Analysis of Single Circulating Tumor Cells (CTCs) to Infer Phenotype and Genome Changes in Response to Therapeutic Pressures in Biliary Tract Cancer

Thesis

How to cite:

For guidance on citations see FAQs.

© 2020 The Author

https://creativecommons.org/licenses/by-nc-nd/4.0/

Version: Version of Record

Link(s) to article on publisher’s website:
http://dx.doi.org/doi:10.21954/ou.ro.0001179d

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.

oro.open.ac.uk
Analysis of single circulating tumor cells (CTCs) to infer phenotype and genome changes in response to therapeutic pressures in biliary tract cancer

Thesis presented for the Degree of Doctor of Philosophy

The Open University, Milton Keynes (UK)
School of Life, Health and Chemical Sciences

Carolina Reduzzi
M.Sc. in Biology Applied to Biomedical Research
Personal identifier: F2823860

March, 2020
ABSTRACT
Biliary tract cancer (BTC) is a highly fatal disease mainly treated with standard chemotherapy, albeit with limited efficacy. New therapeutic options are greatly needed, but the use of targeted treatments is often prevented by the impossibility of obtaining tissue biopsies for molecular characterization. Circulating tumor cells (CTCs) could represent a valuable alternative to tissue biopsies, also allowing a real-time monitoring of disease evolution and therapeutic resistance. Unfortunately, by using conventional CTC detection methods, which are based on the expression of epithelial markers, CTCs have been detected in only a small proportion of BTC patients, thus limiting their implementation in BTC clinical management.

Here, we developed a protocol for the identification of not only the classic epithelial CTCs (eCTCs), but also of non-conventional CTCs (ncCTCs) lacking epithelial and leukocyte markers, but presenting aberrant genomes, and therefore representing bona fide CTCs. CTCs were analyzed in 41 blood samples longitudinally collected from 21 patients with BTC. The detection of ncCTCs in addition to eCTCs resulted in an increase in CTC-positivity from 19% to 83%. The presence of at least 1 eCTC/10 mL of blood at baseline was associated with a significantly shorter median disease-specific survival (9 months vs. 17 months, \( p=0.03 \)). Conversely, ncCTCs were not prognostic but variations in their number during treatment mirrored patient response, supporting their role for treatment monitoring.

The developed workflow also allowed the molecular characterization of single-CTCs. Copy number alteration profiling was performed for 88 single-CTCs collected from 23 BTC patients. Unsupervised clustering analysis revealed a segregation of CTCs according to patient best response and allowed the identification of genomic regions possibly involved in mechanisms of therapeutic resistance.
Overall, our results demonstrate the presence of a novel subpopulation of CTCs in BTC, paving the way for the use of liquid biopsy to improve clinical management of BTC patients.
This study was conducted under the supervision of Drs. Vera Cappelletti and Nadia Zaffaroni at the Biomarkers Unit, headed by Dr. Maria Grazia Daidone, at the Department of Applied Research & Technological Development, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy.
TABLE OF CONTENTS
ABSTRACT ..............................................................................................................3
TABLE OF CONTENTS ...........................................................................................7
LIST OF FIGURES AND TABLES ........................................................................10
1. INTRODUCTION .................................................................................................17
   1.1. Biliary tract cancer .......................................................................................18
       1.1.1. Clinical management of BTC: State of the art .....................................20
       1.1.2. Clinical management of BTC: Future challenges .........................23
   1.2. Circulating tumor cells ...............................................................................29
       1.2.1. First generation of CTC studies: CTC enumeration ......................31
           1.2.1.1. CTC enumeration in patients with BTC .................................34
       1.2.2. Technologies to capture CTC heterogeneity ..................................35
       1.2.3. New generation of CTC studies: CTC characterization at the single-cell level 43
2. SCOPE OF THE THESIS ....................................................................................52
3. MATERIALS AND METHODS ........................................................................55
   3.1. Patient information and clinical sample collection ..............................56
   3.2. Cell lines and culture conditions .............................................................57
   3.3. Flow cytometry analysis .........................................................................58
   3.4. Spike-in procedure .................................................................................59
   3.5. Enrichment methods ..............................................................................59
       3.5.1. ScreenCell® filters® ..........................................................................59
       3.5.2. AutoMACS® Pro separator .................................................................60
       3.5.3. OncoQuick® ....................................................................................60
       3.5.4. Parsortix® .......................................................................................60
   3.6. Post-enrichment procedures ....................................................................62
       3.6.1. Fixation and staining ....................................................................62
       3.6.2. DEPArray™ analysis ......................................................................63
   3.7. Molecular analysis .....................................................................................65
       3.7.1. Whole-genome amplification and quality control assay ...............65
       3.7.2. Mutational profiling ........................................................................65
       3.7.3. Copy number alteration profiling ..................................................66
   3.8. RNA analysis .............................................................................................67
   3.9. Dual-positive cell evaluation by CellSearch® .......................................69
   3.10. Statistical analyses .................................................................................70
4. RESULTS ............................................................................................................71
   4.1. Development of the protocol for CTC analysis ......................................72
## Table of contents

4.1.1. Analytical validity assessment ................................................................. 73
4.1.2. Enrichment methods .................................................................................. 80
  4.1.2.1. ScreenCell® filters ........................................................................... 81
  4.1.2.2. AutoMACS® Pro separator ................................................................. 81
  4.1.2.3. OncoQuick® ..................................................................................... 82
  4.1.2.4. Parsortix® ........................................................................................ 84
4.1.3. Identification strategies and phenotypic evaluation .................................... 86
  4.1.3.1. Positive selection ............................................................................. 86
  4.1.3.2. Negative selection ........................................................................... 87
  4.1.3.3. Phenotypic evaluation ..................................................................... 88
4.1.4. Whole-genome amplification and quality control assay ............................. 90
4.1.5. Copy number alteration profiling ............................................................. 91

4.2. Detection of CTCs in blood samples from patients with BTC ....................... 93
  4.2.1. Detection of eCTCs and ncCTCs ............................................................ 94
  4.2.2. Prognostic role of CTCs ....................................................................... 97
  4.2.3. Treatment monitoring .......................................................................... 99
  4.2.4. Phenotypic characterization .................................................................. 102
  4.2.5. Molecular characterization .................................................................... 105
    4.2.5.1. Mutational profiling ....................................................................... 105
    4.2.5.2. Copy number alteration analysis .................................................... 107
      4.2.5.2.1. Chromosomal instability evaluation ......................................... 107
      4.2.5.2.2. Investigation of alterations involved in therapeutic resistance .... 110
    4.2.5.3. Double-negative cells in healthy donors ........................................ 113

4.3. Exploratory studies ...................................................................................... 114
  4.3.1. Analysis of CTCs’ RNA ........................................................................ 114
  4.3.2. Characterization of dual-positive cells .................................................. 118

5. DISCUSSION ................................................................................................... 122

6. CONCLUSIONS & FUTURE PERSPECTIVES ............................................... 134

7. REFERENCES .................................................................................................. 137

PUBLICATIONS ................................................................................................. 166

DECLARATION OF AUTHORSHIP ..................................................................... 169
LIST OF FIGURES AND TABLES
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Anatomical sub-variants of BTC.</td>
<td>18</td>
</tr>
<tr>
<td>1.2</td>
<td>Algorithm for the management of patients with BTC.</td>
<td>20</td>
</tr>
<tr>
<td>1.3</td>
<td>Molecular spectrum of intrahepatic and extrahepatic CCA</td>
<td>24</td>
</tr>
<tr>
<td>1.4</td>
<td>Polyclonal seeding and tumor evolution in BTC</td>
<td>26</td>
</tr>
<tr>
<td>1.5</td>
<td>Changes occurring in the molecular landscape of advanced BTC</td>
<td>27</td>
</tr>
<tr>
<td>1.6</td>
<td>CTCs as a real-time liquid biopsy</td>
<td>30</td>
</tr>
<tr>
<td>1.7</td>
<td>Clinical validity of CTC enumeration by CellSearch®.</td>
<td>32</td>
</tr>
<tr>
<td>1.8</td>
<td>Prognostic impact of CTCs in patients with BTC.</td>
<td>34</td>
</tr>
<tr>
<td>1.9</td>
<td>Phenotypic changes in CTCs in response to treatment in breast cancer patients.</td>
<td>36</td>
</tr>
<tr>
<td>1.10</td>
<td>New methods for CTC detection.</td>
<td>38</td>
</tr>
<tr>
<td>1.11</td>
<td>Standard workflow for single-CTC sequencing.</td>
<td>39</td>
</tr>
<tr>
<td>1.12</td>
<td>Genomic alterations in single CTCs.</td>
<td>42</td>
</tr>
<tr>
<td>1.13</td>
<td>Dendrogram representing hierarchical clustering and the timeline of sample acquisition.</td>
<td>44</td>
</tr>
<tr>
<td>1.14</td>
<td>Intra- and inter-patient genomic heterogeneity of single CTCs.</td>
<td>46</td>
</tr>
<tr>
<td>1.15</td>
<td>Clonality and genomic alterations in single CTCs.</td>
<td>48</td>
</tr>
<tr>
<td>1.16</td>
<td>CTC CNA-based classifier and clinical outcome.</td>
<td>50</td>
</tr>
<tr>
<td>3.1</td>
<td>Evaluation of DPcells by CellSearch®.</td>
<td>69</td>
</tr>
<tr>
<td>4.1</td>
<td>Workflow for CTC analysis.</td>
<td>72</td>
</tr>
<tr>
<td>4.2</td>
<td>Image gallery showing unspecific staining of cells analyzed with the DEPArray™.</td>
<td>74</td>
</tr>
<tr>
<td>4.3</td>
<td>Flow cytometry analysis of WBCs isolated from HD blood samples collected in EDTA or in Streck tubes and stained with distinct antibodies directed against EpCAM.</td>
<td>76</td>
</tr>
<tr>
<td>4.4</td>
<td>Image gallery showing staining and morphological characteristics of cells analyzed with the DEPArray™.</td>
<td>78</td>
</tr>
<tr>
<td>4.5</td>
<td>DEPArray™ image gallery showing MCF7 cells identified in a spiked-in HD blood sample.</td>
<td>80</td>
</tr>
<tr>
<td>4.6</td>
<td>Graphic representation of OncoQuick® enrichment.</td>
<td>82</td>
</tr>
<tr>
<td>4.7</td>
<td>MCF7 cells identified during DEPArray™ analysis of a spiked-in sample enriched with OncoQuick®.</td>
<td>83</td>
</tr>
<tr>
<td>4.8</td>
<td>Graphic representation of Parsortix® enrichment.</td>
<td>84</td>
</tr>
</tbody>
</table>
**Figure 4.9.** CTCs, WBCs and double-negative cells identified in two blood samples from biliary tract cancer patients. .................................................................89

**Figure 4.10.** Copy number alteration profiles of control DNA samples. .................................................92

**Figure 4.11.** Number of patients enrolled and included in each type of analysis. .................................94

**Figure 4.12.** Identification of eCTCs, ncCTCs and WBCs in clinical samples. ............................................95

**Figure 4.13.** Association between eCTCs and survival. .................................................................97

**Figure 4.14.** Effect of treatment on CTCs in BTC patients.................................................................99

**Figure 4.15.** CTC monitoring during treatment. ......................................................................................101

**Figure 4.16.** Vimentin expression in eCTCs. .........................................................................................102

**Figure 4.17.** MCF7 cells identified in a spiking experiment.................................................................103

**Figure 4.18.** Non-aberrant, circulating double-negative cells expressing CSV .................................104

**Figure 4.19.** Mutational profiling of 9 CTCs recovered from one blood sample from a BTC patient. ........................................................................................................106

**Figure 4.20.** CTCs’ chromosomal instability in responding and non-responding patients. ....................................................................................................................110

**Figure 4.21.** Clustering analysis of CNA profiles of single CTCs isolated from BTC patients. ....................................................................................................................111

**Figure 4.22.** Comparison of CNA frequency in 2 specific chromosome regions, detected in CTCs from clusters 2 and 3. .........................................................................................112

**Figure 4.23.** Detection of HuCCT-1 cells in HD WBCs by RNA analysis. .................................116

**Figure 4.24.** Detection of HuCCT-1 cells in HD WBCs by RNA analysis after transcripts’ pre-amplification ...........................................................................................................117

**Figure 4.25.** DPcells detected in BTC patients. ......................................................................................118

**Figure 4.26.** DPcells and DP-CTCs distribution in clinical samples. ....................................................119

**Figure 4.27.** Combined DPcell- and CTC- stratification in association with progression-free survival. ........................................................................................................................121

**Table 3.1.** Antibodies used to stain surface markers. .................................................................62

**Table 3.2.** Antibodies used to stain intracellular markers. .................................................................63

**Table 4.1.** Antibodies against EpCAM compared in the flow cytometry experiments........75

**Table 4.2.** Recovery rates of samples enriched by ScreenCell® filters. ........................................81

**Table 4.3.** Recovery rates of samples enriched by OncoQuick®..............................................................82

**Table 4.4.** Recovery rates of samples enriched by Parsortix®..............................................................85
Table 4.5. Antibodies added to the protocol and tested for sensitivity and specificity. ..........86
Table 4.6. Patients’ clinico-pathological characteristics. .................................................93
Table 4.7. Number of eCTCs and ncCTCs detected for each analyzed blood sample..........96
Table 4.8. Association between eCTC-status and clinical characteristics of the patients
included in the survival analysis......................................................................................98
Table 4.9. Variations in CTC numbers during treatment. .................................................100
Table 4.10. LST scores of CTCs recovered from patients responding to treatment. ........108
Table 4.11. LST scores of CTCs recovered from patients non-responding to treatment..109
ADO  Allele drop out
ALK  Anaplastic lymphoma kinase
AR   Androgen receptor
BL   Baseline
BTC  Biliary tract cancer
CCA  Cholangiocarcinoma
CI   Confidence interval
cis/gem  Cisplatin plus gemcitabine
CK   Cytokeratins
CNA  Copy number alteration
CSV  Cell surface vimentin
CTC  Circulating tumor cell
DPcell  Dual-positive cell
DSS  Disease-specific survival
DT   During treatment
eCTC  Epithelial CTC
EGFR  Epithelial growth factor receptor
EMT  Epithelial to mesenchymal transition
EOT  End of treatment
EpCAM Epithelial cell adhesion molecule
EphA3  Ephrin type-A receptor 3
FBS  Fetal bovine serum
FGFR  Fibroblast growth factor receptor
FOLFOX  Oxaliplatin, L-folinic acid and 5-fluorouracil
FU   Follow-up
GBC  Gallbladder cancer
GII  Genome integrity index
HD   Healthy donor
HER2  Human epidermal growth factor receptor 2
HR   Hazard ratio
IDH1  Isocitrate dehydrogenase 1
INDELs  Small insertions and deletions
KRT8  Cytokeratin-8
LM-PCR  Ligation-mediated PCR
lp-WGS  Low-pass whole-genome sequencing
LST  Large-scale state transition
MALBAC  Multiple annealing and looping-based amplification
MDA  Multiple displacement amplification
ncCTC  Non-conventional CTC
NGS  Next generation sequencing
OS   Overall survival
PBS  Phosphate-buffered saline
PD   Progressive disease
PFA  Paraformaldehyde
PFS  Progression-free survival
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR</td>
<td>Partial response</td>
</tr>
<tr>
<td>QC</td>
<td>Quality control</td>
</tr>
<tr>
<td>RBC</td>
<td>Red blood cell</td>
</tr>
<tr>
<td>SD</td>
<td>Stable disease</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular endothelial growth factor</td>
</tr>
<tr>
<td>VIM</td>
<td>Vimentin</td>
</tr>
<tr>
<td>WBC</td>
<td>White blood cell</td>
</tr>
<tr>
<td>WGA</td>
<td>Whole-genome amplification</td>
</tr>
</tbody>
</table>
1. INTRODUCTION
1.1. **Biliary tract cancer**

Biliary tract cancer (BTC) consists of different epithelial malignancies arising in any part of the biliary tree and includes cholangiocarcinoma (CCA), gallbladder cancer (GBC) and ampulla of Vater cancer. According to the location, cholangiocarcinomas are subdivided into intrahepatic and extrahepatic CCA (located in the intrahepatic and extrahepatic bile ducts, respectively) and the latter can be further divided into distal and hilar CCA (Figure 1.1) [Tariq N.U. et al., 2019].

![Anatomical sub-variants of BTC.](image)

**Figure 1.1. Anatomical sub-variants of BTC.**

According to the location of the tumor, BTCs are subdivided into gallbladder cancer (GBC), ampulla of Vater cancer (AVC), intrahepatic cholangiocarcinoma (IHC) and extrahepatic cholangiocarcinoma (EHC), further subdivided into perihilar and distal extrahepatic cholangiocarcinoma. [Tariq N. et al., 2019]

BTC accounts for approximately 3% of all gastrointestinal malignancies [Charbel H. and Al-Kawas F.H., 2011] and is the second most common hepatobiliary cancer, after hepatocellular carcinoma [Hennedige T.P. et al., 2014]. Although in most countries it is a
rare disease (< 6 cases per 100,000 people) [Banales J.M. et al., 2016], its incidence is exceptionally high in some Eastern countries (up to 85 cases per 100,000 people for northeast Thailand) due to different geographical risk factors and genetic determinants [Khan S.A. et al., 2019]. In 2017, the global BTC incidence was 211,000 cases, with 174,000 deaths [Global Burden of Disease Collaboration, 2019] and, over the past decades, both its incidence and its mortality have increased worldwide, in particular with regards to intrahepatic CCA [Taylor-Robinson S.D. et al., 2001; Patel T., 2002; Khan S.A. et al., 2002; Bertuccio P. et al., 2013; Global Burden of Disease Cancer Collaboration, 2015; Bertuccio P. et al., 2019].

BTCs are aggressive diseases characterized by a poor prognosis (5-years survival rate = 5-15%, considering all stages) [Anderson C. and Kim R., 2009; Lamarca A. et al., 2020]. Moreover, since they are generally asymptomatic in early stages, most BTCs are diagnosed at metastatic stage, when the 5-year survival rate is only 2% [Tariq N.U. et al., 2019].
1.1.1. Clinical management of BTC: State of the art

Currently, the treatment of BTC is not based on the anatomical subtypes, but solely on the stage of the disease and it essentially consists of surgery and systemic chemotherapy (Figure 1.2) [Valle J.W. et al., 2016].

![Algorithm for the management of patients with BTC.](image)

The ESMO guidelines for the treatment of early-stage, locally-advanced and metastatic biliary tract cancer are reported. MDT, multidisciplinary team; PS, performance status; iCCA, intrahepatic cholangiocarcinoma. [Valle J.W. et al., 2016]

The only potentially curative therapy for BTC is radical surgical resection [Nathan H. et al., 2007], with a 5-year survival rate of 18% [DeOliveira M.L. et al., 2007].
Unfortunately, though, only approximately 20% of patients present an early stage disease at diagnosis and are eligible for surgery [Bridgewater J. et al., 2014]. Moreover, the majority of patients undergoing surgical resection will relapse, predominantly developing liver metastasis [Miyazaki M. et al., 2017]. Until recently, the benefit derived from the use of adjuvant therapies in reducing BTC relapse and improving patient survival was not clear due to the lack of randomized clinical trials dedicated to BTC [Lamarca A. et al., 2020]. In 2019, the BILCAP trial, a randomized phase III trial comparing the use of adjuvant capecitabine to observation alone in BTC patients who had undergone a complete resection with curative intent, was completed and showed an improvement in median overall survival (OS) for patients with BTC receiving capecitabine as adjuvant chemotherapy (36 to 53 months for patients in the observation group vs. capecitabine group, respectively, \( p = 0.028 \)) [Primrose J.N. et al., 2019]. Although the relapse rate remained high in patients treated with capecitabine (60%), the results of the BILCAP study have led to a change in the international guidelines for patients with resected BTC, which now recommend adjuvant capecitabine for 6 months after surgery as standard of care [Shroff R.T. et al., 2019].

However, for the vast majority of BTC patients (including those experiencing recurrence after surgery and those diagnosed with locally-advanced or metastatic disease), palliative systemic treatments represent the only option. Clinical trials have shown that systemic chemotherapy, compared with best supportive care, extends patients’ survival [Glimelius B. et al., 1996, Sharma A. et al. 2010] and, in 2010, cisplatin combined with gemcitabine (cis/gem) became the first-line standard of care treatment for patients with advanced BTC [Banales J.M. et al., 2016], supported by the results of two randomized trials conducted in the United Kingdom (ABC-02, phase III) and in Japan (BT22, phase II). Both the ABC-02 [Valle J.W. et al., 2010] and the BT22 [Okusaka T. et al., 2010] studies compared the use of the cis/gem combination to the use of single-agent gemcitabine and showed the
superiority of the combination over gemcitabine alone, with an increased median OS for the patients treated with cis/gem of 11.7 vs. 8.1 months (hazard ratio (HR) = 0.64; 95% confidence interval (CI): 0.52–0.80) and 11.1 vs. 7.7 months (HR = 0.69; 95% CI: 0.42–1.13) for the ABC-02 and the BT22 study, respectively. These results were further confirmed by a combined meta-analysis of the 2 trials [Valle J.W. et al., 2014] showing a significant improvement in progression-free survival, PFS, (HR = 0.64, 95% CI 0.53–0.76, \( p < 0.001 \)) and OS (HR = 0.65, 95% CI: 0.54–0.78, \( p < 0.001 \)) for cis/gem over gemcitabine alone, independent of patient age and gender, primary tumor site, prior therapy, stage of disease and ethnicity. The meta-analysis also revealed that patients with poor performance status (performance status = 2) derived the least benefit from cis/gem, suggesting that gemcitabine alone could be the best option for these patients.

Nonetheless, the effectiveness of cis/gem treatment is limited (OS < 1 year) and most patients will develop resistance and will undergo a second-line therapy. Unfortunately today there is no established second-line regimen available. In fact, different drugs have been tested over the years to treat patients with BTC progressing to first-line chemotherapy, but a systematic review [Lamarca A. et al., 2014] of 25 studies on a total of 761 patients receiving second-line systemic chemotherapy (including 14 phase II clinical trials, 9 retrospective analyses and 2 case reports) reported disappointing median PFS (3.2 months; 95% CI: 2.7-3.7) and response rate (7.7%; 95% CI: 4.6-10.9) and did not find sufficient evidence to recommend any second-line treatment. However, an ongoing randomized phase III trial (ABC-06) is comparing active symptom control alone or in combination with oxaliplatin, L-folinic acid and 5-fluorouracil (FOLFOX) as second-line therapy for advanced BTC patients. The initial results of the trial, presented last year at the ASCO conference [Lamarca A. et al., 2019], showed an increased survival after 12 months of treatment, going from 10% to 25%, for patients receiving active symptom control alone and for those additionally receiving FOLFOX, respectively. In fact, even though no
second-line therapy is currently recommended, FOLFOX is frequently used as second-line treatment in BTC patients.

1.1.2. **Clinical management of BTC: Future challenges**

Considering the modest therapeutic efficacy of chemotherapy in BTC, new therapies are urgently needed.

The first attempts to improve patient outcome by using targeted agents were focused on the vascular endothelial growth factor (VEGF) and on the epithelial growth factor receptor (EGFR) axes. In fact, both angiogenesis and the overexpression of the EGFR family have been implicated in BTC carcinogenesis [Wehbe H. et al., 2006; Pignochino Y. et al., 2010; Simone V. et al., 2017; Adeva J. et al., 2019]. Different phase II trials have tried to assess the benefit of targeting the VEGF pathway using antibodies (bevacizumab) [Zhu A.X. et al., 2010; Iyer R.V. et al., 2018; Larsen F.O. et al., 2018] or tyrosine kinase inhibitors (sorafenib, cediranib and vandetanib) [Moehler M. et al., 2014; Valle J.W. et al., 2015; Santoro A. et al., 2015] either alone or in combination with chemotherapy, but failed to produce evidence supporting the use of anti-angiogenic agents. Similarly, trials using EGFR inhibitor [Lee J. et al., 2012] or anti-EGFR antibody [Sohal D.P. et al., 2013; Malka D. et al., 2014; Chen J.S. et al., 2015] in combination with chemotherapy did not produce encouraging results. The lack of positive results, though, could be due to the fact that these trials were conducted in unselected populations.

Actually, even though in the past BTCs have been therapeutically managed as a single disease focusing only on the disease stage, recent technological advances (such as next generation sequencing, NGS) have revealed a remarkable molecular complexity of BTC, showing not only the presence of numerous alterations in putative driver and actionable genes [Churi C.R. et al., 2014; Simbolo M. et al., 2014; Jain A. and Javle M., 2016; Javle M. et al., 2016; Abou-Alfa G.K. et al., 2016; Farshidfar F. et al., 2017; Lowery M.A. et al., 2018; Wardell C.P. et al., 2018], but also a different molecular spectrum depending on the
anatomical subtype [Nakamura H. et al., 2015; Jusakul A. et al., 2017]. In particular, alterations in \textit{IDH1/2}, \textit{EPHA1}, \textit{BAP1} and \textit{FGFR2} were more frequently found in intrahepatic CCA, whereas gene fusions involving \textit{PRKACA} or \textit{PRKACB} and genetic aberration in \textit{ARID1A}, \textit{PI3KCA} and the \textit{ERBB} family were detected in extrahepatic CCA (Figure 1.3); GBC was instead characterized by \textit{ERBB3} and \textit{EGFR} mutations [Braconi C. et al., 2019].

![Figure 1.3. Molecular spectrum of intrahepatic and extrahepatic CCA.](image)

The most unique and prevalent genetic alterations found in different anatomical locations are reported. iCCA, intrahepatic cholangiocarcinoma; pCCA, perihilar cholangiocarcinoma; dCCA, distal cholangiocarcinoma; eCCA, extrahepatic cholangiocarcinoma. [Braconi C. et al., 2019]

These results highlighted the high heterogeneity of BTC and the need of including molecular profiling for clinical decisions. Through BTC sequencing analyses, promising new molecular targets have been identified and new trials have been initiated to explore their potential in BTC [Bogenberger J.M. et al., 2018; Rizvi S. et al., 2018]. In particular phase III trials targeting the isocitrate dehydrogenase 1, \textit{IDH1}, and the fibroblast growth factor receptor, \textit{FGFR}, pathways, in selected populations (carrying \textit{IDH1} mutations and
Introduction

*FGFR* fusions, respectively), are ongoing, while other targeted agents (such as anti-HER2 and PARP inhibitors) are in different stages of clinical development [Adeva J. et al., 2019]. However, as the information about BTC genomic aberrations expanded, it also revealed partially conflicting data, possibly due to differences in study populations, use of different detection technologies, and misclassifications [Kendall T. et al., 2019]. Further studies are needed to accomplish more accurate molecular profiling data, and to reach a better understanding of the differences between anatomical subtypes and between subgroups within the subtypes. Moreover, in a recent study on 4 patients with intrahepatic CCA [Walter D. et al., 2017], different samples from the same tumor were analyzed for the detection of private and common mutations to evaluate intra-tumor heterogeneity. Private mutations were identified in 3 out of 4 patients and, across all samples, the mean percentage of private mutations per sample was 12%, indicating a high level of intra-tumor heterogeneity, limiting the potential applicability of personalized medicine in BTC. On the other hand, high intra-tumor heterogeneity could be indicative of sensitivity to immune checkpoint inhibitors [McGranahan N. et al., 2016]. In fact, immunotherapy is a promising therapeutic strategy in BTC and it is currently under evaluation in different phase I and II trials [Tariq N.U. et al., 2019; Kelley R.K. et al., 2020]. Intra-tumor heterogeneity in intrahepatic CCA was investigated also in another study [Dong L-Q. et al., 2018] in which multi-regional whole-exome sequencing was performed for 6 patients and showed the presence, in all patients’ tumors, of different subclones, supporting a parallel evolution model. Moreover, for one patient from whom samples of both the primary tumor and the recurrence were available, the authors observed a subclonal structure also in the recurrence samples and the presence of private mutations at recurrence, suggesting a polyclonal metastatic seeding and clonal evolution (Figure 1.4).
Figure 1.4. Polyclonal seeding and tumor evolution in BTC.

(A) Timeline of the clinical history of patient ICC1239 from whom tissue samples were available at the time of primary resection (1239P) and the time of intrahepatic recurrence (1230R). CT/MRI images of each tumor are reported; the tumors are indicated by the red arrows. (B) Phylogenetic tree showing alterations shared by all biopsies (green), private alterations of the recurrence (orange) and private alterations of the primary tumor (grey). [Adapted from Dong L-Q. et al., 2018]

These data suggest that performing a single biopsy of the primary tumor for therapeutic decisions could not be enough and could be an explanation for the limited efficacy of targeted therapies in BTC, even in highly selected populations: In the BGJ398 phase II study, a FGFR kinase inhibitor was given to selected intrahepatic CCA patients with FGFR alterations or FGFR2 fusions but an overall response rate of only 15% was reached [Javle M. et al., 2018].

Another aspect that has to be taken into consideration is that most sequencing data produced so far were obtained by the analysis of resected tumors, while the majority of patients present with an inoperable tumor at diagnosis. Considering that tumors can evolve over time and in response to therapeutic pressure, and that current evidences suggest a polyclonal metastatic seeding of BTCs, we cannot be sure that the molecular alterations of
early stage tumors reflect the genomic landscape of advanced tumors (Figure 1.5) [Braconi C. et al., 2019].

![Figure 1.5. Changes occurring in the molecular landscape of advanced BTC.](image)

Polyclonal metastatic seeding, emergence of new clones in response to treatment, and evolution of the tumor over time (causing the loss of subclones present in the primary tumor) are mechanisms potentially causing differences in the molecular portraits of primary vs. advanced BTCs. [Braconi C. et al., 2019]

There is, therefore, the need to deepen our knowledge on the molecular characteristics of advanced BTC. Unfortunately, patients with advanced BTC are generally diagnosed by cytology or small biopsies and it is often impossible to obtain enough tumor material to perform genomic characterization analyses. One possibility to overcome the lack of tumor tissue from these patients could be the use of liquid biopsies. In particular, circulating tumor cells (CTCs) could function as an alternative source of tumor material, giving new hints on molecular drivers in advanced BTC for the development of new therapeutic strategies.

CTCs could also be used to study tumor evolution and the development of therapeutic resistance over time, without the need of performing repeated tissue biopsies (an invasive procedure which is even more difficult to perform in BTC patients, due to the anatomical location of the tumor). In fact, BTCs are characterized by a remarkable resistance to both chemotherapy and targeted therapies. This ability is probably related to the high intra-tumor heterogeneity and the presence of multiple subclones which can drive tumor progression through multiple trajectories, but could also be due to the presence of complex
mechanisms of chemoresistance that confer to BTC cells the so-called multi-drug resistance phenotype [Marin J.J.G. et al., 2018]. Given the high heterogeneity of BTCs, the genetic signatures underlying the mechanisms of chemoresistance are also diverse and can vary over time, in response to pharmacological treatment, contributing to the acquired chemoresistance [Fouassier L. et al., 2019]. Understanding the molecular bases of chemoresistance and how these evolve during treatment would allow for the prediction of the potential failure of a given treatment and would support the choice of the best therapy for each patient at a specific time. This can only be attained by getting multiple tumor biopsies over time from the same patient, which is hardly feasible in BTC patients. In this context, the analysis of CTCs could really constitute a pivotal point in the advance of our knowledge on BTC’s molecular changes during disease progression.
1.2. Circulating tumor cells

In patients with solid tumors, CTCs are released from both the primary tumor and the metastatic lesions into the bloodstream during the course of the disease. Different technologies allow the detection and the characterization of CTCs, which are therefore considered a real-time liquid biopsy of tumors [Pantel K. and Alix-Panabieres C. 2010]. The term liquid biopsy includes also the analysis of other tumor-derived elements circulating in the blood, such as circulating tumor DNA (ctDNA), tumor-derived exosomes and microvesicles, tumor-educated platelets, circulating tumor micro RNA, mRNA and non-coding RNA [Junqueira-Neto S. et al., 2019; Keller L. and Pantel K., 2019], each of which can provide different and complementary information. CTCs, in particular, being intact and viable cells, offer the possibility of performing a multilevel analysis of genotype (DNA) and phenotype (RNA and proteins). Moreover, they are a highly selected subpopulation of tumor cells, able to leave the primary tumor and survive in the bloodstream (the majority of CTCs die soon after entering the blood vessels, due to anoikis, attack by cells of the immune system and fluid shear stress [Berezovskaya O. et al., 2005; Huang Q. et al., 2018]), suggesting that CTCs could be representative of the most aggressive clones of the tumor [Pantel K. and Alix-Panabieres C. 2013]. Overall, CTC analysis can potentially be used for i) early detection of cancer, ii) prognostic stratification of patients, iii) identification of therapeutic targets, iv) prediction of response to targeted treatments, v) treatment monitoring and vi) identification of resistance mechanisms (Figure 1.6) [Alix-Panabieres C. and Pantel K. 2016]. Whereas for the latter applications a large amount of data have been produced (and will be discussed subsequently), data regarding the use of CTCs for early detection are scanty and limited to lung cancer. In a study by Ilie and colleagues, CTCs could be detected in a small proportion of high-risk individuals with chronic obstructive pulmonary disease and without clinically detectable lung cancer, up to 4 years before radiological diagnosis [Ilie M et al.,
Nonetheless, in a following study by the same investigators [Marquette C.H. et al., 2020], the detection of CTCs in high-risk patients was unable to predict lung cancer development. The applicability of CTCs for the early detection of cancer is therefore still an open question.

![Diagram of CTCs and their implications](image)

**Figure 1.6. CTCs as a real-time liquid biopsy.**

CTCs can be derived from primary tumor or organs of metastasis. CTCs serve as a liquid biopsy of cancer and reveal important information on therapeutic targets and/or resistance mechanisms, which might be used in the future to stratify patients for such targeted therapies as inhibition of EGFR/HER2 or endocrine therapy.
and to monitor the efficacy of treatment and the development of resistance in real-time. ER+, estrogen receptor positive. [Alix-Panabieres C. and Pantel K., 2013]

However, the study of CTCs is extremely technically challenging because they are rare events diluted in billions of blood cells. In order to detect and analyze CTCs it is therefore necessary to use technologies to enrich and identify them.

1.2.1. First generation of CTC studies: CTC enumeration

The first generation of CTC studies used methodologies for CTC detection based on the expression of epithelial markers (such as epithelial cell adhesion molecule, EpCAM and cytokeratins, CK). These markers are in fact expressed by epithelial tumors but not by white blood cells (WBCs) and allow the discrimination between CTCs, i.e. cells positive for EpCAM/CK and negative for the pan-leukocyte marker CD45, and WBCs, i.e. EpCAM/CK-negative and CD45-positive cells [Alix-Panabieres C. and Pantel K., 2013].

The gold-standard instrument for CTC detection based on EpCAM and CK expression is the CellSearch® system [Allard W.J. et al., 2004], which, starting from 7.5 mL of blood, performs an enrichment step of CTCs using ferrofluids conjugated with antibodies against EpCAM, followed by an identification step in which the expression of CK and CD45 in the enriched cells is evaluated by staining them with fluorescently-labeled antibodies, to allow for CTC identification and counting. By using the CellSearch® system, in 2004 Cristofanilli and colleagues [Cristofanilli M. et al., 2004] demonstrated, for the first time, the prognostic significance of CTCs in metastatic breast cancer patients: Patients with ≥ 5 CTCs per 7.5 mL of blood had a shorter median PFS (2.7 months vs. 7.0 months, \(p < 0.001\)) and OS (10.1 months vs. >18 months, \(p < 0.001\)), than patients presenting < 5 CTCs. Successively, the prognostic relevance of CTC enumeration by CellSearch® was further confirmed by a pooled analysis of 20 studies including data from around 2000 patients with metastatic breast cancer [Bidard F.-C. et al., 2014] and was also demonstrated in patients with other types of tumors, including non-metastatic breast cancer [Rack B. et
al., 2014; Bidard F.-C. et al., 2018], colorectal cancer [Huang X. et al., 2015; van Dalum G. et al., 2015; Bork U. et al., 2015], metastatic prostate cancer [de Bono J.S. et al., 2008], small-cell [Hou J.M., et al., 2012; Tay R.Y., et al., 2019] and non-small-cell [Krebs M.G., et al., 2011; Lindsay C.R., et al., 2019] lung cancer. Moreover, changes in CTC counts during therapy have been shown to have a prognostic significance in patients with metastatic breast [Hayes D.F. et al., 2006], colorectal [Cohen S.J., et al., 2008], prostate [Lorente D. et al., 2016] and lung [Hou J.M., et al., 2012] cancer, thus supporting the use of CTC enumeration for treatment monitoring as well (Figure 1.7) [Cabel L. et al., 2017].

Figure 1.7. Clinical validity of CTC enumeration by CellSearch®.

CTC counts have been shown to be associated with prognosis in patients with breast, colorectal and metastatic prostate cancer. The detection of variations in CTC numbers during therapy can be used for treatment monitoring. [Adapted from Cabel L. et al., 2017]

This evidence led to the approval by the Food and Drug Administration of the CTC enumeration by CellSearch® to predict prognosis and monitor treatment in patients with metastatic breast, colorectal and prostate cancer. In 2010, CTCs were included in the tumor, node, metastasis (TNM) cancer staging as cM₀ (i+), indicating no clinically detectable metastasis but presence of CTCs or disseminated tumor cells in bone marrow or
lymphnodes [Alix-Panabieres C. and Pantel K., 2014], and in 2019 an expert consensus paper [Cristofanilli M. et al., 2019], performing a pooled analysis of almost 2500 patient data, recommended CTC enumeration as a new tool to improve prognostic stratification of metastatic breast cancer in Stage IV_{indolent} and Stage IV_{aggressive} (for patients presenting < 5 CTCs and ≥ 5 CTCs, respectively).

The clinical utility of CTC enumeration was first assessed by the SWOG S0500 study [Smerage J.B. et al., 2014], a trial designed to test whether, in metastatic breast cancer patients undergoing first-line chemotherapy, persistently high levels of CTCs after the first cycle of therapy could be used to identify the patients who would benefit from switching to a new chemotherapy regimen. The study reported no survival improvement for patients who switched therapy, suggesting that CTCs were not helpful in driving therapeutic intervention. However, the negative results of the trial were probably due to a design flaw [Rossi E. and Fabbri F., 2019]. In fact, in metastatic breast cancer, patients not responding to first-line therapy are often markedly resistant to subsequent lines of therapy [Bidard F.-C. and Pierga J.-Y., 2015] and changes in first-line chemotherapy introduced on the base of imaging evaluation did not improve patients’ OS as well [Alunni-Fabbroni M. et al., 2014]. Therefore, the clinical utility of CTCs is still an open question and there are several clinical trials currently evaluating if treatment decisions can be improved by considering either CTC count (CirCe01 and STIC-CTC studies) or CTC phenotype (DETECT III and DETECT IV trials) [Schochter F. et al., 2019]. While most studies are still ongoing, the results of the STIC-CTC trial were presented at the San Antonio Breast Cancer Symposium in 2018 and showed that, in patients with human epidermal growth factor receptor 2 (HER2)-negative hormone receptor-positive metastatic breast cancer, CTC-driven administration of chemotherapy as first-line treatment resulted in a significantly longer PFS, compared to clinically chosen endocrine therapy (HR = 0.62; 95% CI: 0.45–0.84, p = 0.002) [Bidard F.-C. et al., 2019]. Although further studies are
needed, these promising results support the clinical utility of CTCs, at least in selected cohorts of cancer patients.

1.2.1.1. CTC enumeration in patients with BTC

The clinical relevance of CTCs in patients with BTC was investigated in a few studies by using the CellSearch® [Al Utswani O. et al., 2012; Valle J.W. et al., 2015; Yang J.D. et al., 2016; Backen A.C. et al., 2018]. As observed for other malignancies, the presence of CTCs was associated with poor prognosis also in BTC [Valle J.W. et al., 2015; Yang J.D. et al., 2016]. In the ABC-03 study [Valle J.W. et al., 2015], this result was confirmed by using both 1 cell or 2 cells per sample as positivity threshold, showing a significantly higher risk of death for patients presenting 1 or ≥ 2 CTCs, than patients with 0 CTCs (1 CTC vs. 0 CTC, HR = 3.25; 95% CI: 1.81–5.83; ≥ 2 CTCs vs. 0 CTC, HR = 3.00; 95% CI: 1.73–5.22, p < 0.0001), indicating that the presence of even 1 CTC is clinically relevant in this clinical setting (Figure 1.8).

![Figure 1.8. Prognostic impact of CTCs in patients with BTC.](image)

Kaplan-Meier curves for overall survival, by number of CTCs per sample (collected before treatment). Median overall survival were 18.1, 10.3 and 8.7 months for patients presenting 0, 1 and ≥ 2 CTCs, respectively. [Valle J.W. et al., 2015]
However, in the ABC-03 trial, CTCs, although associated with prognosis, were not treatment-predictive [Backen A.C. et al., 2018]. Moreover, the reported CTC detection rates in all the studies were quite low, ranging from 17% - 25% [Al Utswani O. et al., 2012; Yang J.D. et al., 2016] to 47% [Valle J.W. et al., 2015] by using 2 CTCs and 1 CTC as positivity threshold, respectively. These results suggest that in the blood of BTC patients, as already reported for other epithelial cancers, there could be a subpopulation of CTCs which is not detected by CellSearch®, due to a non-epithelial phenotype [Backen A.C. et al., 2018]. Further studies using innovative technologies also capable of detecting these CTCs are therefore needed in order to understand the role of CTCs in BTC.

1.2.2. Technologies to capture CTC heterogeneity

Even though the first strategy used to detect CTCs (based on the expression of epithelial markers) led to the assessment of the clinical validity of CTCs in some cancers (including breast, prostate, colorectal and lung cancer), it was not equally effective in other tumors of epithelial origin (such as ovarian, kidney and biliary tract cancer) or in demonstrating CTCs’ clinical utility [Raimondi C. et al., 2014]. This could be due to the presence of CTCs with a non-epithelial phenotype, as suggested by the fact that CellSearch® detects CTCs only in a proportion of advanced cancer patients, and by the reported presence of CK-negative/CD45-negative cells in patients’ blood [Mego M. et al., 2011; Gazzaniga P. et al., 2011; Wang L. et al., 2016]. CTCs can, in fact, lose the expression of epithelial markers through the epithelial to mesenchymal transition (EMT), a process involved in carcinoma invasion and metastasis [Paterlini-Brechot P. and Benali N.L., 2006; Lee J.M. et al., 2006]. EMT is not an “all or nothing” process, but rather induces a variety of phenotypic changes with different degrees of expression of epithelial and mesenchymal markers [Christiansen J.J. and Rajasekaran A.K., 2006; Nieto M.A. et al., 2016]. CTCs
with a mesenchymal [Zhang L. et al., 2013; Satelli A. et al., 2015; Xu L. et al., 2017] or a mixed epithelial-mesenchymal [Lecharpentier A. et al., 2011; Giordano A. et al., 2012; Yokobori T. et al., 2013; Morrow C.J. et al., 2016; Bulfoni M. et al., 2016] phenotype have in fact been reported in cancer patients, indicating a much higher phenotypic heterogeneity in CTCs than originally thought. Moreover, it has been demonstrated that CTCs’ phenotype can change over the course of the disease, and these changes can be related to treatment resistance [Yu M. et al., 2013; Tsao S.C. et al., 2018]. In particular, Yu and colleagues, by evaluating the expression of epithelial and mesenchymal markers in CTCs from 10 patients with metastatic breast cancer undergoing treatment, reported an increase in CTCs with mesenchymal features after treatment in patients undergoing progression, whereas, in patients responding to treatment, a decrease in the CTC number and/or in the proportion of mesenchymal CTCs was observed, suggesting an association between mesenchymal CTCs with disease progression (Figure 1.9).

Figure 1.9. Phenotypic changes in CTCs in response to treatment in breast cancer patients.

For 10 patients with breast cancer (responding to therapy, upper part, and with progressive disease, lower part), CTCs were analyzed at 2 time points (before and after the initiation of treatment). The detected CTCs were divided, based on the expression of epithelial and mesenchymal markers into 5 classes, going from completely mesenchymal CTCs (M), to completely epithelial CTCs (E), including 3 intermediate phenotypes.
Introduction

(M>E; M=E; E>M). For each patient, the fractionation of CTCs, according to E/M ratios, is reported for both time points. The breast cancer subtype, the treatment regimen (chemotherapy or targeted therapies) and the number of days under treatment are also reported. [Adapted from Yu M. et al., 2013]

Based on the ascertained phenotypic heterogeneity of CTCs, over the last several years a wide variety of new technologies have been developed to try to detect all CTC subpopulations [Millner L.M. et al., 2013; Ferreira M.M. et al., 2016; Batth I.S. et al., 2019]. All the different methodologies exploited 3 main strategies: i) to expand the range of markers used for positive selection, by additionally including mesenchymal markers as the cell surface vimentin, CSV, [Satelli et al., 2015], stem cell markers as CD133 [Guo W. et al., 2018], and cancer specific antigens as HER2 and the prostate-specific membrane antigen, PSMA [Galletti G. et al., 2014; Kirby B.J. et al., 2012]; ii) to perform a negative depletion of blood cells using antibodies against CD45 [Bluemke K. et al., 2009; Giordano A. et al., 2012]; iii) to use physical properties for distinguishing between CTCs and WBCs such as size, density and deformability, through filtration [Desitter I. et al., 2011; Ma Y.-C. et al., 2013], density gradient centrifugation [Muller V. et al., 2005; Hodgkinson C.L. et al., 2014], or the use of microfluidic devices [Che J. et al., 2016; Miller M.C. et al., 2018].

New CTC detection technologies also moved forward from the simple enumeration and focused on the characterization of detected cells as well, to allow both a deeper understanding of the role of the different CTC subpopulations and to increase the sensitivity of the detection methods (Figure 1.10). The abovementioned detection strategies have, in fact, an increased sensitivity but at the same time can detect a high number of false positive events (such as normal cells expressing positive selection markers; circulating CD45-negative cells that are not CTCs, as endothelial cells; and WBCs with physical characteristic similar to CTCs). Due to the specificity issue of these methods, characterization becomes of major importance to identify true CTCs [Alix-Panabieres C. and Pantel K., 2014].
The new methods for CTC enrichment are based on physical properties, positive selection or negative selection. Generally, the methods include also a characterization step analyzing CTCs’ DNA, RNA or proteins, as well as their functional properties. [Cabel L. et al., 2017]

With regards to characterization approaches, recent technological advances (such as instruments for single-cell isolation and NGS) have allowed the development of methods for the genomic analysis of CTCs at the single-cell level (Figure 1.11) [Lim S.B. et al., 2019]. These methods include, after the enrichment, a step for the isolation of single cells. Cells can be individually isolated by laser capture microdissection [Park E.S. et al., 2018] and fluorescence-activated cell sorting (FACS) [Lambros M.B. et al., 2018], or by using specific instruments such as the DEPArray™ which isolates single cells by exploiting dielectrophoresis [Peeters D.J. et al., 2013], microscopic manipulators as the CellCelector™ [Lampignano R. et al., 2017; Reinhardt F. et al., 2019] and microfluidic devices [Yeo T. et al., 2016; Valihrach L. et al., 2018].
Figure 1.11. Standard workflow for single-CTC sequencing.

After enrichment from whole blood, CTCs are individually isolated and undergo WGA. The quality of WGA products is checked, and high quality amplified DNA can be sequenced for the detection of small-scale and large-scale alterations. Quality checks are performed on the sequencing library to assess the possibility of amplification bias. LCM, laser capture microdissection; FACS, fluorescence-activated cell sorting; DOP-PCR, degenerate oligonucleotide primed PCR; LA-PCR, linker-adapter PCR; MDA, multiple displacement amplification; MALBAC, multiple annealing and looping-based amplification cycles; SNVs, single nucleotide variants; MSI, microsatellite instability; CNVs, copy number variations, LSTs, large-scale state transitions; ADO, allelic drop out; QC, quality control. [Lim S.B. et al., 2019]

Independently from the type of isolation method used, all isolated cells will undergo the whole-genome amplification (WGA) in order to be analyzed. WGA is based on 2 approaches: PCR and multiple displacement amplification (MDA), and the different WGA methods employ either only one of the 2 (PCR-based and MDA-based) or a combination of both mechanisms (hybrid) [Blainey P.C., 2013]. PCR-based methods always require a fragmentation of the genome before the amplification, due to the limited processivity of the polymerase (generally a Taq DNA polymerase), and use, for the amplification step, either random primers that anneal across the genome in a non-deterministic manner (degenerate oligonucleotide primed PCR, DOP-PCR), or specific sequence primers that are ligated at the end of the template DNA fragments (linker adapter, or ligation-mediated PCR, LA-PCR or LM-PCR). In the Ampli1™ method (an LM-PCR WGA approach), the use of nonrandom primers is coupled with a deterministic pattern of fragmentation of the genome,
thanks to the use of a restriction endonuclease, allowing a defined and reproducible fragmentation and priming pattern and a high genomic coverage [Czyz Z.T. and Klein C.A., 2015]. MDA-based methods instead use highly processive DNA polymerases (such as the Phi29) which perform an isothermal rolling circle amplification of the template DNA, yielding a high quantity of DNA (20-30 µg per single cell) and producing large amplicons of > 10 kb [Dean F.B. et al., 2002]. Hybrid approaches, as the PicoPLEX® [Kamberov E. et al., 2004] and the multiple annealing and looping-based amplification (MALBAC) [Zong C. et al., 2012], combine an initial MDA-based pre-amplification step, followed by the amplification of the produced amplicons by PCR. All the WGA approaches, unfortunately, have biases which can result in a non-homogeneous amplification of all amplicons (due to random events occurring during the initial steps of the amplification, or to the intrinsic characteristics of the amplicons, such as the GC content) and in the production of artifacts related to the polymerase activity [Sabina J. and Leamon J.H., 2015]. These factors can result in allele drop out (ADO), loss of coverage, low coverage uniformity, or allelic imbalances. Each method is characterized by advantages and limitations, as reported by studies which have compared different commercially available WGA kits based on LM-PCR (Ampli1™), MDA (REPLI-g®), or a combination of both (PicoPLEX® and MALBAC), sometimes obtaining contrasting results [Czyz Z.T. et al., 2015; Babayan A. et al., 2016; Borgstrom E. et al., 2017; Deleye L. et al., 2017]. Overall, REPLI-g® (MDA-based WGA) showed the least sensitivity and coverage uniformity and the highest ADO rate [Czyz Z.T. et al., 2015; Babayan A. et al., 2016; Borgstrom E. et al., 2017], although in the study by Deleye and colleagues it was reported as an efficient method, comparable to the other approaches [Deleye L. et al., 2017]. LM-PCR-based (Ampli1™) and hybrid (PicoPLEX® and MALBAC) methods have shown similar performance in representing the template DNA, with Ampli1™ being reported as the one with the lowest ADO rate [Borgstrom E. et al., 2017] and the most
suited for the detection of single nucleotide polymorphisms (SNPs) and small insertions-deletions (INDELs) [Babayan A. et al., 2016], whereas hybrid approaches seemed to be the best choice for copy number alteration (CNA) analysis [Babayan A. et al., 2016]. However, the applicability of Ampli1™ for CNA analysis was demonstrated by studies showing the possibility of detecting known CNAs in single cells undergoing Ampli1™ WGA [Klein C.A. et al., 1999; Schmidt-Kittler O. et al., 2003; Mohlendick B. et al., 2013]. WGA is not always successful due to the possibility of losing cells during the isolation procedure and to the quality of the isolated cells' DNA. Therefore, after WGA, quality control (QC) assays are performed to assess the DNA yield and the length of amplified fragments. In particular, Polzer and colleagues [Polzer B. et al., 2014] developed a QC assay based on a multiplex-PCR approach which assigns a score called genome integrity index (GII) to WGA products. The GII is predictive of successful application of different downstream sequencing analyses: Good quality samples can undergo any type of sequencing analysis (including Sanger sequencing, array comparative genomic hybridization (aCGH) platforms or genome-wide NGS) for the detection of a variety of genomic alterations (Figure 1.12) including small-scale alterations (single nucleotide variants (SNVs), INDELs and microsatellite instability), and large-scale alterations (copy number variations (CNVs), chromosomal breakpoints or large-scale state transitions (LSTs), and chromosomal rearrangements).
Figure 1.12. Genomic alterations in single CTCs.

List and characteristics of the type of genomic alterations that can be detected in CTCs by single-cell sequencing. [Adapted from Lim S.B. et al., 2019]

The methods for CTC characterization at the single-cell level opened a new chapter of liquid biopsy research, aimed at characterizing and monitoring changes in tumor heterogeneity in individual patients to further understand the biology of tumor evolution.
1.2.3. **New generation of CTC studies: CTC characterization at the single-cell level**

The studies focused on the analysis of CTCs at the single-cell level have tried to answer 3 main questions:

- Do CTCs recapitulate tumor spatial heterogeneity?
- Can we use CTCs to study the evolution of tumor heterogeneity during the course of the disease?
- Can CTC molecular characterization provide information on treatment resistance?

In 2014 Lohr and colleagues conducted the first proof-of-concept study showing the possibility to recapitulate tumor heterogeneity by analyzing CTCs [Lohr J.G. et al., 2014]. The authors performed whole-exome sequencing of 19 single CTCs, 1 metastatic biopsy and 9 spatially different primary tumor biopsies from a single patient with metastatic prostate cancer. Fifty-six metastatic trunk mutations (mutations present in both the metastasis and the primary tumor) were identified and 73% of these were detected in CTCs as well. CTCs also carried 51% of all the mutations detected in the metastatic site. Moreover, by hierarchical clustering it was possible to assess the similarities between the samples collected from different anatomical locations (primary tumor, lymph node and blood) over the entire course of the disease (from the time of the radical prostatectomy to the blood collection, 5.3 years later), allowing to reconstruct an evolutionary tree of the tumor (Figure 1.13). The 9 primary tumor biopsies showed high intra-tumor heterogeneity, but there was one region that closely resembled CTCs and the metastatic biopsy, suggesting that this particular region of the primary tumor could be the one responsible for the metastatic spread.
Somatic mutations in the listed genes were detected in all individual biopsies of tissue (early trunk) or in all biopsies belonging to only one of the two branchpoints (the nonsynonimius mutations are indicated in bold, with asterisks). The pathology blocks from which tissue biopsies were obtained are represented in pink (drawn to scale) and, within each block, the area with histological presence of tumor is shown by the dotted line. The colors highlight the site where each individual biopsy of the dendrogram was taken. Below, a schematic representation of the regions of the prostate from which the pathology blocks were obtained, is reported. [Adapted from Lohr J.G. et al., 2014]

These results suggest that CTCs can be used to study tumor heterogeneity and evolution and that their analysis can be more useful than tissue profiling in detecting the tumor subclones actively involved in dissemination, since, at the primary tumor site, such clones could be diluted or be present only in a particular region.
Later studies comparing mutations and CNAs detected in single CTCs and tissue biopsies from the same patients with prostate [Lambros M.B. et al. 2018], breast [De Luca F. et al., 2016; Paoletti C. et al., 2018], colorectal [Gao Y. et al., 2017] and lung [Chemi F. et al., 2019] cancer, further confirmed the findings of Lohr and colleagues. In particular, Chemi and colleagues [Chemi F. et al., 2019] showed that, in a patient with non-small-cell lung cancer, CTCs collected at the time of surgery were molecularly more similar to the metastasis developed 10 months later (91% of overlapping mutations), than to the primary tumor (79% of overlapping mutations), whereas only 46% of primary tumor mutations were detected in CTCs, suggesting that CTCs detected at surgery derived from a subclone of the primary tumor that was the precursor of the metastasis.

Nonetheless, the aforementioned studies analyzing the molecular characteristics of single CTCs also demonstrated a high level of heterogeneity between CTCs collected from the same blood draw. In the study by Lambros and colleagues [Lambros M.B. et al. 2018], 185 CTCs collected from 14 patients with prostate cancer were analyzed for CNA profiling and showed distinct degrees of intra-patient heterogeneity that varied from patient to patient (Figure 1.14). Notably, while some patients presented very homogeneous CTCs (Figure 1.14, left side), other patients had CTCs with very different CNA profiles (Figure 1.14, right side). There were also patients presenting different CTC subtypes, such as patient 13, for whom one subset of CTCs had homogeneous CNA profiles resembling the alterations found in a metastatic tissue biopsy, a second subset presented CNAs similar to the prostatectomy biopsy, and a third subset showed very complex and non-clonal CNA patterns.
Figure 1.14. Intra- and inter-patient genomic heterogeneity of single CTCs.

Unsupervised hierarchical clustering heat map of single CTCs collected from each patient, based on CNA profiles. Each column is a CNA profile of a single CTC, showing gains/amplifications (pink/red) and losses/deletions (light-blue/dark-blue) for the entire genome (chromosomes are indicated on the left). CTCs collected from the same patient are indicated by colors in the top row. The heat maps of each patient are organized by intra-patient diversity from left to right. [Adapted from Lambros M.B. et al., 2018]

These results further highlight the importance of performing CTC analysis at the single-cell level, since this is the only way to truly unravel intra-patient heterogeneity, which could not be detected by bulk analyses.
Intra-patient heterogeneity of CTCs can also change over time. Su and colleagues [Su Z. et al., 2019] monitored this by comparing CNA profiles of single CTCs collected from a patient with small-cell lung cancer at 4 different time points (before and during first-line chemotherapy, before second-line chemotherapy and during third-line chemotherapy). CTCs, which were highly homogeneous at the first time point, became progressively heterogeneous and this was possibly due to allelic losses of the initially uniform CNAs, leading to the emergence of new subclones. Overall, these studies support the use of CTCs to assess and study intra-tumor heterogeneity. However, whereas the marked diversity between CTCs on one hand carries a large amount of information about tumor evolution, on the other hand it makes the interpretation of this information much more difficult, in particular when the number of available CTCs is limited.

Other studies focused more on investigating whether the genomic profiling of single CTCs could be used to better understand therapeutic response and resistance. Dago and colleagues [Dago A.E. et al., 2014] assessed both phenotypic and genomic changes of single CTCs collected from one patient with castration-resistant prostate cancer, undergoing chemotherapy followed by androgen deprivation therapy with abiraterone. CTCs were analyzed in 4 blood samples collected before starting chemotherapy (draw 1), after progression to chemotherapy, before starting abiraterone (draw 2), after 3 weeks of treatment with abiraterone, when the patient was responding (draw 3) and after 9 weeks of treatment, during disease progression (draw 4). CTC enumeration and the evaluation of androgen receptor (AR) expression were performed using the CellSearch®. Forty-one single CTCs were isolated and analyzed for CNAs, and compared with the CNA profile of 1 bone metastasis biopsy taken at the time of diagnosis, 5 months before draw 1. By unsupervised clustering of the 41 CTCs, 3 clones were identified (Figure 1.15 A).
Figure 1.15. Clonality and genomic alterations in single CTCs.

A) Unsupervised hierarchical clustering of CNA profiles of 1 metastatic biopsy and 41 single CTCs collected at 4 time points (draw 1-4) from a single patient with castration-resistant prostate cancer. Each column is the CNA profile of a single CTC. Gains and losses are indicated in red and blue, respectively. The colors in the top row indicate the blood sample from which the CTC was isolated (draw 1, yellow; draw 2, orange; draw 3, purple; draw 4, black). The green sample represents the metastatic biopsy. B) Plot of the AR amplification events for the 3 clones. Each line is a single CTC. [Adapted from Dago A.E. et al., 2014]

Clone A included most of the CTCs collected from draw 1 and 2, before abiraterone treatment. These CTCs were genetically similar to the metastatic biopsy and were AR-positive. Most CTCs also presented AR amplification (which was not detected in the metastasis). After starting treatment with abiraterone, when the patient was still responding (draw 3), CTCs expressed little or no AR and showed near-normal CNA profiles (clone B), suggesting that abiraterone targeted the androgen-dependent subpopulation of CTCs. However, when the patient progressed to abiraterone, a new subclone of CTCs emerged.
(clone C), showing again the expression of AR but distinct CNA profiles with regards to clone A. Interestingly, these CTCs presented AR amplification as CTCs in clone A. But whereas in clone A the AR amplification events were very heterogeneous among the different CTCs, in clone C all CTCs presented the same amplification profile with a higher amplification level (Figure 1.15 B), suggesting that this new abiraterone-resistant clone could represent a novel lineage, induced by selection pressure, deriving from a single resistant cell. The authors also identified, in clone C, the amplification of MYC as a possible bypass mechanism for AR-independent resistance, supporting the hypothesis that CTC genomic characterization can give hints on treatment resistance mechanisms.

The use of CTCs for detecting alterations involved in therapeutic resistance was exploited by other investigators. For instance, Paolillo and colleagues [Paolillo C. et al., 2017] investigated the presence of ESR1 mutations in single CTCs collected from patients with hormone receptor-positive metastatic breast cancer receiving endocrine therapy. In 2 out of 3 patients, the activating mutation Y537S was detected in CTCs and mirrored treatment failure. Similarly, Pailler and colleagues [Pailler E. et al., 2019] analyzed single CTCs collected from patients with anaplastic lymphoma kinase (ALK)-rearranged non-small-cell lung cancer, progressing to treatment with ALK inhibitors. By performing targeted sequencing of 48 cancer-related genes and 14 ALK mutations, the authors could detect multiple mutations in different genes involved in both ALK-dependent and ALK-independent pathways, suggesting the presence, also within the same patient, of heterogeneous subclones of tumor cells undertaking different therapeutic resistance mechanisms, that were not detectable in the corresponding tissue biopsy. Overall these results show that CTC molecular characterization can be more informative than tissue analysis for investigating therapeutic resistance mechanisms. However, they are still anecdotal and have not yet demonstrated that CTC molecular profiling can improve clinical practice.
A stronger support for the clinical relevance of single-CTC molecular characterization was provided by the study by Carter and colleagues [Carter L. et al., 2017]. By analyzing the CNA profiles of 88 single CTCs collected from 13 chemotherapy-naive patients with small-cell lung cancer, the authors developed a CNA-based classifier able to distinguish between patients with chemorefractory (i.e. undergoing disease progression within 3 months after the end of first-line chemotherapy) and chemosensitive disease. The classifier’s ability of identifying chemorefractory and chemosensitive patients was validated in an independent cohort of 18 patients, in which 83.3% of patients were correctly classified (Figure 1.16 A)

![Figure 1.16. CTC CNA-based classifier and clinical outcome.](image)

A) Clinical status and CTC CNA-based classification of 31 patients with small-cell lung cancer, divided in a training set (13 patients) and a testing set (18 patients). The accuracy of the CNA-based classifier is reported for the training set, the testing set and the entire cohort. B) Kaplan-Meier curves for progression-free survival and overall survival of the 31 patients according to the clinical classification (left, top and bottom) and to the CNA-based classification of patients (right, top and bottom) [Adapted from Carter L. et al., 2017]

Moreover, the patients classified as chemorefractory by the CTC CNA-based classifier showed a significantly shorter PFS than patients classified as chemosensitive (median
PFS = 2.8 vs. 5.8 months, $p = 0.0166$), supporting the clinical validity of the classifier (Figure 1.16 B). Overall, these results demonstrate that molecular characterization of CTCs before treatment initiation can give information on tumor chemosensitivity and can potentially help predicting which patients will respond to chemotherapy.

The field of single-CTC molecular profiling is still under development and many more studies are required to prove its clinical validity. However, the results obtained so far definitely support the potential of single-CTC analysis to provide a real-time assessment of the disease, of its heterogeneity and its evolution, offering the possibility to better understand therapeutic resistance.
2. SCOPE OF THE THESIS
Most BTC patients are diagnosed at an advanced stage, when surgery, the only potentially curative therapy, is not feasible. Chemotherapy with cisplatin and gemcitabine is the standard first-line treatment for patients with advanced biliary tract cancer, but unfortunately most patients develop resistance to this treatment and the median overall survival is less than one year. For patients with progressive disease who need to undergo a second-line therapy there is, currently, no standard treatment.

There is, therefore, a great need for new therapies in BTC. Molecular characterization data suggest that many patients carry actionable mutations and would benefit from targeted therapies. However, the use of targeted therapies is limited due to the difficulty to obtain tissue biopsies for molecular profiling.

CTCs could represent an alternative source of tumor material to perform molecular profiling, fostering personalized therapy in BTC. Moreover, CTC characterization at the single-cell level could be used to monitor the evolution and the response to treatment of the disease over time.

However, the application of CTC analysis in BTC is hindered by the fact that, by using conventional CTC detection approaches based on epithelial marker expression, CTCs are detected only in a small proportion of patients, probably due to the presence of CTCs that do not express epithelial markers. New approaches that also allow for the detection of these non-epithelial CTCs could therefore help in implementing the use of CTC analysis in BTC patients.

Based on these considerations, the ultimate goal of this study was to assess if a multilevel characterization of single CTCs could provide information on tumor molecular features and on response/resistance to treatment in patients with BTC.
Therefore, the specific aims of the thesis work were:

1) to develop a method for the detection not only of epithelial CTCs but also of those lacking epithelial markers, allowing their characterization at the single-cell level;

2) to assess if CTCs with different phenotypes have a different clinical relevance;

3) to test whether single-CTC molecular characterization could be used to perform a molecular profiling of the tumor and to study the evolution of the disease in response to treatment.
3. MATERIALS AND METHODS
3.1. **Patient information and clinical sample collection**

This was a prospective, monocentric, observational study conducted at Fondazione IRCCS Istituto Nazionale dei Tumori (Milan, Italy). For this study, 24 patients with a confirmed diagnosis of metastatic/unresectable BTC were consecutively recruited between January 2015 and March 2017. Due to the explorative nature of the study, no statistical hypothesis was postulated. The number of enrolled patients was consistent, however, with the entropy-based approach to sample size in translational clinical trials as proposed by Piantadosi and colleagues [Piantadosi S., 2005].

Patients have been treated and followed up as per clinical practice, with frequent clinical evaluations and tumor assessment with chest/abdomen CT scans and/or MRIs performed every 2-3 months. The treatment efficacy was assessed as per RECIST v1.1. Clinical information was collected from medical records and included demographic data, tumor anatomical location, tumor extension, and treatment history. The patients’ vital status was updated at the end of June 2018.

All CTC evaluations were carried out without the knowledge of the patient’s clinical status.

Samples of peripheral venous whole blood (10 mL) were drawn in EDTA tubes (K$_2$EDTA BD Vacutainer®, Becton Dickinson, Franklin Lakes, NJ, USA), stored at 4 °C protected from light and processed within 1 hour for CTC enrichment (the first mL of blood was discarded to avoid skin epithelial cell contamination). Blood samples were longitudinally collected at times corresponding to baseline (BL), *i.e.* before initiation of a new treatment line, during treatment (DT) close to clinical and imaging evaluations, at the end of treatment (EOT) and at subsequent follow-up (FU) or new treatment lines.

All subjects have signed a written informed consent form accepting participation in this study, which was approved by the local ethical board in November 2014 (INT 177/14) and subsequently reconfirmed in January 2018.
3.2. **Cell lines and culture conditions**

The human breast cancer cell lines MCF7, MDA-MB-231 and T-47D were acquired from American Type Culture Collection and cultured in DMEM-F12 medium (Lonza, Basel, Switzerland) supplemented with 10% fetal bovine serum (FBS) (Gibco, Thermo Fisher Scientific, Waltham, MA, USA). The human dermal fibroblast cell line NHDF was acquired from Lonza and cultured in FBM™ Basal Medium (Lonza) supplemented with the FGM™-2 SingleQuots™ kit (Lonza) according to the manufacturer’s instructions. The human lymphoma cell lines DOHH-2, RAJI, SU-DHL-4 and NAMALWA were kindly provided by Dr. Massimo Di Nicola (Fondazione IRCCS Istituto Nazionale dei Tumori) and cultured in RPMI 1640 medium (Lonza) supplemented with 1% HEPES, 1% L-glutamine and 10% FBS (Gibco). The human prostate cancer cell line PC-3 was acquired from American Type Culture Collection and cultured in RPMI 1640 medium (Lonza) supplemented with 10% FBS (Gibco). The human intrahepatic CCA cell lines HuCCT-1 and HuH28, and the extrahepatic CCA cell line EGI-1 were kindly provided by Prof. Mario Strazzabosco (Università Milano-Bicocca, Monza, Italy) and were cultured in RPMI 1640 medium (Lonza) supplemented with 10% FBS (Gibco).

Cells were grown at 37 °C in a humidified 5% CO₂ atmosphere as a monolayer (MCF7, MDA-MB-231, T-47D, NHDF, HuCCT-1, HuH28, EGI-1 and PC-3) or in suspension (DOHH-2, RAJI, SU-DHL-4 and NAMALWA). For each cell line, authenticity was verified by short tandem repeat analysis performed by the Integrated Biology and Bioinformatics Platform Unit at our institute.
3.3. Flow cytometry analysis

Flow cytometry analysis was performed to test the specificity of 2 antibodies against EpCAM: anti-human EpCAM PE-conjugated, clone 0.N.276 (Santa Cruz Biotechnology, Dallas, TX, USA) and anti-human EpCAM FITC-conjugated, clone HEA-125 (Miltenyi Biotec, Bergisch Gladbach, Germany).

Two blood samples (10 mL each) were collected from a healthy donor (HD) in EDTA tubes (K₂EDTA BD Vacutainer®, Becton Dickinson) and in Streck tubes (Cell-Free DNA BCT®, Streck, La Vista, NE, USA), discarding the first mL of blood to avoid skin epithelial cell contamination. WBCs were collected from whole blood by density gradient centrifugation using HISTOPAQUE® 1083 (Sigma-Aldrich, St. Louis, MO, USA) following the manufacturer’s instructions and counted using a Bürker chamber. Two aliquots of 2x 10⁶ WBCs, collected from the EDTA tube sample, were stained with the anti-EpCAM PE-conjugated antibody (Santa Cruz Biotechnology) and with the anti-EpCAM FITC-conjugated antibody (Miltenyi Biotec), respectively, according to the manufacturers’ instructions. Similarly, two aliquots of 2x 10⁶ WBCs, collected from the Streck tube sample, were stained with the 2 antibodies. Two aliquots of 2x 10⁶ WBCs, from the EDTA tube sample and from the Streck tube sample, respectively, were collected but not stained and were used as negative controls.

WBCs were analyzed using the BD FACSCalibur™ flow cytometer (Becton Dickinson). The stained samples were compared to the control samples to evaluate the presence of unspecific staining detected in the PE and the FITC channels.
3.4. **Spike-in procedure**

For spiking experiments, adherent cells were detached from the culture flask with trypsin (Lonza), highly diluted in the appropriate culture medium to obtain a single-cell suspension, and put into a culture dish. Cells were visualized under an inverted optical microscope, individually picked using a micropipette and directly spiked into HD blood samples. In case of blood collected in EDTA tubes, the spiked-in samples were immediately processed for subsequent analyses, whereas, in case of blood collected in Streck tubes, the spiked-in samples were kept at room temperature for 24 to 48 hours before undergoing subsequent analyses.

For the spiking experiments using cells labeled with the cell tracker, cells were incubated for 20-30 min at 37 °C with the CellTracker™ Green CMFDA Dye (Thermo Fisher Scientific, Waltham, MA, USA), following the manufacturer’s instructions, before being detached from the culture flask for the spike-in procedure. The CMFDA fluorescent probe was visualized in the FITC channel using a fluorescence microscope.

3.5. **Enrichment methods**

3.5.1. **ScreenCell® filters®**

Spiked-in HD blood samples (9 mL) collected in either Streck or EDTA tubes were processed with the ScreenCell® CC kit (ScreenCell, Sarcelles, France), according to the manufacturer’s instructions, for tumor cell enrichment. Each filtration unit can filter 3 mL of blood, so, for each spiked-in sample, 3 filtration units were used. After filtration of the blood through the filtration unit, enriched cells that were trapped on the isolation support (i.e. a membrane with pores of 6.5 μm diameter) were detached from it through repeated washings using a micropipette, following the protocol PR_A17 provided by the manufacturer. Collected cells could then undergo downstream fixation (necessary only in the case of live cells, for blood samples collected in EDTA tubes) and staining.
3.5.2. **AutoMACS® Pro separator**

Spiked-in HD blood samples collected in EDTA tubes (10 mL) were incubated for 10 min at room temperature with the Lysing Buffer (BD Pharm Lyse™, Becton Dickinson) 1:10 for the lysis of red blood cells (RBCs). Lysed blood samples were then incubated for 15 min at 4°C with CD45 MicroBeads (MicroBeads conjugated with monoclonal anti-human CD45 antibodies, Miltenyi Biotec), and processed with the AutoMACS® Pro separator (Miltenyi Biotec) using the DepleteS program, following the manufacturer’s instructions. Enriched cells collected from the instrument underwent downstream fixation and staining.

3.5.3. **OncoQuick®**

Whole blood collected from HD in EDTA tubes (15 mL) and spiked-in with tumor cells was transferred into an OncoQuick® 50 mL-polypropylene tube (Greiner Bio-One, Kremsmünster, Austria) and centrifuged at 1600 g for 20 min at 4 °C. The fraction that remained above the porous barrier of the OncoQuick® tube (containing the enriched cells) was collected and washed with phosphate-buffered saline (PBS) containing 0.5% bovine serum albumin, as indicated by the manufacturer, before undergoing fixation and staining.

3.5.4. **Parsortix®**

Spiked-in HD blood samples (10 or 20 mL) or clinical samples (10 mL) collected in EDTA tubes were transferred into a 50 mL polypropylene tube (Corning™ Falcon Conical Centrifuge Tubes, Thermo Fisher Scientific) and loaded into the Parsortix® instrument (Parsortix PR1 Cell Separation System, ANGLE plc, Guildford, UK). The program PX2_S99F and the 6.5 µm Cell Separation Cassette (ANGLE plc) were used to process the samples, as suggested by the manufacturer. At the end of the enrichment, the flow direction inside the cassette was inverted and the cells trapped in the cassette were washed...
out with PBS. For samples spiked-in with tumor cells labeled with the CellTracker™ Green CMFDA Dye (Thermo Fisher Scientific), enriched cells were harvested in 96-well or 24-well plates and incubated for 20 min with the NucBlue™ Live ReadyProbes™ Reagent (Hoechst 33342) (Thermo Fisher Scientific), according to the manufacturer’s instructions, for nuclear staining. The cells were then visualized and counted using a fluorescence microscope. For clinical samples, enriched cells were instead harvested in a 1.5 mL Protein LoBind Tube (Eppendorf, Hamburg, Germany) and underwent subsequent fixation and staining.
3.6. **Post-enrichment procedures**

3.6.1. **Fixation and staining**

Cells enriched from blood samples collected in EDTA tubes were fixed with 2% paraformaldehyde (PFA) for 20 min at room temperature (according to protocol W_BIO_008, provided by the DEPArray™ manufacturer) before undergoing the staining protocol. The fixation step was not performed if cells were enriched from blood samples collected in Streck tubes.

Subsequently, the cells were stained with fluorescently-labeled antibodies directed against surface markers. All antibodies (described in Table 3.1) were used according to the manufacturers’ instructions.

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Clone</th>
<th>Conjugate</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>EpCAM</td>
<td>0.N.276</td>
<td>PE</td>
<td>Santa Cruz Biotechnology*</td>
</tr>
<tr>
<td>EpCAM</td>
<td>HEA-125</td>
<td>PE</td>
<td>Miltenyi Biotec</td>
</tr>
<tr>
<td>EGFR</td>
<td>423103</td>
<td>PE</td>
<td>R&amp;D Systems (Minneapolis, MN, USA)</td>
</tr>
<tr>
<td>CD14</td>
<td>M5E2</td>
<td>APC</td>
<td>BD Biosciences Pharmigen (San Diego, CA, USA)</td>
</tr>
<tr>
<td>CD16</td>
<td>3G8</td>
<td>APC</td>
<td>BD Biosciences Pharmigen</td>
</tr>
<tr>
<td>CD45</td>
<td>5B1</td>
<td>APC</td>
<td>Miltenyi Biotec</td>
</tr>
<tr>
<td>CSV</td>
<td>84-1</td>
<td>AF488**</td>
<td>Abnova (Taipei, Taiwan)</td>
</tr>
</tbody>
</table>

Table 3.1. Antibodies used to stain surface markers.

For each antibody, the antigen, the clone, the conjugated fluorophore and the manufacturer are indicated. *This antibody was tested during technical validity assessments but was not included in the staining protocol for clinical samples. **The antibody anti-CVS was purchased as not-labeled and subsequently labeled with AF488 using the Alexa 488 Labeling Kit (Immunological Sciences, Rome, Italy).
After the staining of surface markers, the cells were permeabilized using Inside Perm (part of the Inside Stain Kit, Miltenyi Biotec) for 10 min at room temperature and stained for intracellular markers using the antibodies reported in Table 3.2, following the manufacturers’ instructions.

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Clone</th>
<th>Conjugate</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>CK</td>
<td>C11</td>
<td>PE</td>
<td>Abcam (San Francisco, CA, USA)</td>
</tr>
<tr>
<td>CK</td>
<td>AE1/AE3</td>
<td>PE</td>
<td>NSJ Bioreagents (San Diego, CA, USA)</td>
</tr>
<tr>
<td>VIM</td>
<td>V9</td>
<td>AF488</td>
<td>Santa Cruz Biotechnology</td>
</tr>
</tbody>
</table>

Table 3.2. Antibodies used to stain intracellular markers.
For each antibody, the antigen, the clone, the conjugated fluorophore and the manufacturer are indicated.

Finally, cells’ nuclei were stained with 1 μg/mL Hoechst 33342 (Sigma-Aldrich) for 5 min at room temperature.

During spiking experiments, a protocol including RBC lysis was tested. In that case, RBC lysis was performed after enrichment and before fixation, using the RBC Lysing Buffer (BD Pharm Lyse™, Becton Dickinson) for 15 min at room temperature.

3.6.2. DEPArray™ analysis
After fixation and staining, the samples were analyzed with the DEPArray™ system (Menarini Silicon Biosystems, Bologna, Italy), an automated platform which allows an image-based selection and a dielectrophoresis-based isolation of single cells [Di Trapani M. et al., 2018].

The DEPArray™ platform was used following the manufacturer’s instructions. Briefly, the fixed and labeled cells were washed twice with the SB115 buffer (Menarini Silicon Biosystems). After the washings, the cells were resuspended in 13 μL of SB115 buffer and
were loaded into the DEPArray™ A300K DS cartridge (Menarini Silicon Biosystems). The cartridge was then inserted into the DEPArray™ V2 System for sample analysis. The instrument performed the scan of the chip containing the sample and the acquisition of images for each cell, using an integrated fluorescence microscope. The settings used for the chip scan were first assessed and defined during the technical validation experiments, and were then identically applied to each clinical sample. The cells present in the sample were visualized using the CellBrowser™ software for the evaluation of the fluorescence labeling and of morphological characteristics. The cells of interest were selected by an operator and individually collected into 0.2 mL PCR tubes. The recovered cells underwent a volume reduction step (following the protocol W_BIO_002 provided by Menarini Silicon Biosystems) to manually remove the SB115 buffer, leaving the cells in around 1 µL of PBS. The cells were then stored at -20 °C until WGA.
3.7. **Molecular analysis**

3.7.1. **Whole-genome amplification and quality control assay**

Recovered single cells or pools of WBCs were subjected to WGA using the Ampli1™ WGA kit (Menarini Silicon Biosystems), a LM-PCR-based WGA method. The quality of the amplified DNA was assessed with the Ampli1™ QC kit (Menarini Silicon Biosystems), a PCR assay allowing the amplification of 4 human genomic targets of different lengths, starting from 1 μL of Ampli1™ WGA product. The successful amplification of the targets was assessed by the detection of the 4 amplicons by capillary gel electrophoresis using the Agilent 2100 Bioanalyzer instrument (Agilent Technologies, Santa Clara, CA, USA) and the Agilent DNA 1000 kit (Agilent Technologies). The detection of 0 to 4 amplicons was translated into a GII ranging between 0 and 4. The samples having a GII \( \geq 2 \) were subjected to the following genomic analyses.

3.7.2. **Mutational profiling**

For this analysis, Ampli1™ WGA products were shipped to Menarini Silicon Biosystems for DNA sequencing and bioinformatic analysis.

Amplified DNA was subjected to mutational profiling by targeted sequencing using the Ion AmpliSeq Custom Cancer Hotspot Panel v2 (Thermo Fisher Scientific), which was modified to consider the enzymatic fragmentation of the input DNA performed during Ampli1™ WGA. Sequenced data were analyzed with TorrentSuite v5.2. Variants were called using the “AmpliSeqSomatic” pipeline in the regions defined by “Ampli1™ Cancer Hotspot Panel custom vBeta”, setting parameters as recommended for AmpliSeq data. This configuration provides for a set of parameters suitable for calling variants at a lower level of stringency (allele frequencies: >2% for SNP, INDEL and HOTSPOTs). Variants were annotated with Variant Effect Predictor (VEP) v83, a tool developed by ENSEMBL group ([http://www.ensembl.org/info/docs/tools/vep/index.html](http://www.ensembl.org/info/docs/tools/vep/index.html)).
3.7.3. **Copy number alteration profiling**

CNA profiling was performed through low-pass whole-genome sequencing (lp-WGS) of Ampli1™ WGA products. The barcoded libraries for the sequencing were prepared using the Ampli1™ LowPass kit (Menarini Silicon Biosystems). The sequencing was performed by the *Integrated Biology and Bioinformatics Platform Unit* at our institute, using the Ion S5™ System and Ion 530 chips (Thermo Fisher Scientific), following the manufacturer’s instructions.

Bioinformatic analyses were performed as follows. WGS sequences were aligned to the human reference genome (hg19) using the TMAP aligner tool on Torrent_Suite 5.4.0. CNAs were predicted by using Control-FREEC 11.0 with the following settings: coefficientofVariation = 0.05, mateOrientation = 0, sex = XY or XX. Control-FREEC produced different window sizes according to the sequencing depth in each sample. “Gain” and “loss” calls were filtered out when the Wilcoxon Rank-Sum Test and Kolmogorov-Smirnov p-values were greater than 0.05.

Comparison of single cells’ CNA profiles was performed considering copy number log-ratio values evaluated for non-overlapping equal regions size of 500 kb hierarchically clustered using Euclidean distance and Ward linkage method; as implemented in the R package ([CRAN.R-project.org](https://CRAN.R-project.org)). The most significantly and differentially altered regions between groups of CTCs from responders and non-responders were identified by GISTIC2.0 [Mermel C.H. et al., 2011] and subsequently validated using Fisher’s exact test. Large scale state transition scores [Greene S.B. et al., 2016] were calculated using the Genomic.Instability R package ([https://github.com/SilvestriMR/Genomic.Instability](https://github.com/SilvestriMR/Genomic.Instability)).
3.8. RNA analysis

The RNA was extracted from spiked-in samples after Parsortix® enrichment, using the Agencourt® RNAAdvace® Cell v2 kit (Beckman Coulter, Brea, CA, USA) following the manufacturer’s instructions. The extracted RNA was eluted in 20 μL of nuclease-free water and divided into 2 aliquots: Half of the sample was analyzed without pre-amplification; the other followed the protocol including pre-amplification. All the extracted RNA was subjected to reverse transcription to cDNA using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems, Thermo Fisher Scientific), leading to a final volume of 20 μL.

For the samples not undergoing pre-amplification, all produced cDNA (20 μL) was divided into 4 aliquots and used for the amplification and for the detection, through real-time quantitative PCR, of 4 transcripts of interest: EPCAM, cytokeratin-8 (KRT8), CD45 and ACTB. The real-time PCR was performed based on the TaqMan™ technology using the TaqMan™ Fast Universal PCR Master Mix (Applied Biosystems) and the following TaqMan® Gene Expression Assays (Applied Biosystems): EPCAM, Hs00901885_m1; KRT8, Hs01630795_s1; CD45, Hs00365634_g1; and ACTB, Hs03023943_g1. For each transcript, the reaction mix was prepared using 5 μL of the produced cDNA, 15 μL of PCR Master Mix, 1.5 μL of the appropriate assay, and 8.5 μL of nuclease-free water. Each reaction mix was split into technical duplicates.

For the samples undergoing the protocol including the pre-amplification, the TaqMan® PreAmp Master Mix Kit (Applied Biosystems) was used, following the manufacturer’s instructions. First, the produced cDNA (12.5 out of 20 μL) was subjected to the pre-amplification of the 3 target transcripts (EPCAM, KRT8 and CD45) by using a pool of the TaqMan® Gene Expression Assays mentioned before, and the TaqMan® PreAmp Master Mix (Applied
Ten pre-amplification cycles were performed in a thermal cycler (GeneAmp® PCR System 9700, Applied Biosystems), according to the manufacturer’s instructions. Subsequently, the pre-amplified cDNA products were used for the real-time PCR reaction using the same TaqMan® Gene Expression Assays and the TaqMan™ Gene Expression Master Mix (Applied Biosystems). For each target, the reaction mix was prepared using 15 µL of pre-amplification products (diluted 1:5 with TE buffer), 30 µL of Master Mix, 3 µL of the appropriate assay and 12 µL of nuclease-free water. Each reaction mix was split into technical triplicates. Amplification reactions were carried out in MicroAmp® Fast Optical 96-Well Reaction Plates (Applied Biosystems) using a QuantStudio™ 12K Flex Real-time PCR System (Applied Biosystems) equipped with the QuantStudio™ 12K Flex software v1.2.3 (Applied Biosystems).
3.9. **Dual-positive cell evaluation by CellSearch®**

Dual-positive cells (DPcells) were counted by reviewing old image galleries of samples that had been processed with the CellSearch® for the detection of CTCs. DPcells were defined as cells presenting a nuclear staining and a clearly positive signal in both the PE and the APC channels (Figure 3.1 A). Cells were defined as DPcells if the pattern of the signals in the PE and the APC channels were different and did not seem suspicious of auto-fluorescence or spillover between channels (Figure 3.1 B).

<table>
<thead>
<tr>
<th>CK/DAPI</th>
<th>CK-PE</th>
<th>DAPI</th>
<th>CD45-APC</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
</tbody>
</table>

**Figure 3.1. Evaluation of DPcells by CellSearch®.**

A. Representative images of DPcells showing a nucleus and the expression of both CK and CD45. B. Examples of suspicious events of auto-fluorescence and spillover between channels that were not considered as DPcells.
3.10. **Statistical analyses**

To assess differences between groups, the chi-square or the Fisher exact tests were used as appropriate.

For BTC patients, disease-specific survival (DSS) was calculated from the date of enrollment to the date of death, or censored at the date of the last follow-up for living patients (median follow-up = 20 months, range = 16-30 months). For breast cancer patients treated at Northwestern University, PFS was calculated as the time between the enrollment date and the date of disease progression, or the date of the last follow-up for patients not developing disease progression (median follow-up = 4 months, range = 0.5-14.5 months). DSS and PFS were estimated using the Kaplan-Meier method and compared across the groups using the log-rank test.
4. RESULTS
4.1. Development of the protocol for CTC analysis

The first aim of the project was to develop a protocol able to detect the heterogeneous subpopulation of CTCs in patients’ blood samples. As previously described, protocols for CTC detection can vary greatly, but they are generally based on a few basic steps: blood sample collection, CTC enrichment, staining, CTC identification and molecular characterization (Figure 4.1).

![Workflow for CTC analysis](image.png)

Figure 4.1. Workflow for CTC analysis.

The steps of the workflow (blue boxes) and issues specifically addressed for the development of our protocol (white boxes) are indicated.

To develop our protocol, we first defined a general workflow that could be feasible in our institute (based on the available technologies), and then, for each step, we evaluated different possibilities in order to define the final CTC detection protocol.

We started our technical development from a workflow including:

- blood collection in Streck or EDTA tubes: Streck tubes contain a preservative which stabilizes nucleated blood cells, allowing CTCs’ to remain stable for up to 7 days (compared to 4 hours for EDTA tubes, which contain no preservatives);

- CTC enrichment by filtration (ScreenCell® CC kit);

- staining for nuclei, epithelial (EPCAM) and leukocyte (CD45) markers, using a fluorescent DNA dye and fluorescently-labeled antibodies, following a protocol provided by the DEPArray™ manufacturer;
- fixation with paraformaldehyde (PFA) 2% following a protocol provided by DEPArray™ manufacturer (only in case of blood collection in EDTA tubes);
- analysis with the DEPArray™ platform for CTC identification and single-cell isolation for downstream molecular characterization.

The workflow was designed to be applied to blood samples from patients with any type of solid tumor, therefore its performance was tested using cancer cell lines from different tumors, not only BTC cell lines.

4.1.1. Analytical validity assessment

The analytical validity of the initial workflow was tested, to assure its capability of effectively detecting and recovering CTCs.

CTCs are extremely rare and, during sample processing, cell loss due to manipulation can seriously impede CTC recovery. In our workflow, the steps causing the highest level of sample manipulation were fixation and staining procedures. Twelve aliquots of 500,000 cells/each from 6 different cell lines (MCF7, MDA-MB-231, DOHH-2, RAJI, SU-DHL-4, NAMALWA) underwent the fixation protocol and were counted again at the end. On average, only 5% of cells were lost during fixation, with only one outlier case showing a 47% loss. We repeated a similar experiment using 3 aliquots of 100,000 cells/each from 3 different cell lines (MCF7, DOHH-2, NHDF) to test the staining protocol; the mean cell loss was 6%.

Once it was ascertained that the cell loss due to manipulation was minimal, we assessed the analytical validity of the entire workflow by analyzing 3 healthy donor (HD) blood samples (9 mL): Sample A and B were spiked-in with 150 tumor cells/sample from 2 breast cancer cell lines expressing EPCAM (MCF7 and T-47D for sample A and B, respectively); sample C contained no tumor cells. For these experiments, blood samples were collected in Streck tubes. After blood filtration with ScreenCell® filters and staining
for EpCAM, CD45 and nuclei, samples were analyzed using the DEPArray™ for tumor cell identification. In samples A and B it was neither possible to identify any tumor cells by staining (EPCAM-positive and CD45-negative cells), nor by morphology (MCF7 and T-47D cells are bigger than blood cells). Conversely, we detected cells positive for both EPCAM and CD45, suggesting a possible unspecific staining for CD45 (Figure 4.2 A, B). Unexpectedly, similar cells were also found in sample C, which did not contain tumor cells (Figure 4.2 C), indicating an unspecific staining of WBCs for EPCAM.

![Figure 4.2. Image gallery showing unspecific staining of cells analyzed with the DEPArray™.](image)

Cells were identified in HD blood samples, drawn in Streck tubes, spiked-in with MCF7 (A), T-47D (B) or no tumor cells (C), enriched by filtration, fluorescently-stained for EpCAM (PE channel), CD45 (APC channel) and nucleus (DAPI channel) and analyzed. All cells are positive for both EpCAM and CD45.
The unspecific staining could be due to the low specificity of the antibody, or to the fixation of the cells with the Streck tube preservative. To test this, we compared, by flow cytometry, the staining of HD blood collected with Streck tubes vs. EDTA tubes, using the anti-EpCAM antibody used in the previous experiments and a new antibody against EpCAM (different clone and different manufacturer, Table 4.1).

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Clone</th>
<th>Manufacturer</th>
<th>Conjugate</th>
</tr>
</thead>
<tbody>
<tr>
<td>EpCAM-Old</td>
<td>0.N.276</td>
<td>Santa Cruz Biotechnology</td>
<td>PE</td>
</tr>
<tr>
<td>EpCAM-New</td>
<td>HEA-125</td>
<td>Miltenyi Biotec</td>
<td>FITC</td>
</tr>
</tbody>
</table>

Table 4.1. Antibodies against EpCAM compared in the flow cytometry experiments.

The table shows, for the old (first row) and the new (second row) antibody against EpCAM, the clone, the manufacturer and the conjugated fluorophore.

Blood samples were collected in EDTA or Streck tubes and WBCs were isolated by density gradient centrifugation. Isolated WBCs were stained with one of the 2 anti-EpCAM antibodies or with no antibody as negative control. Staining was evaluated by flow cytometry analysis (Figure 4.3).

As already observed in the spiking experiment, the antibody EpCAM-Old was not specific, in fact 16% and 40% of WBCs were stained in the sample collected with EDTA and Streck tubes, respectively (Figure 4.3 A). This problem was instead not observed using the new antibody: Only a very low percentage of WBCs were positively stained for EpCAM both in samples collected in EDTA and in Streck tubes (<1% and 2%, respectively, Figure 4.3 B). In both cases, though, unspecific staining was less for EDTA samples, suggesting that the preservative contained in Streck tubes can interfere with antigen-antibody reactivity. For this reason we decided to collect blood in EDTA tubes and to use the EpCAM-New antibody for subsequent samples.
Results

Figure 4.3. Flow cytometry analysis of WBCs isolated from HD blood samples collected in EDTA or in Streck tubes and stained with distinct antibodies directed against EpCAM.

(A) EpCAM-Old antibody. Variable fractions of WBCs are stained for EpCAM both in blood collected with EDTA tubes (left, 16%) and with Streck tubes (right, 40%). (B) EpCAM-New antibody. No or very low unspecific staining is observed. The dotted lines indicate the positivity thresholds of staining intensities defined on negative controls. FSC, forward scatter. [Adapted from Reduzzi C. et al., 2017].

The analytical validity of the refined workflow was tested with 2 new spiking experiments: For each experiment, 50 MCF7 cells were spiked into 3 mL of HD blood collected in EDTA tubes, enriched by filtration, stained, fixed and analyzed with the DEPArray™ platform.

As mentioned before, when using EDTA tubes, a fixation step with PFA 2% has to be added after enrichment. To evaluate if this type of fixation could also affect antibodies’ reactivity, in the first experiment the staining was performed before the fixation of the cells, while in the second experiment the order was inverted. In the two samples we clearly
identified 13 and 11 MCF7 cells (EpCAM-positive and CD45-negative nucleated cells), respectively. No unspecific staining of WBCs was observed (Figure 4.4).

However, during the analysis of the samples, a new issue was noticed: The presence of red blood cells (RBCs) in the brightfield channel interfered with the morphological evaluation of the cells by masking them partially or, in some cases, completely (Figure 4.4, red arrows).

To overcome this new problem, two similar spiking experiments (50 MCF7 spiked into 3 mL of HD blood, each) were repeated by following the same workflow, but introducing, after enrichment and before fixation, an RBC lysis step. In the two experiments, the recovery rate of tumor cells decreased by 2- and 4-fold, therefore this option was abandoned.
Cells were identified in a HD blood sample, spiked-in with 50 MCF7 cells. It is possible to distinguish tumor cells (upper part) from WBCs (lower part) both by marker expression (all MCF7 cells are positive in the PE channel and negative in the APC channel while all WBCs are EpCAM- and CD45+) and by size (tumor cells appear bigger than leukocytes). In some cases, the morphological evaluation in the brightfield channel is impaired by the presence of RBCs (indicated by red arrows). [Reduzzi C. et al., 2017].

### Table

<table>
<thead>
<tr>
<th></th>
<th>DAPI</th>
<th>EpCAM-PE</th>
<th>CD45-APC</th>
<th>brightfield</th>
<th>Merged DAPI/PE (blue/green)</th>
<th>Merged DAPI/APC (blue/red)</th>
</tr>
</thead>
</table>

**Figure 4.4.** Image gallery showing staining and morphological characteristics of cells analyzed with the DEPAarray™.
The solution to this issue came from another experiment that was performed in parallel. Since the markers conventionally used for CTC identification are both EpCAM and CK, an antibody against CK was tested to be added to the workflow. CK are located intracellularly, so, in order to stain them with the fluorescently-labeled antibody, it was also necessary to perform a permeabilization step. After assessing the antibody’s sensitivity on MCF7 cells by observation using a fluorescence microscope, its specificity was also tested by spiking MCF7 cells into 2 HD blood samples (50 cells/sample). The 2 samples were processed with the previously defined protocol, substituting EpCAM staining with CK staining and adding the permeabilization step. No unspecific staining was observed and both tumor cells and WBCs could easily be identified. During the analyses we also noticed the absence of RBCs in the brightfield channel in both samples. The detergent-containing buffer required for permeabilization effectively eliminated RBCs, without affecting tumor cell recovery-rates (13 and 8 tumor cells were detected in the 2 samples).

Two final HD blood samples were spiked-in with 50 MCF7 cells and processed with the entire workflow, including staining for EpCAM and CK and the permeabilization step. Since EpCAM and CK are both epithelial markers and are used together as CTC identification markers, the antibodies against EpCAM and CK were both PE-conjugated, in order to visualize the two epithelial markers in the same channel.

These experiments confirmed the overall analytical validity of the workflow; in fact it was possible to identify 14 and 11 tumor cells in the two samples, respectively, and no unspecific staining was observed (Figure 4.5).
Figure 4.5 DEPArray™ image gallery showing MCF7 cells identified in a spiked-in HD blood sample.

MCF7 cells are correctly stained for epithelial markers (PE channel) but not for CD45 (APC channel). In the first row, 2 WBCs are also present in the dielectrophoretic cage and show a correct staining (PE- and APC+). No RBCs are present (brightfield channel).

4.1.2. Enrichment methods

Enrichment is one of the most crucial steps in CTC analysis and choosing the wrong enrichment method can drastically decrease the number of detectable CTCs. As already described, many different technologies exist, exploiting different strategies. In this study, taking into consideration the heterogeneity of CTCs and the fact that conventional (epithelial marker-based) strategies were not successful in BTC patients, we decided to focus only on methodologies not based on the expression of epithelial markers. In particular, 4 technologies were tested: ScreenCell® filters; AutoMACS® Pro separator; OncoQuick®; and Parsortix®. To test their efficiency, HD blood samples collected in EDTA tubes were spiked-in with 50, 25 or 10 MCF7 cells, and were enriched with the four mentioned technologies. The enriched fractions were then fixed with PFA 2% and stained for EpCAM, CD45, CK, and nuclei as previously described. The number of recovered target cells was counted during DEPArray™ analysis and adjusted considering the void volume of the DEPArray™ cartridge.
4.1.2.1. ScreenCell® filters

This technology enriches CTCs based on cell size; whole blood is filtered through a membrane with pores of 6.5 µm diameter, which allows the passage of most blood cells and traps larger cells, such as CTCs. We performed 8 independent experiments and the mean recovery rate was 34% (Table 4.2).

<table>
<thead>
<tr>
<th>Number of spiked-in MCF7 cells</th>
<th>50</th>
<th>25</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>EpCAM/CK+ &amp; CD45- nucleated cells identified</td>
<td>11 13 8 13 14 11</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>% recovery corrected for DEPArray™ cartridge void volume</td>
<td>31 37 22 37 39 31</td>
<td>22</td>
<td>56</td>
</tr>
<tr>
<td>Mean recovery</td>
<td></td>
<td></td>
<td>34%</td>
</tr>
</tbody>
</table>

Table 4.2. Recovery rates of samples enriched by ScreenCell® filters.

The number of MCF7 cells identified in 8 spiking experiments are reported. To calculate the recovery rate, the number of expected cells was corrected by taking into account the void volume of the DEPArray™ cartridge. [Adapted from Reduzzi C. et al., 2017].

4.1.2.2. AutoMACS® Pro separator

This instrument exploits immunomagnetic beads coated with antibodies and separation columns to allow both positive and negative selection for selected markers. In our experiments, we used beads coated with antibodies against CD45 to perform a negative selection of the cells lacking this marker. Two experiments were performed; in both cases 50 MCF7 cells were spiked into HD blood, but none could be detected after enrichment. Although positive results have been published by investigators using this technology coupled with the DEPArray™ platform [Bulfoni M. et al., 2016], we decided to discard the use of AutoMACS® Pro separator which could not work efficiently in our workflow (probably due in part to logistical issues linked to the accessibility of the instrument).
4.1.2.3. OncoQuick®

OncoQuick® enriches CTCs by density gradient centrifugation: A porous barrier is inserted in a 50 mL-polypropylene tube above a specifically-developed separation medium which is able to separate most blood cells from CTCs (Figure 4.6).

![Figure 4.6. Graphic representation of OncoQuick® enrichment.](image)

By centrifugation in the OncoQuick® tube, erythrocytes and the majority of leukocytes are collected below the porous barrier, while tumor cells form a ring above the barrier, together with a small fraction of leukocytes and platelets. [OncoQuick® Instruction Manual]

Nine experiments were performed by spiking 50, 25 and 10 MCF7 cells into HD blood and the mean recovery rate by OncoQuick® enrichment was 81% (Table 4.3).

<table>
<thead>
<tr>
<th>Number of spiked-in MCF7 cells</th>
<th>50</th>
<th>25</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>EpCAM/CK+ &amp; CD45- nucleated cells identified</td>
<td>26</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>% recovery corrected for DEPArray™ cartridge void volume</td>
<td>72</td>
<td>61</td>
<td>100</td>
</tr>
<tr>
<td>Mean recovery</td>
<td>81%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3. Recovery rates of samples enriched by OncoQuick®.

The number of MCF7 cells identified in 9 spiking experiments are reported. To calculate the recovery rate, the number of expected cells was corrected by taking into account the void volume of the DEPArray™ cartridge. [Adapted from Reduzzi C. et al., 2017].
Despite the high sensitivity, which was a great advantage of this technology, the spiking experiments also highlighted a quite low specificity of the method (\textit{i.e.} a high number of contaminating WBCs). This might not be an issue in other CTC detection protocols, but for our workflow it constituted an important flaw. In fact, the maximum capacity of the DEPArray™ cartridge, where the enriched sample is loaded after fixation and staining, is 40,000 cells. An overload of the cartridge results in the impossibility of isolating single cells. Most of the samples processed with OncoQuick® contained a number of WBCs that exceeded the maximum capacity of the cartridge. Therefore, MCF7 cells could not be isolated because they could not be individually trapped inside the dielectrophoretic cages of the DEPArray™ cartridge (Figure 4.7).

A possible way to overcome this issue could be to split each sample into multiple cartridges, but this solution would imply a substantial increase in processing time and cost.

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
DAPI & brightfield & EpCAM/CK -PE & CD45-APC & Merged DAPI/PE (blue/red) & Merged DAPI/ACP (blue/yellow) & Merged DAPI/ACP/PE (blue/yellow/red) \\
\hline
\end{tabular}
\end{table}

\textbf{Figure 4.7.} MCF7 cells identified during DEPArray™ analysis of a spiked-in sample enriched with OncoQuick®.

In each row (showing one dielectrophoretic cage of the cartridge) it is possible to see one MCF7 (PE+ and APC-) surrounded by multiple contaminating WBCs.
4.1.2.4. Parsortix®

This technology is based on cell-size and deformability. Whole blood flows inside a microfluidic cassette that lets most of the blood cells pass through, whereas it retains large and non-deformable cells (Figure 4.8). The enriched cells can then be collected in a tube by reversing the flow and washing the cassette with PBS. The instrument allows for a first washing with ~200 µL of PBS and an optional second washing with ~1 mL of PBS.

![Graphic representation of Parsortix® enrichment.](https://angleplc.com/library/product-information/#brochure)

Since the reproducibility of our workflow had already been tested with the previous experiments (Table 4.2 and 4.3), the spiking experiments using Parsortix® were performed focusing only on the enrichment step. In particular, tumor cells were labeled with a fluorescent cell tracker before being spiked into HD blood samples and then counted, after harvesting from the Parsortix® cassette, using a fluorescence microscope. For each sample, the first wash of the cassette (200 µL) was harvested in a 96-well plate and the second (1 mL) in a 24-well plate for comparison. We performed 14 spiking experiments using MCF7 cells (for comparison with the other methods), one prostate cancer cell line
(PC-3) and three cholangiocarcinoma cell lines (EGI-1, HuCCT-1, HuH28). The recovery changed depending on the cell line (as expected, considering the different physical characteristics of the cell lines), ranging from 75% to 90% (Table 4.4).

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Spiked cells, n</th>
<th>Recovered cells, n</th>
<th>Recovery,</th>
<th>Mean recovery, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st wash</td>
<td>2nd wash</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCF7</td>
<td>50</td>
<td>35</td>
<td>0</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>21</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>PC-3</td>
<td>50</td>
<td>37</td>
<td>0</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>18</td>
<td>0</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>8</td>
<td>0</td>
<td>80</td>
</tr>
<tr>
<td>EGI-1</td>
<td>50</td>
<td>37</td>
<td>0</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>20</td>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>7</td>
<td>0</td>
<td>70</td>
</tr>
<tr>
<td>HuCCT-1</td>
<td>50</td>
<td>39</td>
<td>0</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>23</td>
<td>0</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>HuH28</td>
<td>50</td>
<td>43</td>
<td>0</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>24</td>
<td>0</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>8</td>
<td>0</td>
<td>80</td>
</tr>
</tbody>
</table>

Table 4.4. Recovery rates of samples enriched by Parsortix®.

The number of tumor cells recovered by Parsortix® in 14 spiking experiments using 5 different cancer cell lines (MCF7, PC-3, EGI-1, HuCCT-1, HuH28) are reported. The mean recovery rate varies depending on the cell line, ranging from 75 % to 90 %. [Adapted from Reduzzi C. et al., 2017].

Overall, the mean recovery rate across the cell lines was 81%. These experiments also showed that 99% of recovered tumor cells were washed out of the cassette with the first wash, making the second wash unnecessary.

While the sensitivity of the Parsortix® was very similar to the one of OncoQuick®, its specificity proved to be substantially higher: After loading the Parsortix®-enriched cells into the DEPArray™ cartridge, the total number of detected cells, including WBCs, was on average around 1,500, therefore allowing for single-cell recovery without the need to split the sample. These results established Parsortix® as the best enrichment method to be used in our workflow.
4.1.3. Identification strategies and phenotypic evaluation

CTC identification can be performed using 2 main strategies: positive selection by identifying the cells expressing tumor markers (conventionally EpCAM and CK); and negative selection by excluding all the cells expressing blood cell markers (usually CD45). During the development of the workflow, both strategies were implemented by expanding the cocktail of antibodies used for the staining in order to increase sensitivity. All new antibodies were tested, following manufacturers’ instructions, *in vitro* and/or through spiking experiments using different positive and negative controls to assess their sensitivity and specificity (Table 4.5), before being added to the workflow.

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Positive control</th>
<th>Negative control</th>
<th>Conjugate</th>
</tr>
</thead>
<tbody>
<tr>
<td>CK (AE1/AE3)</td>
<td>HuCCT-1</td>
<td>HD’s WBCs</td>
<td>PE</td>
</tr>
<tr>
<td>EGFR</td>
<td>MDA-MB-231</td>
<td>HD’s WBCs</td>
<td>PE</td>
</tr>
<tr>
<td>CD14</td>
<td>HD’s WBCs</td>
<td>MCF7</td>
<td>APC</td>
</tr>
<tr>
<td>CD16</td>
<td>HD’s WBCs</td>
<td>MCF7</td>
<td>APC</td>
</tr>
<tr>
<td>VIM</td>
<td>MDA-MB-231</td>
<td>MCF7</td>
<td>AF488</td>
</tr>
</tbody>
</table>

Table 4.5. Antibodies added to the protocol and tested for sensitivity and specificity.

The antigen recognized by each antibody and the cells used as positive and negative controls are reported (for the antibody against CK, the specific clone is indicated in parenthesis). The antibodies were PE-conjugated, APC-conjugated or AF488-conjugated, according to their use for positive selection, negative selection or phenotypic characterization, respectively.

4.1.3.1. Positive selection

At first, the identification of CTCs in our workflow was based on EpCAM and CK expression. The antibody against CK that we tested during the analytical validation was the one suggested by the DEPArray™ manufacturer, who also provided us with the staining protocol. It is a pan-CK antibody directed against clone C11 and detects CK 4, 5, 6, 8, 10,
13, 18. Nonetheless, during the development of the protocol, we decided to add, to the set of identification markers (visualized in the PE channel), a second antibody directed against clone AE1/AE3 which recognizes CK 1, 3-6, 8, 10, 14-16 and 19. This second antibody was chosen because the antibodies against clone AE1/AE3 are currently used in our institute for staining CK in BTC tissue samples for anatomical pathology tests, and it is also reported in literature to be particularly efficient for the detection of micrometastasis in patients with BTC [Natarajan S. et al., 2005].

In a previous study performed using a different CTC detection protocol in breast cancer patients, we observed that CTC detection was increased when EGFR was added to the set of markers used for enrichment [Fina E. et al., 2015]. Therefore, we introduced an antibody against EGFR into the staining cocktail of the workflow. As EGFR was used as a positive selection marker, the selected antibody was PE-conjugated.

4.1.3.2. Negative selection

Although improving positive selection can increase CTC detection, this strategy will always be limited by the fact that it is impossible to predict the entire arsenal of markers that the different CTC subpopulations can express. From this perspective, performing a negative selection of all WBCs seems more fruitful.

Unfortunately, by using only CD45 for negative selection, we observed a significant number of cells negative for both CD45 and epithelial markers (double-negative cells) also in HD blood samples (around 200 cells per sample), indicating a low (or lack of) expression of CD45 in a minority of blood cells, or a limited sensitivity of the antibody. To increase the WBC staining efficiency of our protocol, two antibodies against CD14 and CD16 (mainly expressed by monocytes and natural killers, respectively) were included in the staining step. The staining for the 3 WBC markers (CD45, CD14 and CD16) together decreased the number of observed double-negative cells to between 0 and 20 cells per
Results

Although it was still not possible to define whether these cells were CTCs or normal cells, their reduced number allowed for their recovery for subsequent molecular analysis to evaluate their possible tumor origin.

4.1.3.3. Phenotypic evaluation

Considering the known role of the EMT in CTCs, we hypothesized that the double-negative cells that we observed in our samples could be CTCs with a mesenchymal phenotype. To investigate this possibility, as well as to evaluate the presence of CTCs with a mixed epithelial-mesenchymal phenotype, an antibody against the mesenchymal marker vimentin (VIM) was tested on MDA-MB-231 cells (a triple-negative breast cancer cell line with a mesenchymal-like phenotype), and subsequently added to the staining protocol. Unfortunately, VIM is also expressed by blood cells and could therefore not be used as a selection marker, but only for the phenotypic characterization of the selected putative CTCs. For this reason an AF488-conjugated antibody was used, in order to visualize VIM in a different channel (FITC channel) from the ones used for positive and negative selection markers (PE and APC channel, respectively).

Overall, the improved staining protocol allowed the detection of 3 types of cells: epithelial CTCs (eCTCs, PE-positive and APC-negative); WBCs (PE-negative and APC-positive); and double-negative cells (PE-negative and APC-negative). eCTCs’ and double-negative cells’ phenotype could be further characterized for VIM expression (Figure 4.9).
Figure 4.9. CTCs, WBCs and double-negative cells identified in two blood samples from biliary tract cancer patients.

(A) Three epithelial CTCs (eCTCs, i.e. PE+ and APC- nucleated cells) identified based on the expression of epithelial and blood cell markers (PE and APC channel, respectively). The phenotype of these cells was further characterized by the evaluation of VIM expression (FITC channel); the cells showed a mixed phenotype going from a more mesenchymal (CTC 1) to a more epithelial (CTC 3) one. It is possible to see 2 WBCs (PE- and APC+ nucleated cells) next to CTC 2 and CTC 3 (B) Four double-negative cells. In the FITC channel the heterogeneous expression of VIM is shown (cell 7 has the highest level of expression, while cell 6 is negative). [Adapted from Reduzzi C. et al., 2017]
4.1.4. Whole-genome amplification and quality control assay

To perform any kind of molecular characterization on single cells, it is mandatory to perform a WGA step to increase the amount of available DNA. We assessed our capability of performing WGA and the quality of the amplified DNA by using the cells recovered during spiking experiments.

At the end of the DEPArray™ processing, single CTCs were collected in 0.2 mL PCR-tubes in a drop (around 20 µL) of SB115 buffer, a proprietary buffer developed by the DEPArray™ manufacturer to allow for the correct formation of the dielectrophoretic cages inside the cartridge. Before starting the WGA it was therefore necessary to perform a volume reduction procedure to leave the cell in 1 µL of PBS. This procedure, which was performed manually, could have potentially caused the loss of the recovered cell.

After volume reduction, the cells were lysed to allow WGA, which was followed by a quality assessment through a PCR-based approach. This way, we assigned a genome integrity index (GII) to each amplified DNA. The GII value could range from 0 to 4: 0 indicated a failure in the WGA, probably due to the loss of the cell during volume reduction; 1-2 corresponded to a low-quality sample; 3-4 was indicative of high-quality amplified DNA. The quality of the DNA is predictive of successful downstream molecular analyses: High-quality samples can undergo genome-wide analysis, whereas samples with a GII = 1 can only be used for gene-specific assays. Samples with a GII = 2 can be analyzed by genome-wide assays but the success rate will be lower than for samples with a GII = 3-4 [Polzer et al. 2014].

The WGA was performed on 286 samples (single cells and small pools of cells) recovered during spiking experiments. We obtained high-quality DNA (GII = 3-4) in 73% of the samples, low-quality (GII = 1-2) in 21% and no amplification (GII = 0) in 6%. Overall, both the low rate of cell loss during volume reduction and the quality of the obtained DNA
were satisfying and predicted that, by using this approach, a high fraction of samples could be molecularly characterized.

4.1.5. Copy number alteration profiling

By low-pass whole-genome sequencing (lp-WGS) it is possible to obtain CNA profiling of single cells at a lower cost than microarrays and high-throughput WGS, thus allowing the analysis of a higher number of samples. Compared with deep sequencing strategies, lp-WGS produces only a fraction of the data per sample and relies on computational methods to fill in the missing information. So, while it does not provide information on small-scale alterations (INDELs), it instead gives an overview of the macro-aberrations present in the entire genome of a cell, thus indicating whether a cell is normal (flat, diploid profile) or cancerous (aberrant profile). This approach was particularly interesting for us because it allowed for the profiling of all of the identified double-negative cells (putative CTCs), to ascertain their tumor origin.

Since this type of sequencing was new in our institute, the technical validity of this pipeline (including library preparation, lp-WGS and bioinformatic analysis) was tested on tumor and normal control DNA provided by the manufacturer, before analyzing patients’ samples. Five samples were processed and were correctly identified as aberrant or normal (Figure 4.10). Three pools of 10 WBCs isolated from 3 BTC patients were also analyzed and showed flat CNA profiles.

CNA profiling by lp-WGS was then performed on all the double-negative cells isolated from patients, which were then re-classified as normal double-negative cells or bona fide non-conventional CTCs (ncCTCs).
Figure 4.10. Copy number alteration profiles of control DNA samples.

For each sample (rows), the chromosomes are displayed on the x-axis while the copy number is on the y-axis. The 2 upper samples show numerous gains (in red) and losses (in blue), while the other 3 samples present a flat and diploid profile (in black).
4.2. Detection of CTCs in blood samples from patients with BTC

To evaluate the clinical relevance of the detection of eCTCs and ncCTCs by using our workflow, we recruited 24 patients with intrahepatic CCA (n = 13), extrahepatic CCA (n = 9) and GBC (n = 2) at our institute. Most patients (n = 22) were undergoing first-line chemotherapy (Table 4.6).

<table>
<thead>
<tr>
<th>Cases, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
</tr>
<tr>
<td>- female</td>
</tr>
<tr>
<td>- male</td>
</tr>
<tr>
<td><strong>Age</strong></td>
</tr>
<tr>
<td>- ≤ 50</td>
</tr>
<tr>
<td>- 51-69</td>
</tr>
<tr>
<td>- ≥ 70</td>
</tr>
<tr>
<td><strong>Metastatic at diagnosis</strong></td>
</tr>
<tr>
<td>- Yes</td>
</tr>
<tr>
<td>- No</td>
</tr>
<tr>
<td><strong>Tumor anatomical location</strong></td>
</tr>
<tr>
<td>- Intrahepatic</td>
</tr>
<tr>
<td>- Extrahepatic</td>
</tr>
<tr>
<td>- Gallbladder</td>
</tr>
<tr>
<td><strong>Previous treatment</strong></td>
</tr>
<tr>
<td>- Yes</td>
</tr>
<tr>
<td>- Cisplatin + Gemcitabine</td>
</tr>
<tr>
<td>- Gemcitabine</td>
</tr>
<tr>
<td>- Radiotherapy + Capecitabine</td>
</tr>
<tr>
<td>- No*</td>
</tr>
<tr>
<td><strong>Treatment regimen</strong></td>
</tr>
<tr>
<td>- First-line</td>
</tr>
<tr>
<td>- Cisplatin + Gemcitabine</td>
</tr>
<tr>
<td>- FOLFOX</td>
</tr>
<tr>
<td>- Second-line</td>
</tr>
<tr>
<td>- FOLFOX</td>
</tr>
<tr>
<td>- FGFR inhibitor</td>
</tr>
<tr>
<td><strong>Previous surgery</strong></td>
</tr>
<tr>
<td>- Yes</td>
</tr>
<tr>
<td>- No</td>
</tr>
</tbody>
</table>

Table 4.6. Patients’ clinico-pathological characteristics.

*Patients receiving previous treatment 6 or more months before recruitment were considered as untreated.
[Adapted from Reduzzi et al., 2019]
Blood samples (10 mL) were collected at different time points: at baseline (BL, before starting a new line of treatment); during treatment (DT, close to disease re-evaluation); at the end of treatment (EOT); and after the end of treatment (follow-up, FU). The collected samples were used to  

1. assess the presence of eCTC and ncCTC,  
2. evaluate the prognostic significance of CTCs, and  
3. perform CTCs’ molecular characterization. 

The patients included in each type of analysis and the reasons for exclusions are shown in Figure 4.11.

![Diagram showing the flow of patient analysis](image)

**Figure 4.11. Number of patients enrolled and included in each type of analysis.**

Twenty-four patients were recruited. For 21 patients, eCTC and ncCTC assessment could be performed after CNA profiling of double-negative cells was completed, of these, 18, for whom a baseline sample was available, were included in survival analysis. All patients for whom at least 1 CTC underwent successful molecular profiling were included in the CNA analysis. [Adapted from Reduzzi C. et al., 2019]

### 4.2.1. Detection of eCTCs and ncCTCs

We evaluated the presence of eCTCs and ncCTCs in 41 samples collected at different times (BL, DT, EOT, FU) from 21 patients. eCTC counts were obtained by phenotypic evaluation only, whereas ncCTCs were identified from the recovered double-negative cells by CNA profiling (Figure 4.12). Overall, 19 eCTCs and 437 double-negative cells were
Results

detected. CNA analysis was performed on 367 double-negative cells (70 cells (19%) could not be evaluated due to low quality DNA after WGA, \textit{i.e.} GII = 0 or 1) and, among these, 73 ncCTCs were identified (Table 4.7). By using 1 CTC/sample as positivity threshold, 8 samples were positive for eCTCs (19%, median eCTC number = 1.5, range 1-5) and 31 were positive for ncCTCs (76%, median ncCTC number = 2, range 1-7).

![Figure 4.12. Identification of eCTCs, ncCTCs and WBCs in clinical samples.](image)

(A) Two nucleated cells (DAPI channel). The bottom/left cell is a classical eCTC, expressing epithelial markers (PE channel) and lacking leukocyte markers (APC channel). The top/right cell is a WBC, showing the opposite staining pattern (PE- and APC+). (B) Three double-negative nucleated cells (PE-, APC- and DAPI+) with a similar morphology (brightfield channel, BF). The CNA profiles corresponding to each cell are reported on the right side and reveal the tumor origin for the first two cells from the top, whereas the third cell has a normal diploid CNA profile. [Reduzzi C. et al., 2019]

Overall, 3 samples (7%) had only eCTCs, 26 samples (64%) had only ncCTCs, 5 samples (12%) had both types of CTCs, and 7 samples (17%) had no CTCs. Therefore, the majority of the samples (64%) that were positive solely for ncCTCs would have been considered CTC-negative by conventional CTC identification criteria. Thus, by also looking at ncCTCs, the overall CTC-positivity rate increased by 4.3-fold, from 19% to 83%. In particular, we detected CTCs in almost all patients (20 out of 21, considering all time points).
<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Anatomic location</th>
<th>Timing</th>
<th>eCTCs (n)</th>
<th>Sequenced double-negative cells (n)</th>
<th>ncCTCs (n)</th>
<th>CTC phenotype (e/nc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BT03</td>
<td>I</td>
<td>BL cis/gem</td>
<td>0</td>
<td>20</td>
<td>2</td>
<td>-/+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EOT</td>
<td>0</td>
<td>9</td>
<td>7</td>
<td>-/+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL FOLFOX</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>-/+</td>
</tr>
<tr>
<td>BT10</td>
<td>I</td>
<td>DT</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>-/-</td>
</tr>
<tr>
<td>BT11</td>
<td>E</td>
<td>DT</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>-/+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>-/-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FU</td>
<td>0</td>
<td>12</td>
<td>4</td>
<td>-/-</td>
</tr>
<tr>
<td>BT13</td>
<td>E</td>
<td>DT</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>-/+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FU</td>
<td>4</td>
<td>13</td>
<td>4</td>
<td>+/-</td>
</tr>
<tr>
<td>BT15</td>
<td>E</td>
<td>DT</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>-/+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL capcitabine</td>
<td>0</td>
<td>6</td>
<td>3</td>
<td>-/+</td>
</tr>
<tr>
<td>BT16</td>
<td>I</td>
<td>BL FOLFOX</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-/-</td>
</tr>
<tr>
<td>BT19</td>
<td>GB</td>
<td>BL cis/gem</td>
<td>0</td>
<td>6</td>
<td>4</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL FOLFOX</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>++/</td>
</tr>
<tr>
<td>BT20</td>
<td>I</td>
<td>BL cis/gem</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>-/-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL FOLFOX</td>
<td>0</td>
<td>28</td>
<td>0</td>
<td>-/-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EOT</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>-/+</td>
</tr>
<tr>
<td>BT21</td>
<td>GB</td>
<td>BL capcitabine</td>
<td>0</td>
<td>5</td>
<td>3</td>
<td>-/+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>-/-</td>
</tr>
<tr>
<td>BT22</td>
<td>I</td>
<td>BL cis/gem</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>-/+</td>
</tr>
<tr>
<td>BT23</td>
<td>I</td>
<td>BL cis/gem</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>-/+</td>
</tr>
<tr>
<td>BT24</td>
<td>E</td>
<td>BL cis/gem</td>
<td>0</td>
<td>15</td>
<td>2</td>
<td>-/+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT</td>
<td>0</td>
<td>25</td>
<td>1</td>
<td>-/+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EOT</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>-/+</td>
</tr>
<tr>
<td>BT25</td>
<td>E</td>
<td>DT</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>-/+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EOT</td>
<td>0</td>
<td>18</td>
<td>2</td>
<td>-/+</td>
</tr>
<tr>
<td>BT26</td>
<td>I</td>
<td>BL cis/gem</td>
<td>2</td>
<td>7</td>
<td>1</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>-/-</td>
</tr>
<tr>
<td>BT27</td>
<td>E</td>
<td>BL FOLFOX</td>
<td>5</td>
<td>8</td>
<td>0</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FU</td>
<td>0</td>
<td>30</td>
<td>4</td>
<td>-/+</td>
</tr>
<tr>
<td>BT28</td>
<td>I</td>
<td>BL cis/gem</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>+/-</td>
</tr>
<tr>
<td>BT29</td>
<td>E</td>
<td>BL FOLFOX</td>
<td>0</td>
<td>13</td>
<td>2</td>
<td>-/+</td>
</tr>
<tr>
<td>BT30</td>
<td>I</td>
<td>BL cis/gem</td>
<td>0</td>
<td>24</td>
<td>3</td>
<td>-/+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT</td>
<td>0</td>
<td>16</td>
<td>6</td>
<td>-/+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EOT</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>-/-</td>
</tr>
<tr>
<td>BT33</td>
<td>I</td>
<td>BL cis/gem</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>+/-</td>
</tr>
<tr>
<td>BT34</td>
<td>E</td>
<td>BL cis/gem</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>-/+</td>
</tr>
</tbody>
</table>

Table 4.7. Number of eCTCs and ncCTCs detected for each analyzed blood sample.
Each row corresponds to one sample. The patient ID, the anatomic location of the tumor (intrahepatic, I, extrahepatic, E, and gallbladder, GB) and the timing of the blood draw (baseline, BL, during treatment, DT,
end of treatment, EOT, and follow-up, FU) are reported in the first, second and third column, respectively. Samples from the same patient are reported chronologically, some patients (BT03, BT16, BT19, BT20, BT21) were followed across multiple treatment lines. Indicated for each sample are: the number of detected epithelial CTCs (eCTCs); the number of double-negative cells that were analyzed by lp-WGS; the number of ncCTCs confirmed by CNA profiling; and the sample category according to CTCs’ phenotype. Cis/gem, cisplatin plus gemcitabine. [Adapted from Reduzzi C. et al., 2019]

4.2.2. Prognostic role of CTCs

To investigate the prognostic impact of CTCs, we considered only patients from whom a blood sample drawn before treatment initiation (BL sample) was evaluated. First, we assessed the prognostic role of eCTCs only. For this analysis, three additional BL samples for patients BT15, BT16 and BT25, evaluated for eCTCs only (before the use of CD14 and CD16 as negative selection markers), were included in the survival analysis, for a total of 18 patients. The median follow-up was 20 months and at least 1 eCTC was detected in 4 patients (25%). Despite the limited number of patients, an association between eCTC and disease-specific survival (DSS) was observed (Figure 4.13).

![Figure 4.13. Association between eCTCs and survival.](image)

Kaplan-Meier’s analysis of DSS for eCTC-positive and eCTC-negative patients. DSS estimates according to eCTC at baseline: ≥ 1 CTC/10 mL of blood (dotted line) vs. no CTC/10 mL of blood (solid line). [Reduzzi C. et al., 2019]
The median DSS was 17 months and 9 months for eCTC-negative and eCTC-positive patients, respectively. The relation between eCTCs and patients’ clinicopathological features was also investigated and no significant association was observed (Table 4.8).

<table>
<thead>
<tr>
<th></th>
<th>Overall (n = 18)</th>
<th>eCTC+ve (n = 4)</th>
<th>eCTC-ve (n = 14)</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>50.0</td>
<td>50.0</td>
<td>50.0</td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>50.0</td>
<td>50.0</td>
<td>50.0</td>
<td></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 50</td>
<td>16.7</td>
<td>0</td>
<td>21.5</td>
<td>p = 0.213</td>
</tr>
<tr>
<td>51-69</td>
<td>66.6</td>
<td>50.0</td>
<td>71.5</td>
<td></td>
</tr>
<tr>
<td>≥ 70</td>
<td>16.7</td>
<td>50.0</td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td><strong>Metastatic at diagnosis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>50.0</td>
<td>75.0</td>
<td>43.0</td>
<td>p = 0.576</td>
</tr>
<tr>
<td>No</td>
<td>50.0</td>
<td>25.0</td>
<td>57.0</td>
<td></td>
</tr>
<tr>
<td><strong>Tumor anatomical location</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrahepatic</td>
<td>50.0</td>
<td>75.0</td>
<td>43.0</td>
<td>p = 0.753</td>
</tr>
<tr>
<td>Extrahepatic</td>
<td>38.9</td>
<td>25.0</td>
<td>43.0</td>
<td></td>
</tr>
<tr>
<td>Gallbladder</td>
<td>11.1</td>
<td>0</td>
<td>14.0</td>
<td></td>
</tr>
<tr>
<td><strong>Previous systemic treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>22.2</td>
<td>0</td>
<td>28.6</td>
<td>p = 0.524</td>
</tr>
<tr>
<td>No</td>
<td>77.8</td>
<td>100</td>
<td>71.4</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cis/gem</td>
<td>77.8</td>
<td>75.0</td>
<td>78.6</td>
<td>p = 1.000</td>
</tr>
<tr>
<td>FOLFOX</td>
<td>16.7</td>
<td>25.0</td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td>Capecitabine</td>
<td>5.6</td>
<td>0</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td><strong>Performance status (ECOG)</strong></td>
<td></td>
<td></td>
<td></td>
<td>p = 1.000</td>
</tr>
<tr>
<td>1</td>
<td>11.1</td>
<td>0</td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>88.9</td>
<td>100</td>
<td>85.7</td>
<td></td>
</tr>
<tr>
<td><strong>Previous surgery</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>38.9</td>
<td>25.0</td>
<td>42.9</td>
<td>p = 1.000</td>
</tr>
<tr>
<td>No</td>
<td>61.1</td>
<td>75.0</td>
<td>57.1</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.8. Association between eCTC-status and clinical characteristics of the patients included in the survival analysis.

The presence of eCTC was related to: patients’ gender and age; the presence of metastasis at diagnosis; the anatomical location of the tumor; the eventual previous systemic treatment; the treatment the patient was undergoing; the ECOG performance status; and the previous surgery. No significant association was observed. *Fisher exact test [Adapted from Reduzzi C. et al., 2019]

On the contrary, survival analysis with respect to ncCTCs, performed on 15 patients, did not show an association with survival.
These results support the already reported prognostic significance of eCTCs, but do not suggest a prognostic relevance for ncCTCs.

4.2.3. Treatment monitoring

To determine if the presence and phenotype of CTCs are affected by treatment, we compared eCTCs and ncCTCs detection in samples collected at BL from untreated patients (BL untreated, n = 13), at BL from patients who were starting a new treatment but had already received a previous one (BL pretreated, n = 6), and during or at the end of treatment (DT/EOT, n = 19), (Figure 4.14).

![Figure 4.14. Effect of treatment on CTCs in BTC patients.](image)

Each bar represents samples drawn at a different time point: at baseline from untreated patients (BL untreated, left bar); at baseline from pretreated patients (BL pretreated, middle bar); and during or at the end of treatment (DT/EOT, right bar). On the y-axis, the detection frequencies of eCTCs and ncCTCs are reported. [Reduzzi C. et al., 2019]
At baseline in untreated patients, eCTCs and ncCTCs were detected in 30.8% and 84.6% of patients, respectively. ncCTCs were therefore 2.75-fold more frequent than eCTCs and no sample was CTC-negative. At baseline in samples drawn from pretreated patients, CTCs were not detected in 1 sample and eCTCs were detected in 1 sample in association with ncCTCs. Four samples were positive for ncCTCs only. In samples collected DT/EOT, ncCTCs were also detected in the majority of cases (57.9%), but in this case they were 6.3-fold more frequent than eCTCs (which were present in only 10.4% of samples). Moreover, in 6 samples (32%) we did not detect any CTCs. The higher frequency of CTC-negative samples at DT/EOT (as compared to BL) was the most relevant effect induced by treatment. Interestingly, all the patients that were CTC-negative at DT/EOT were classified as responders (i.e. showing a partial response, PR, or a stable disease, SD, based on RECIST) at the time of CTC evaluation, suggesting a possible relation between ncCTC presence and patient response to treatment.

To look further into the possibility of using CTCs for treatment monitoring, we focused on the patients for whom CTC evaluation was performed at both BL and DT/EOT (8 patients, reported in Table 4.9). For 4 of these patients, a third blood sample, collected 1 to 4 months after the end of treatment, was also included in the analysis.

<table>
<thead>
<tr>
<th></th>
<th>BT03</th>
<th>BT17</th>
<th>BT19</th>
<th>BT20</th>
<th>BT21</th>
<th>BT24</th>
<th>BT26</th>
<th>BT30</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL</td>
<td>nc</td>
<td>e</td>
<td>nc</td>
<td>e</td>
<td>nc</td>
<td>e</td>
<td>nc</td>
<td>e</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

**Table 4.9. Variations in CTC numbers during treatment.**

For each patient (columns), the number of ncCTCs and eCTCs detected at the different time points is shown: baseline (BL); during/at the end of treatment (DT/EOT); during follow-up after the end of treatment (FU). n.a., not available. [Reduzzi C. et al., 2019]
Considering the low abundance of eCTCs in these samples, we focused on ncCTCs and related their presence with patients’ response to treatment (Figure 4.15).

![Figure 4.15. CTC monitoring during treatment.](image)

Variations in ncCTC number evaluated in samples longitudinally collected from 8 patients (reported in Table 4.9) at baseline (BL), during/at the end of treatment (DT/EOT) and during follow-up (FU). Each line represents one patient and the color indicates the patient’s response to treatment: stable disease/partial response (SD/PR) = responding, in blue; progressive disease (PD) = non-responding, in red. [Reduzzi C. et al., 2019]

At the time of the DT/EOT CTC evaluation (the second blood sample), all patients were responding to the therapy, showing either a stable disease or a partial response. Accordingly, in 5 patients (BT17, BT19, BT20, BT21 and BT24) the number of ncCTCs decreased from BL to DT/EOT (with 3 patients passing from ncCTC-positive to ncCTC-negative status). One patient (BT26) had only 1 ncCTC at BL and still presented 1 ncCTC at DT/EOT. The remaining 2 patients, however, experienced an increase of ncCTCs: Patient BT03 passed from 2 to 7 ncCTCs and patient BT30 from 3 to 6 ncCTCs, at BL and DT/EOT, respectively. Nonetheless, patient BT03 rapidly progressed (1 month
after DT/EOT CTC evaluation) and in the FU sample (1 month after the clinical assessment of progressive disease, PD) a high number of ncCTCs (n = 6) was still detected. On the contrary, patient BT30 was still classified as in PR around 1 month after the end of treatment (when CTCs were re-evaluated) and the patient had in fact become CTC-negative. FU samples were also collected from patients BT19 and BT20 shortly after disease progression. Both patients were CTC-negative at DT/EOT evaluation (when they were responding to treatment). At the FU CTC evaluation, patient BT20 was still CTC-negative despite the PD, whereas for patient BT19 an increase in ncCTCs was detected 1 month after the clinical assessment of PD. Overall, these results suggested that changes in CTCs during treatments can mirror patient response to therapy.

4.2.4. Phenotypic characterization

To better define the phenotype of ncCTCs and investigate whether they were CTCs with a mesenchymal phenotype, the staining for the mesenchymal marker VIM was evaluated in 35 samples. Of the 53 ncCTCs identified in the evaluated samples, the majority (36 cells, 68%) showed VIM expression. In the same samples, 19 eCTCs were also identified, and 10 could be evaluated for VIM expression (for 9 cells it was impossible to clearly assess VIM staining due to the presence of VIM-positive WBCs partially covering them). All eCTCs resulted as VIM-positive (Figure 4.16).

![Figure 4.16. Vimentin expression in eCTCs.](image)

Two representative eCTCs identified in a patient blood sample showing both the staining for epithelial markers (PE channel) and for VIM (FITC channel). A few WBCs, positive for CD45 (ACP channel), are also
Results

103

present in the dielectrophoretic cages, but in these cases their presence does not hinder the evaluation of eCTCs’ VIM expression.

Although CTCs with a mixed epithelial-mesenchymal phenotype have been reported in literature, the fact that 100% of the detected eCTCs were also VIM-positive seemed suspicious for a possible technical artifact. So, to verify if the observed VIM expression was actually a biological phenomenon, rather than an artifact connected to a spillover between the PE and the FITC channels of the fluorescence microscope, we performed a control experiment by spiking MCF7 cells into a HD blood sample. The sample was processed with the same workflow used for clinical samples, but in this case the antibody for VIM was not included in the staining in order to have no true signal in the FITC channel and to evaluate if a false signal in FITC could be detected for PE-positive cells (MCF7 cells). All detected MCF7 cells were PE-positive and FITC-negative (Figure 4.17), indicating that the signal observed in clinical samples for eCTCs was not a technical artifact. Based on these results, a partial EMT seems to be a frequent phenomenon in CTCs in BTC, possibly helping their migration into the bloodstream.

Nonetheless, the antibody against VIM was replaced with an antibody against cell surface vimentin (CSV). This was done because we wanted a mesenchymal marker able to
Results
discriminate between ncCTCs and normal double-negative cells (in order to decrease the number of cells undergoing lp-WGS), and because CSV, unlike intracellular VIM, is mainly expressed by cancer cells and is reported to be a more specific tumor cell marker [Satelli A. et al., 2015^a; Satelli A. et al., 2015^b]. We performed the staining for CSV in 6 clinical samples: In 2 samples no CTCs were detected; in the other 4 we identified 21 ncCTCs, of which only 3 were CSV-positive. Although, as expected, a reduced staining of cells was observed, we also recovered double-negative cells that expressed CSV but resulted as normal cells by CNA profiling. This suggests that this antibody is not specific for tumor cells only. Interestingly, 2 of the detected CSV-positive normal cells showed a non-rounded morphology (Figure 4.18). Unfortunately, no eCTCs were present in the samples that we had analyzed, so we do not have data regarding the staining of this subpopulation of CTCs for CSV.

<table>
<thead>
<tr>
<th>DAPI</th>
<th>brightfield</th>
<th>PE</th>
<th>APC</th>
<th>FITC</th>
<th>DAPI/PE</th>
<th>DAPI/APC</th>
<th>DAPI/FITC</th>
</tr>
</thead>
</table>

Figure 4.18. Non-aberrant, circulating double-negative cells expressing CSV.

Two double-negative cells (PE- and APC-) collected from the blood of 2 patients, showing positive CSV staining (FITC channel). Although the cells are CD45-negative (APC channel), and have a non-rounded morphology, they both showed a flat CNA profile.
4.2.5. Molecular characterization

4.2.5.1. Mutational profiling

After the identification of eCTCs by positive selection and ncCTCs by negative selection combined with CNA analysis, our aim was to combine phenotypic and molecular characterization by performing single-cell mutational profiling of each CTC.

Mutational profiling was performed by targeted sequencing of 50 cancer-associated genes on 20 CTCs isolated from 7 patients with extrahepatic (n = 4), intrahepatic (n = 2) and gallbladder disease (n = 1), plus one pool of WBCs for each patient, as normal controls. Most of the identified mutations were private mutations (i.e. present in only 1 CTC) and this limited the reliability and the biological interpretation of the data. In fact, since the samples underwent WGA, which can introduce errors during DNA replication, private mutations found in only one CTC could not be considered reliable. Among the cells that were analyzed, we found only 1 mutation shared by multiple CTCs from the same patient (with extrahepatic BTC), from whom 9 CTCs were analyzed (Figure 4.19). This was a heterozygous mutation of KIT (p.Met541Leu), a gene which is not frequently mutated in BTC.
Figure 4.19. Mutational profiling of 9 CTCs recovered from one blood sample from a BTC patient.

The mutations detected in each ncCTC (top), and in each eCTC (bottom) are shown in red. Only the mutation in KIT is shared by multiple CTCs, while all the other mutations are unique to individual CTCs. These mutations were not detected in the WBC control sample (first row).

These results indicate the possibility of performing this type of analysis on CTCs isolated from BTC patients but also highlight its limitations, especially for samples presenting few CTCs. Therefore, considering the low median number of CTCs present in our clinical samples and the inaccessibility of tumor tissue to validate the detected mutations, we decided not to continue analyzing CTCs and to eventually perform mutational profiling only in selected patients with a high number of CTCs. Instead, we decided to focus on CNA profiling data, which were more reliable and had already been used for double-negative cells’ confirmatory analyses.
4.2.5.2. Copy number alteration analysis

To fully exploit the information carried by CNA profiles for the molecular characterization of the collected CTCs, two different approaches were used. The first was aimed at measuring the chromosomal instability of the cells; the second was focused on the investigation of alterations involved in therapeutic resistance. CNA profiles were available for 88 CTCs (15 eCTCs and 73 neCTCs) collected from 23 BTC patients.

4.2.5.2.1. Chromosomal instability evaluation

To measure the chromosomal instability of the cells, we used the large-scale state transition score (LST score), which is calculated as the number of chromosomal breaks between adjacent regions of 10 Mb [Greene S.B. et al., 2016]. An LST score was assigned to each CTC (median LST score = 9, range 0-40) and the scores were related to patients’ clinical outcome, i.e. best response to treatment = PR/SD (Table 4.10) vs. PD (Table 4.11).
Table 4.10. LST scores of CTCs recovered from patients responding to treatment.

For every patient (Pt.), the analyzed CTCs are reported (ID). CTCs’ phenotype (epithelial/non-conventional, e/nc), the time point of CTCs collection (Time: baseline, BL; during treatment, DT; end of treatment, EOT; follow-up, FU) and the patient’s best response to the considered treatment line (Best resp.: partial response, PR; stable disease, SD) are also indicated. The last column (LST) shows the LST score for each CTC. [Adapted from Reduzzi C. et al., 2019]
Table 4.11. LST scores of CTCs recovered from patients non-responding to treatment.

For every patient (Pt.), the analyzed CTCs are reported (ID). CTCs’ phenotype (epithelial/non-conventional, e/nc), the time point of CTCs collection (Time: baseline, BL; during treatment, DT; end of treatment, EOT; follow-up, FU) and the patient’s best response to the considered treatment line (Best resp.: disease progression, PD; not evaluable, NE) are also indicated. The last column (LST) shows the LST score for each CTC. [Adapted from Reduzzi C. et al., 2019]

To evaluate if CTCs’ chromosomal instability could predict patients’ response to treatment, we compared the LST scores of CTCs collected at BL from patients responding (best response = PR or SD) vs. not responding (best response = PD) to treatment (Figure 4.20). CTCs collected from responding patients showed an overall slightly lower
LST score than CTCs collected from non-responding patients, with median LST scores of 8.5 vs. 12.5, respectively.

Figure 4.20. CTCs' chromosomal instability in responding and non-responding patients.

Median LST scores in CTCs collected at BL from 6 patients undergoing progression (non-responding, noResp, n = 18 CTCs) and 13 patients achieving a response (resp, n = 28 CTCs). The median LST scores are 12.5 and 8.5 for noResp and resp CTCs, respectively (p = 0.31, Tukey test). [Reduzzi C. et al., 2019]

4.2.5.2.2. Investigation of alterations involved in therapeutic resistance

To further explore the informative content of the CNAs detected in the CTCs, we performed a clustering analysis of the CNA profiles obtained from the 88 collected CTCs (Figure 4.21). The CTCs clustered into 4 clusters. Interestingly, patient individuality did not play a major role in the grouping of the CTCs. In fact, even though this analysis was affected by a non-homogeneous distribution of CTCs across patients, among the 9 patients presenting at least 4 CTCs, the clustering of CTCs was driven by patient individuality for only 5 patients.
Figure 4.21. Clustering analysis of CNA profiles of single CTCs isolated from BTC patients.

By hierarchical clustering it is possible to distinguish 4 clusters of CTCs (1-4, right side). Reference bars on the left indicate anatomical location of the tumor (intrahepatic, I; extrahepatic, E; gallbladder, GB), timing of
Results

blood collection (baseline, BL; during treatment, DT; end of treatment, EOT; follow-up, FU) and treatment-line best response (disease progression, PD; partial response/stable disease, PR/SD; not evaluable, NE). Numbers in brackets indicate the total number of CTCs within each category. [Reduzzi C. et al., 2019]

The anatomical location of the primary tumor and the blood collection time also did not drive CTC clustering ($p = 0.090$ and 0.216, respectively). Conversely, an association between patient best response and grouping of CTCs was observed, with 87% of CTCs collected from patients who did not respond to treatment (best response = PD) falling into clusters 3 and 4 ($p = 0.00041$). These results suggest that clusters 3 and 4 could be enriched in specific genomic alterations associated with treatment resistance.

To test this hypothesis we performed a pairwise comparison of the CNAs among the four clusters. We identified 2 chromosome regions that were more frequently altered in CTCs included in cluster 2 than in cluster 3, which were enriched in CTCs from responders and non-responders, respectively (Figure 4.22). These results suggest that the identified regions could include genes involved in mechanisms of resistance to therapy. However, the data are still preliminary and do not allow us to draw hypotheses on any specific gene.

![Figure 4.22. Comparison of CNA frequency in 2 specific chromosome regions, detected in CTCs from clusters 2 and 3.](image)

For each chromosome region (3p11.1 and 10q22.2) it is reported the frequency of CNAs (either gains, in red; losses, in blue; and no alteration, in white) in CTCs belonging to cluster 2 (left, enriched in CTCs from...
4.2.5.3. Double-negative cells in healthy donors

Not all the double-negative cells that we collected from patients were confirmed as ncCTCs by CNA analysis, indicating that a small proportion of non-malignant blood cells can present a low expression of common WBCs markers. To verify if this phenomenon also occurs in healthy individuals, 3 blood samples from 3 different HDs were processed with our workflow. In the samples from HD 1, 2 and 3, we detected 3, 9 and 13 double-negative cells, respectively, indicating that this type of cell is also present in individuals without a clinically diagnosed tumor.

To validate our definition of ncCTCs, CNA analysis was performed on 12 double-negative cells derived from the HDs. In particular, 3, 4 and 5 double-negative cells collected from HD 1, 2 and 3, respectively, underwent CNA analysis, together with WBCs (collected as single cells and pools) from each HD. Unexpectedly, while all cells collected from HD 2 and 3 (9 double-negative cells, 3 single WBCs and 1 pool of WBCs) showed a flat CNA profile, one double-negative cell from HD 1 presented CNAs. However, one pool of WBCs from the same donor also presented an altered CNA profile, suggesting that the alterations could be due to a particular condition of this HD, or to a technical problem during WGA or CNA analysis. While more investigations about this HD are ongoing, for future samples, we decided to analyze 1 WBC pool for each patient as well, to be used as a control for both possibilities: technical biases and presence of CNAs in WBCs.
4.3. **Exploratory studies**

4.3.1. **Analysis of CTCs’ RNA**

During the development of the workflow, we thought to use epithelial and mesenchymal markers to phenotypically characterize CTCs, but when we analyzed clinical samples we realized that this was not appropriate, since 32% of the identified ncCTCs did not express VIM. Unfortunately, the fluorescence microscope of the DEPArray™ platform only has 4 fluorescence channels, which were already dedicated to nuclei, epithelial markers, WBC markers and VIM detection, so we could not evaluate the expression of any additional markers through staining. To overcome this limitation and to be able to assess the expression of other tumor markers, we aimed at developing a parallel approach for mRNA analysis. Starting from a double volume of blood (20 mL instead of 10 mL), the samples were processed with the Parsortix®. Subsequently, the enriched cells were split in half: The first aliquot was processed with the developed workflow for eCTC and ncCTCs assessment; the second aliquot was lysed and subjected to RNA extraction for the detection of transcripts of interest by real-time PCR.

To test the feasibility and the sensitivity of this approach, different numbers of tumor cells from two BTC cell lines (EGI-1 and HuCCT-1) were spiked into WBCs collected from one HD blood sample by density gradient centrifugation (OncoQuick®), to resemble a sample obtained after Parsortix® processing (few CTCs in around 2,000 WBCs). In particular, for each cell line, 4 aliquots of 4,000 WBCs were spiked-in with 0, 10, 20 and 100 tumor cells (aliquots 1 to 4, respectively). Immediately after the spike-in, the aliquots were lysed and subsequently underwent RNA extraction. The RNA extracted from each aliquot was eluted in 20 µL of water and then split into 2 parts of 10 µL each (aliquots 1A-B, 2A-B, 3A-B and 4A-B). Therefore, for each cell line we obtained 8 RNA samples approximately containing the RNA of 2,000 WBCs and of different numbers of tumor cells: 0 (aliquot 1A-B), 5 (aliquot 2A-B), 10 (aliquot 3A-B) and 50 (aliquot 4A-B). The RNA extracted
Results

from aliquots 1A, 2A, 3A and 4A (for both cell lines) was retro-transcribed to cDNA and analyzed by real-time PCR (in technical duplicates) for the detection of epithelial/tumor transcripts (EPCAM and cytokeratin-8, KRT8), and of CD45 and ACTB as internal controls (Figure 4.23). EPCAM and KRT8 were not detected in aliquots 1A, which contained 0 tumor cells, but were instead detected in all of the other samples at different amplification cycles, consistent with the number of tumor cells (at higher cycles for aliquots 2A, containing 5 tumor cells, and at lower cycles for aliquots 4A, containing 50 tumor cells). CD45 and ACTB were detected in all samples at very similar amplification cycles, indicating the technical reliability of the method.

The RNA extracted from aliquots 1B, 2B, 3B and 4B (for both cell lines) was processed with a very similar protocol, but in this case a pre-amplification step of the target transcripts (EPCAM, KRT8 and CD45) was added before the real-time PCR. By increasing the RNA amount through the pre-amplification, it is in fact possible to analyze the presence of up to 16 gene transcripts, compared to the 4 to 5 genes for the protocol without pre-amplification, allowing for a much deeper characterization of the cells. In samples containing HuCCT-1 cells, the pre-amplification did not affect EPCAM and KRT8 detection, nor did it affect the discrimination between samples containing 5, 10 or 50 cells (Figure 4.24 A, B). The detection of CD45 also remained very homogeneous across samples, notwithstanding the pre-amplification (Figure 4.24 C). Similar results, for both protocols, were obtained with EGI-1 cells (although with a lower discrimination capability between samples containing 5 and 10 cells).
Figure 4.23. Detection of HuCCT-1 cells in HD WBCs by RNA analysis.

Amplification plots showing the detection of EPCAM (A), KRT8 (B), ACTB and CD45 (C) in 4 aliquots of HD WBCs spiked-in with 50, 10, 5, and 0 HuCCT-1 cells (samples were analyzed in technical duplicates). The y-axis shows the fluorescence detected by the instrument (ΔRn = normalized reporter signal minus background fluorescence); the x-axis shows the number of amplification cycles. The horizontal lines are the detection thresholds calculated automatically by the instrument.
Figure 4.24. Detection of HuCCT-1 cells in HD WBCs by RNA analysis after transcripts’ pre-amplification.

Amplification plots showing the detection of EPCAM (A), KRT8 (B) and CD45 (C) in 4 aliquots of HD WBCs spiked-in with 50, 10, 5, and 0 HuCCT-1 cells (samples were analyzed in technical triplicates). The y-axis shows the fluorescence detected by the instrument ($\Delta$Rn = normalized reporter signal minus background
Results

fluorescence); the x-axis shows the number of amplification cycles. The horizontal lines are the detection thresholds calculated automatically by the instrument.

These results suggest that, by using the developed approach, it is possible to detect tumor cell transcripts with a sensitivity of 5 CTCs per 10 mL of blood. New experiments are ongoing to test the detection of the transcripts of interest in HD samples, in order to set reliable detection thresholds, before starting to analyze clinical samples.

4.3.2. Characterization of dual-positive cells

During the analysis of BTC patients’ blood samples for the detection of eCTCs and ncCTCs, we noticed the presence of cells presenting the expression of both epithelial and WBC markers (Figure 4.25). We defined these as dual-positive cells (DPcells). Since these cells do not meet the criteria for CTC identification (because they express WBC markers) we had initially excluded them from our analyses.

![Figure 4.25. DP cells detected in BTC patients.](image)

Three DP cells (rows) collected from the blood of a patient with BTC are reported. The cells show the nuclear staining (DAPI channel), a round morphology (brightfield) and the simultaneous expression of epithelial markers (PE channel) and of leukocyte markers (APC channel). [Adapted from Reduzzi C. et al., 2020]
However, considering that some of the DPcells seemed markedly positive for epithelial markers, we decided to focus our attention on them as well, and to ascertain their tumor origin by exploiting CNA profiling (as already performed for double-negative cells).

Thirty-five single-DPcells were isolated from 10 blood samples collected from 9 BTC patients, and were subjected to WGA and CNA analysis. Fifteen DPcells (43%) showed an altered CNA profile, indicating them as tumor cells, and were thus defined as DP-CTCs.

The number of DP-CTCs and normal DPcells identified for each blood sample (Figure 4.26) suggests that DP-CTCs are found more frequently in patients presenting high numbers of DPcells. In fact, among the 10 considered samples, only the 3 samples presenting only 1 DPcell did not contain any DP-CTCs.

![Figure 4.26. DPcells and DP-CTCs distribution in clinical samples.](image)

The number of DPcells detected and analyzed by CNA profiling and the number of aberrant DP-CTCs are reported for 10 samples collected from 9 BTC patients.

To further investigate the clinical relevance of DPcells, we wanted to correlate their presence with patient clinical outcome. However, our patient cohort (considering patients with DPcells evaluation at BL) was too small. To rapidly perform DPcell counting in a larger cohort of patients, we thought to use data derived from samples processed with the
CellSearch®. In fact, even though DPcells are excluded by the CellSearch® identification criteria, they are detected by the instrument (because they express epithelial markers) and it is possible to count them by reviewing old image galleries of processed samples, without the need for collecting new ones. Unfortunately, we do not have a CellSearch® in our institute and very few groups have used it to analyze blood samples from patients with BTC.

While we are still trying to obtain CellSearch® image galleries from other investigators working on BTC, thanks to our collaboration with the laboratory of Prof. Massimo Cristofanilli at the Northwestern University of Chicago (where I spent 6 months as visiting scholar during the final year of my PhD), we obtained some interesting preliminary data on DPcells’ clinical significance in a cohort of patients with metastatic breast cancer. By reviewing the images of samples that were collected from 82 patients and processed with the CellSearch®, we observed an association between DPcell presence (positivity threshold = 1 DPcell per sample) and shorter PFS: median PFS = 4.27 vs. 6.70 months for DPcell-positive (n = 40) vs. DPcell-negative (n = 42) patients, respectively ($p = 0.046$, log-rank test). Interestingly, by combining CTC- with DPcell-stratification (Figure 4.27), we observed that among patients presenting < 5 CTCs (CTC-negative patients, n = 51), those that were DPcell-positive (n = 21) experienced a shorter PFS with respect to DPcell-negative ones (median PFS = 5.67 vs. 9.33 months, $p = 0.041$, log-rank test). Conversely, no difference was observed between median PFS of CTC-positive/DPcell-negative and CTC-positive/DPcell-positive patients: 3.57 and 3.63 months, respectively. These data, though preliminary, support the clinical relevance of DPcells.
Figure 4.27. Combined DPcell- and CTC- stratification in association with progression-free survival.

Kaplan-Meier’s analysis of PFS according to positivity for CTCs and DPcells for 82 patients with metastatic breast cancer. Positivity thresholds = 5 cells/7.5 mL of blood for CTCs and 1 cell/7.5 mL of blood for DPcells.
5. DISCUSSION
BTC is a deadly disease with limited therapeutic options. Although it is possible to perform surgery with curative intent in the early-stage disease, most patients are diagnosed with advanced BTC, when the only option is systemic chemotherapy. The standard first-line treatment for advanced BTC is cis/gem, which has limited efficacy, and there is no standard second-line treatment.

Molecular profiling has not only revealed that BTCs carry multiple potentially actionable mutations, but has also shown a great heterogeneity between different anatomical subtypes (i.e. intrahepatic CCA, extrahepatic CCA and GBC). These results support the possibility of personalized therapeutic approaches in BTC patients, but also highlight the need of performing the molecular analysis of the tumors, which is not feasible for many patients due to tumor tissue inaccessibility. Moreover, the observed intra-tumor heterogeneity can represent another obstacle in the implementation of personalized medicine in BTC, since a single tissue biopsy may not recapitulate the overall tumor heterogeneity.

The analysis of CTCs at the single-cell level therefore represents a valuable option in BTC, offering an alternative and easy-to-get source of tumor material, and allowing for a better understanding of intra-tumor subclonal composition. Moreover, since CTCs can be repeatedly assessed over time through simple blood draws, they can be used to monitor disease evolution in response to treatment.

So far, no study has performed single-cell molecular profiling of CTCs in BTC. In fact, the few studies that have analyzed CTCs in BTC patients were focused on their enumeration. Moreover, these studies detected only CTCs expressing epithelial markers, without considering other non-epithelial CTC subpopulations.

In this thesis work, we developed a method to isolate both eCTCs (expressing epithelial markers) and bona fide ncCTCs (lacking epithelial and blood cell markers but presenting an aberrant genome), allowing for their molecular characterization at the single-cell level. Spiking experiments performed using different cell lines and that tested each step of the
workflow, showed the analytical validity of the method. Moreover, when the method was applied to clinical samples from patients with BTC, we observed that the presence of at least 1 eCTC at baseline was associated with shorter DSS, confirming the already reported prognostic significance of eCTCs in BTC [Valle J.W. et al., 2015; Yang J.D. et al., 2016] and thus supporting the clinical validity of our workflow.

Nonetheless, by also considering ncCTCs, an overall 4.3-fold increase in CTC-positivity was attained, going from 19% (considering only eCTCs) to 83% (also considering ncCTCs). These results show that eCTCs are only a small subpopulation of CTCs in BTC patients and highlight the importance of using detection methods that are not solely based on epithelial markers’ expression. By using our protocol, we also obtained similar results in patients with renal cell carcinoma [Cappelletti V. et al., 2020], another cancer type for which a very low CTC-detection rate (16%) was reported when using epithelial marker-based methods [Gradilone A. et al., 2011]. Thus, our marker-independent approach can increase the capability of detecting CTCs in other types of carcinoma as well.

In our cohort of BTC patients, the increase in CTC detection obtained using our protocol, and especially the fact that we could detect CTCs in 100% of untreated patients at BL, support the feasibility of using CTCs as a surrogate for tissue material to perform mutational profiling for identifying new therapeutic targets. However, the mutational analysis results showed a high level of intra-patient heterogeneity, with most mutations being present in only 1 CTC. This finding, which was also reported by other studies performing single-CTC mutational profiling in different malignancies [Lohr J.G. et al., 2014; De Luca F. et al., 2016; Paoletti C. et al., 2018], severely limits the reliability and the interpretation of the obtained data. This limitation is connected to the WGA step, which can produce errors during the DNA replication. Thus, a mutation detected in a single CTC after WGA can be considered truly reliable only if it is also detected in another sample from the same patient, such as in another CTC or in a tumor tissue biopsy. Unfortunately,
in our cohort of BTC patients, the low number of CTCs detected per sample, the great heterogeneity among CTCs, and the inaccessibility of tumor tissue hindered the implementation of this type of analysis. Nonetheless, our results support the technical feasibility of the method, which could be applied, in the future, to patients presenting high numbers of detected CTCs. A possible way to expand the applicability of this approach to a larger number of patients would be to increase the number of detected CTCs by processing a higher volume of blood. Another solution could be to avoid amplification artifacts by improving DNA analysis techniques by using, for example, third-generation sequencing methods which do not require WGA [Heather J.M. and Chain B., 2016; Ericson N.G. et al., 2019].

Whereas the developed approach still presents limitations for the study of subclonal heterogeneity by mutational profiling of single CTCs, it did allow us to detect CTCs during treatment and to assess whether eCTCs and ncCTCs could give information about patient response. Until today, only one study [Backen A.C. et al., 2018] has investigated the potential of CTC assessment for treatment monitoring in BTC. That study, which only evaluated eCTCs, did not observe a role of CTCs in monitoring treatment response. Here, instead, we could also analyze ncCTCs, and we observed a decrease in the proportion of eCTCs during treatment as compared to baseline, accompanied by an increase in the proportion of ncCTCs. This supports the importance of considering the ncCTC subpopulation, especially for analyses performed during treatment, and could be an explanation for the negative results obtained by Backen and colleagues in monitoring treatment response with CTCs. In our cohort of 8 patients with available samples collected during therapy, we indeed observed that fluctuations of ncCTC numbers mirrored treatment outcome, and in 2 cases also anticipated the clinical assessment of disease progression. Although we recognize that these are preliminary results with limitations due to the small size of the patient cohort, we think that they suggest a role of ncCTCs in
treatment resistance. In particular, we assume that a decrease or the disappearance of ncCTCs during treatment are associated with therapeutic response, whereas ncCTCs’ increase is associated with resistance and disease progression.

In BTC, there is the need for reliable biomarkers helping treatment monitoring. Currently, the serum marker Ca19.9, which is mostly used as a diagnostic marker, is also considered for treatment monitoring but it lacks specificity and falls far from the ideal standards for biomarkers as molecules whose dynamic expression varies according to response to treatment [Saengboonmee C. et al., 2018]. Here, due to poor overlapping in blood drawing time points, we have not performed a direct comparison between Ca19.9 and ncCTCs for treatment monitoring. However, in a limited series of cases where CA19.9 and CTC detection timings were close, we could observe that ncCTCs were more reliable for treatment monitoring than Ca19.9.

Based on the observed relation between ncCTCs and resistance, and the known association between EMT and cancer progression [Lee J.M. et al., 2006], we hypothesized that ncCTCs are CTCs that underwent EMT, losing, through this process, the expression of epithelial markers. Actually, other studies have reported the presence of mesenchymal CTCs in different types of carcinomas and their association with treatment resistance or worse clinical outcome [Yu M. et al., 2013; Satelli A. et al., 2015a; Satelli A. et al., 2015b; Xu L. et al., 2017], although mainly through CTCs’ phenotypic evaluation. Here, the malignant nature of all ncCTCs was instead confirmed by CNA profiling, and VIM expression analysis demonstrated the mesenchymal phenotype of 68% of ncCTCs, thus supporting our hypothesis for the majority of ncCTCs.

Surprisingly, all evaluable eCTCs showed VIM expression as well. This finding further supports the importance of EMT in BTC dissemination, suggesting that eCTCs had also undergone EMT, although without completely switching to a mesenchymal phenotype, but rather assuming a mixed epithelial/mesenchymal one. CTCs with a mixed phenotype have
been reported in other malignancies, and have also been associated with poor survival in metastatic breast cancer [Bulfoni M. et al., 2016] and with the presence of metastases in prostate cancer [Xu L. et al., 2017]. The clinical relevance of CTCs with a mixed phenotype is in keeping with the results obtained in animal models, in which cancer cells arrested in a mesenchymal state have proved to be highly invasive and able to enter into the bloodstream but unable to form metastasis in distant organs, a capability that was dependent on the re-acquisition of an epithelial phenotype through the reverse EMT process [Tsuji T. et al., 2008; Ocaña O.H. et al., 2012; Tsai J.H. et al., 2012; Brabletz T., 2012]. These findings support the importance of plasticity in cancer progression and underline the importance of studying CTCs with a mixed phenotype (which are assumed to be the ones with the highest plasticity) in cancer patients. Nonetheless, most studies on CTCs, especially the first generation studies, only looked at epithelial markers, thus potentially underestimating the prevalence of CTCs with a mixed phenotype within the epithelial subset. This type of CTCs could, in fact, be more frequent than originally supposed, as suggested by our results as well as by two studies analyzing CTCs in patients with non-small-cell lung cancer. In fact, one study [Morrow C.J. et al., 2016] reported that, in one considered patient, CTCs with a mixed phenotype were the most frequent subpopulation (47%) as compared to CTCs expressing only epithelial (23%) or mesenchymal (30%) markers, and in the other study [Lecharpentier A. et al., 2011] all the CTCs that expressed CK (detected in six patients) also expressed VIM. These results, together with our findings, although obtained from a limited number of patients, definitely suggest that further investigations on the prevalence and the clinical relevance of CTCs with a mixed epithelial/mesenchymal phenotype in cancer patients, are required in order to better understand the role of plasticity in cancer progression.

Notwithstanding the observed expression of VIM in all eCTCs and in 68% of ncCTCs, which supports the role of EMT in BTC, there was a proportion of ncCTCs that showed no
expression of VIM, indicating the presence of heterogeneous subpopulations within ncCTCs. The sole evaluation of VIM expression, as performed in this study, is therefore not sufficient to fully characterize ncCTCs, and a deeper analysis of ncCTCs’ phenotype is needed to better understand the features of the different ncCTC subpopulations and the role of these cells in BTC patients.

The optimal solution to achieve a comprehensive characterization of ncCTCs would be to perform a combined DNA and RNA analysis of single ncCTCs. Although in 2002 Klein and colleagues had already reported the possibility of analyzing the genome and the transcriptome of the same single cell [Klein C.A. et al., 2002] and, more recently, other groups have developed methods for simultaneous DNA and RNA analysis from single cells [Magbanua M.J.M. et al., 2018; Kong S.L. et al., 2019], this goal still remains extremely technically challenging, and more studies are required before validated kits will become commercially available.

As an alternative way to obtain insights on ncCTCs’ gene expression, we developed an approach to detect tumor transcripts in cells collected after the enrichment step. This method, applied in parallel to the workflow for single-CTC isolation and DNA profiling, will give hints on genes expressed by the detected CTCs. Nonetheless, we recognized the limits of this approach since it performs a pooled analysis of cells, not allowing for the discrimination of differences among cells. Moreover, it is not applied to the exact same blood sample undergoing single-CTC profiling, but to a second blood sample. Although both samples are collected at the same time, they might not be identical with regards to the CTC content.

Another weakness of the phenotypic selection of putative ncCTCs performed in this study is its low specificity. In fact, in order to detect 73 ncCTCs in our cohort, we had to collect and analyze 437 double-negative cells. This makes our protocol very expensive and time consuming, definitely not applicable to clinical practice. Unfortunately, in the analyzed
Discussion

Cells, we could not identify, by eye inspection, any morphological or nuclear feature that would distinguish between ncCTCs and normal double-negative cells. The use of the tumor-specific mesenchymal marker CSV, which could potentially help in increasing the specificity of the selection process of putative ncCTCs, is currently under investigation. Nonetheless, based on the observed presence of ncCTCs not expressing VIM, we can hypothesize that ncCTCs negative for CSV exist as well, and their identification will not be improved through the use of CSV. Therefore, after expanding our case series, all the data regarding double-negative cells which are produced by DEPArray™ processing (including morphological and nuclear staining features such as diameter, circularity, staining mean intensity and integral intensity) will be analyzed in comparison with the results of CNA profiling by exploiting deep learning algorithms to extract biophysical features related to malignancy [Chen C.L. et al., 2016]. This will potentially allow for the definition of criteria to better distinguish between aberrant and normal double-negative cells during DEPArray™ analysis, to hopefully avoid the isolation and the molecular characterization of a high number of normal double-negative cells.

CNA profiling of single CTCs not only provides evidence for CTC malignancy but also offers the opportunity to explore intra-patient tumor heterogeneity to collect information on tumor evolution. Here, this was used to investigate correlations between genomic instability and clinical outcome and to identify specific genome alterations associated with treatment resistance.

Genomic instability is a hallmark of cancer [Hanahan D. and Weinberg R.A., 2011] and it has been associated with poor prognosis in cancer patients [Carter S.L. et al., 2006; Walther A. et al., 2008]. Its role is sometimes controversial, though, going from improving cancer cells’ biological fitness to causing detrimental effects with an impact on cancer cells’ viability [Birkbak N.J. et al., 2011]. A precise measurement of chromosomal instability by performing bulk tissue biopsy analyses is difficult, due to intra-tumor
heterogeneity and to the presence of contaminating normal cells. Thus, the assessment of chromosomal instability in single CTCs could provide new insights on its role in tumor progression and therapeutic resistance. Here, we used the LST score [Greene S.B. et al., 2016] as a surrogate to measure chromosomal instability of single CTCs. Our data demonstrate the technical feasibility of this approach and suggest the presence of differences in LST scores in CTCs collected at baseline from patients responding vs. non-responding to treatment, supporting a predictive role of the LST score. However, the data produced here are still not enough to allow drawing any conclusions. Studies on larger cohorts of patients are needed to confirm our preliminary findings and to unravel the role of chromosomal instability in tumors, in the complex balance between improved fitness/therapeutic resistance and decreased viability.

CNA profiling data were also used to investigate differences and similarities across single CTCs collected from 23 patients, at different time points during treatment. By unsupervised clustering of the CNA profiles of 88 single CTCs, different degrees of intra-patient variability were observed. Similar results have already been reported in a study on 14 patients with metastatic castration-resistant prostate cancer where CNA profiles of 185 single CTCs were analyzed [Lambros M.B. et al., 2018]. Also in this study CTCs’ intra-patient heterogeneity varied across patients, with some patients presenting homogeneous CTCs, and others presenting CTCs with very different CNA profiles. Moreover, Lambros and colleagues observed that, for two patients from whom multiple tissue biopsies collected at different time points were available, some CTCs were more similar to the primary tumor and others were more similar to the metastatic biopsies, thus recapitulating the evolution of the disease. Unfortunately, this type of analysis could not be performed in our study, since tissue biopsies adequate for molecular studies were not available. We could instead rely on multiple blood draws from the same patient, which neither suggested a common evolution during treatment nor any effect linked to patients’
individuality. Future studies also including tissue biopsies will potentially help to better understand intra-tumor heterogeneity and tumor evolution in BTC.

Furthermore, the clustering analysis revealed that CNA profiles of CTCs collected from patients responding to therapy were different from those of CTCs derived from non-responding patients, possibly suggesting an association between CNAs and resistance mechanisms. Actually, in small-cell lung cancer such an association has already been proposed by Carter and colleagues [Carter L. et al., 2017], who developed a classifier, based on CNAs of single CTCs, able to predict which patients will initially respond or not to first-line chemotherapy (i.e. chemosensitive and chemorefractory patients, respectively), by analyzing CTCs collected before the start of treatment.

Collectively, our results and those obtained by Carter and colleagues support the hypothesis that CTC genomic analysis can be used for a timely recognition of patients who are more likely to not respond to treatment, and also suggest the potential of identifying, by CTCs’ CNA profile analysis, genome regions containing genes that might be involved in resistance mechanisms. Here, by comparing CNAs detected in CTCs included in the 4 clusters obtained by clustering analysis, we identified 2 chromosomal regions that were differently altered in cluster 2 (enriched in CTCs from responders) and in cluster 3 (enriched in CTCs from non-responders). Among the genes included in these regions is EPHA3, a gene encoding for the ephrin type-A receptor 3 (EphA3) which is a receptor tyrosine kinase belonging to the ephrin receptor family, a group of receptors reported to have a complex role in cancer, both favoring and hindering it [Pasquale E.B., 2010]. Mutations in EPHA3 have been reported in ampulla of Vater cancer and pancreatic ductal adenocarcinoma [Corbo V. et al., 2010], colorectal cancer [Bardelli A. et al., 2003] and lung cancer [Davies H. et al., 2005], although their functional relevance is not fully understood. Nonetheless, EphA3 offers the advantage of being druggable and, in fact, an anti-EphA3 monoclonal antibody has already been tested in patients with hematological
malignancies where EphA3 is frequently overexpressed [Swords R.T. et al., 2016]. EphA3 represents, therefore, an interesting gene, possibly involved in therapeutic resistance in BTC and with the potential of becoming a new treatment target; worthy of being further investigated in future pre-clinical studies.

Another interesting result obtained in this thesis study is the detection of DP-CTCs, which are circulating cells expressing both epithelial and WBC markers, presenting an aberrant genome. Circulating cells with a similar phenotype have also been reported in patients with other cancer types [reviewed in Sutton T.L. et al., 2019; and in Reduzzi C. et al., 2020], and a recent paper suggested that they are hybrids deriving from the fusion of tumor cells with macrophages [Gast C.E. et al., 2018]. Thus, the proof of the malignancy of DP-CTCs that we obtained in BTC, but also in other cancer types [Reduzzi C et al., 2020], paves the way for new studies focused on the role of heterotypic cell fusion in cancer.

This finding also highlighted a crucial issue in CTC studies, which is CTC definition. Originally, CTCs have been defined based on their phenotype. More recently, however, the observed phenotypic heterogeneity of CTCs suggested that we should favor a genome-based definition of CTCs: circulating cells presenting various phenotypes and genomic aberrations. Here, for instance, we showed that not all cells lacking WBC markers were aberrant and, on the other hand, we detected CTCs positive for WBC markers (i.e. DP-CTCs). Moreover, recent studies, performing CNA profiling of single CTCs identified using conventional criteria (circulating cells positive for epithelial markers and lacking CD45), demonstrated the presence of “CTCs” with flat profiles, which are probably normal epithelial cells that had entered the bloodstream [Lambros M.B. et al., 2018; Chemi F. et al., 2019]. This evidence clearly shows the limits of a definition based on the phenotype and emphasizes the importance of evaluating genomic aberrations to identify CTCs, since, in the circulation, it is possible to find CTCs with unconventional and unexpected phenotypic characteristics but aberrant genomes, and vice versa.
On the other hand, here we showed that it also is possible to detect aberrations in WBCs (probably due to technical artifacts related to WGA and NGS analysis). Moreover, we cannot forget that the purely phenotypic identification of CTCs, performed by the CellSearch® platform, proved to be clinically valid.

It still remains an open question, therefore, which definition of CTC is the best and the most clinically relevant one. Only studies performing a multilevel (phenotypic and genotypic) characterization of a high number of single circulating cells will eventually solve the dilemma between phenotype and genotype in CTC identification and analysis.
6. CONCLUSIONS & FUTURE PERSPECTIVES
CTCs hold promise in BTC clinical management. Although the results presented here are preliminary and need confirmation in larger cohorts of patients, they suggest that, by using approaches allowing the detection of both epithelial and non-epithelial CTCs, it is possible to use CTCs to i) identify patients with poor prognosis, ii) monitor patients’ response to treatment, and iii) unravel treatment resistance mechanisms, thus discovering new putative treatment targets to be further validated.

This study also highlighted some limitations of the current workflow that need to be improved. In particular, i) mutational profiling is currently not feasible for all patients and needs to be improved by processing a larger volume of blood or by using innovative WGA techniques, ii) ncCTCs do not constitute a uniform subpopulation of CTCs and require further characterization, and iii) the number of normal double-negative cells that were selected as putative ncCTCs is too high and prevents the implementation of the current workflow in clinical practice, thus, an optimization of the selection process is needed.

Overall, the results obtained here lay the foundation for future studies to:

- confirm the role of eCTCs and ncCTCs in a larger cohort of patients;
- better characterize ncCTCs phenotype by using the developed RNA analysis assay;
- test the use of CSV as a more specific marker for mesenchymal CTCs;
- assess the role of EphA3 in resistance to chemotherapy through *in vitro*/*in vivo* studies using BTC cell lines;
- investigate the clinical relevance of DPcells in BTC and in other malignancies.

Moreover, for all patients included in this study, plasma samples were collected to allow the analysis of cell-free DNA. The analysis is ongoing and the results will be combined to CTC data.
Finally, in this study we focused on CTCs circulating as single cells, but it is well known that CTCs can circulate also as clusters (which were reported to have a higher metastatic potential than single CTCs [Aceto N. et al., 2014]). Currently, no data on the prevalence of CTC-clusters in BTC are available. Therefore, new approaches will be implemented to analyze the presence and the clinical relevance of CTC-clusters in BTC patients.
7. REFERENCES


References


References


Global Burden of Disease Cancer Collaboration. (2019) Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017:


References


with locally advanced / metastatic biliary tract cancers (ABC) previously-treated with cisplatin/gemcitabine (CisGem) chemotherapy. Journal of Clinical Oncology 37:15_suppl, 4003-4003.


References


PUBLICATIONS
Publications related to my PhD project


Other publications


- Appierto V, Di Cosimo S, **Reduzzi C,** Pala V, Cappelletti V, Daidone MG. (2017) *How to study and overcome tumor heterogeneity with circulating biomarkers: The


DECLARATION OF AUTHORSHIP
I hereby certify that the thesis I am submitting is entirely my own original work, except for the following parts:

- Recruitment and clinical assessment of patients was performed by Monica Niger (Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy);
- Flow cytometry analysis and AutoMACS® Pro separator sample processing were carried out in collaboration with Giulia Bertolini (Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy);
- Mutational profiling of single CTCs (including DNA sequencing and bioinformatic analysis) was performed by Menarini Silicon Biosystems;
- Low-pass whole genome sequencing for CNA parofiling was carried out by the institutional Integrated Biology and Bioinformatics Platform Unit;
- Bioinformatic analyses were performed by Marco Silvestri (Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy).