL of 4% PFA (ThermoScientific, 218957) for 10 minutes at room temperature. The solution was aspirated, and the cells were washed with 500 µL of PBS, repeating 3x with 5 minutes between each wash. 500

µL of 0.1% Triton X-100 in PBS was added to each well to permeabilize the cells for 10 minutes at room temperature. Following aspiration, the cells were washed with 500 µL of PBS, 3x with 5 minutes between each wash.

Blocking buffer preparation. First, 50 mL of 1X TBST solution was made using 5 mL of 10X TBS (Fisher Bioreagents, 191513), 45 mL of sterile water, and 50 µL of Tween-20 (Fisher Bioreagents, 190656). Taking a 20 mL aliquot of the 1X TBST solution, 1 µL of 0.01% Triton-X, 0.2 g of 2% BSA (Sigma Aldrich, 10735078001) and 200 µL of FBS (Invitrogen, 12483-020) were added to make the blocking buffer. To dissolve the BSA powder in solution, a small aliquot of the 1X TBST was added to the solute separately before adding it to the rest of the solution. 500 µL of the blocking buffer was then added to each well and left to rest for 1 hour at room temperature.

Primary antibody staining. The selected antibodies **(Supplemental Table 2)** were diluted at the indicated ratio in blocking buffer and 100 µL was added to each well. Cells were left to incubate overnight at 4^{o}C.

Secondary antibody staining. The primary antibody solution was aspirated, and the cells were washed 3x with 500 µL PBS, keeping 5 minutes between each wash. 100 µL of the diluted secondary antibody solution in PBS was added to each well. Cells were left to incubate for 1 hour at room temperature. From this point on, the plate was kept in the dark.

Slide preparation. The cells were then washed with 500 µL PBS, 3x with 5 minutes in between each wash, and then a quick wash with 500 µL of sterile water. Coverslips were mounted onto labelled slides (Fisherbrand, 12-550-15), using mounting medium with DAPI (Vectashield Hardset Antifade, H-1500-10) and forceps. The slides were left to dry overnight at 4^{o}C in a slide box.

Confocal imaging and analysis. Images of fluorescently stained fibroblasts were taken using a Zeiss LSM 880 confocal microscope with the 20x objective. Measurement of fluorescence intensity was performed using ImageJ software developed by the National Institute of Health.

2.3.3. Fluorescence-activated cell sorting (FACS)

Sample preparation. Fibroblasts were passaged as mentioned above and resuspended in 1 mL of PBS. An aliquot of the sample was taken, and heat shocked at 65°C in a water bath for 5 min for the viability control. The rest of the sample was aliquoted into labeled blue capped-FACS tubes (Falcon, 352235): main *sample (3 tubes)*, *fluorescence minus one (5 tubes)*, *viability control (1 tube)* and *unstained control (1 tube)*. 500 µL PBS was added to each tube and pipetted up and down. The samples were centrifuged at 300g for 5 min at 4°C. While the samples were being centrifuged, the viability dye **(Supplemental Table 2)** diluted at 1:1000 in PBS was prepared on ice in the dark. The samples were decanted to get rid of the supernatant, keeping only the pellet at the bottom. 100 µL of Aqua blue 405nm viability solution was added to each tube and left to incubate for 30 min on ice in the dark. 500 µL of PBS was added to each tube to wash the dye, and then spun down at 300g for 5 min at 4°C. While cells were spinning, antibodies **(Supplemental Table 2)** were diluted in 100 µL of PBS on ice in the dark in microtubes accordingly: (1) all markers, (2) FMO minus FAP, (3) FMO minus PDGFRa, (4) FMO minus PDGFRb, (5) FMO minus CD74, (6) FMO minus PDPN, (7) FMO minus viability, (8) single stain FAP, (9) single stain PDGFRa, (10) single stain PDGFRb, (11) single stain CD74, (12) single stain PDPN. 100 µL of antibody solution was added to their appropriate tube, and incubated for 30 min on ice in the dark. 500 µL of PBS was added to each tube to wash the dye, and then spun down at 300g for 5

min at 4°C. The supernatant was decanted, keeping only the cells at the bottom. The cells were then resuspended in FACS buffer **(Supplemental Table 2)**.

Flow cytometry and analysis. For each sample, 30,000-50,000 events were acquired using the BD FACS Canto machine. Analysis of flow cytometry results were performed using FlowJo software developed by BD Biosciences.

2.4. 3D PDO AND CAF co-culture model

2.4.1. Establishment and maintenance of PDOs

Passaging PDOs. Once the organoids were mature, passaging was performed under sterile conditions. The growth medium was removed and replaced with 500 µL of Matrigel digestion buffer: 1:100 dilution of 100mg/ml stock Collagenase/Dispase (Sigma, 10269638001) with AD-DF+++ medium. The dome was broken by pipetting up and down. The plate was then incubated at 37°C for 1 hour. After digestion, the media was collected in a 15 mL tube and centrifuged at 500g for 5 min at 22°C. The supernatant was then removed and replaced with 1 mL of 0.25% trypsin. The tube was put into a 37°C water bath for 5 min. The cells were quickly vortexed until the organoids were broken down into single cells. To inactivate the trypsin, 4 mL of 5% FBS-PBS was added and gently vortexed before centrifuging at 500 g for 5 min at 22°C. The supernatant was removed and resuspended in 1 mL of AD-DF+++ medium. Cell viability and count was assessed as mentioned previously for fibroblasts.

Plating PDOs. The required volume (32,000 cells per dome) was aliquoted into a 1.5 mL microtube, then centrifuged at 500 g, 5 min, 4°C. The supernatant was removed, and the cells were resuspended in Matrigel, then plated as domes on a 24-well plate, as previously mentioned for organoid propagation.

Freezing PDOs. Cells were resuspended in CryoStor media (Stemcell) and transferred into CryoPure (Sarstedt, 9081911) 1.8 mL cryovials. The cryovials were placed in CoolCell freezing containers (Corning, 432002) for storage at -80^{o}C.

2.4.2. PDO-CAF co-culture

Tumor-derived organoids and fibroblasts were individually passaged as previously mentioned. The co-culture model was performed at a 2:1 (fibroblast to organoid) ratio, in a 96-well plate. 10,000 single cells from organoids were mixed with 20,000 fibroblasts for each well and mixed in a 1.5 mL microtube. The cells were centrifuged at 450g for 5 min at 4°C. The supernatant was removed and then resuspended in a 1:7 Purecol:Matrigel solution. The volume of Matrigel was first added to the pellet before adding Purecol. 30 µL of the suspension was added to each well in a 96-well plate, then was left to incubate at 37 °C, 3% O_2 for 30 minutes. 100 µL of complete growth medium was then added to each well. Co-culture models were left to grow for 3-5 days for subsequent drug screening or marker staining.

2.5. Brightfield imaging

Brightfield imaging of organoids and were captured using the EVOS M7000 microscope. The images were taken at 4x (650 µm), 10x (275 µm), 20x (150 µm), and 40x (75 µm) objectives.

2.6. Drug screening

2.6.1. Drug stocks

Docetaxel (Cayman, 11637) and 5-FU (Cayman, 14416) were diluted in DMSO (Sigma, SHBL2891), and Cisplatin (Cayman, 13119) was diluted in 0.9% saline. The ratios used for DCF were 1:1:10, respectively.

2.6.2. Drug treatment

CAF-PDO co-culture models were plated on 96-well plates as previously mentioned, in 10 wells, followed by control wells with organoids only (30,000 cells) and fibroblasts (30,000 cells). Following PDO and CAF growth, after 3-4 days, the cells were supplemented with 100 µL of complete gastric media, and were then treated with aliquots of Docetaxel, Cisplatin, and 5-FU (DCF) using the Tecan D300e digital dispenser, according to the company protocol. The drug dispenser was put into the sterile tissue culture hood, a T8+ Dispensehead Cassette (Tecan, 30097370) was loaded onto the dispenser, and a protocol with 10 different concentrations and normalization fluids (DMSO + 0.01% Tween 20) was implemented. After drug addition, the cells were incubated for 72 hours in the 37 °C/3% O_2 incubator, and were then imaged to compare control wells to drug-treated wells.

Drug screening on fibroblasts was conducted to determine their response to DCF. 15,000 fibroblasts were plated in 30 wells on the 96-well plate. The protocol above was implemented for the plating and drug treatment. Following drug treatment, cell viability was assessed using a CellTiter Glo 3D Cell Viability Assay (Promega, G9683), and the luminescence protocol which measured the cell metabolic ATP activity, was obtained using the Varioskan Lux plate reader (ThermoScientific). GraphPad Prism software was used to generate dose-response curves and IC50 concentrations.

2.7. Statistical analysis

Statistical analyses and graphs were conducted using GraphPad Prism software (v9.2.0) and Excel. Statistical significance was determined by a P-value of < 0.05. Statistical details are elaborated under each figure legend.

CHAPTER 3: RESULTS

3.1. Clinicopathological characteristics of select patient cohort

3.1.1. Characteristics of patient cohort

A cohort of 40 GEA patients was selected from a large biobank consisting of >350 GEA patient samples treated between January 2018 and February 2023 **(Table 1)**. The main criteria for selecting this cohort were the tumor type (adenocarcinoma), tumor site (esophageal/EGJ), treatment with a docetaxel-based triplet neoadjuvant chemotherapy, tumor pathological response to neoadjuvant chemotherapy, availability of Tbio fibroblasts, Tbio PDOs, and scRNA-seq data **(Figure 1A and Supplemental Table 3)**. With regards to the latter three criteria, not all patients had available Tbio fibroblasts, PDOs, or scRNA-seq, but were included if they had at least one of the three. Other characteristics that were noted include patient age, sex, mismatch repair protein status (microsatellite stable, MSS; or MSI-H), HER2 status (positive or negative), Response Evaluation Criteria in Solid Tumors (RECIST) (complete response, CR; partial response, PR; stable disease, SD) tumor grade (well, moderately, or poorly differentiated), clinical score (TNM), and pathological score (TNM).

The patients were categorized by their pathological response to chemotherapy: 14 patients (TRG0/1 good), 12 patients (TRG2 moderate), and 14 patients (TRG3 poor). The median age of patients was 69.7 years (range, 50-86 years). Most patients within this cohort were male (31 males vs. 9 females), with all patients having a tumor at the distal esophagus or EGJ site. 50% of patients presented with a poorly differentiated tumor and 35% patients had a stable disease, whereby no reduction of the tumor was seen following neoadjuvant chemotherapy. Recurrence of the tumor and/or metastasis occurred among 25% of patients, of which 60% were pathological poor responders.

Table 1. Clinical characteristics of patient cohort. 40 GEA patients separated based on pathological response following treatment with docetaxel-based triplet neoadjuvant therapy.

* refers to the combination of DCF with other treatments, such as avelumab. Other treatments provided include carboplatin+paclitaxel and FOLFOX.

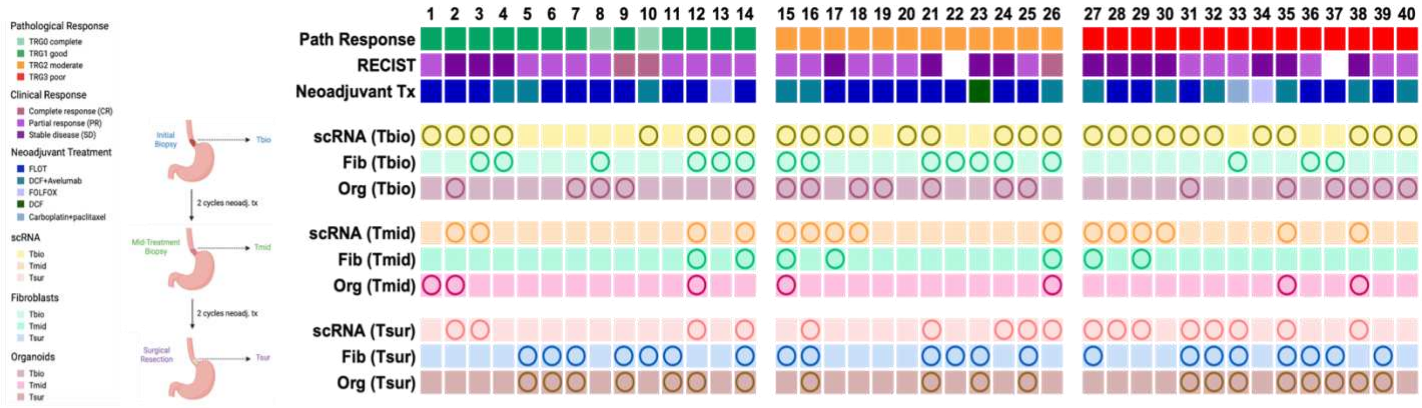
Clinical Characteristics	Total (n = 40)	Good Response (n = 14)	Moderate Response (n = 12)	Poor Response (n = 14)
Age, mean (range)	69.7 (50– 86)	67.5 (50 – 79)	71.1 (61 – 76)	70.7 (54 – 86)
Sex				
Male	31 (78%)	11 (79%)	9 (75%)	11 (79%)
Female	9 (22%)	3 (21%)	3 (25%)	3 (21%)
Tumor site				
Distal esophagus/EGJ	40 (100%)	14 (100%)	12 (100%)	14 (100%)
Neoadjuvant Treatment				
DCF*	13 (30%)	3 (21%)	4 (33%)	6 (43%)
FLOT	24 (62%)	10 (72%)	8 (67%)	6 (43%)
Other	3 (8%)	1 (7%)	0 (0%)	2 (14%)
Tumor Differentiation				
Poorly differentiated	20 (50%)	5 (36%)	7 (58%)	8 (57%)
Moderately differentiated	12 (30%)	2 (14%)	4 (33%)	6 (43%)
Well differentiated	6 (15%)	5 (36%)	1 (8%)	0 (0%)
No tumor at resection	1 (2.5%)	1 (7%)	0 (0%)	0 (0%)
Unknown	1 (2.5%)	1 (7%)	0 (0%)	0 (0%)
Treatment response, pathology				
TRG0 complete	2 (32%)	2 (14%)	0 (0%)	0 (0%)
TRG1 good	12 (22%)	12 (86%)	0 (0%)	0 (0%)
TRG2 moderate	12 (24%)	0 (0%)	12 (100%)	0 (0%)
TRG3 poor	14 (22%)	0 (0%)	0 (0%)	14 (100%)
Treatment response, clinical				
Stable disease	14 (35%)	3 (22%)	4 (34%)	7 (50%)
Partial response	21 (53%)	9 (64%)	6 (50%)	6 (43%)
Complete response	3 (7%)	2 (14%)	1 (8%)	0 (0%)
Unknown	2 (5%)	0 (0%)	1 (8%)	1 (7%)
HER2				
Negative	37 (93%)	13 (93%)	11 (92%)	13 (93%)
Positive	2 (5%)	0 (0%)	1 (8%)	1 (7%)
Unknown	1 (2%)	1 (7%)	0 (0%)	0 (0%)
Metastasis Sites				
0	30 (75%)	13 (93%)	9 (75%)	6 (50%)
≥1	10 (25%)	1 (7%)	3 (25%)	6 (50%)

3.1.2. Patient survival outcome and clinical progression

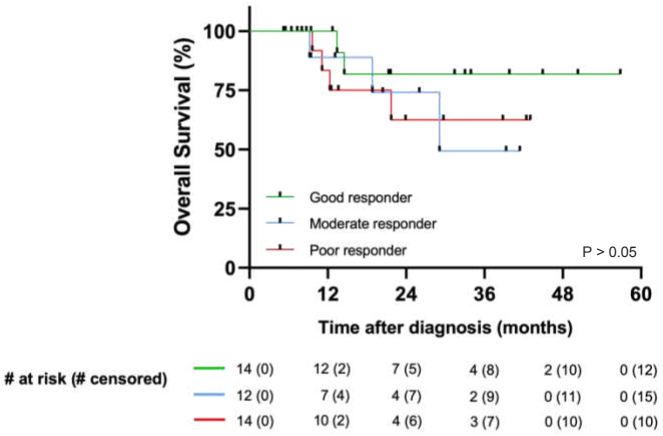
OS and disease-free survival (DFS) of patients within this cohort were determined and compared to pathological response to neoadjuvant chemotherapy **(Figure 1B and 1C)**. Given that patients were admitted to the clinic at variable timepoints, length of follow-up was not uniform. OS was determined by the time at which the patient was diagnosed with GEA until follow-up time end or death. **Figure 1B** demonstrates a non-significant difference between the survival of 14 good responders, 12 moderate responders, and 14 poor responders (p=0.2967). Similarly, **Figure 1C** demonstrates the difference in DFS between pathological good, moderate, and poor responder patients. A non-significant difference can be seen between 14 good responders, 12 moderate responders, and 14 poor responders (p=0.1618) to neoadjuvant chemotherapy.

The clinical progression of patients within this cohort was followed from the start of their diagnosis to their last follow-up or death, separated according to pathological response to neoadjuvant chemotherapy **(Figure 1D)**. The number assigned to each patient corresponds to the number denoted on **Supplementary Table 3**. In particular, the clinical timeline highlights the occurrence of metastasis and disease recurrence events among poor responder patients compared to good responders.

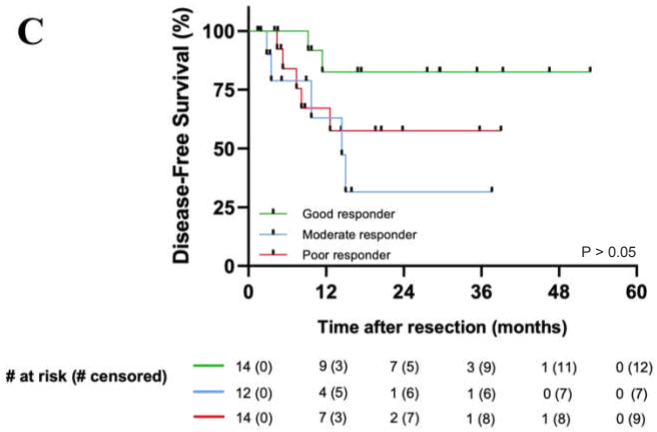
A


Pathological Response
TRG0 complete
TRG1 good
TRG2 moderate
TRG3 poor
Clinical Response
Complete response (CR)
Partial response (PR)
Stable disease (SD)
Neoadjuvant Treatment
FLOT
DCF+Avelumab
FOLFOX
DCF
Carboplatin+paclitaxel
scRNA
Tbio
Tmid
Tsur
Fibroblasts
Tbio
Tmid
Tsur
Organoids
Tbio
Tmid
Tsur
Initial Biopsy
Tbio
2 cycles neoadj. tx
Mid-Treatment Biopsy
Tmid
2 cycles neoadj. tx
Surgical Resection
Tsur
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
Path Response
RECIST
Neoadjuvant Tx
scRNA (Tbio)
Fib (Tbio)
Org (Tbio)
scRNA (Tmid)
Fib (Tmid)
Org (Tmid)
scRNA (Tsur)
Fib (Tsur)
Org (Tsur)

B


Overall Survival (%)
100
75
50
25
0
Good responder
Moderate responder
Poor responder
P > 0.05
0 12 24 36 48 60
Time after diagnosis (months)
at risk (# censored)
14 (0) 12 (2) 7 (5) 4 (8) 2 (10) 0 (12)
12 (0) 7 (4) 4 (7) 2 (9) 0 (11) 0 (15)
14 (0) 10 (2) 4 (6) 3 (7) 0 (10) 0 (10)

C


Disease-Free Survival (%)
100
75
50
25
0
Good responder
Moderate responder
Poor responder
P > 0.05
0 12 24 36 48 60
Time after resection (months)
at risk (# censored)
14 (0) 9 (3) 7 (5) 3 (9) 1 (11) 0 (12)
12 (0) 4 (5) 1 (6) 1 (6) 0 (7) 0 (7)
14 (0) 7 (3) 2 (7) 1 (8) 1 (8) 0 (9)

D

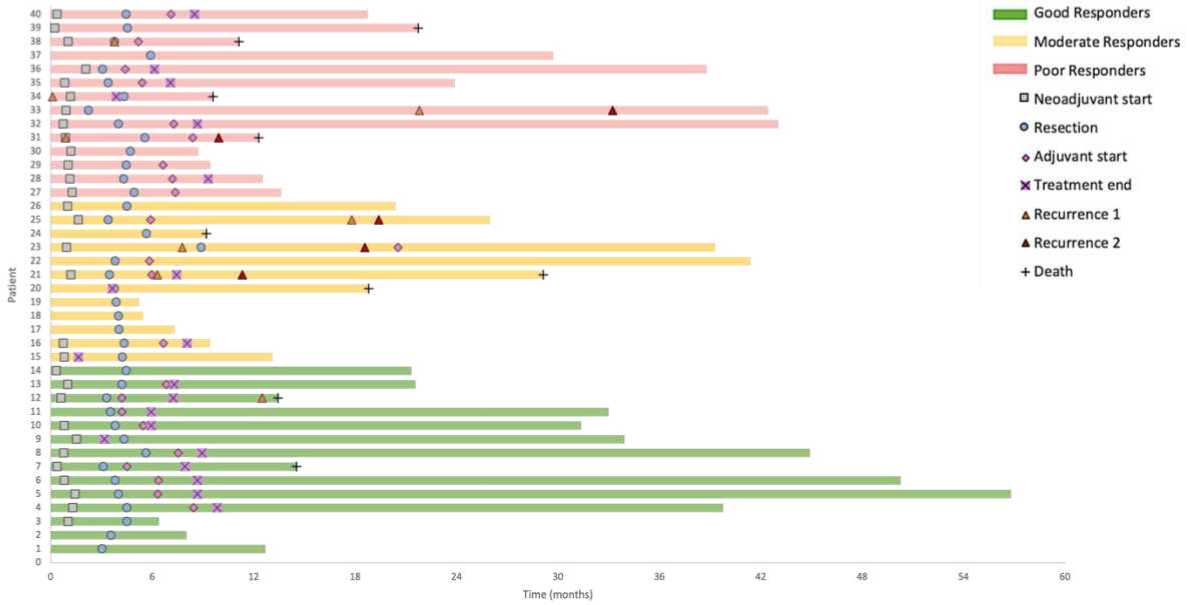

Patient
Time (months)
0 6 12 18 24 30 36 42 48 54 60
Good Responders
Moderate Responders
Poor Responders
Neoadjuvant start
Resection
Adjuvant start
Treatment end
Recurrence 1
Recurrence 2
Death

Figure 1. Clinical survival and progression of patient cohort based on pathological response to neoadjuvant docetaxel-based triplet chemotherapy. (A) Experimental design of study. Tumor sample was collected at initial endoscopic biopsy (Tbio), followed by mid-treatment endoscopic biopsy (Tmid) after two cycles of neoadjuvant chemotherapy, and finally at surgical resection (Tsur) after completion of two more cycles of neoadjuvant chemotherapy. Pathological and clinical response, and availability of scRNA-seq analysis, primary fibroblasts, and patient-derived organoids were noted for all 40 patients within the cohort, indicated by the presence of a circle symbol. Undefined information is denoted by a blank space. (B) Kaplan Meier plot for overall survival (%) of select patients as a function of time after diagnosis (in months) according to pathological response (TRG0/1 good, green; TRG2 moderate, blue; TRG3 poor, red). (C) Kaplan Meier plot for disease-free survival (%) of patient cohort based on time after resection (in months), distinguished by pathological response (TRG0/1 good, green; TRG2 moderate, blue; TRG3 poor, red). Statistical analysis was performed using a log-rank test for trend and significance was defined by a p-value < 0.05. The number of patients at risk and the number of patients censored at 12-month increments is denoted beneath the survival curves. (D) Swimmer plot demonstrating patient clinical progression as a function of overall survival (in months) according to pathological response (TRG0/1 good, green; TRG2 moderate, yellow; TRG3 poor, red).

3.2. Validation of known CAF markers based on review of literature in treatment naïve biopsy patient-derived fibroblasts

3.2.1. Phenotypic and functional heterogeneity of CAFs reported in the literature

CAF heterogeneity has been reported in several other cancer types, including pancreatic and breast, and their various functions have been extensively reported in the literature. **Figure 2** represents a comprehensive review of commonly reported CAF markers in the literature associated to their functions in the TME. Given that subtyping and phenotypic characterization of CAFs remains unclear in GEA, this review was performed as a preliminary step to select known CAF

markers for further validation of their expression in primary fibroblasts from GEA patients **(Supplementary Table 4)**.

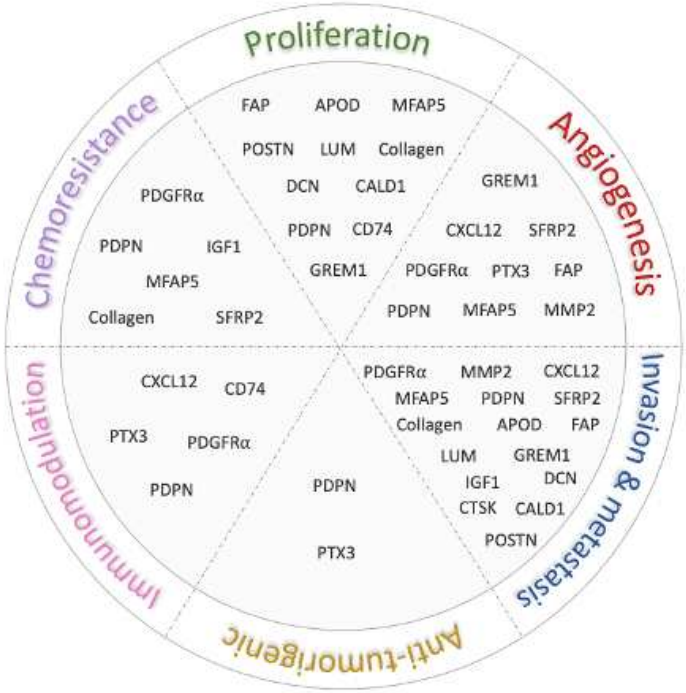


Figure 2. Heterogenous functions of commonly reported CAF markers in the TME. Extensive literature review on 18 known CAF markers relevant in other cancer types associated to six common functions of CAFs in the TME. Certain markers have multiple reported functions.

3.2.2. Validation of known CAF markers in GEA primary fibroblasts identifies significant difference between good and poor responders to chemotherapy

Ten CAF markers were selected for further validation in treatment naïve biopsy fibroblasts using IF staining. **Figure 3A** demonstrates the quantified expression of these CAF markers following imaging with confocal microscopy. Comparison of good (n=1), moderate (n=1),

moderate with metastasis (n=1), and poor (n=1) pathological response identified three markers, VIM, PDGFRb, and S100A4, that were significantly differentially expressed in GEA primary fibroblasts. **Figure 3B** shows the corresponding confocal images of the DEGs between patients of varying pathological response. Decreased VIM expression was associated with poor pathological response, while PDGFRb and S100A4 were found to be upregulated in poor-responder CAFs. Of note, fibroblasts from adjacent normal esophageal tissue were used as a control. Furthermore, differences in CAF morphology and marker expression can be seen not only between patients, but also within a patient's sample.

In addition, FACS was performed on treatment-naïve biopsy fibroblasts to sort based on the expression of five known CAF surface markers **(Figure 4)**. CAFs were first gated from debris based on the forward and side scatter, and then gated on single cells. Live cells were selected to then perform further analysis of the markers. Individual gates were placed for each marker based on their fluorescence-minus-one (FMO) control. This gating strategy was applied to all primary fibroblast samples **(Figure 4A)**. When looking at the percent of fibroblasts expressing each marker in **Figure 4B**, CD74, FAP, and podoplanin (PDPN) were found to be significantly different between good responders (n=3) and poor responders (n=2). CD74 and FAP expression was present in a greater percentage of good-responder fibroblasts compared to moderate and poor responders, while a higher percentage of poor-responder fibroblasts expressed PDPN. The intensity of marker expression was further evaluated in **Figure 4C**, demonstrating a significant upregulation for PDPN in moderate and poor responder patients' fibroblasts compared to good responders' fibroblasts.

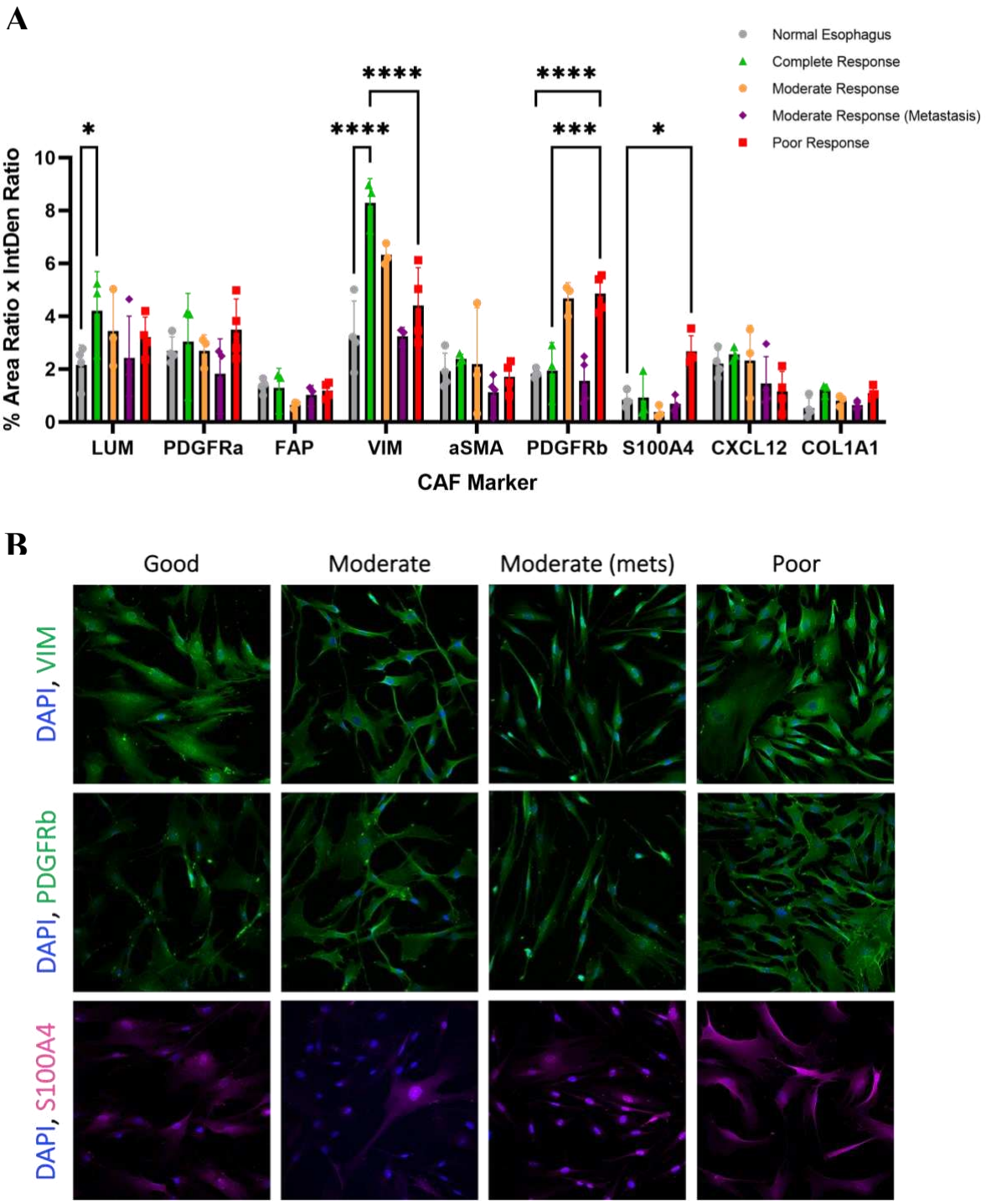
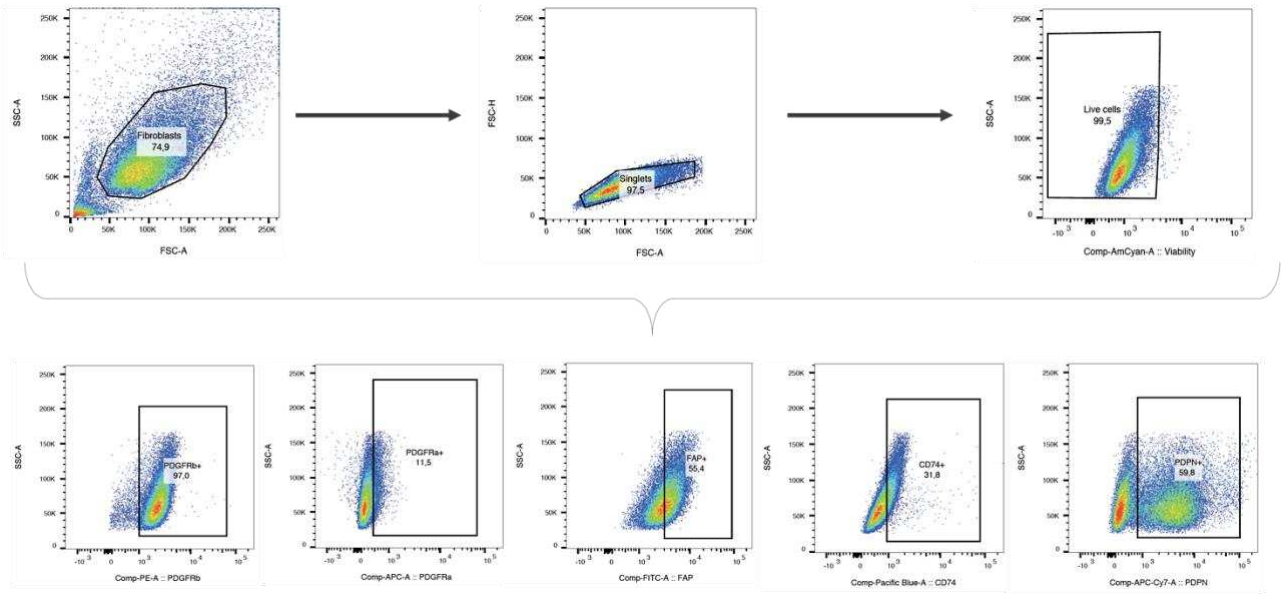


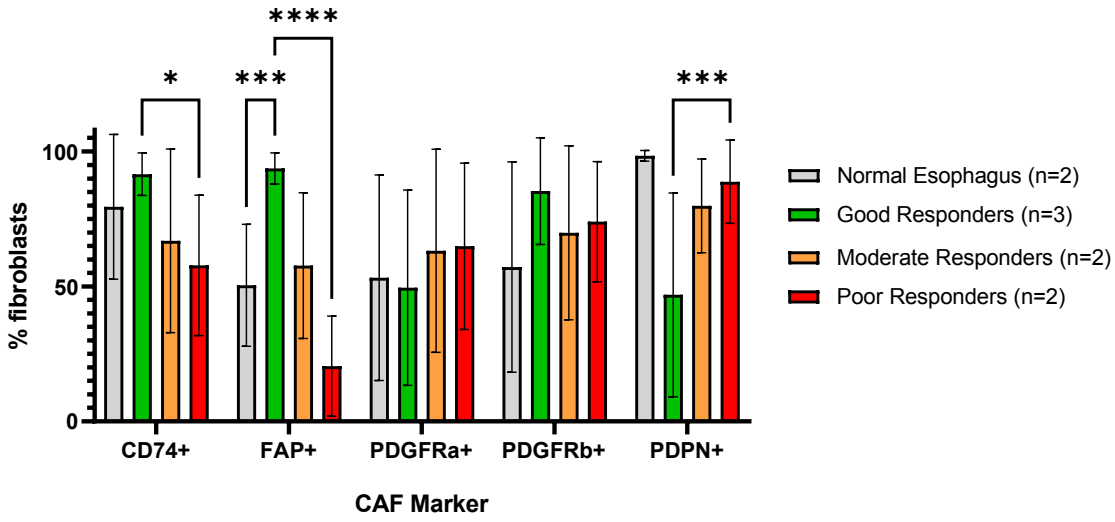
Figure 3. Expression of known CAF markers in primary fibroblasts from treatment naïve biopsy according to pathological response. (A) Quantification of immunofluorescence on treatment naïve biopsy primary fibroblasts (n=5) for nine selected markers, based on response to neoadjuvant docetaxel-based triplet chemotherapy. Marker expression is defined as the (ratio of

the %area of the marker to the %area of DAPI) x (ratio of integrated density of marker to the max integrated density). Statistical analysis was performed using a Two-way Anova test and significance was defined as: *, $p<0.05$; ***, $p<0.001$; ****, $p<0.0001$. (B) Comparison of differentially expressed CAF markers (Vimentin, PDGFRb, S100A4) using immunofluorescence between TRG1 complete, TRG2 moderate, TRG2 moderate with metastasis, TRG3 poor responders to neoadjuvant docetaxel-based triplet chemotherapy.

A


SSC-A
FSC-A
Fibroblasts
74,9
FSC-H
Singlets
97,5
Live cells
99,5
Comp-AmCyan-A :: Viability
PDGFRb+
97,0
Comp-PE-A :: PDGFRb
PDGFRa+
11,5
Comp-APC-A :: PDGFRa
FAP+
55,4
Comp-FITC-A :: FAP
CD74+
31,8
Comp-Pacific Blue-A :: CD74
PDPN+
59,8
Comp-APC-Cy7-A :: PDPN

B


****
*
***
***
% fibroblasts
100
50
0
CD74+
FAP+
PDGFRa+
PDGFRb+
PDPN+
CAF Marker
Normal Esophagus (n=2)
Good Responders (n=3)
Moderate Responders (n=2)
Poor Responders (n=2)

C

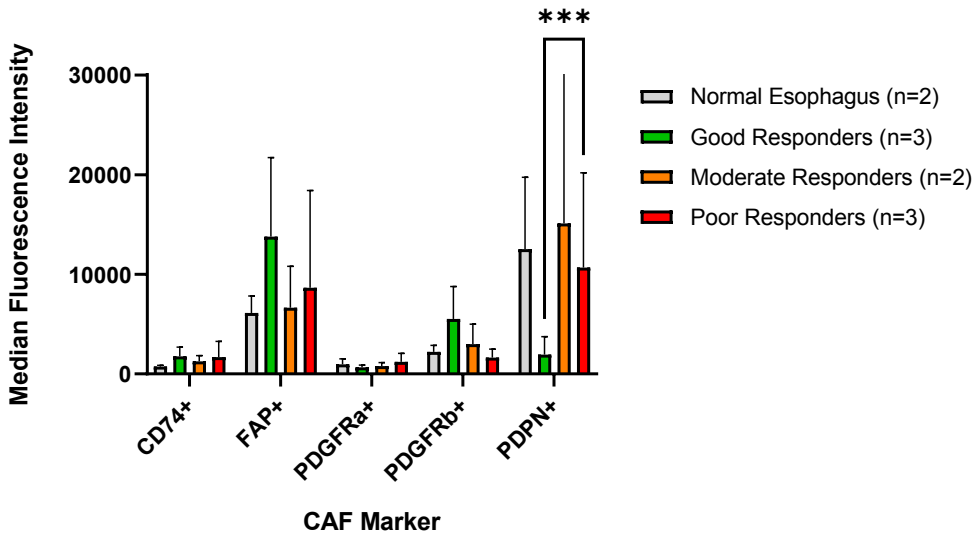

***
Median Fluorescence Intensity
30000
20000
10000
0
CD74+
FAP+
PDGFRa+
PDGFRb+
PDPN+
CAF Marker
Normal Esophagus (n=2)
Good Responders (n=3)
Moderate Responders (n=2)
Poor Responders (n=3)

Figure 4. FACS using known CAF markers on treatment naïve biopsy fibroblasts as a function of pathological response. (A) Gating strategy of a representative TRG3 poor responder GEA patient. (B) % fibroblasts expressing select markers (CD74, FAP, PDGFRa, PDGFRb, and PDPN) according to pathological response (TRG0/1 good, TRG2 moderate, TRG3 poor). (C) Mean fluorescence intensity to define marker expression among select patients according to response. Statistical analysis was performed using Two-way Anova and significance was defined as: *, $p<0.05$; ***, $p<0.001$; ****, $p<0.0001$.

3.3. Identification of two major CAF subpopulations in GEA using single-cell RNA sequencing

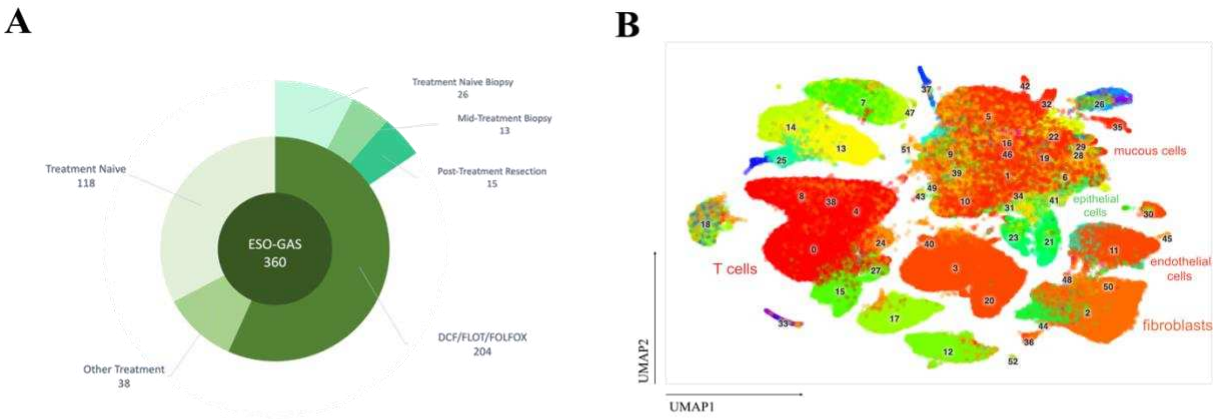
3.3.1. Introduction to single-cell RNA sequencing atlas

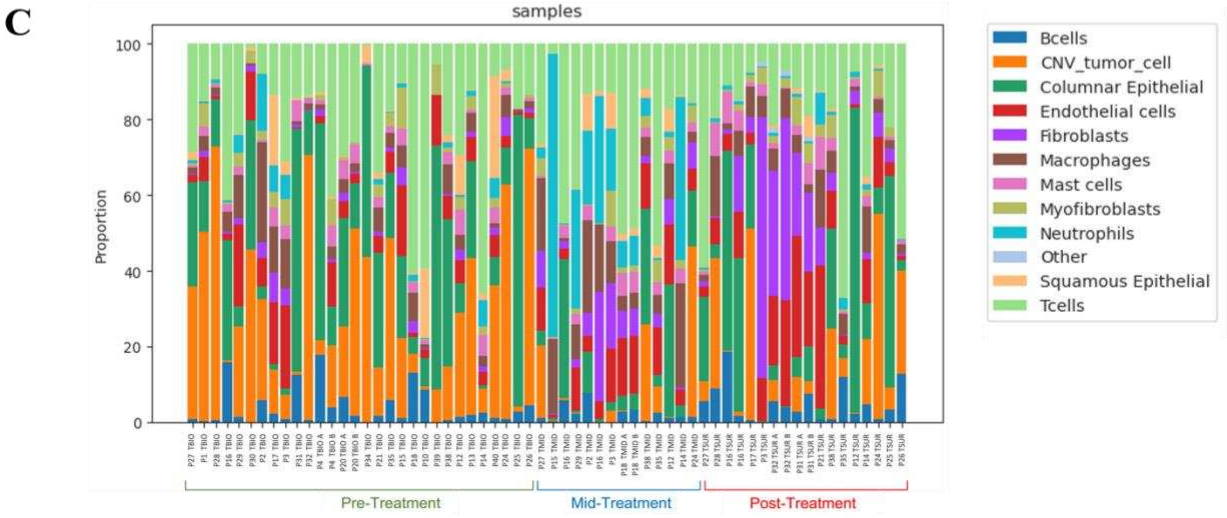
scRNA-seq was performed on tumor-adjacent normal, tumor biopsy, and tumor resection samples to characterize the specific phenotypic and molecular heterogeneity in GEA. From approximately 360 GEA patient samples biobanked by EGCDB, 204 patients received neoadjuvant docetaxel-based triple chemotherapy, of which 46 samples were sequenced **(Figure 5A)**. Following quality control and normalization, 204,047 cells were categorized by correlating each cell's transcriptome to a reference database of known cell types to define the query cell's lineage and to then cluster them accordingly. **Figure 5B** portrays the major cell type clusters, with 9,199 fibroblasts located at cluster 4.

34 patient samples were further evaluated by looking into the proportion of each cell type, comparing between pre-treatment tumor biopsy, mid-treatment tumor biopsy, and post-treatment tumor resection samples **(Figure 5C)**. Extensive heterogeneity can be seen between patient samples of differing pathological response, but also between treatment timepoints for the same patient. Interestingly, the proportion of fibroblasts (purple) was observed to be increased in post-

treatment surgical resection samples for several patients; this was also seen when comparing the combined samples' cell type proportions between pre-, mid-, and post-treatment, notably for good and poor pathological responders **(Figure 5D, top)**. In addition, epithelial cells confirmed to be tumor cells (orange) based on copy-number variation profile (CNV), was found to be of greater proportion in moderate and poor pathological responders compared to good. The proportion of tumor cells was found to generally decrease in mid- and post-treatment samples **(Figure 5D, top)**. Neutrophil infiltration was consistently observed to increase in mid-treatment samples across pathological response, compared to their respective pre- and post-treatment sample **(Figure 5C and 5D, top)**.

Cell type proportions were further compared between pre-, mid-, and post-treatment samples according to clinical response to neoadjuvant chemotherapy **(Figure 5D, bottom)**. Of note, only one patient was classified as CR. Likewise to the comparison made according to pathological response, fibroblast proportion was found to increase in post-treatment resection samples for PR and SD patients compared to pre- and mid-treatment biopsy samples. Additionally, neutrophil populations were found to increase as well in the mid-treatment samples.


A
Treatment Naive Biopsy
26
Mid-Treatment Biopsy
13
Post-Treatment Resection
15
Treatment Naive
118
ESO-GAS
360
DCF/FLOT/FOLFOX
204
Other Treatment
38
B
mucous cells
epithelial cells
endothelial cells
fibroblasts
T cells
UMAP2
UMAP1


C
samples
Proportion
Pre-Treatment
Mid-Treatment
Post-Treatment
Bcells
CNV_tumor_cell
Columnar Epithelial
Endothelial cells
Fibroblasts
Macrophages
Mast cells
Myofibroblasts
Neutrophils
Other
Squamous Epithelial
Tcells

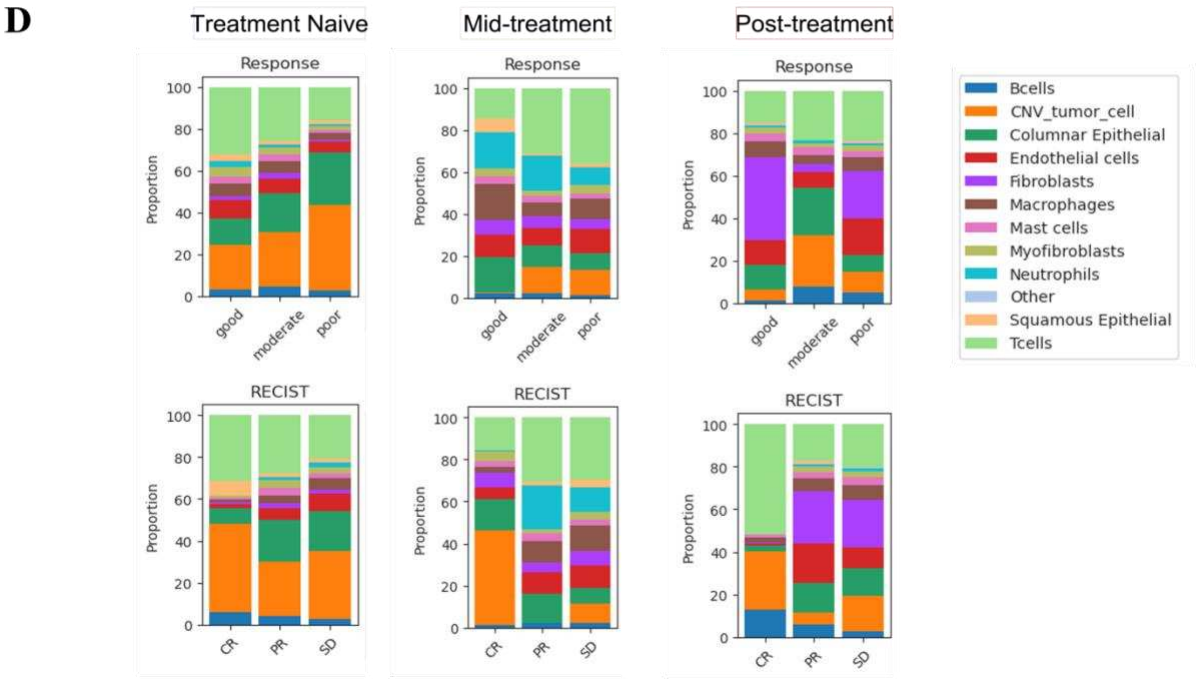

D
Treatment Naive
Mid-treatment
Post-treatment
Response
RECIST
Proportion
good
moderate
poor
CR
PR
SD
Bcells
CNV_tumor_cell
Columnar Epithelial
Endothelial cells
Fibroblasts
Macrophages
Mast cells
Myofibroblasts
Neutrophils
Other
Squamous Epithelial
Tcells

Figure 5. Overview of single-cell RNA sequencing atlas reveals transcriptomic landscape in GEA. (A) Available patients for scRNA-seq from an ESO-GAS biobank of 360 patients. Out of 204 patients who received docetaxel-based triplet neoadjuvant chemotherapy, 26 treatment naïve biopsy, 13 mid-treatment biopsy, and 15 post-treatment resection samples were sequenced. (B) Uniform manifold approximation and projection (UMAP) plot consisting of 204,047 tumor and tumor-adjacent cells from 46 patients at different treatment timepoints (treatment naïve biopsy, mid-treatment, post-treatment & treatment naive resection), clustered by major cell types. (C) Major cell type proportions (%) for 34 patients according to pathological response (TRG0/1 good, TRG2 moderate, TRG3 poor). (D) Summarized cell type proportions over treatment timepoints (pre-treatment, mid-treatment, post-treatment) according to pathological response (TRG0/1 good, TRG2 moderate, TRG3 poor) and clinical response (complete response; CR, partial response; PR, stable disease; SD).

3.3.2. scRNA-seq identifies two distinct populations of CAFs from treatment naïve tumor biopsy samples

CAFs from treatment-naïve tumor biopsy samples were investigated to identify the presence of fibroblast subpopulations in GEA, as previously elucidated in other cancer types. Initially, 14 distinct fibroblast clusters were identified **(Figure 6A)**, each demonstrating a unique gene signature as presented via a heat map in **Figure 6B**. Gene Ontology (GO) analysis was performed to identify the molecular function and biological processes associated with each cluster, clusters were then grouped them based on commonalities in their gene signatures **(Supplementary Table 5)**. As seen in pancreatic and breast cancers, two distinct populations were identified in GEA tumor samples, iCAFs (clusters 3, 4, 6, 7, and 16) and myCAFs (clusters 0, 1, 2, 5, 9, 11, 15, 17, 19) as seen in **Figure 6C**. Each subpopulation was defined by their distinct gene signature; iCAFs were distinguished by their expression of inflammatory genes such as *CXCL14*, *RARRES2,*

MMP2, PDGFRA, PDPN, C1R, C1S, CTHRC1, SOD2, and *S100A13*, while myCAF markers included genes involved in ECM organization and reminiscent of pericytes, such as *NDUFA4L2, RGS5, MCAM, NOTCH3, ACTA2, ADIRF, FRZB, MGP,* and *COL4A1* **(Supplementary Table 6)**. PDGFRa, a commonly reported CAF marker that was previously validated in our treatment naïve biopsy fibroblasts, was found to be enriched in iCAFs at the transcriptomic level compared to myCAFs **(Figure 6E, top panel)**. Similarly, matrix metalloproteinase 2 (MMP2), involved in ECM remodelling following inflammatory signals, was also shown to be distinct to iCAFs relative to myCAFs **(Figure 6E, second panel)**. Importantly, MMP2 has been implicated in invasion of tumor cells along with CAF infiltration, and further promotes proinflammatory pathways via signalling with toll-like receptors (TLRs) on antigen presenting cells[88,89]. In contrast, ACTA2 and regulator of G protein signaling 5 (RGS5), commonly implicated with ECM reorganization and vasculature development, respectively, were found to be distinctive to myCAFs **(Figure 6E, bottom panels).**

Individual subsets of clusters within the myCAF and iCAF clusters were further defined and characterized based on the GO analysis. Cluster 0, Cluster 9, and Cluster 15 (ecm-myCAFs) were associated with ECM organization and response to wounding; Cluster 1 and Cluster 19 (angiogenic-myCAFs) with vasculature development; Cluster 2 and Cluster 5 (contractile-myCAFs) with smooth muscle cell contraction pathways; Cluster 3 (metastatic-iCAFs) with cell migration and ECM-associated pathways; Cluster 4 (ecm-iCAFs) with ECM-associated proteins in inflammation; Cluster 6 and Cluster 16 (wound-iCAFs) with wound healing and mineral absorption; Cluster 7 (adipogenic-iCAFs) with lipid metabolism and immune response; Cluster 11 (IFNγ-iCAFs) with interferon gamma signalling; and Cluster 17 (acto-myCAFs) with actin cytoskeleton organization.

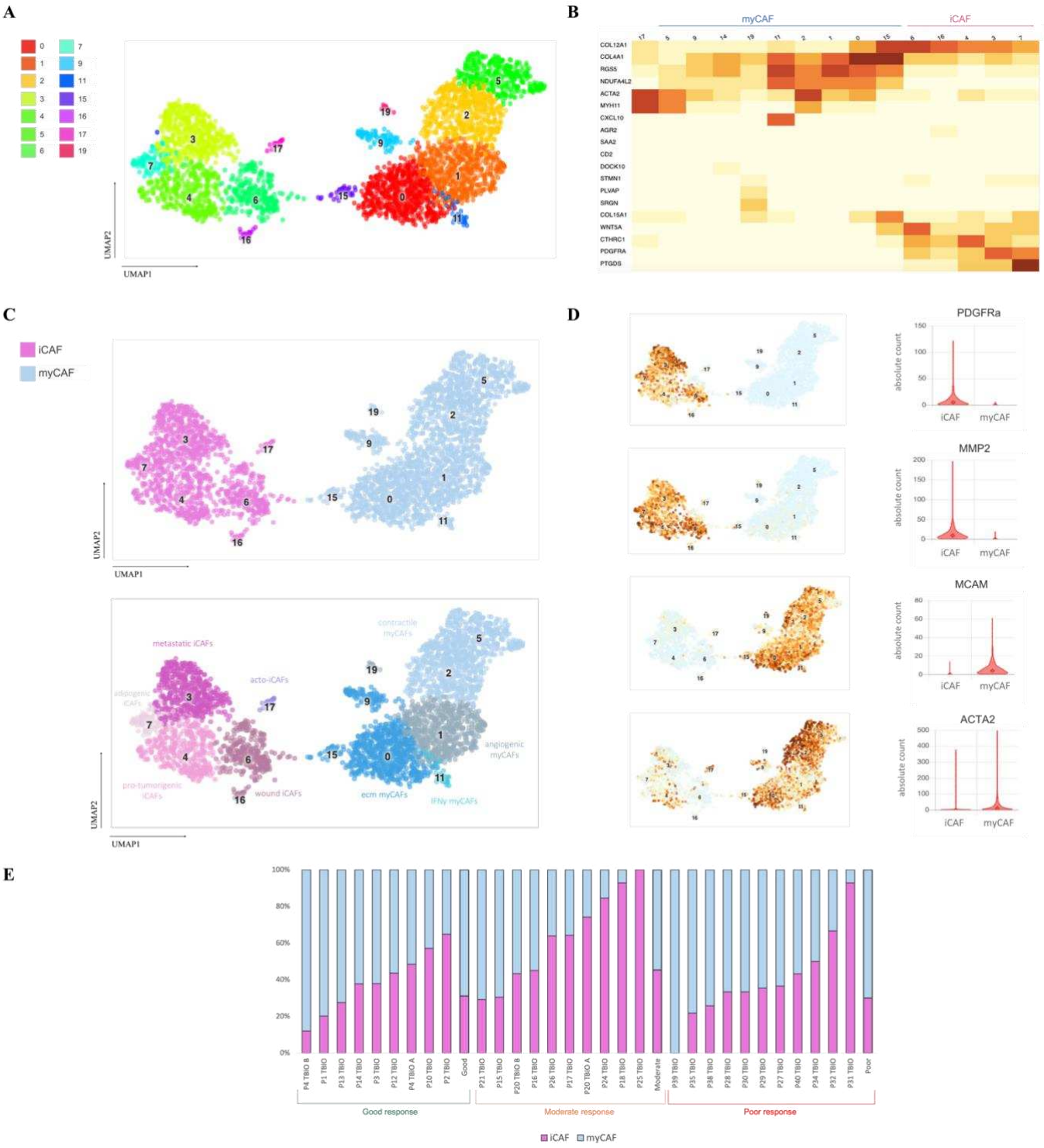

A
B
C
D
E
UMAP1
UMAP2
myCAF
iCAF
metastatic iCAFs
acto-iCAFs
adipogenic iCAFs
pro-tumorigenic iCAFs
wound iCAFs
contractile myCAFs
angiogenic myCAFs
ecm myCAFs
IFNγ myCAFs
PDGFRa
MMP2
MCAM
ACTA2
absolute count
Good response
Moderate response
Poor response

Figure 6. scRNA-seq atlas identifies two major subpopulations of CAFs among treatment naïve biopsy samples. (A) UMAP plot consisting of 3,556 CAFs from 26 treatment naïve biopsy samples, segregated into 14 Leiden clusters according to transcriptomic gene expression. (B) Heatmap of differentially expressed genes according to each of the Leiden clusters. Dark brown color is associated with a significantly differentially expressed gene relative to other clusters, light yellow defined a non-differentially expressed gene. (C) Identification of two major CAF subpopulations, inflammatory CAFs (iCAFs) and myofibroblast CAFs (myCAFs). (D) Classification of clusters within the iCAF and myCAF populations according to gene ontology analysis of their molecular function. (E) Differentially expressed genes between iCAF and myCAF clusters. Top to bottom: Platelet-derived growth factor alpha (PDGFRa) in iCAFs, matrix metalloproteinase 2 (MMP2) in iCAFs, Regulator of G protein signaling 5 (RGS5) in myCAFs, and Actin alpha 2 smooth muscle (ACTA2) in myCAFs. Violin plots on the right compare expression of gene between iCAF and myCAF populations. Dark brown refers to higher gene expression, light yellow defines low gene expression. (F) Proportion of iCAF (pink) to myCAF (blue) populations in 30 individual patient samples.

3.3.3. Association of CAF subpopulations with patient clinical characteristics

CAFs derived from the 26 treatment-naïve tumor biopsy samples were analyzed for their relationship to clinical characteristics. Tumor stage (T0-T4) and metastasis status showed a lack of correlation to any specific CAF subcluster **(Figure 7A and 7B)**. Similarly, pathological response and clinical response to neoadjuvant chemotherapy did not show significant association to any CAF subcluster **(Figure 7D and 7E)**. Cluster 6, wound-iCAFs, showed a population of CAFs derived primarily from poorly differentiated tumors; however, this association was not significant **(Figure 7C)**. Individual CAF subclusters for 17 treatment naïve biopsy samples (>30 cells) were observed according to their pathological response **(Figure 7F)**. Heterogeneous CAF populations were observed between patients of the same pathological response group, as well as

differing groups. Given the extensive heterogeneity, no significant association could be made between a particular cluster and pathological response. Cluster 3 (metastatic-iCAFs), cluster 4 (ecm-iCAFs), cluster 0 (ecm-myCAFs), cluster 1 (angiogenic-myCAFs), cluster 2 and cluster 5 (contractile-myCAFs) were found to be present in most samples to varying extents.

With the differentiation of two distinct CAF subpopulations, the proportion of iCAFs to myCAFs was observed in **Figure 6F** for 30 treatment-naïve tumor biopsy samples. The proportion of the two subpopulations varies significantly from patient to patient but was found to have no significant association with pathological response **(Figure 7G)**. myCAF populations comprised >50% of the CAF populations present among patient samples across all pathological response groups.

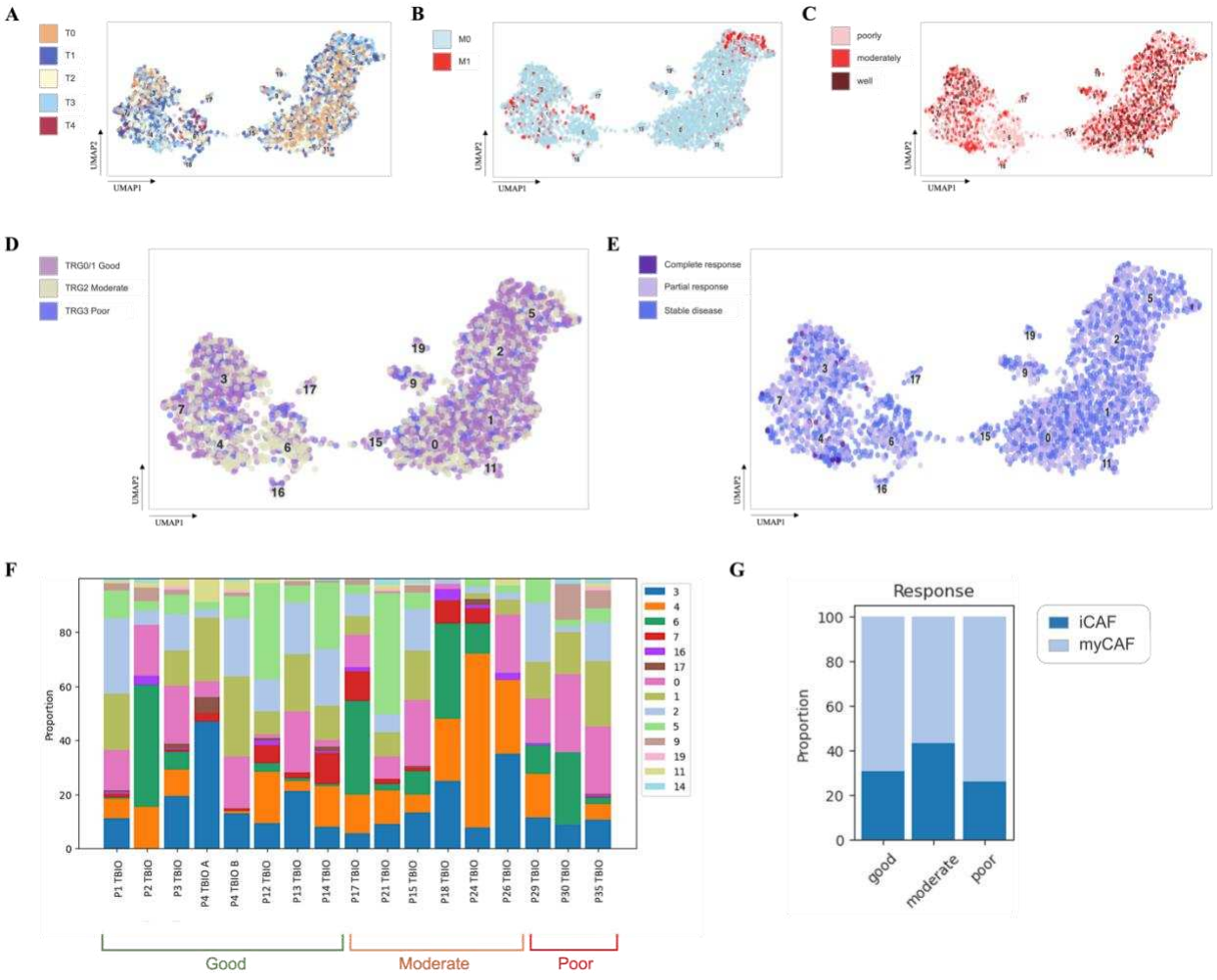


Figure 7. Treatment naïve biopsy CAFs from GEA patients associated to their clinical characteristics. UMAP plot of 26 treatment naïve samples according to (A) tumor stage, T0 (no tumor) to T4 (larger tumor); (B) metastasis occurrence, M0 (no metastasis) and M1 (metastasis); (C) tumor grade (well, moderately, and poorly); (D) pathological response, TRG01 (good), TRG2 (moderate), TRG3 (poor); and (E) clinical response, complete response (CR), partial response (PR), and stable disease (SD). (F) Proportion of CAF Leiden clusters for 17 treatment naïve biopsy samples (>30 cells) according to pathological response. (G) Proportion of iCAF (dark blue) to myCAF (light blue) comparison according to pathological response.

3.3.4. Differentially expressed genes (DEGs) associated with treatment response between iCAFs and myCAFs

Due to the lack of association of distinct CAF subpopulations to treatment response and the lack of DEGs between good, moderate, and poor clinical and pathological responders when looking at the CAFs globally, we proceeded to investigate the iCAF and myCAF populations individually. It was found that within the iCAF population, two genes, CC motif chemokine ligand 20 (CCL20) and chordin-like 1 (CHRDL1), were differentially expressed between good and moderate/poor pathological responders **(Figure 8A)**. CCL20, a chemoattractant for tumor-associated macrophages (TAMs) and T lymphocytes to create a protumorigenic and immunosuppressive TME, was found to be downregulated in good pathological responders, relative to moderate and poor pathological responders[90–92]. In contrast, CHRDL1, an antagonist of bone morphogenetic proteins (BMPs) which are secreted cytokines from the TGFβ family, was significantly upregulated in good pathological responders compared to moderate and poor pathological responders. No DEGs were found between CR, PR, and SD clinical responders within the iCAF population.

Within the myCAF population, nine genes, including, interferon stimulated exonuclease gene 20 (ISG20) and chromosome 20 open reading frame 27 (C20orf27), were found to be differentially expressed between PR and SD clinical response classes **(Figure 8B and Supplemental Figure 2)**. Given that only one patient was a CR clinically, it was not considered in the comparison. All differentially expressed genes between the two clinical response groups were significantly downregulated in the PR group relative to SD. ISG20, associated with activation of the innate immune defense, has been shown to be involved in angiogenesis and tumor progression in renal cell carcinoma through upregulation of matrix metalloproteinase 9 (MMP9)[93],

and C20orf27, a gene present on chromosome 20 involved activating the TGFβR-TAK1-NFκB pathway via interaction with phosphatases, both showed lower expression in patients who responded partially to chemotherapy[94].

Additionally, the expression of 85 genes was plotted on a heatmap separating iCAF from myCAF and further divided according to pathological response **(Figure 8C)** and clinical response **(Figure 8D)**. FAP showed lower expression in myCAFs compared to iCAFs, but also in poor responder iCAFs compared to good and moderate responder iCAFs. Similarly, GREM1 and POSTN showed lower expression in myCAFs and poor responder iCAFs relative to good and moderate responder iCAFs. Although these differences were seen between myCAFs and iCAFs when correlating to clinical response, no association could be found between the three clinical response groups. COL15A1 was found to be mildly expressed among good and moderate responders' myCAFs, but showed little to no expression among myCAFs from poor responders.

The DEGs found within iCAFs between good and poor pathological responders, and within myCAFs between PR and SD patients, can therefore be correlated to their involvement in tumorigenesis and response to chemotherapy.

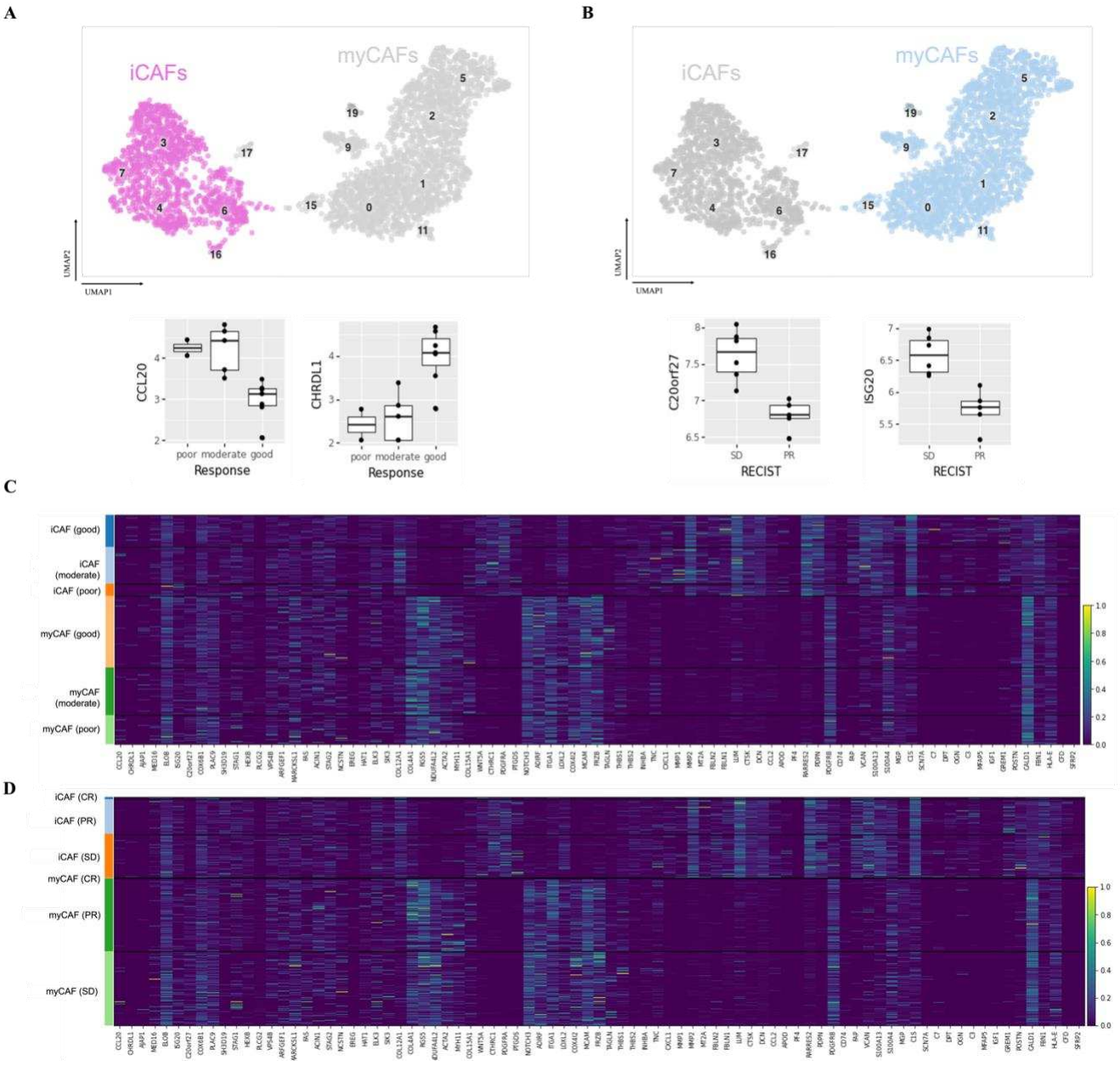


Figure 8. Differential expression of CAF markers in iCAFs and myCAFs according to pathological and clinical response to neoadjuvant chemotherapy. (A) Differentially expressed genes, CCL20 and CHRDL1, in iCAF population, according to pathological response (TRG01 good vs. TRG2 moderate vs. TRG3 poor). (B) Differentially expressed genes, AJAP1, MED16, and ELOB, in myCAF populations, according to clinical response (stable disease vs. partial response). Heatmap showing expression of 85 genes between iCAF and myCAF populations associated to (C) pathological response and (D) clinical response.

3.4. Investigation of GEA CAFs across three treatment timepoints unveils dynamic biomarkers associated to chemotherapy response

CAFs from tumor samples collected at three treatment timepoints (initial endoscopic biopsy, Tbio; mid-treatment endoscopic biopsy, Tmid; and post-treatment surgical resection, Tsur) were assessed to identify differences in CAF subpopulations following chemotherapy **(Figure 9A)**. Tsur CAFs formed a distinct cluster (top left cluster; clusters 1, 2, 9, 6, 7, 8, 19) from Tbio and Tmid CAFs, which were found to occupy overlapping populations (top and bottom right clusters). When observed against pathological response **(Figure 9B)** and clinical response **(Figure 9C)**, 62.4% of CAFs from the Tsur-specific cluster were derived from poor-responder patient samples, and 62.8% were derived from patients who showed clinical PR. Furthermore, the Tsur-specific cluster demonstrated a unique gene signature involved mainly in ECM remodelling and inflammatory signalling, as seen through the heatmap in **Figure 9D**, with enrichment of complement component 7 (C7), dermatopontin (DPT), osteoglycin (OGN), and complement component 3 (C3). Sodium Voltage Gated Channel Alpha Subunit 7 (SCN7A) was found to be enriched in the Tsur-specific CAF cluster, and similar to other DEGs in this cluster, it was found to be expressed in low amounts in certain iCAF clusters previously defined **(Figure 9E)**.

Analysis of DEGs between Tbio and Tmid treatment timepoints was performed in association with pathological and clinical response. **Figure 9F and Supplemental Figure 2** demonstrate differential analysis among fibroblasts with an iCAF signature over the two treatment timepoints, correlating to pathological response (TRG0/1 good vs TRG2 moderate + TRG3 poor). 16 genes, including stromal antigen 1 and 2 (STAG1 and STAG2), and histone acetyltransferase 1 (HAT1), were shown to distinctively shift in their expression profile in correlation to their pathological response. STAG1 and STAG2, subunits part of the cohesion complex involved in

sister chromatid cohesion, DNA repair, and replication[95], were both found to increase in expression from Tbio to Tmid in good responder patients, but decreased in moderate+poor responders across the two timepoints. In contrast, HAT1 was shown to decrease from treatment naïve biopsy to mid-treatment in good responders, but increased in expression among moderate and poor responders following initial treatment with chemotherapy.

When investigating fibroblasts with a myCAF signature between the two timepoints **(Figure 9E)**, inhibitor of DNA binding 3 (ID3) was found to decrease significantly following treatment with chemotherapy in pathological good responder patients, while NADH:Ubiquinone Oxidoreductase Subunit B4 (NDUFB4), among several other DEGs **(Supplemental Figure 3B)**, was found to decrease in SD patients from treatment naïve to mid-treatment samples, but slightly increase in patients with PR.

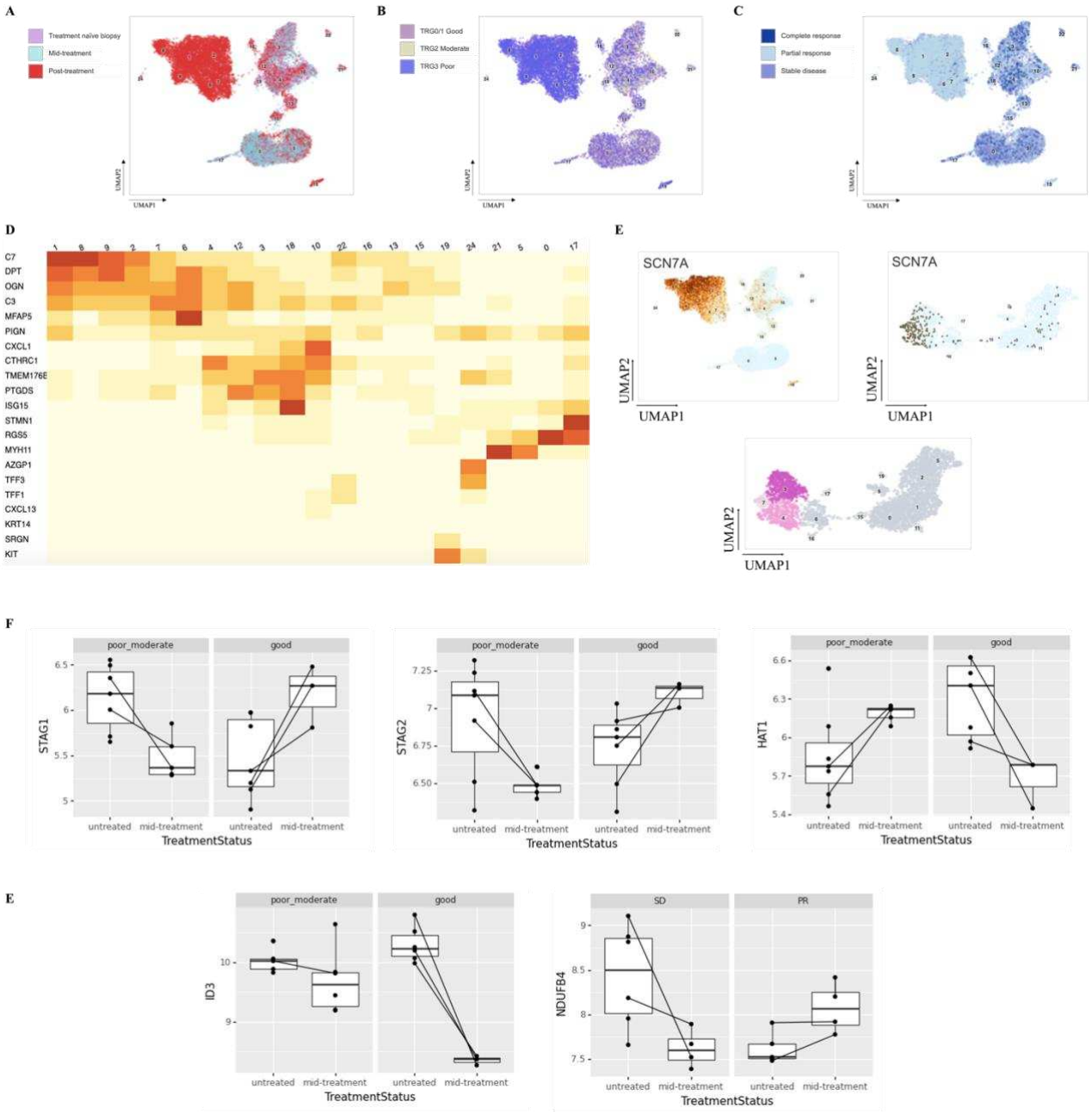


Figure 9. CAFs from three treatment timepoints (treatment naïve biopsy, mid-treatment, and post-treatment resection) demonstrate dynamic biomarkers associated to therapy response. UMAP plot of 20,294 CAFs from 28 patient samples according to (A) treatment timepoint (treatment naïve biopsy, purple; mid-treatment, blue; post-treatment resection, red); (B) pathological response (TRG0/1 good, purple; TRG2 moderate, beige; TRG3 poor, blue); (C) clinical response (complete response, dark blue; partial response, light blue; stable disease, grey

blue). (D) Heat map of differentially expressed genes according to each Leiden cluster. (E) Left: Differentially expressed gene, SCN7A, in post-treatment resection compared to treatment naïve and mid-treatment biopsies. Right: SCN7A expression in treatment naïve biopsy samples. Bottom: Origin of post-treatment differentially expressed genes from metastatic, adipogenic, and pro-tumorigenic iCAFs. (F) Differentially expressed genes, STAG1, STAG2, and HAT1, according to pathological response in iCAFs over two treatment timepoints: treatment naïve (Tbio) to mid-treatment (Tmid) endoscopic biopsies. (G) ID3 is differentially expressed between Tbio and Tmid in myCAFs according to pathological response (good vs. moderate+poor), while NDUFB4 is associated to clinical response (PR vs. SD). Analysis was performed among 22 samples, or 5 paired observations.

3.5. Establishment of a 3D PDO-CAF co-culture model

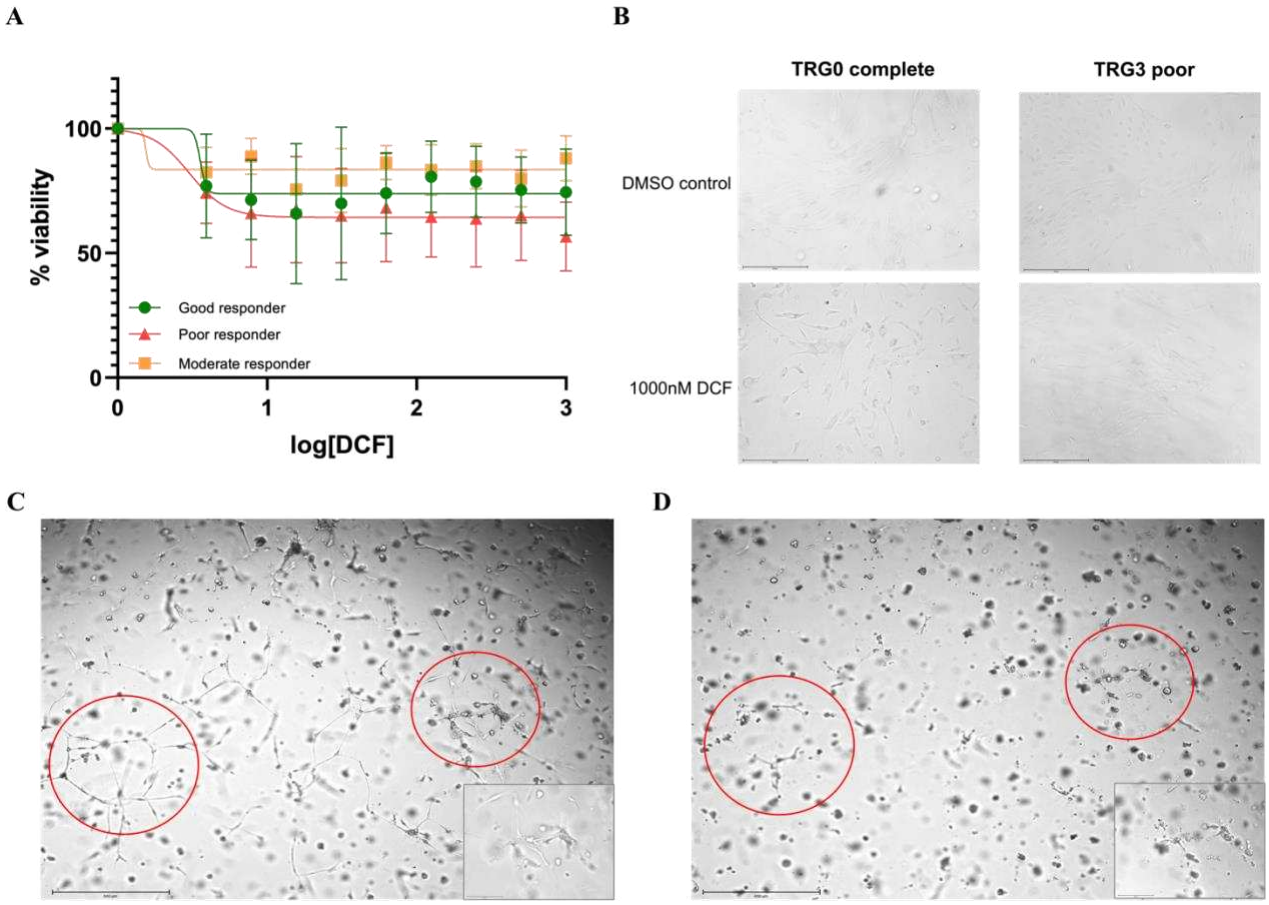


Figure 10. DCF treatment of primary fibroblasts and PDO-CAF co-culture. (A) Chemosensitivity curve following drug screen on treatment naïve biopsy fibroblasts, comparing TRG0 complete, TRG2 moderate, and TRG3 poor responder patients. Cell viability (%) was observed at ten different concentrations (log [DCF]). (B) Brightfield images of fibroblasts from TRG0 complete and TRG3 poor responder patients following 3 days of growth. DMSO control is shown on top, and max DCF concentration (1000 nM) treatment is shown on the bottom. (Figures C and D) TRG0 good responder matching PDO-CAF co-culture model followed by DCF treatment. Brightfield image after (C) 3 days of growth, and (D) 3 days of drug treatment. Highlighted areas demonstrate change in PDO-CAF interactions following drug treatment. Scale used for imaging: 4x, 600 µm; and 20x, 150 µm.

Following the identification of CAF subtypes and distinct markers associated to treatment response and timepoints, the next step was to develop a 3D PDO-CAF co-culture model to recapitulate the tumor-stromal interactions observed within the TME. The goal is to then use this model to perform drug screens with the standard-of-care treatment to understand the role of stromal interactions with the tumor in regard to its response to therapy.

As a proof of concept, fibroblasts from differing pathological responses were treated with DCF after 3-4 days of growth in a Matrigel-Purecol solution to mimic the ECM-rich environment. Cell viability was assessed after 72h following drug treatment, and it was confirmed that fibroblasts demonstrate a lack of sensitivity to the DCF, regardless of pathological response **(Figure 10A)**. The area under the curve (AUC) values were comparable for the three curves: good responder, AUC=0.768; moderate responder, AUC=0.699; poor responder, AUC=0.845, all responding resistantly to the chemotherapy. Furthermore, despite the morphological differences observed between fibroblasts from the TRG0 complete responder patient and the TRG3 poor responder patient, a similar sensitivity was seen following drug treatment when observed under the microscope **(Figure 10B)**.

To establish the PDO-CAF co-culture model, matching PDOs and CAFs from a TRG0 complete responder patient were cultured together in the Matrigel-Purecol solution and allowed to grow for 3-4 days. Fibroblasts developed close interactions with the PDOs, which appear as domes, as well as with one another, creating a well-connected matrix **(Figure 10C)**. The co-culture was then treated with 1000 nM of DCF for 72h, resulting in the death of PDOs, as seen through the exploding of the cell and release of intercellular components. Despite the death of the PDOs and some fibroblasts, stromal-stromal and stromal-tumor interactions were still present in certain areas

as highlighted in red **(Figure 10D)**, pointing to the possibility of chemoresistant, tumor-protective interactions.

CHAPTER 4: DISCUSSION

Given the poor outcome of GEA patients, with 40% of patients presenting to the clinic with innately resistant tumors to the most effective standard-of-care treatment to date, and recurrence of disease occurring among 50% of initial responders[34], it has become vital that the underlying mechanisms behind chemotherapy resistance are understood in GEA and novel biomarkers are identified to reliably distinguish between chemosensitive from chemoresistant patients prior to treatment provision. The implication of CAFs in the TME and its role in tumor progression, invasion, and response to therapy has been extensively studied in other cancer types[52,59]. Molecular and functional heterogeneity of CAFs, like other cell types within the TME, poses a challenge for effective response to treatment. The use of scRNA-seq technology has emerged as a ground-breaking tool to tackle the question of CAF heterogeneity. Several studies have performed the characterization of major CAF subtypes within the TME according to gene signature and function, allowing for the identification of hallmark CAF markers and association of distinct populations to therapy resistance[60–63,96].

Unfortunately, CAF markers specific to GEA are still unknown and CAF subtypes have yet to be elucidated. Commonly reported CAF markers were first investigated in this study to determine their expression and specificity to GEA CAFs. Indeed, these markers were found to be expressed by our patients' treatment naïve biopsy fibroblasts, with VIM, FAP, PDPN and S100A4 demonstrating a significant difference in expression between good and poor pathological responders. PDGFRb was found to be significantly differentially expressed between good and poor pathological responders when evaluated with IF; however, FACS analysis provided opposing results. Previous studies on lung adenocarcinoma and PDAC associated high VIM expression with tumor invasion and overall poorer patient survivals[97–99]. Our findings contrasted with these studies,

such that VIM showed greater expression in good pathological responders and decreased with poorer response. Furthermore, VIM expression was high even at the transcriptomic level among tumor-adjacent normal fibroblasts when referring to our scRNA-seq atlas **(Supplemental Figure 4A)**, implying that VIM may not serve as a specific CAF marker in GEA. Studies have reported FAP as a hallmark marker to characterize subpopulations, as seen in breast cancer, such that the CAF-S1 cluster was defined by a FAP^{hi} signature and was implicated in creating an immunosuppressive environment[62,63]. However, studies regarding the prognostic value of FAP have shown contrasting results, such that another study in breast cancer reported the association of FAP expression with longer survival[100]. From our findings, it was found that the % of cells expressing FAP and the overall intensity of expression decreased with poor pathological response when assessed using FACS. scRNA-seq traced FAP expression to a subset of treatment-naïve CAFs and was not expressed among normal fibroblasts **(Supplemental Figure 4B)**. Interestingly, increased expression of PDPN was correlated with poor response, which has also been previously reported in ESCC and pancreatic cancer[101,102]. Finally, S100A4, a calcium binding protein, was also found to be increased in expression in poor responder CAFs. This protein has been implicated in other cancers for promoting invasion and metastasis via EMT[103]. However, S100A4 can be seen to be expressed at the transcriptomic level in tumor-adjacent normal fibroblasts **(Supplemental Figure 4C)**. Although these markers demonstrated differential expression between good and poor pathological responders, the lack of specificity to tumor fibroblasts necessitated the identification of specific markers for GEA.

This study is one of the first to identify two major subpopulations of CAFs in GEA, iCAFs and myCAFs, using scRNA-seq data from 26 treatment naïve endoscopic biopsy samples with their associated pathological and clinical response to neoadjuvant chemotherapy. Reminiscent of

the CAF subpopulations identified in pancreatic and breast cancers, iCAFs were defined by a primarily inflammatory gene signature, demonstrating enrichment of cytokines and chemokines, such as PDGFRα and CXCL14, while myCAFs from our findings presented with differential expression of genes related to ECM reorganization, contractility, and TGFβ signalling, such as ACTA2 and MYH11. In addition, the myCAFs expressed pericyte lineage markers such as RGS5 and NOTCH3[66]. The presence of apCAF populations were reported in PDAC and breast cancer, with an enrichment of MHC class II family genes, such as CD74[61,63]. Although CD74-expressing CAFs were present, a distinct clustered population could not be identified. iCAF and myCAF populations were further subdivided by their molecular function, as determined through GO analysis.

Four subsets of iCAFs were identified along with five subsets of myCAFs, each with their unique gene profile. Metastatic-iCAFs (cluster 3), wound-iCAFs (cluster 6 and 16), and adipogenic iCAFs (cluster 7) would likely contribute to the development of an immunosuppressive environment, given the release of inflammatory signals, such as CCL2, involved in recruiting MDSCs and T lymphocytes[73]. The gene signature of myCAFs, specifically ecm-myCAFs (cluster 0, 9, 15) alludes to its role in ECM deposition and remodelling, such that a stiff ECM is created and causes increased interstitial pressure, all of which promote tumor proliferation and migration, and create a physical barrier that prevents the entry of immune cells, thereby compounding on the immunosuppressive effects of iCAFs[57,59]. Furthermore, myCAFs promote angiogenesis and neovascularization, as seen through the perivasculature gene signature expressed by angiogenic-myCAFs (cluster 1 and 19). The role of IFNγ signalling by IFNγ-myCAFs (cluster 11) is two-fold, such that it has been shown to be involved in both anti-tumor, immune-stimulating functions, as well as pro-tumorigenic, immunosuppressing functions[104]. IFNγ is commonly released by

immune cells, such as natural killer cells, T lymphocytes, Tregs, and APCs, to facilitate cross talk within the TME[105]. On one hand, IFNγ can help to regress tumorigenesis by activation of the JAK/STAT signalling pathway, which has been shown to inhibit metastasis. IFNγ can also sensitize Tregs such that they lose their immunosuppressive activity[104,105]. IFNγ signalling has been shown to induce the release of pro-inflammatory factors, such as CXCL9 and CXCL11, on tumor cells, endothelial cells, and fibroblasts, all of which can promote T lymphocyte recruitment. However, binding of CXCL11 to its receptor, CXCR7, can further promote tumor progression and neovascularization[104,105].

When investigating the newly identified GEA CAF subpopulations according to pathological and clinical response to neoadjuvant chemotherapy, there was a lack of association between myCAF and iCAF populations and chemoresponse. This may be explained by the extensive intra- and inter-patient CAF heterogeneity, as seen with the varying proportions of myCAF to iCAF populations among the samples **(Figures 6F and 7F).** Despite the lack of correlation of a particular cluster to treatment response, DEGs across pathological and clinical response were identified within the iCAF and myCAF populations. CCL20, which has been shown to create a protumorigenic and immunosuppressive TME through lymphocyte and TAM recruitment[90–92], was upregulated in iCAFs from moderate+poor pathological responders compared to good responders. CHRDL1, which functions as an antagonist to BMPs involved in TGFβ signalling, was upregulated in our good pathological responders' iCAFs. It has been further associated to positive clinical outcomes among breast cancer patients, such that BMPs have been shown to demonstrate pro-tumorigenic activity[106]. In myCAFs, several genes were found to be downregulated in PR group compared to the SD group. Of interest, ISG20 has been shown to be involved in angiogenesis and tumor progression in renal cell carcinoma through upregulation of

matrix metalloproteinase 9 (MMP9)[93], while C20orf27 was found to promote cell proliferation and tumor growth in colorectal cancer[94]. These markers have great prognostic value and can therefore serve as specific CAF biomarkers that provide insight into the patients' likely response to the standard-of-care treatment.

Our study further identified dynamic biomarkers across two treatment timepoints, treatment naïve biopsy and mid-treatment biopsy, according to pathological and clinical response. STAG1 and STAG2, subunits found on the cohesin complex, were both found to decrease in expression in iCAFs from moderate+poor pathological responders from initial to mid-treatment biopsy, while the opposite was seen in good responder patients. Despite the commonalities in sequence and function, STAG2, but not STAG1, has been implicated as a tumor suppressor gene, as mutations in this gene have been reported in other cancer types[107,108]. STAG2 loss in melanoma cells resulted in DNA binding to switch to STAG1, and further resulted in the enrichment of IFNγ response. It was found that STAG2 targets interferon regulatory factor 9 (IRF9), a major component in type 1 interferon signalling, and STAG2 loss was associated with IRF9 activation and upregulated expression of PD-L1 on cancer cells, thus promoting immune evasion and tumor progression[109]. It is therefore hypothesized that STAG2 loss among poor pathological responders' iCAFs across treatment time is related to increased IFNγ signalling and fostering an immunosuppressive environment.

Furthermore, HAT1 was found to increase in expression in moderate/poor pathological responders' iCAFs across treatment time, while the opposite was seen among good responders' iCAFs. HAT1 overexpression has been implicated in gemcitabine resistance in pancreatic cancer through regulation of long noncoding RNA PVT1 (lncRNA PVT1) and enhancer of zeste 2 polycomb repressive complex 2 subunit (EZH2)[110]. In addition, HAT1 overexpression was

similarly observed in pancreatic cancer, with correlation to poor prognosis, and was found to transcriptionally regulate PD-L1, again promoting tumor immune evasion[111].

When investigating myCAFs for dynamic changes in marker expression with association to treatment time and chemoresponse, ID3 was found to be significantly downregulated in good responders following chemotherapy, and a slight decrease was seen in moderate/poor responders. ID3 loss has been implicated in increasing DNA damage in cisplatin-treated pancreatic cancer cells[112]. This may explain our findings as ID3 loss in good pathological responders' myCAFs may increase DNA damage and tumor cell death following neoadjuvant therapy consisting of cisplatin.

The findings of this study have significant implications given the identification of potential CAF biomarkers in GEA, which to date have been unknown. Furthermore, the classification of distinct CAF subpopulations provides further insight into the elaborate functional and molecular heterogeneity of CAFs in GEA. The markers identified in this study can provide avenues to developing more targeted-therapies that can complement the current standard-of-care, while serving as biomarkers to detect the possibility of chemoresistance among patients, preventing them from undergoing futile and exhaustive cycles of chemotherapy. Additionally, IHC staining of GEA tumors by pathology using CAF markers such as CHRDL1, which has been found from this study to be downregulated in poor responders at the time of pre-treatment biopsy, can be performed to identify patients that potentially may develop chemoresistance in the future.

This study is unique with its use of such a large biobank of tumor and tumor-adjacent samples from >350 patients, as well as scRNA-seq data for 46 patients, of which treatment-naïve CAFs from 26 patients have their matching mid- and/or post-treatment CAFs, as well as information regarding their pathological and clinical response. Additionally, the development of

PDOs for the use of *in vitro* drug screens is incredibly valuable due to the ability of PDOs to recapitulate the unique genomic and histological features of the patient's tumor.

Some limitations to this study include: (1) the differing number of CAFs coming from each patients' scRNA-seq sample, such that some patients had very few cells relative to others, which may shift the clustering; (2) the comparison of endoscopic biopsy samples to surgical resection samples using scRNA-seq, as surgical resection samples consists of a larger tumor mass collected and therefore a greater number of cells; (3) the small sample size when comparing commonly reported CAF markers using IF and FACS; and (4) difficulty comparing OS and DFS between patients who were diagnosed at varying points. A larger sample size for scRNA-seq analysis and collection of multiple tumor samples from endoscopic biopsy and surgical resection would allow for normalization of this variation in cell number between samples.

In addition, it is important to consider that transcriptomic expression of genes may often not correlate with proteomic expression, and therefore it is necessary to validate the markers identified in this study in the patients' primary fibroblasts by IF, FACS, and ELISA, as well as staining of markers in primary tissue samples. Furthermore, the establishment of the PDO-CAF co-culture model will allow for a deeper understanding of the effects of stromal-stromal and stromal-tumor interactions on a patient's chemoresponse, by performing heterotypic recombinant co-cultures between PDOs and CAFs of differing pathological responses and subjecting them to a drug screen to determine if chemoresponse is altered. The markers identified in this study will be further validated in an organ-on-a-chip model that is currently established in our lab, which incorporates TME interactions between CAFs, PDOs, and TILs, as well as microfluidic flow to mimic the movements of gastrointestinal tract.

Overall, the findings of this study reveal the unique CAF heterogeneity in GEA and the presence of two major subpopulations within the TME that are implicated in cultivating a pro-tumorigenic, immunosuppressive environment ideal for tumor survival and progression. Several markers associated to pathological and clinical response were identified, as well as potential dynamic biomarkers across the treatment timepoints.

BIBLIOGRAPHY

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **71**, 209–249 (2021).
2. Domper Arnal, M. J., Ferrández Arenas, Á. & Lanas Arbeloa, Á. Esophageal cancer: Risk factors, screening and endoscopic treatment in Western and Eastern countries. *World J. Gastroenterol. WJG* **21**, 7933–7943 (2015).
3. Van Cutsem, E., Sagaert, X., Topal, B., Haustermans, K. & Prenen, H. Gastric cancer. *The Lancet* **388**, 2654–2664 (2016).
4. Bass, A. J. *et al.* Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **513**, 202–209 (2014).
5. Kim, J. *et al.* Integrated genomic characterization of oesophageal carcinoma. *Nature* **541**, 169–175 (2017).
6. Ma, J., Shen, H., Kapesa, L. & Zeng, S. Lauren classification and individualized chemotherapy in gastric cancer (Review). *Oncol. Lett.* **11**, 2959–2964 (2016).
7. Hayakawa, Y., Sethi, N., Sepulveda, A. R., Bass, A. J. & Wang, T. C. Oesophageal adenocarcinoma and gastric cancer: should we mind the gap? *Nat. Rev. Cancer* **16**, 305–318 (2016).
8. Wong, H. H. & Chu, P. Immunohistochemical features of the gastrointestinal tract tumors. *J. Gastrointest. Oncol.* **3**, 262–284 (2012).
9. Siewert, J. R. & Stein, H. J. Classification of adenocarcinoma of the oesophagogastric junction. *Br. J. Surg.* **85**, 1457–1459 (1998).
10. Nagaraja, A. K., Kikuchi, O. & Bass, A. J. Genomics and Targeted Therapies in Gastroesophageal Adenocarcinoma. *Cancer Discov.* **9**, 1656–1672 (2019).
11. Cristescu, R. *et al.* Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat. Med.* **21**, 449–456 (2015).
12. Sue-Ling, H. M. *et al.* Gastric cancer: a curable disease in Britain. *BMJ* **307**, 591–596 (1993).
13. Hofstetter, W. *et al.* Treatment Outcomes of Resected Esophageal Cancer. *Ann. Surg.* **236**, 376–385 (2002).
14. Siewert, J. R., Böttcher, K., Stein, H. J. & Roder, J. D. Relevant prognostic factors in gastric cancer: ten-year results of the German Gastric Cancer Study. *Ann. Surg.* **228**, 449–461 (1998).

15. Siewert, J. R. *et al.* Prognostic relevance of systematic lymph node dissection in gastric carcinoma. *BJS Br. J. Surg.* **80**, 1015–1018 (1993).
16. Ferri, L. E. *et al.* Perioperative docetaxel, cisplatin, and 5-fluorouracil (DCF) for locally advanced esophageal and gastric adenocarcinoma: a multicenter phase II trial. *Ann. Oncol.* **23**, 1512–1517 (2012).
17. Cunningham, D. *et al.* Perioperative Chemotherapy versus Surgery Alone for Resectable Gastroesophageal Cancer. *N. Engl. J. Med.* **355**, 11–20 (2006).
18. Van Cutsem, E. *et al.* Phase III Study of Docetaxel and Cisplatin Plus Fluorouracil Compared With Cisplatin and Fluorouracil As First-Line Therapy for Advanced Gastric Cancer: A Report of the V325 Study Group. *J. Clin. Oncol.* **24**, 4991–4997 (2006).
19. Ridwelski, K. *et al.* Combination chemotherapy with docetaxel and cisplatin for locally advanced and metastatic gastric cancer. *Ann. Oncol.* **12**, 47–51 (2001).
20. Al-Batran, S.-E. *et al.* Perioperative chemotherapy with fluorouracil plus leucovorin, oxaliplatin, and docetaxel versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4): a randomised, phase 2/3 trial. *The Lancet* **393**, 1948–1957 (2019).
21. Bang, Y.-J. *et al.* Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *The Lancet* **376**, 687–697 (2010).
22. Gravalos, C. & Jimeno, A. HER2 in gastric cancer: a new prognostic factor and a novel therapeutic target. *Ann. Oncol.* **19**, 1523–1529 (2008).
23. Waddell, T. *et al.* Epirubicin, oxaliplatin, and capecitabine with or without panitumumab for patients with previously untreated advanced oesophagogastric cancer (REAL3): a randomised, open-label phase 3 trial. *Lancet Oncol.* **14**, 481–489 (2013).
24. Lordick, F. *et al.* Capecitabine and cisplatin with or without cetuximab for patients with previously untreated advanced gastric cancer (EXPAND): a randomised, open-label phase 3 trial. *Lancet Oncol.* **14**, 490–499 (2013).
25. Van Cutsem, E. *et al.* A randomized, open-label study of the efficacy and safety of AZD4547 monotherapy versus paclitaxel for the treatment of advanced gastric adenocarcinoma with FGFR2 polysomy or gene amplification. *Ann. Oncol.* **28**, 1316–1324 (2017).

26. Shah, M. A. *et al.* A Randomized Phase II Study of FOLFOX With or Without the MET Inhibitor Onartuzumab in Advanced Adenocarcinoma of the Stomach and Gastroesophageal Junction. *The Oncologist* **21**, 1085–1090 (2016).
27. Fuchs, C. S. *et al.* Ramucirumab monotherapy for previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (REGARD): an international, randomised, multicentre, placebo-controlled, phase 3 trial. *The Lancet* **383**, 31–39 (2014).
28. Wilke, H. *et al.* Ramucirumab plus paclitaxel versus placebo plus paclitaxel in patients with previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (RAINBOW): a double-blind, randomised phase 3 trial. *Lancet Oncol.* **15**, 1224–1235 (2014).
29. Casak, S. J. *et al.* FDA Approval Summary: Ramucirumab for Gastric Cancer. *Clin. Cancer Res.* **21**, 3372–3376 (2015).
30. Janjigian, Y. Y. *et al.* Nivolumab plus chemotherapy versus chemotherapy as first-line treatment for advanced gastric cancer/gastroesophageal junction cancer/oesophageal adenocarcinoma (CheckMate 649): a multicentre, randomised, open-label, phase 3 trial. *Lancet Lond. Engl.* **398**, 27–40 (2021).
31. Diaz, L. A. *et al.* Pembrolizumab versus chemotherapy for microsatellite instability-high or mismatch repair-deficient metastatic colorectal cancer (KEYNOTE-177): final analysis of a randomised, open-label, phase 3 study. *Lancet Oncol.* **23**, 659–670 (2022).
32. Shitara, K. *et al.* Efficacy and Safety of Pembrolizumab or Pembrolizumab Plus Chemotherapy vs Chemotherapy Alone for Patients With First-line, Advanced Gastric Cancer. *JAMA Oncol.* **6**, 1–10 (2020).
33. Wainberg, Z. A. *et al.* Efficacy of Pembrolizumab Monotherapy for Advanced Gastric/Gastroesophageal Junction Cancer with Programmed Death Ligand 1 Combined Positive Score ≥10. *Clin. Cancer Res.* **27**, 1923–1931 (2021).
34. Yoon, H. H. *et al.* Adverse Prognostic Impact of Intratumor Heterogeneous HER2 Gene Amplification in Patients With Esophageal Adenocarcinoma. *J. Clin. Oncol.* **30**, 3932–3938 (2012).
35. Lubner, S. J., Uboha, N. V. & Deming, D. A. Primary and acquired resistance to biologic therapies in gastrointestinal cancers. *J. Gastrointest. Oncol.* **8**, 499–512 (2017).

36. Blangé, D., Stroes, C. I., Derks, S., Bijlsma, M. F. & van Laarhoven, H. W. M. Resistance mechanisms to HER2-targeted therapy in gastroesophageal adenocarcinoma: A systematic review. *Cancer Treat. Rev.* **108**, 102418 (2022).

37. Baxter, M. A., Middleton, F., Cagney, H. P. & Petty, R. D. Resistance to immune checkpoint inhibitors in advanced gastro-oesophageal cancers. *Br. J. Cancer* **125**, 1068–1079 (2021).

38. Gao, J.-P., Xu, W., Liu, W.-T., Yan, M. & Zhu, Z.-G. Tumor heterogeneity of gastric cancer: From the perspective of tumor-initiating cell. *World J. Gastroenterol.* **24**, 2567–2581 (2018).

39. Greaves, M. & Maley, C. C. Clonal evolution in cancer. *Nature* **481**, 306–313 (2012).

40. Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**, 81–94 (2018).

41. Baccelli, I. & Trumpp, A. The evolving concept of cancer and metastasis stem cells. *J. Cell Biol.* **198**, 281–293 (2012).

42. Zhou, H.-M., Zhang, J.-G., Zhang, X. & Li, Q. Targeting cancer stem cells for reversing therapy resistance: mechanism, signaling, and prospective agents. *Signal Transduct. Target. Ther.* **6**, 1–17 (2021).

43. Pisco, A. O. & Huang, S. Non-genetic cancer cell plasticity and therapy-induced stemness in tumour relapse: 'What does not kill me strengthens me'. *Br. J. Cancer* **112**, 1725–1732 (2015).

44. Brock, A., Chang, H. & Huang, S. Non-genetic heterogeneity — a mutation-independent driving force for the somatic evolution of tumours. *Nat. Rev. Genet.* **10**, 336–342 (2009).

45. Shi, W.-J. & Gao, J.-B. Molecular mechanisms of chemoresistance in gastric cancer. *World J. Gastrointest. Oncol.* **8**, 673–681 (2016).

46. Mansoori, B., Mohammadi, A., Davudian, S., Shirjang, S. & Baradaran, B. The Different Mechanisms of Cancer Drug Resistance: A Brief Review. *Adv. Pharm. Bull.* **7**, 339–348 (2017).

47. Luan, S. *et al.* Advances in Drug Resistance of Esophageal Cancer: From the Perspective of Tumor Microenvironment. *Front. Cell Dev. Biol.* **9**, 664816 (2021).

48. Bejarano, L., Jordão, M. J. C. & Joyce, J. A. Therapeutic Targeting of the Tumor Microenvironment. *Cancer Discov.* **11**, 933–959 (2021).

49. Kim, R. *et al.* Early Tumor–Immune Microenvironmental Remodeling and Response to First-Line Fluoropyrimidine and Platinum Chemotherapy in Advanced Gastric Cancer. *Cancer Discov.* OF1–OF18 (2021) doi:10.1158/2159-8290.CD-21-0888.

50. Sahai, E. *et al.* A framework for advancing our understanding of cancer-associated fibroblasts. *Nat. Rev. Cancer* **20**, 174–186 (2020).

51. Xing, F., Saidou, J. & Watabe, K. Cancer associated fibroblasts (CAFs) in tumor microenvironment. *Front. Biosci. J. Virtual Libr.* **15**, 166–179 (2010).

52. Ping, Q. *et al.* Cancer-associated fibroblasts: overview, progress, challenges, and directions. *Cancer Gene Ther.* **28**, 984–999 (2021).

53. Kanzaki, R. & Pietras, K. Heterogeneity of cancer-associated fibroblasts: Opportunities for precision medicine. *Cancer Sci.* **111**, 2708–2717 (2020).

54. Dongre, A. & Weinberg, R. A. New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* **20**, 69–84 (2019).

55. Yu, Y. *et al.* Cancer-associated fibroblasts induce epithelial–mesenchymal transition of breast cancer cells through paracrine TGF-β signalling. *Br. J. Cancer* **110**, 724–732 (2014).

56. Yeon, J. H. *et al.* Cancer-derived exosomes trigger endothelial to mesenchymal transition followed by the induction of cancer-associated fibroblasts. *Acta Biomater.* **76**, 146–153 (2018).

57. Kobayashi, H. *et al.* Cancer-associated fibroblasts in gastrointestinal cancer. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 282–295 (2019).

58. Louault, K., Li, R.-R. & DeClerck, Y. A. Cancer-Associated Fibroblasts: Understanding Their Heterogeneity. *Cancers* **12**, 3108 (2020).

59. Biffi, G. & Tuveson, D. A. Diversity and Biology of Cancer-Associated Fibroblasts. *Physiol. Rev.* **101**, 147–176 (2021).

60. Öhlund, D. *et al.* Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J. Exp. Med.* **214**, 579–596 (2017).

61. Elyada, E. *et al.* Cross-Species Single-Cell Analysis of Pancreatic Ductal Adenocarcinoma Reveals Antigen-Presenting Cancer-Associated Fibroblasts. *Cancer Discov.* **9**, 1102–1123 (2019).

62. Costa, A. *et al.* Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell* **33**, 463-479.e10 (2018).

63. Kieffer, Y. *et al.* Single-Cell Analysis Reveals Fibroblast Clusters Linked to Immunotherapy Resistance in Cancer. *Cancer Discov.* **10**, 1330–1351 (2020).

64. Kim, J. *et al.* Single-cell analysis of gastric pre-cancerous and cancer lesions reveals cell lineage diversity and intratumoral heterogeneity. *Npj Precis. Oncol.* **6**, 1–11 (2022).

65. Grunberg, N. *et al.* Cancer-Associated Fibroblasts Promote Aggressive Gastric Cancer Phenotypes via Heat Shock Factor 1–Mediated Secretion of Extracellular Vesicles. *Cancer Res.* **81**, 1639–1653 (2021).

66. Kumar, V. *et al.* Single-Cell Atlas of Lineage States, Tumor Microenvironment, and Subtype-Specific Expression Programs in Gastric Cancer. *Cancer Discov.* **12**, 670–691 (2022).

67. Orimo, A. *et al.* Stromal Fibroblasts Present in Invasive Human Breast Carcinomas Promote Tumor Growth and Angiogenesis through Elevated SDF-1/CXCL12 Secretion. *Cell* **121**, 335–348 (2005).

68. Goulet, C. R. *et al.* Cancer-associated fibroblasts induce epithelial–mesenchymal transition of bladder cancer cells through paracrine IL-6 signalling. *BMC Cancer* **19**, 137 (2019).

69. Glabman, R. A., Choyke, P. L. & Sato, N. Cancer-Associated Fibroblasts: Tumorigenicity and Targeting for Cancer Therapy. *Cancers* **14**, 3906 (2022).

70. Tan, H.-X. *et al.* CXCR4/TGF-β1 mediated hepatic stellate cells differentiation into carcinoma-associated fibroblasts and promoted liver metastasis of colon cancer. *Cancer Biol. Ther.* **21**, 258–268 (2019).

71. Kinoshita, H. *et al.* Interleukin-6 Mediates Epithelial–Stromal Interactions and Promotes Gastric Tumorigenesis. *PLoS ONE* **8**, e60914 (2013).

72. Winkler, J., Abisoye-Ogunniyan, A., Metcalf, K. J. & Werb, Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat. Commun.* **11**, 5120 (2020).

73. Yang, X. *et al.* FAP Promotes Immunosuppression by Cancer-Associated Fibroblasts in the Tumor Microenvironment via STAT3–CCL2 Signaling. *Cancer Res.* **76**, 4124–4135 (2016).

74. Chen, X. & Song, E. Turning foes to friends: targeting cancer-associated fibroblasts. *Nat. Rev. Drug Discov.* **18**, 99–115 (2019).

75. Hessmann, E. *et al.* Fibroblast drug scavenging increases intratumoural gemcitabine accumulation in murine pancreas cancer. *Gut* **67**, 497–507 (2018).

76. Ma, J., Song, X., Xu, X. & Mou, Y. Cancer-Associated Fibroblasts Promote the Chemo-resistance in Gastric Cancer through Secreting IL-11 Targeting JAK/STAT3/Bcl2 Pathway. *Cancer Res. Treat. Off. J. Korean Cancer Assoc.* **51**, 194–210 (2019).

77. Qiao, Y. *et al.* IL6 derived from cancer-associated fibroblasts promotes chemoresistance via CXCR7 in esophageal squamous cell carcinoma. *Oncogene* **37**, 873–883 (2018).

78. Zhang, H. *et al.* Cancer-associated fibroblasts mediated chemoresistance by a FOXO1/TGFβ1 signaling loop in esophageal squamous cell carcinoma. *Mol. Carcinog.* **56**, 1150–1163 (2017).

79. Resovi, A. *et al.* Soluble stroma-related biomarkers of pancreatic cancer. *EMBO Mol. Med.* **10**, e8741 (2018).

80. Torres, S. *et al.* Proteome Profiling of Cancer-Associated Fibroblasts Identifies Novel Proinflammatory Signatures and Prognostic Markers for Colorectal Cancer. *Clin. Cancer Res.* **19**, 6006–6019 (2013).

81. Calon, A. *et al.* Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. *Nat. Genet.* **47**, 320–329 (2015).

82. Underwood, T. J. *et al.* Cancer-associated fibroblasts predict poor outcome and promote periostin-dependent invasion in oesophageal adenocarcinoma. *J. Pathol.* **235**, 466–477 (2015).

83. Scott, A. M. *et al.* A Phase I dose-escalation study of sibrotuzumab in patients with advanced or metastatic fibroblast activation protein-positive cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **9**, 1639–1647 (2003).

84. Hofheinz, R.-D. *et al.* Stromal Antigen Targeting by a Humanised Monoclonal Antibody: An Early Phase II Trial of Sibrotuzumab in Patients with Metastatic Colorectal Cancer. *Oncol. Res. Treat.* **26**, 44–48 (2003).

85. Alvarez, R. *et al.* Stromal disrupting effects of nab-paclitaxel in pancreatic cancer. *Br. J. Cancer* **109**, 926–933 (2013).

86. Feng, R. *et al.* Nab-paclitaxel interrupts cancer-stromal interaction through C-X-C motif chemokine 10-mediated interleukin-6 downregulation in vitro. *Cancer Sci.* **109**, 2509–2519 (2018).

87. West, H. *et al.* Atezolizumab in combination with carboplatin plus nab-paclitaxel chemotherapy compared with chemotherapy alone as first-line treatment for metastatic non-

squamous non-small-cell lung cancer (IMpower130): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol.* **20**, 924–937 (2019).

88. Muniz-Bongers, L. R. *et al.* MMP2 and TLRs modulate immune responses in the tumor microenvironment. *JCI Insight* **6**, (2021).

89. Peng, K., Zhang, Y., Liu, D. & Chen, J. MMP2 is a immunotherapy related biomarker and correlated with cancer-associated fibroblasts infiltrate in melanoma. *Cancer Cell Int.* **23**, 26 (2023).

90. Liu, B. *et al.* Tumor-associated macrophage-derived CCL20 enhances the growth and metastasis of pancreatic cancer. *Acta Biochim. Biophys. Sin.* **48**, 1067–1074 (2016).

91. Samaniego, R. *et al.* CCL20 Expression by Tumor-Associated Macrophages Predicts Progression of Human Primary Cutaneous Melanoma. *Cancer Immunol. Res.* **6**, 267–275 (2018).

92. Walch-Rückheim, B. *et al.* Stromal Fibroblasts Induce CCL20 through IL6/C/EBPβ to Support the Recruitment of Th17 Cells during Cervical Cancer Progression. *Cancer Res.* **75**, 5248–5259 (2015).

93. Xu, T. *et al.* ISG20 serves as a potential biomarker and drives tumor progression in clear cell renal cell carcinoma. *Aging* **12**, 1808–1827 (2020).

94. Gao, J. *et al.* C20orf27 Promotes Cell Growth and Proliferation of Colorectal Cancer via the TGFβR-TAK1-NFκB Pathway. *Cancers* **12**, 336 (2020).

95. Romero-Pérez, L., Surdez, D., Brunet, E., Delattre, O. & Grünewald, T. G. P. STAG Mutations in Cancer. *Trends Cancer* **5**, 506–520 (2019).

96. Hu, H. *et al.* Three subtypes of lung cancer fibroblasts define distinct therapeutic paradigms. *Cancer Cell* **39**, 1531-1547.e10 (2021).

97. Maehira, H. *et al.* Vimentin Expression in Tumor Microenvironment Predicts Survival in Pancreatic Ductal Adenocarcinoma: Heterogeneity in Fibroblast Population. *Ann. Surg. Oncol.* **26**, 4791–4804 (2019).

98. Richardson, A. M. *et al.* Vimentin Is Required for Lung Adenocarcinoma Metastasis via Heterotypic Tumor Cell–Cancer-Associated Fibroblast Interactions during Collective Invasion. *Clin. Cancer Res.* **24**, 420–432 (2018).

99. Berr, A. L. *et al.* Vimentin is Required for Tumor Progression and Metastasis in a Mouse Model of Non-Small Cell Lung Cancer. 2020.06.04.130963 Preprint at https://doi.org/10.1101/2020.06.04.130963 (2020).

100. Ariga, N., Sato, E., Ohuchi, N., Nagura, H. & Ohtani, H. Stromal expression of fibroblast activation protein/seprase, a cell membrane serine proteinase and gelatinase, is associated with longer survival in patients with invasive ductal carcinoma of breast. *Int. J. Cancer* **95**, 67–72 (2001).

101. Chuang, W.-Y. *et al.* Concordant podoplanin expression in cancer-associated fibroblasts and tumor cells is an adverse prognostic factor in esophageal squamous cell carcinoma. *Int. J. Clin. Exp. Pathol.* **7**, 4847–4856 (2014).

102. Shindo, K. *et al.* Podoplanin expression in cancer-associated fibroblasts enhances tumor progression of invasive ductal carcinoma of the pancreas. *Mol. Cancer* **12**, 168 (2013).

103. Fei, F., Qu, J., Zhang, M., Li, Y. & Zhang, S. S100A4 in cancer progression and metastasis: A systematic review. *Oncotarget* **8**, 73219–73239 (2017).

104. Gocher, A. M., Workman, C. J. & Vignali, D. A. A. Interferon-γ: teammate or opponent in the tumour microenvironment? *Nat. Rev. Immunol.* **22**, 158–172 (2022).

105. Jorgovanovic, D., Song, M., Wang, L. & Zhang, Y. Roles of IFN-γ in tumor progression and regression: a review. *Biomark. Res.* **8**, 49 (2020).

106. Cyr-Depauw, C. *et al.* Chordin-Like 1 Suppresses Bone Morphogenetic Protein 4-Induced Breast Cancer Cell Migration and Invasion. *Mol. Cell. Biol.* **36**, 1509–1525 (2016).

107. Kon, A. *et al.* Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nat. Genet.* **45**, 1232–1237 (2013).

108. Thota, S. *et al.* Genetic alterations of the cohesin complex genes in myeloid malignancies. *Blood* **124**, 1790 (2014).

109. Chu, Z. *et al.* STAG2 regulates interferon signaling in melanoma via enhancer loop reprogramming. *Nat. Commun.* **13**, 1859 (2022).

110. Sun, Y. *et al.* Histone acetyltransferase 1 promotes gemcitabine resistance by regulating the PVT1/EZH2 complex in pancreatic cancer. *Cell Death Dis.* **12**, 1–13 (2021).

111. Fan, P. *et al.* Overexpressed histone acetyltransferase 1 regulates cancer immunity by increasing programmed death-ligand 1 expression in pancreatic cancer. *J. Exp. Clin. Cancer Res.* **38**, 47 (2019).

112. Xu, J., Palestino Dominguez, M. & Alewine, C. Loss of ID3 in pancreatic cancer cells increases DNA damage without impairing MDC1 recruitment to the nuclear foci. *Cancer Commun.* **42**, 269–272 (2021).

APPENDICES

Supplemental Table 1. Media recipes

Fibroblast Medium		
Component	**Final Concentration**	**For 500 mL**
DMEM+Glutamax-1 (Invitrogen 10569-010)		500 mL
Gentamicin (Invitrogen 15710-064) (10 mg/mL)	50 ug/mL (1:200)	2.5 mL
EGF (100 ug/mL)	5 ng/mL (1:20,000)	25 uL
Insulin (Invitrogen 12585-014) (4 mg/mL)	5 ug/mL (1:800)	625 uL
BPE (Wisent 002-011-IL) (5mL)	(1:400)	1.25 mL
Normocin (Invivogen ANT-NR2) (50 mg/mL)	100 ug/mL (1:500)	1 mL
FBS	10%	50 mL
AD-DF+++ medium		
Component	**Final Concentration**	**For 150 mL**
Advanced DMEM/F12 (Invitrogen 12634-010)		145.5 mL
Glutamax-1 (Invitrogen 35050-061) (200 mM)	2 mM (1:100)	1.5 mL
HEPES (Invitrogen 15630-080) (1M)	10 mM (1:100)	1.5 mL
Pen/Strep (Sigma P4333) (100x)	100 U/mL	1.5 mL
IntestiCult medium		
Component	**Final Concentration**	**For 100 mL**
Component A		50 mL
Component B		50 mL
Pen/Strep (Sigma P4333) (100x)	100 U/mL	1 mL
ROCK inhibitor (50 mM)	10 uM (1:1000)	20 uL

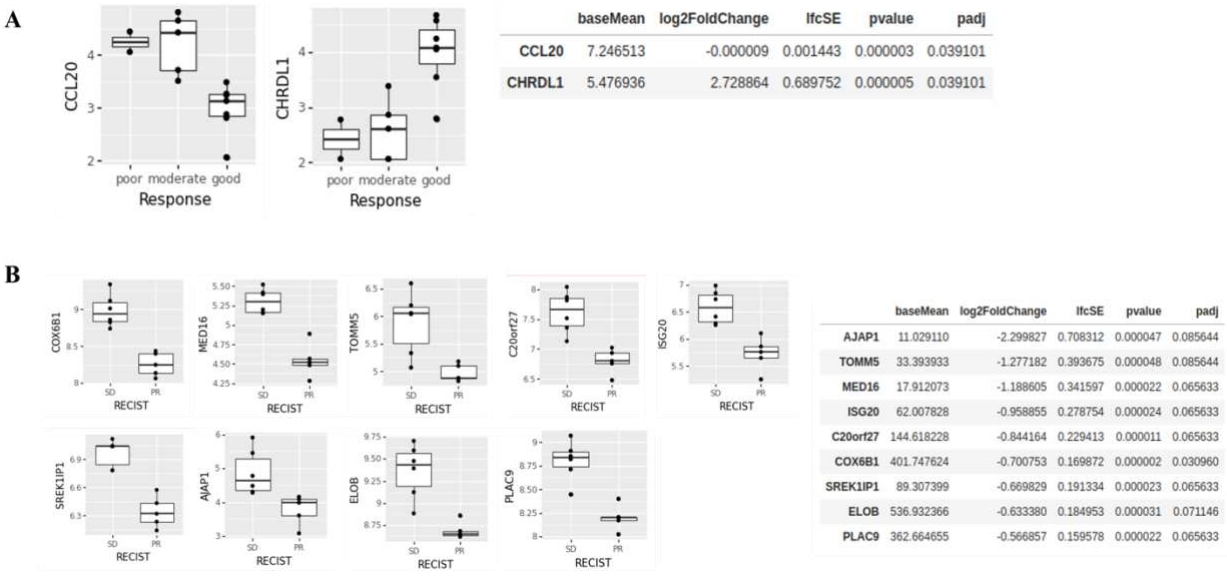
Complete Growth Media		
Component	**Final Concentration**	**For 50 mL**
Intesticult medium		25 mL
Gastric medium (homemade recipe)		25 mL
FACS Buffer		
Component	**Final Concentration**	**For 50 mL**
EDTA	2mM	47.5 mL
BSA	1%	2.5 mL

Supplemental Table 2. Antibodies for immunofluorescence and flow cytometry

Antibody	**Company**	**Catalog No.**
Immunofluorescence		
LUM	Abcam	ab168348
FAP	Abcam	ab53066
S100A4	Abcam	ab124805
ACTA2	ThermoFisher	MA5-11547
CXCL12	Novus Bio	JJ09-20
PDGFRa	R&D Systems	AF307-SP
PDGFRb	R&D Systems	AF385-SP
DCN	R&D Systems	MAB143-SP
COL1A1	R&D Systems	MAB6220-SP
VIM	Sigma Aldrich	V4630
Phalloidin 647	ThermoFisher	A22287
Donkey anti-Rabbit 555	ThermoFisher	A-31572
Donkey anti-Mouse 555	ThermoFisher	A-31570

Donkey anti-Mouse 488	ThermoFisher	A-21202
Donkey anti-Goat 488	ThermoFisher	A-11055
FACS		
FAP, AF488	R&D Systems	FAB3715G-025
PDGFRa, AF647	BD BioSciences	562798
PDGFRb, PE	BD BioSciences	558821
CD74, BV421	BD BioSciences	743731
PDPN, APC/Cy7	Biolegend	337029
Aqua blue 405 viability stain	ThermoFisher	L34957

Supplemental Figure 1. Differentially expressed genes in iCAF and myCAF populations according to chemoresponse

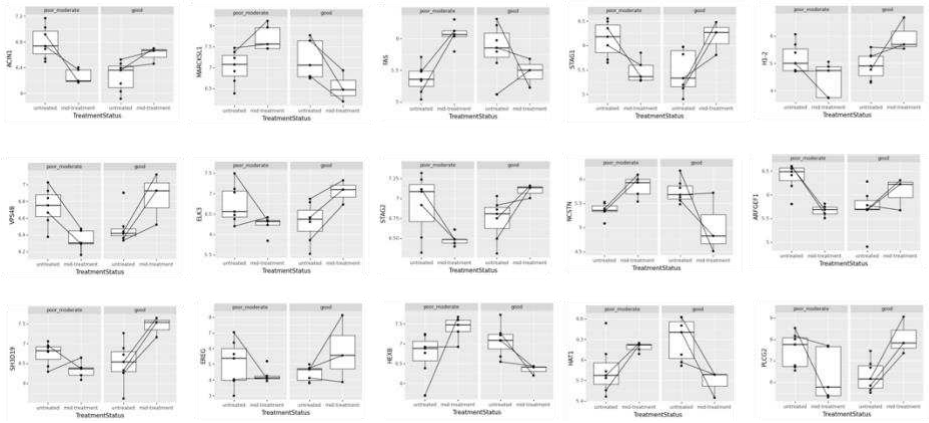


	baseMean	log2FoldChange	lfcSE	pvalue	padj
CCL20	7.246513	-0.000009	0.001443	0.000003	0.039101
CHRDL1	5.476936	2.728864	0.689752	0.000005	0.039101

	baseMean	log2FoldChange	lfcSE	pvalue	padj
AJAP1	11.029110	-2.299827	0.708312	0.000047	0.085644
TOMM5	33.393933	-1.277182	0.393675	0.000048	0.085644
MED16	17.912073	-1.188605	0.341597	0.000022	0.065633
ISG20	62.007828	-0.958855	0.278754	0.000024	0.065633
C20orf27	144.618228	-0.844164	0.229413	0.000011	0.065633
COX6B1	401.747624	-0.700753	0.169872	0.000002	0.030960
SREK1IP1	89.307399	-0.669829	0.191334	0.000023	0.065633
ELOB	536.932366	-0.633380	0.184953	0.000031	0.071146
PLAC9	362.664655	-0.566857	0.159578	0.000022	0.065633

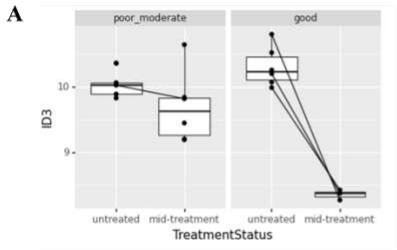
(A) DEGs, CCL20 and CHRDL1, in iCAFs according to pathological response. **(B)** DEGs in myCAFs accordings to clinical response.

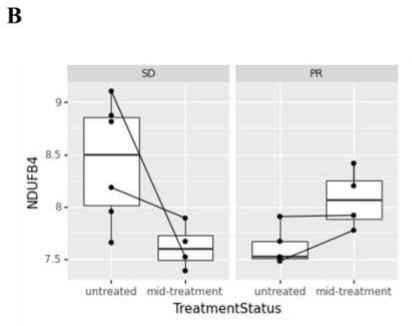
Supplementary Figure 2. DEGs in iCAFs across two timepoints, treatment naïve biopsy and mid-treatment biopsy, according to pathological response to neoadjuvant chemotherapy.



	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
SH3D19	100.451842	1.382174	0.329786	17.887150	0.000023	0.075957
STAG1	50.451882	1.613639	0.376141	18.784959	0.000015	0.075957
HEXB	122.233067	-1.483304	0.354735	16.832127	0.000041	0.088220
PLCG2	152.464625	2.873628	0.776533	13.932940	0.000189	0.096447
H1-2	29.575683	2.319303	0.636603	13.635391	0.000222	0.096447
VPS4B	89.722334	0.916359	0.231587	15.802261	0.000070	0.096447
ARFGEF1	57.450006	1.182274	0.306363	15.067995	0.000104	0.096447
MARCKSL1	145.258969	-1.457689	0.376431	14.458103	0.000143	0.096447
FAS	45.912874	-1.288417	0.333803	14.643464	0.000130	0.096447
ACIN1	83.651435	0.930433	0.239669	15.288020	0.000092	0.096447
STAG2	105.588650	0.922306	0.251713	13.624753	0.000223	0.096447
NCSTN	40.992153	-1.466002	0.389172	13.703920	0.000214	0.096447
EREG	35.958843	4.424474	1.193746	13.634814	0.000222	0.096447
HAT1	60.509638	-1.116804	0.286559	14.813895	0.000119	0.096447
ELK3	90.182142	1.382858	0.371592	14.077309	0.000175	0.096447

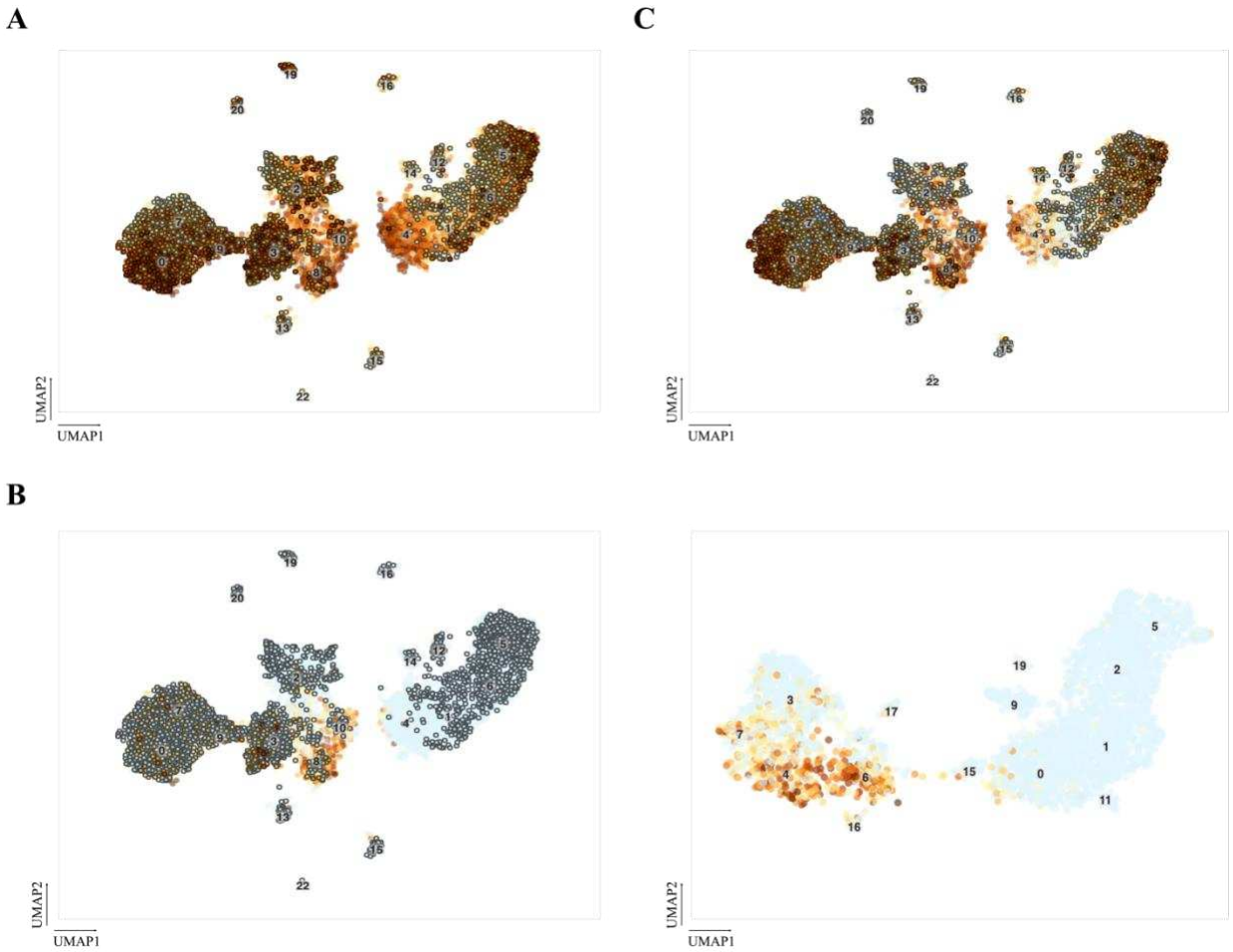
Supplementary Figure 3. DEGs in myCAFs across two timepoints, treatment naïve biopsy and mid-treatment biopsy, according to (A) pathological response, and (B) clinical response.





	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
NDUFB4	255.286192	1.440383	0.370180	14.946946	0.000111	0.035741
COX17	170.165002	1.310323	0.336804	14.954295	0.000110	0.035741
RPL32	1951.964596	0.816181	0.205352	15.705816	0.000074	0.035741
COX6A1	279.049134	0.984817	0.255043	14.807900	0.000118	0.035741
RPL34	2488.011253	0.856785	0.243411	12.308644	0.000451	0.081315
GABARAPL2	157.705812	0.939773	0.284586	10.932265	0.000997	0.081315
MRPL33	176.047615	1.195316	0.352082	11.400862	0.000734	0.081315
RPL26	1147.515361	1.318740	0.401423	10.946122	0.001103	0.081315
MTRNR2L31	1076.387401	-2.142894	0.653915	10.569766	0.001150	0.081315
RPS23	1567.735370	1.144196	0.344010	10.942579	0.000940	0.081315
RPL41	5018.062378	0.970775	0.297379	10.567386	0.001151	0.081315
RPL38	824.225828	1.084735	0.315772	11.691753	0.000828	0.081315
SOD1	311.659784	0.703050	0.203664	11.863623	0.000572	0.081315
COX8C	540.604698	1.067193	0.317674	11.170347	0.000831	0.081315
COX6B1	422.859627	0.912477	0.276574	10.825997	0.001001	0.081315
SEC62	306.782160	0.953856	0.276928	11.807582	0.000590	0.081315
RPS25	990.031831	1.249512	0.365442	11.546619	0.000879	0.081315
RPS27A	2182.411965	0.790186	0.251309	9.829074	0.001718	0.096894
RSL24D1	169.459617	0.804011	0.285725	9.961706	0.001596	0.096894
TCEAL4	228.035024	0.838172	0.266439	9.869669	0.001680	0.096894
ATP5PD	263.278719	0.952489	0.299444	10.058319	0.001517	0.096894
PRDX1	193.073798	1.348975	0.429137	9.766985	0.001777	0.096894

Supplemental Figure 4. Validation of commonly reported markers in scRNA-seq atlas.



(A) VIM expression in normal (highlighted) vs. tumor cells. (B) Left: FAP expression in normal (highlighted) vs. tumor cells. Right: FAP expression in treatment naïve biopsy fibroblasts. (C) S100A4 expression in normal (highlighted) vs. tumor cells. Dark brown indicates high expression, light yellow indicates low expression, blue indicates no expression.