

CNA profiling of single CTCs in locally advanced esophageal cancer patients during therapy highlights unexplored molecular pathways.

Giulia Gallerani (✉ giulia.gallerani@irst.emr.it)

IRCCS Istituto Romagnolo per lo Studio dei Tumori (IRST) "Dino Amadori" <https://orcid.org/0000-0003-0121-9621>

Tania Rossi

IRCCS Istituto Romagnolo per lo Studio dei Tumori (IRST) "Dino Amadori"

Martina Valgiusti

IRCCS Istituto Romagnolo per lo Studio dei Tumori (IRST) "Dino Amadori"

Davide Angeli

IRCCS Istituto Romagnolo per lo Studio dei Tumori (IRST) "Dino Amadori"

Pietro Fici

IRCCS Istituto Romagnolo per lo Studio dei Tumori (IRST) "Dino Amadori"

Sara De Fanti

"Alma Mater Research Institute on Global Challenges and Climate Change (Alma Climate)"

Erika Bandini

Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori

Claudia Cocchi

Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori <https://orcid.org/0000-0002-7736-9039>

Giovanni Luca Frassinetti

IRCCS Istituto Romagnolo per lo Studio dei Tumori (IRST) "Dino Amadori"

Massimiliano Bonafè

University of Bologna

Francesco Fabbri

Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori

Article

Keywords: esophageal cancer, circulating tumor cells, CNA profiling

Posted Date: August 31st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-850357/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Cancers on December 19th, 2021. See the published version at <https://doi.org/10.3390/cancers13246369>.

Abstract

Locally advanced esophageal cancer (EC) is an aggressive disease with a dismal prognosis. The role of circulating tumor cells (CTC) in the EC metastatic process is still an unaddressed question. Here we monitored the evolution of CTCs spread in 11 patients affected by locally advanced EC who were undergoing neoadjuvant chemo-radio therapy followed by surgery. We developed an *ad hoc* assay named 'Grab-all' assay using DEPArray system. Longitudinal monitoring allowed the identification of CTCs expressing epithelial, mesenchymal or mixed phenotype markers, in at least one time-point per patient. Through single cell copy number aberration (CNA) analysis, we revealed that individual CTCs from relapsed patients displayed higher degree of genomic imbalance compared to disease-free patients' ones. Post-neoadjuvant therapy CTCs show genomic aberrations previously reported in EC, namely 23/23 amplifications and 13/19 deletions. Notably the phenomenon was not restricted to the group of relapsed patients. In-depth analysis showed that networks enrichment in all time-points CTCs were inherent to innate immune system. At variance, transcription/gene regulation, post-transcriptional and epigenetic modifications were uniquely affected in CTCs of relapsed patients. Collectively, our data add clues to the comprehension of the role of CTCs in EC aggressiveness: chromosomal aberrations on genes related to innate immune system behave as relevant to the onset of CTC-status, whilst pathways of transcription / gene regulation, post-transcriptional and epigenetic modifications seem linked to patients' outcome.

Introduction

Esophageal cancer (EC) represents a significant burden on human health being the sixth most common cause of cancer-associated death globally ¹ with a 5-year survival rate that is less than 20%, including West and Asian countries ². Even for those patients (> 50%) deemed fit for curative interventions with aggressive (despite non-targeted) neoadjuvant therapy followed by surgery, the outcomes are still poor with a 5-year survival rate of < 45% ³. While EC treatments use a similar criterion, the underlying biological differences of esophageal tumor types are becoming evident ⁴. Furthermore in gut adenocarcinomas widespread genomic disruption due to copy number alterations (CNA) differs across cancers and cancer types ^{5,6}.

EC is notable among cancer types for its high degree of chromosomal instability ⁷ as well as genomic disruption ⁵, which is due to CNA involving whole chromosomes or chromosome arms ⁶⁻⁸.

Unlike the mutational status, CNAs events are restricted to EC and are not present on its precursor lesion, Barrett's esophagus ⁹; this makes CNA analysis highly suitable for the early detection of esophageal cancer cells. Furthermore, CNAs were adopted by Noorani et al to suggest the "diaspora" theory for EC dissemination, their high-resolution perspective study revealed that metastases directly resemble the primary site ¹⁰. In this scenario, the analysis of circulating tumor cells (CTCs) represents a key player to understand EC aggressiveness related to tumor spreading. Previously we and others illustrated the presence of CTCs in non-metastatic cancers, proving that tumor spreading starts in cancer early stages ¹¹⁻¹³.

Esophageal CTCs in the blood of non-metastatic patients were found during multimodal treatment ¹⁴ and their presence was correlated with poor clinical outcome on patients that underwent esophagectomy solely ^{12,15}.

In-depth CTCs analysis of different cancer types, such as breast and lung cancers, provided new insights about tumor spreading as well as identifying CNA signatures on relapsed patients ¹⁶. It is therefore mandatory to investigate CTCs in EC too.

To better understand the role of CTCs in locally advanced non-metastatic EC, we designed a perspective study that necessitated multiple blood biopsies from individual patients drawn at different time points during treatment (neoadjuvant chemo-radiotherapy followed by surgery) (Fig. 1A).

Here we set up a multi-antibodies analysis named Grab-all assay to assess CTCs for the epithelial (EpCAM/E-Cadherin/Cytokeratins) and mesenchymal/stem (N-Cadherin/CD44v6/ABCG2) cell phenotypes. All identified CTCs were sorted as single cells, subjected to whole genome amplifications and low-resolution whole genome sequencing to obtain the CNAs of individual CTCs (Fig. 1B).

Our data indicate that the Grab-all assay improves the identification of CTCs in peripheral blood of non-metastatic EC patients. CTC CNA profiling at the single-cell level in combination with a longitudinal study design highlights differences in the extent of genomic imbalance between CTCs isolated from relapsed patients as opposed to those without evidence of disease, as well as in relation to the time when the liquid biopsy was taken.

Results

Grab-all assay setup

To examine CTCs phenotypes, we combined the Oncoquick marker-independent CTC enrichment^{13,17} with the novel Grab-all assay. This procedure uses an array of antibodies against mesenchymal and epithelial antigens, which come under an Epithelial Tag (E-tag consisting of EpCam, E-cadherin, Keratins 4, 5, 6, 8, 10, 13, 18) or a Stem-Mesenchymal Tag (SM-tag consisting of N-Cadherin, CD44v6, ABCG2), as well as CD45 as a leukocyte marker. Our Grab-all assay was validated on the cell line MDA-MB436, which expresses both epithelial and mesenchymal markers (Fig. 1S). Grab all assay sensitivity was also evaluated in MCF10A cells undergoing TGF-beta1-induced epithelial mesenchymal transition (EMT). In particular, these epithelial mammary cells are devoid of any SM-tag positivity in standard culture conditions, but after 4 days of TGF-beta1 exposure, when morphological changes of EMT began to appear, a decrease of the expression of E-tag became overt (Fig.S2). Moreover, after six days of TGF-beta1 treatment, SM-tag weakly positive cells can be detected, with a continuous increase up to day 10 of treatment. Intermediate cells, with dual epithelial and mesenchymal phenotypes, as well as cells that are negative for both the E-tag and SM-tag were all present at late stages of EMT.

Detection efficiency of Grab-all assay on metastatic esophageal cancer patients

First, we intended to verify the performance of the Grab-all assay on blood of patients with a high disease burden (primary tumor and multiple metastases, both unresected). Purposely, we applied the Grab-all assay on peripheral blood samples from 4 metastatic esophageal cancer patients undergoing clinical treatment (Table 1). CTCs were detected in all blood samples (Table 1 and Fig. 2) and the Grab-all assay identified CTCs expressing E-tag, or SM-tag or bi-phenotypic E- SM-tag.

TABLE 1

Clinical-pathological characteristics of metastatic esophageal cancer blood donors and phenotypic characterization of identified CTCs.

Patient	Sex	Age	Histological Type	Tumor Grade	Metastatic sites	First line chemotherapy	Total Epithelial CTCs	Total Mesenchymal/stem CTCs
M1	M	71	ESCC	G3	Liver	5-Fluorouracil	1	24
M2	M	41	EAC	G3	Liver, lymph nodes, bones, lung, peritoneum	FOLFOX	2	0
M3	M	43	EAC	G3	lymph nodes, brain, bones	FOLFOX	5	2
M4	M	73	EAC	G3	Liver	Cisplatin/ 5-Fluorouracil	7	0

Patient characteristics

Then, the Grab-all assay was employed on peripheral blood samples from a total of 11 patients diagnosed with locally advanced non-metastatic esophageal cancer: 8 (72.7%) adenocarcinomas (EAC), 3 (27.3%) squamous carcinomas. Patient characteristics are summarized in Table 2. All the enrolled patients were treated with neoadjuvant chemo-radiotherapy with surgery intent, detailed treatments are listed in Table 2. Eight out of 11 patients (72.7%) underwent surgery. Three remaining patients (27.3%) did not undergo surgery due to disease progression. All the surgery patients showed a tumor reduction after neoadjuvant therapy as revealed by yp stage compared to c stage.

TABLE 2

Clinical characteristics of enrolled patients affected with non-metastatic esophageal cancer (EC). Histological type: EAC esophageal adenocarcinoma and ESCC esophageal squamous carcinoma. Disease stage at diagnosis as clinical classification (cTNM). Pathological classification (ypTNM) is shown for the patients who underwent surgery. Only patients defined as «stable disease» or responding to therapy by CT/PET scan were eligible for surgery. Post therapy CT scan defined as partial response (PR), stable disease (SD) or not evaluated (NE). N/A as not available.

Patient	Sex	Age	Histological Type	Tumor Grade	Disease stage at diagnosis	Neoadjuvant chemotherapy	Concomitant Radiotherapy	Post-therapy CT	Surgical resection	ypT Stage
NM1	F	62	EAC	G2	cTxN2M0	Cisplatin/ 5-Fluorouracil	48.60 Gy in 27 fractions	PR	No	N/A
NM2	M	72	EAC	N/A	cT3N2M0	Cisplatin/ 5-Fluorouracil	50 Gy in 25 fractions	PR	Yes	T0N0M0
NM3	M	71	EAC	G3	cT2N2M0	Carbotaxol	41.4Gy in 23 fractions	NE	No	N/A
NM4	M	62	EAC	G3	cT4bN2M0	Epirubicin/ Oxaliplatin/ 5-Fluorouracil	50Gy in 25 fractions	SD	Yes	T3N0M0
NM5	M	73	EAC	G2	cT3N1M0	Carbotaxol	45 Gy in 25 fractions	SD	Yes	T0N0M0
NM6	M	45	ESCC	N/A	N/A	Cisplatin/ 5-Fluorouracil	45 Gy in 25 fractions	PR	Yes	T3NxM0
NM7	F	54	ESCC	G2	cTxN2M0	Carbotaxol	45 Gy in 25 fractions	PR	No	N/A
NM8	M	51	EAC	G3	cT4aN3M0	Cisplatin/ 5-Fluorouracil	50.4 Gy in 28 fractions	PR	Yes	T3N0M0
NM9	F	55	EAC	G3	cT3N2M0	Cisplatin/ 5-Fluorouracil	50.4 Gy in 28 fractions	PR	Yes	T1N1M0
NM10	M	76	EAC	G3	cT3N1M0	Carbotaxol	50 Gy in 25 fractions	PR	Yes	T3N1M0
NM16	M	60	ESCC	G2	cT3N2/3M0	Carbotaxol	50.4 Gy in 28 fractions	PR	Yes	T2N3M0

CTCs identification in the blood of non-metastatic patients during treatment

CTCs were detected in every patient (Table 3), albeit expressing distinct epithelial, mesenchymal and mixed epithelial/mesenchymal phenotypes, which were evenly distributed across the timepoints. Notably, CTCs from patient NM4 showed a change in phenotype before and after chemo-radiotherapy: all CTCs disclosed an epithelial phenotype before chemotherapy and turned out to show mixed epithelial/mesenchymal phenotype after treatment (Fig. 3).

TABLE 3

CTC counts during patients' treatment course and clinical outcomes (relapsed or disease free). Blood draws were performed before therapy, one month after the end of neoadjuvant therapy and one month after surgery. The second and the third blood draws were performed concomitantly to CT scan.

CTC count					
Before therapy (A)		After Neoadjuvant therapy (B)		After surgery (C)	
20		27		10	
Relapsed	Disease Free	Relapsed	Disease Free	Relapsed	Disease Free
15	5	13	14	1	9

We found CTCs in six patients (54.5%) before they underwent neoadjuvant therapy (Fig. 4), but we could no longer detect CTCs in three of them (NM5, NM7, NM8) after therapy. All patients underwent surgery, but patient NM7 displayed CTCs even after surgery, and then died due to relapse. Conversely, patients NM5 and NM8 maintained a CTC-negative status and remained disease-free. Another group of CTCs-positive patients before neoadjuvant therapy (NM3, NM4, and NM16) maintained CTC positivity after treatment, with the exception of NM3 who died of sepsis. Despite surgery, patients NM4 and NM16 relapsed and died.

Concerning the five patients (NM1, NM2, NM6, NM9, NM10) who were CTCs-negative at diagnosis, we detected CTCs after neoadjuvant therapy in four of them (NM1, NM6, NM9, NM10). With the exception of NM1, who relapsed and died after neoadjuvant therapy, the remaining three patients underwent surgery but NM6 died from neck metastases before the third blood biopsy. Unexpectedly, we detected CTCs in patients NM2 and NM10 one month after surgery but these patients remain disease free. Finally, NM9, who was CTC-negative, had a disease relapse after two years from surgery.

Copy number alterations in CTCs

Since CTC counts did not provide any information about clinical outcome, we opted for an in-depth single-cell analysis of CNAs of isolated CTCs. We first validated genomic analyses on OE33 esophageal cancer cell lines ex vivo spiked-in peripheral blood samples of healthy donors and stained with Grab-all assay. We sorted single and pooled OE33 cells using the DEPArray platform. Then, we performed WGA on 2 single cells and a pool of 4 cells. Afterwards, samples were quality-tested to verify their genomic integrity. We generated NGS libraries and performed low-coverage whole genome sequencing as previously reported^{18,19}. WGA amplification of the three samples resulted in clear aberrant CNA profiles (Fig.S3 A). Clustering analysis between single and pooled cells showed high levels of concordance (> 95%) (Fig.S3 B). The genomic content of every DEPArray-sorted single CTC from patients was next assessed. All sorted CTCs (n = 57) conveyed an altered CNA profile, supporting further the validity of the Grab-all assay in identifying actual CTCs. In patient NM8 we identified CD45-negative cells without any specific tumor phenotype, subsequent CNA analysis revealed a normal diploid genome indicating that these cells were non-cancer related cells of unknown origin (Fig.S4).

In order to compare and quantitatively assess the genomic imbalance of CTCs relative to a matched diploid genome, we calculated a Jaccard similarity index (JI) defined as the ratio of shared to all aberrations between single CTCs and leucocytes for each patient. JI values range between 0 and 1, with 0 indicating maximal genomic imbalance and 1 indicating complete overlap with a normal diploid genome¹³.

CTC JIs were lower than 0.54 indicating that CTCs share less than 54% of CNA similarity with a normal diploid genome. The distribution of the CTC JIs across cell populations were visualized and compared using violin plots.

Then, we aimed at comparing the genomic imbalance of CTCs isolated from patients at different clinical stages. To this purpose, we plotted the distributions of CNA JIs of CTCs before neoadjuvant treatment, after neoadjuvant therapy and after surgery CTCs (Fig. 5A). CTCs show significantly different values of genomic imbalance across time points (Kruskal-Wallis p-value = 3.1×10^{-6}). After neoadjuvant therapy, CTCs showed a marked change in the profile and peak of the distribution, compatible with reduced genomic imbalance, compared to before treatment CTCs (Wilcoxon p-values of 4.2×10^{-6}) and after surgery CTCs (Wilcoxon p-values of

0.00029). Conversely, genomic imbalance of before treatment CTCs and after surgery CTCs were not statistically different (Wilcoxon p-values of 0.59).

We then examined disease-free and relapsed patients CTCs (Fig. 5B). Marked differences in genomic aberrations were found in the two CTCs populations (Kruskal-Wallis test p value = 1.3×10^{-14}). A large proportion of relapsed patients CTCs show substantial genomic imbalance, with low dispersion of JI values (JI = 0.12) around the median value. At variance, disease-free patient CTCs convey low genomic imbalance and widely dispersed distribution of CTC JI values (JI = 0.19).

Finally, we investigated whether the difference between relapsed and disease-free patients emerged during treatment. We split the data from these two classes of patients in pre- and post- neoadjuvant therapy groups, which were found to be significantly different (Fig. 5C). However, both before and after neoadjuvant therapy CTCs from patients who eventually relapsed were characterized by with higher genomic imbalance compared to CTCs from disease-free patients (Wilcoxon p value 5.4×10^{-11} and 2.2×10^{-10} respectively).

Investigation of previously reported esophageal cancer-specific chromosomal aberration in CTCs

A set of specific chromosomal aberrations, known as focal copy number changes, have been previously reported in esophageal cancer⁵. We searched our CTC whole genome sequencing dataset for a list of 23 chromosomal amplifications and 19 chromosomal deletions (Fig. 6). The aim of this analysis was to characterize CTC CNAs in greater details – against known EC-specific genomic aberrations – and to determine whether focal copy number changes were unevenly across the groups (before neoadjuvant treatment, after neoadjuvant therapy and after surgery) or differed according to the outcome (disease-free versus relapse).

All the known chromosomal amplifications were detected in our dataset, while six out of 19 known deletions were absent (Fig. 6A, C). Genes included in amplified and deleted regions are listed in Fig. 6B and D, respectively.

All chromosomal amplifications were more frequent in CTCs from patients after neoadjuvant treatment, whereas their frequency was reduced in CTCs from patients before neoadjuvant treatment and even more in CTCs after surgery (Fig. 6A). Unexpectedly, focal copy number amplifications were more abundant in CTCs isolated from patients who attained disease-free status, whereas CTCs from patients who eventually relapsed had very few amplifications.

Chromosomal deletions were more frequent in CTCs collected from patients before neoadjuvant therapy compared to CTCs after neoadjuvant therapy or surgery, being almost absent in these latter. Chromosomal deletions were also more frequently observed in CTCs from disease free patients than from relapsed patients (Fig. 6C). Four chromosomal deletions, namely 5q11.2, 4q35.1, 7q36.3, 2q33.3, were exclusively present in CTCs of relapsed patients.

Enrichment analysis:

Finally, to identify functional pathways of CNAs, we performed enrichment and network analysis.

According to GO biological process domain, enriched terms network (adjusted $p < 0.05$) of all CTCs are illustrated in Fig. 7. GO terms are represented as circles (nodes), in which percentage according to the timepoint of the CTCs is reported as a color (red, orange and yellow), the size of the nodes indicates the number of enriched CTCs, and edges (grey lines) represent at least ten common genes overlapping the two terms.

CTCs from all time points shared terms linked to the activation and modulation of innate immune system (yellow), as well as chemotaxis and migration of immune system cells (violet) and cellular response to cytokine stimuli (GO:0071345), thus setting up the main axis of the network. Strikingly, all the above-mentioned terms exhibited aberrant interferon-related genes (Table S1)¹³. Sensory perception of bitter taste (pink) associated terms were also enriched in all timepoints, although disconnected from the main axis (Fig. 7A).

Interestingly, the main axis of the network branches out with terms emerging only on CTCs identified post neoadjuvant therapy, in particular skin development (blue), nervous system development (GO:0007399) and modulation of the transcription and macromolecules biosynthesis (red and green). Of note, enriched terms of response to divalent ions are enriched mostly in after-surgery CTCs (Fig. 7A)

We next sought to investigate differences between enriched networks of CNAs in relapsed patients compared to disease free patients CTCs (Fig. 7B).

The main axis of the network is maintained, with six new emerged terms only in relapsed patients CTCs. In particular three terms (green) are located at the center of the network: positive regulation of transcription, DNA-templated (GO:0045893), positive regulation of transcription from RNA-polymerase II promotor (GO:0045944) and regulation of gene expression (GO:0010468). The other three terms did not share common genes: cytidine to uridine editing (GO:0016554), DNA cytosine deamination (GO:0070383) and DNA cytidine deamination (GO:0070383).

In Table 4, the mostly associated genes of these six terms that have been associated with esophageal cancer or other cancers are reported. In particular, these genes are associated with more aggressive tumor phenotype (SAL2), migration and invasion (TOX4, PRMT5, AJUBA) development and progression (ZNF219, CEBPE). TOX4 and APOBEC3 gene clusters have also been associated with genomic instability²⁰.

TABLE 4

List of mostly associated genes of the six terms distinctive of CTCs from relapsed patients.

GENE	TERM	CTC RELAPSED (n = 29)	CTC DISEASE FREE (n = 28)	REFERENCE
APOBEC3 cluster gene	cytidine to uridine editing (GO:0016554) DNA cytosine deamination (GO:0070383)	24.1%	7.1%	26,45
SALL2	regulation of gene expression (GO:0010468) positive regulation of transcription, DNA-templated (GO:0045893) positive regulation of transcription from RNA polymerase II promoter (GO:0045944)	27.6%	0%	46
TOX4	regulation of gene expression (GO:0010468) positive regulation of transcription, DNA-templated (GO:0045893)	27.6%	0%	47,48
ZNF219	regulation of gene expression (GO:0010468)	27.6%	0.0%	49
PRMT5	regulation of gene expression (GO:0010468)	27.6%	3.6%	50
AJUBA	regulation of gene expression (GO:0010468)	24.1%	3.6%	51
NRL	regulation of gene expression (GO:0010468) positive regulation of transcription, DNA-templated (GO:0045893) positive regulation of transcription from RNA polymerase II promoter (GO:0045944)	24.1%	3.6%	52
PPARGC1A	regulation of gene expression (GO:0010468) positive regulation of transcription, DNA-templated (GO:0045893) positive regulation of transcription from RNA polymerase II promoter (GO:0045944)	41.4%	10.7%	53
MAP2K2	regulation of gene expression (GO:0010468) positive regulation of transcription, DNA-templated (GO:0045893)	27.6%	7.1%	54
BAZ1A	regulation of gene expression (GO:0010468)	24.1%	7.1%	55
DENND4A	regulation of gene expression (GO:0010468)	24.1%	7.1%	56
CHD8	positive regulation of transcription, DNA-templated (GO:0045893) positive regulation of transcription from RNA polymerase II promoter (GO:0045944)	27.6%	0.0%	57
SUPT16H	positive regulation of transcription, DNA-templated (GO:0045893) positive regulation of transcription from RNA polymerase II promoter (GO:0045944)	27.6%	0.0%	
CEBPE	positive regulation of transcription, DNA-templated (GO:0045893) positive regulation of transcription from RNA polymerase II promoter (GO:0045944)	24.1%	3.6%	58

Discussion

In this study we leveraged single cell analysis to investigate the features of CTCs from locally advanced EC patients during clinical management. The presence of these rare tumor cells involved in metastatic cascade have been demonstrated as valid prognostic marker of patients' outcome¹², although the molecular characterization of CTCs is still lacking.

Our findings, albeit conducted on a small cohort, demonstrate that EC spreads epithelial as well as mesenchymal CTCs carrying high burden of genomic imbalance, that entails chromosomal aberrations in innate immune system-related *loci*. The identification and in-depth analysis of chromosomal aberrations in CTCs is the first step towards a more complete understanding of the genetic events that influence CTC's metastatic potential²¹, which will in turn improve our knowledge of EC metastatic cascade.

Through the application of the method we established, named Grab-all assay, we increased the number of identified CTCs with an epithelial phenotype: beside the well-established EpCAM²² and CKs²³ we included E-cadherin^{24,25} as a target. We also opted to include antibodies to identify CTCs with a stem/mesenchymal phenotype using molecular targets as N-Cadherin^{26,27}, CD44v6²⁸⁻³⁰ and ABCG2^{31,32} which allowed to identify both mesenchymal and mixed phenotype CTCs³³. Overall, Grab-all assay maximized the number of CTCs for downstream analysis. Although most of the CTCs had an epithelial phenotype, the presence of CTCs with mixed phenotype and mesenchymal phenotypes strongly suggests the activation of EMT mechanisms in EC spreading³⁴. Due to the small cohort of enrolled patients, it was not possible to associate total CTCs count or CTC phenotype to the clinical outcome. However, by exploiting whole genome copy number analysis on single CTCs, we noticed that the neoadjuvant therapy has a selective action against CTCs, likely promoting a higher genomic imbalance. In fact, post-surgery CTC CNA is similar to that of before therapy CTCs¹³. Considering genomic disruption levels of disease-free CTCs and relapsed patients CTCs, we observed a highly significant difference between the two groups. In particular, the latter CTCs displayed a higher level of CNA than the former, and it was present both before and after therapy. These findings lead us to hypothesize that EC aggressiveness may depend upon a CTC subpopulation with a highly aberrant chromosomal set up that is not affected by neoadjuvant therapy.

Considering that chromosomal amplification is an early event in esophageal cancer tumorigenesis³⁵ and our findings on CTC CNAs, we can therefore assume that chromosomal aberrations located in this sub-population of CTCs may confer tumor aggressiveness. Initially, we focused on the presence of previously-reported esophageal-cancer specific chromosomal aberrations. Interestingly, all 23 chromosomal amplifications are represented, and 13 out of 16 deletions emerged mainly in post-neoadjuvant therapy CTCs. These findings suggest that neo-adjuvant therapy may trigger cellular mechanisms that lead to chromosomal amplification in regions where genes such as GATA6 / 4, MET, CDK6, MYB, PRKC1, MCL1, EGFR are present. These data confirm previous findings on the presence of GATA6 amplification after neoadjuvant therapy on primary tissue of EC³⁶. Copy number abnormalities of MET³⁷, CDK6 and EGFR³⁸ were also observed in esophageal cancer tissues.

Remarkably, the enriched-network generated in all time points CTCs revealed aberrations in pathways related to the innate immune system. Strikingly, we previously identified similar findings in early breast cancer CTCs¹³. Hence, our data underpin a contribution of genome regions harboring genes of the innate immune system (interferon, in particular) in the acquisition of the CTC functional status¹³. In addition, network analysis showed that CTCs survived to neoadjuvant therapy had additional terms of nervous and epidermal system development and transcription modulation, suggesting that these genomic modifications may confer a selective advantage.

By exploring differences between CTCs of relapsed and disease-free patients, six pathways involving transcription / gene regulation, post-transcriptional and epigenetic modifications have emerged. This investigation revealed genes described in the literature, which play roles in esophageal cancer and other types of cancer as TOX4 and APOBEC3 gene cluster related to genomic instability, PRMT5, AJUBA promoting tumor migration and invasion, ZNF219, CEBPE and SAL2 genes associated to tumor progression and aggressiveness, respectively.

In conclusion, our work demonstrates for the first time that CTCs of locally advanced EC have a mixed phenotype with high levels of genomic disruption, particularly those CTCs found in relapsed patients. Furthermore, we have identified a set of aberrant genes common on relapsed patient CTCs that could play a role in the metastatic process and cancer recurrence.

Further experiments and analysis should be directed towards identified gene pathways and probing their likely mechanisms of action, using both in vitro and in vivo models. In-depth CTC investigation in EC holds the potential to increase knowledge of this disease aggressiveness features and pave the way to new and tailored therapeutic strategies.

Methods

Patients recruitment and sample collection

The study had been prospectively approved by the Research Ethics Committee from IRCCS Istituto Romagnolo per lo Studio dei Tumori (IRST) "Dino Amadori" (Italy) (reference, IRSTB017) according to the recommendations of the declaration of Helsinki, patient consented to the project were enrolled at IRCCS-IRST between 2013 and 2017.

Eligible patients were newly diagnosed with metastatic or locally advanced EC with treatment protocols of first line chemotherapy and neoadjuvant chemo-radiotherapy followed by surgery respectively. Patients were staged and managed standard treatment protocols were in accordance with international guidelines³⁹. Clinical data including age, tumor staging, treatments protocols and clinical outcomes were collected. Disease-free patients were defined as those who resulted negative for tumor detection at CT scan after surgery for at least five years.

Metastatic patients' blood was collected at study entry and at CT scan evaluations. Blood drawn of non-metastatic patients was collected before the administration of neoadjuvant therapy, after neoadjuvant therapy at CT-scan evaluation and a month after surgery concurrently with CT-scan. All blood drawn volumes intended for CTC analysis ranged between 18–20 ml; first 3 ml of blood were not included in CTC enrichment to avoid epithelial cells contamination.

Cell Culture

The MDA MB 436 breast cancer cell line and MCF10A epithelial breast cell line were purchased from ATCC. EC cell line OE33 was kindly donated by Prof. Nickolas Stoecklein (University Hospital Düsseldorf). MDA-MB 436 cell line was cultured in Leibovitz's L-15 medium (GIBCO, USA, USA) with 10 ug/ml insulin (Sigma-Aldrich, USA Aldrich), 100U/ml penicillin-streptomycin (GIBCO, USA, USA), 2mM glutamine (GIBCO, USA, USA) complemented with 10% FBS (GIBCO, USA), in a non-free gas exchange condition.

MCF10A cells were cultured in DMEM/F-12, GlutaMAX™ medium (GIBCO, USA) complemented with 5% Horse Serum (New Zealand origin) (GIBCO, USA), 100 U/ml penicillin-streptomycin (GIBCO, USA), 2 mM L glutamine (Invitrogen, USA), 100 mg/ml human epidermal growth factor (EGF) (Merck Millipore, UK), 1 mg/ml hydrocortisone (Sigma-Aldrich, USA), 1mg/ml cholera toxin (Sigma-Aldrich, USA Aldrich) and 10 mg/ml insulin (Sigma-Aldrich, USA Aldrich), 100U/ml penicillin-streptomycin (GIBCO, USA) as mentioned⁴⁰.

OE33 cell line was cultured in RPMI 1640 medium (GIBCO, USA) complemented with 10% FBS (GIBCO, USA), 100U/ml penicillin-streptomycin (GIBCO, USA).

All cell lines were maintained as a monolayer in a 37°C incubator with 5% CO₂ and sub-cultured at 70–75% confluence (approximately twice weekly).

We induced epithelial-to-mesenchymal transition (EMT) on MCF10A cell line by TGF beta treatment as reported⁴¹. Briefly, cells were treated with 10ng/ml of TGF beta1 (Proprtech) every 48 hours for ten days. Morphological changes observation and images were performed by light microscopy Axiovert 200 (Zeiss).

Grab all assay and DEPArray analyses were performed on cells collected at four time points: before treatment, after 4, 6 and 10 days.

CTC enrichment, Grab-all assay and single cell isolation

Blood samples (18-20ml per sample) were collected in EDTA vacutainers (BD) and processed within 3 hours. Label-independent enrichment of CTCs was carried out by density gradient centrifugation with OncoQuick® as we previously used^{17,42}.

Next, Grab-all assay was performed on CTC samples. In details, samples were fixed in 2% paraformaldehyde (Miltenyi Biotech, USA) and incubated 30 minutes at 4 C° with cell membrane directed antibodies: EpCAM-PE (clone HEA-125) (Miltenyi Biotech, USA) 1:10; E-Cadherin-PE (clone 67A4) (Miltenyi Biotech, USA) 1:10; N-Cadherin AlexaFluor 647 (clone 8C11) (BD Pharmingen, USA) 1:20; CD44v6-APC (clone 2F10) (R&D Systems, USA) 1:20; ABCG2 (clone 5D3) (R&D Systems, USA) 1:10; CD45-AlexaFluor 488 (clone GA90) (Life Technologies, USA) 1:10. We performed intracytoplasmic staining with pan-CKs-PE (clone C11) (Aczon, Italy) 1:10, Hoechst 33342 (1 µg/mL, Life Technologies, USA) diluted in a solution of 1X PBS/1% BSA (Sigma-Aldrich, USA-Aldrich) 0.3% Triton™ X-100 (Bio-rad, The Netherlands).

Stained samples were resuspended in 14 ul of SB115 Buffer (Menarini-Silicon Biosystems SpA, Italy), loaded into DEPArray™ chip (Menarini-Silicon Biosystems SpA, Italy) and analyzed by DEPArray platform.

In order to avoid any technical bias related to acquisition of images, first we defined time exposure of DEPAarray acquisition. We stained leukocytes only with CD45 antibody acquiring images of all fluorescent channels to define the maximum exposure time of FITC channel for saturation and to verify the absence of any autofluorescence signals on PE and APC channels. We followed the same approach to define epithelial and stem-mesenchymal tags (PE and APC channels respectively) by staining a mix of MDA MB 436 and leukocytes cells. MDA MB 436 cells were stained with epithelial plus CD45 antibodies and subsequently with stem/mesenchymal plus CD45 antibodies. We configured time exposure of channels to prevent autofluorescence signals especially on FITC channel as well as saturation signals.

All acquired patient's cell images were manually reviewed, cells classified as CTCs had nucleus rounded- or spindle- shaped morphology and were negative for CD45. Epithelial CTCs were positive only for epithelial tag (EpCAM/CKs/E-Cadherin), mesenchymal/stem CTCs were positive only for stem/mesenchymal tag (N-Cadherin/CD44v6/ABCG2) and bi-phenotypic CTCs when cells were positive for both tags. All routable CTCs were sorted as single cells and single leukocytes were sorted as reference diploid chromosome set for further molecular analysis. Cells were sorted into 0,2ml tubes, and stocked at -80°C after manual volume reduction and 1X PBS (GIBCO, USA) wash until whole genome amplification (WGA).

Single CTC WGA and quality control (QC):

We performed WGA using Ampli1 WGA Kit (Menarini-Silicon Biosystems) according to manufacturer's instructions except for the cell lysis step, which was conducted overnight into thermal cycler. DNA samples were stocked at -20°C until use.

All the WGA products were quality checked adopting Ampli1™ QC Kit (Silicon Biosystems) as per manufacturer's instructions. Briefly, we performed a multiplex PCR to amplify four amplicons of different lengths. The amplification of the regions was assessed on a 2% agarose gel using Chemidoc instrument (Bio-rad, The Netherlands).

Library preparation and next generation sequencing (NGS)

DNA libraries for low-pass whole genome sequencing were generated using a streamline kit of WGA products named *Ampli1™* LowPass Kit for Ion Torrent (Menarini-Silicon Biosystems) as reported¹⁸. WGA products of some samples were re-amplified by Ampli1™ ReAmp/ds following manufacturer's protocol for the generation of appropriate libraries. Concisely, 10 µl of WGA products were linked to barcoded adaptors for Ion Torrent System by a PCR amplification. Next, libraries were quantified by Qubit™ 3.0 Fluorometer (Invitrogen, USA) and fragment length was assessed using Bioanalyzer High Sensitivity DNA Kit (Agilent). We combined libraries in an equimolar pool, and selected fragments of the pooled libraries ranging between 300 bp to 450 bp using E-Gel™ SizeSelect™ II Agarose Gels, 2% (Invitrogen, USA) on E-Gel™ iBase™ and E-Gel™ Safe Imager™ (Invitrogen, USA).

Whole genome sequencing of single CTCs and matched leukocytes was carried out on Ion PGM™ and Ion GeneStudio™ S5 System (Thermo Fisher), loading 8 pooled samples on 318 and 520 chip, respectively. Sequencing was conducted following the guidelines reported in the protocol (emulsion PCR for 400-base-length libraries and 525-flows sequencing). For sequencing experiments conducted on Ion S5, chip preparation and emulsion PCR were performed on the Ion Chef instrument, using Ion 510™ & Ion 520™ & Ion 530™ Kit reagents.

CNA calling from low-resolution whole genome sequencing:

The output BAM files were given as input at Control-FREEC for CNA calling. We used two different non-parametric tests (Mann-Whitney and Kolmogorov-Smirnov) to assess the statistical significance of each putative CNA. The whole genome has been subdivided in windows (bin), whose size depends on the quality of the sequencing. Then, we generated plots showing copy numbers mapped for each bin, highlighting the points which are included in CNAs with a significance lower than 0.05 in both the non-parametric tests we applied.

Generation of Jaccard Indexes and statistical analysis

We computed a Jaccard Index (JI) for the detection of similarity, for what concerns CNA content, among each CTC against five leucocytes which have been analyzed with the same workflow. JI has been computed applying the following formula:

$$JI(A, B) = \frac{A \cap B}{A \cup B}$$

in which A is the set of CNAs present in the CTC sample and B is the set of CNAs in a leucocyte. CTC samples have been grouped for access time, response and both of these conditions. Then, we assessed the presence of statistical differences among the groups. When comparing more than two CTC groups (i.e. access time and access time with response), we applied the Kruskal-Wallis test. Moreover, for each group pair in each condition, we applied the Mann-Whitney U test.

Identification of recurrent esophageal cancer aberrations on single CTCs

Starting from a list of cytogenetic bands which were of specific interest for the disease⁵, we extracted the number of CTCs with at least a CNA in the specific region, with a further discrimination for access time and patient response. A similar procedure has been applied to a list of genes included in the same cytogenetic bands.

Enrichment and network generation:

For each sample, we extracted the list of genes located inside a significant CNA detected by the previously described pipeline. From this list, using R package *EnrichR*⁴³ we performed an enrichment analysis exploring GO Biological Process dataset and we filtered out terms with adjusted p-value higher than 0.05. Finally, for each dataset, we generated networks using R package *igraph*⁴⁴. Inside these networks, nodes are the enriched terms while edges give the information about the presence of more than 10 shared genes among the two linked nodes. Moreover, the size of each node is related to the number of genes, while the sections of the pie chart inside nodes are proportional to the percentage of samples of access time.

Declarations

Acknowledgments

We thank the cancer patients and healthy volunteers for donating samples that made these investigations possible. We also thank the supporters of the TRACe crowdfunding campaign. We are also most grateful to Alessandro Coatti and Marcella Tazzari for their constructive comments on the manuscript.

Conflict of Interest

The authors declare no competing interests.

Author contributions:

G.G., T.R., C.C., conducted CTC phenotypic analysis; R.T., F.P., D.F.S., B.E. conducted single-cell genomic analysis; A.D performed the bioinformatics analysis; V.M., F.G.L. recruited and consented patients and collected blood samples; G.G., R.T., B.M., F.F. conceived and directed the study, interpreted the data; G.G., R.T. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Funding:

T.R. and C.C. were funded via a Fellowship research scholar from Associazione Annastaccatolisa ONLUS (Montecatini Terme, Pistoia, Italy). This work was supported by TRACe crowdfunding campaign and by 5X1000 funds by Italian Ministry of Health.

References

1. Smyth EC, Lagergren J, Fitzgerald RC, Lordick F, Shah MA, Lagergren P *et al.* Oesophageal cancer. *Nat Rev Dis Prim* 2017; 3: 17048.
2. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2015. *CA Cancer J Clin* 2015; 65: 5–29.
3. Shapiro J, van Lanschot JJB, Hulshof MCCM, van Hagen P, van Berge Henegouwen MI, Wijnhoven BPL *et al.* Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): Long-term results of a randomised controlled trial. *Lancet Oncol* 2015; 16: 1090–1098.
4. Kim J, Bowlby R, Mungall AJ, Robertson AG, Odze RD, Cherniack AD *et al.* Integrated genomic characterization of oesophageal carcinoma. *Nature* 2017; 541: 169–175.
5. Dulak AM, Schumacher SE, Van Lieshout J, Imamura Y, Fox C, Shim B *et al.* Gastrointestinal adenocarcinomas of the esophagus, stomach, and colon exhibit distinct patterns of genome instability and oncogenesis. *Cancer Res* 2012; 72: 4383–4393.

6. Nijhawan D, Zack TI, Ren Y, Strickland MR, Lamothe R, Schumacher SE *et al.* Cancer Vulnerabilities Unveiled by Genomic Loss. *Cell* 2012; 150: 842–854.
7. Nones K, Waddell N, Wayte N, Patch AM, Bailey P, Newell F *et al.* Genomic catastrophes frequently arise in esophageal adenocarcinoma and drive tumorigenesis. *Nat Commun* 2014; 5. doi:10.1038/ncomms6224.
8. Secrier M, Li X, de Silva N, Eldridge MD, Contino G, Bornschein J *et al.* Mutational signatures in esophageal adenocarcinoma define etiologically distinct subgroups with therapeutic relevance. *Nat Genet* 2016; 48. doi:10.1038/ng.3659.
9. Ross-Innes CS, Becq J, Warren A, Cheetham RK, Northen H, O'Donovan M *et al.* Whole-genome sequencing provides new insights into the clonal architecture of Barrett's esophagus and esophageal adenocarcinoma. *Nat Genet* 2015; 1–11.
10. Noorani A, Li X, Goddard M, Crawte J, Alexandrov LB, Secrier M *et al.* Genomic evidence supports a clonal diaspora model for metastases of esophageal adenocarcinoma. *Nat Genet* 2020; 52: 74–83.
11. Maltoni R, Fici P, Amadori D, Gallerani G, Cocchi C, Zoli M *et al.* Circulating tumor cells in early breast cancer: A connection with vascular invasion. *Cancer Lett* 2015; 367: 43–48.
12. Reeh M, Effenberger KE, Koenig AM, Riethdorf S, Eichstädt D, Vettorazzi E *et al.* Circulating Tumor Cells as a Biomarker for Preoperative Prognostic Staging in Patients With Esophageal Cancer. *Ann Surg* 2015; 261: 1124–1130.
13. Rossi T, Gallerani G, Angeli D, Cocchi C, Bandini E, Fici P *et al.* Single-Cell NGS-Based Analysis of Copy Number Alterations Reveals New Insights in Circulating Tumor Cells Persistence in Early-Stage Breast Cancer. *Cancers (Basel)* 2020; 12: 2490.
14. Kuvendjiska J, Bronsert P, Martini V, Lang S, Pitman M, Hoepfner J *et al.* Non-Metastatic Esophageal Adenocarcinoma: Circulating Tumor Cells in the Course of Multimodal Tumor Treatment. *Cancers (Basel)* 2019; 11: 397.
15. WOESTEMEIER A, GHADBAN T, RIETHDORF S, HARMS-EFFENBERGER K, KONCZALLA L, UZUNOGLU FG *et al.* Absence of HER2 Expression of Circulating Tumor Cells in Patients with Non-Metastatic Esophageal Cancer. *Anticancer Res* 2018; 38: 5665–5669.
16. Carter L, Rothwell DG, Mesquita B, Smowton C, Leong HS, Fernandez-Gutierrez F *et al.* Molecular analysis of circulating tumor cells identifies distinct copy-number profiles in patients with chemosensitive and chemorefractory small-cell lung cancer. *Nat Med* 2016. doi:10.1038/nm.4239.
17. Fabbri F, Carloni S, Zoli W, Ulivi P, Gallerani G, Fici P *et al.* Detection and recovery of circulating colon cancer cells using a dielectrophoresis-based device: KRAS mutation status in pure CTCs. *Cancer Lett* 2013; 335: 225–31.
18. Ferrarini A, Forcato C, Buson G, Tononi P, Del Monaco V, Terracciano M *et al.* A streamlined workflow for single-cells genome-wide copy-number profiling by low-pass sequencing of LM-PCR whole-genome amplification products. *PLoS One* 2018; 13: 1–24.
19. Polzer B, Medoro G, Pasch S, Fontana F, Zorzino L, Pestka A *et al.* Molecular profiling of single circulating tumor cells with diagnostic intention. *EMBO Mol Med* 2014; 6: 1371–86.
20. Shinohara M, Io K, Shindo K, Matsui M, Sakamoto T, Tada K *et al.* APOBEC3B can impair genomic stability by inducing base substitutions in genomic DNA in human cells. *Sci Rep* 2012; 2: 806.
21. Diamantopoulou Z, Castro-Giner F, Aceto N. Circulating tumor cells: Ready for translation? *J Exp Med* 2020; 217: 8–10.
22. Sieuwerts AM, Kraan J, Bolt J, Van Der Spoel P, Elstrodt F, Schutte M *et al.* Anti-epithelial cell adhesion molecule antibodies and the detection of circulating normal-like breast tumor cells. *J Natl Cancer Inst* 2009; 101: 61–66.
23. Joosse S a., Hannemann J, Spötter J, Bauche A, Andreas A, Müller V *et al.* Changes in keratin expression during metastatic progression of breast cancer: Impact on the detection of circulating tumor cells. *Clin Cancer Res* 2012; 18: 993–1003.
24. Gall TMH, Frampton AE. Gene of the month: E-cadherin (CDH1). *J Clin Pathol* 2013; 66: 928–32.
25. Peinado H, Portillo F, Cano A. Transcriptional regulation of cadherins during development and carcinogenesis. *Int J Dev Biol* 2004; 48: 365–375.
26. Khan MA, Chen HC, Zhang D, Fu J. Twist: a molecular target in cancer therapeutics. *Tumour Biol* 2013; 34: 2497–2506.
27. Li L, Bennett SAL, Wang L. Role of E-cadherin and other cell adhesion molecules in survival and differentiation of human pluripotent stem cells. *Cell Adh Migr* 2012; 6: 59–70.
28. Chen Y, Fu Z, Xu S, Xu Y, Xu P. The prognostic value of CD44 expression in gastric cancer: a meta-analysis. *Biomed Pharmacother* 2014; 68: 693–7.
29. Hu B, Luo W, Hu R-T, Zhou Y, Qin S-Y, Jiang H-X. Meta-Analysis of Prognostic and Clinical Significance of CD44v6 in Esophageal Cancer. *Medicine (Baltimore)* 2015; 94: e1238.
30. Orian-Rousseau V, Chen L, Sleeman JP, Herrlich P, Ponta H. CD44 is required for two consecutive steps in HGF/c-Met signaling. *Genes Dev* 2002; 16: 3074–86.

31. Robey RW, Polgar O, Deeken J, To KW, Bates SE. ABCG2: Determining its relevance in clinical drug resistance. *Cancer Metastasis Rev* 2007; 26: 39–57.
32. Tsunoda S, Okumura T, Ito T, Kondo K, Ortiz C, Tanaka E *et al.* ABCG2 expression is an independent unfavorable prognostic factor in esophageal squamous cell carcinoma. *Oncology* 2006; 71: 251–258.
33. Yu M, Bardia A, Wittner BS, Stott SL, Smas ME, Ting DT *et al.* Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* 2013; 339: 580–4.
34. Prins MJD, Ruurda JP, Lolkema MP, Sitarz R, ten Kate FJW, van Hillegersberg R. The role of biological markers of epithelial to mesenchymal transition in oesophageal adenocarcinoma, an immunohistochemical study. *J Clin Pathol* 2015; 68: 529–535.
35. Murugaesu N, Wilson GA, Birkbak NJ, Watkins TBK, McGranahan N, Kumar S *et al.* Tracking the genomic evolution of esophageal adenocarcinoma through neoadjuvant chemotherapy. *Cancer Discov* 2015; 5: 821–832.
36. Plum PS, Löser H, Zander T, Essakly A, Bruns CJ, Hillmer AM *et al.* GATA binding protein 6 (GATA6) is co-amplified with PIK3CA in patients with esophageal adenocarcinoma and is linked to neoadjuvant therapy. *J Cancer Res Clin Oncol* 2021; 147: 1031.
37. Aparicio T, Cozic N, de la Fouchardière C, Meriaux E, Plaza J, Mineur L *et al.* The Activity of Crizotinib in Chemo-Refractory MET-Amplified Esophageal and Gastric Adenocarcinomas: Results from the AcSé-Crizotinib Program. *Target Oncol* 2021; 16: 381–388.
38. Bandla S, Pennathur A, Luketich JD, Beer DG, Lin L, Bass AJ *et al.* Comparative Genomics of Esophageal Adenocarcinoma and Squamous Cell Carcinoma. *Ann Thorac Surg* 2012; 93: 1101–1106.
39. Lordick F, Mariette C, Haustermans K, Obermannová R, Arnold & D, Committee G. Oesophageal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up † on behalf of the ESMO. 2016. doi:10.1093/annonc/mdw329.
40. Soule HD, Maloney TM, Wolman SR, Line EC, Peterson WD, Brenz R *et al.* Isolation and Characterization of a Spontaneously Immortalized Isolation and Characterization of a Spontaneously Immortalized Human Breast. *Cancer Res* 1990; 50: 6075–6086.
41. Fici P, Gallerani G, Morel A-P, Mercatali L, Ibrahim T, Scarpi E *et al.* Splicing factor ratio as an index of epithelial-mesenchymal transition and tumor aggressiveness in breast cancer. *Oncotarget* 2016; 8: 2423–2436.
42. Maltoni R, Gallerani G, Fici P, Rocca A, Fabbri F. CTCs in early breast cancer: A path worth taking. *Cancer Lett* 2016; 376: 205–210.
43. Kuleshov M V., Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 2016; 44: W90–W97.
44. Csárdi G, Nepusz T. The igraph software package for complex network research. *undefined* 2006.
45. Roper N, Gao S, Maity TK, Banday AR, Zhang X, Venugopalan A *et al.* APOBEC Mutagenesis and Copy-Number Alterations Are Drivers of Proteogenomic Tumor Evolution and Heterogeneity in Metastatic Thoracic Tumors. *Cell Rep* 2019; 26: 2651–2666.e6.
46. Luo J, Wang W, Tang Y, Zhou D, Gao Y, Zhang Q *et al.* mRNA and methylation profiling of radioresistant esophageal cancer cells: the involvement of Sall2 in acquired aggressive phenotypes. *J Cancer* 2017; 8: 646–656.
47. Lobbardi R, Pinder J, Martinez-Pastor B, Theodorou M, Blackburn JS, Abraham BJ *et al.* TOX regulates growth, DNA repair, and genomic instability in T-cell acute lymphoblastic leukemia. *Cancer Discov* 2017; 7: 1336–1353.
48. Petty AP, Garman KL, Winn VD, Spidel CM, Lindsey JS. Overexpression of carcinoma and embryonic cytotrophoblast cell-specific Mig-7 induces invasion and vessel-like structure formation. *Am J Pathol* 2007; 170: 1763–1780.
49. Meier T, Timm M, Montani M, Wilkens L. Gene networks and transcriptional regulators associated with liver cancer development and progression. *BMC Med Genomics* 2021; 14: 41.
50. Zhang L-W, Wang B, Yang J-X, Yang H. Circ-PRMT5 stimulates migration in esophageal cancer by binding miR-203. *Eur Rev Med Pharmacol Sci* 2020; 24: 9965–9972.
51. Shi X, Chen Z, Hu X, Luo M, Sun Z, Li J *et al.* AJUBA promotes the migration and invasion of esophageal squamous cell carcinoma cells through upregulation of MMP10 and MMP13 expression. *Oncotarget* 2016; 7: 36407–36418.
52. Garancher A, Lin CY, Morabito M, Richer W, Rocques N, Larcher M *et al.* NRL and CRX Define Photoreceptor Identity and Reveal Subgroup-Specific Dependencies in Medulloblastoma. *Cancer Cell* 2018; 33: 435–449.e6.
53. Chen Z, Wang Y, Wang J, Kang M, Tang W, Chen S. Assessment of PPARGC1A, PPARGC1B, and PON1 Genetic Polymorphisms in Esophageal Squamous Cell Carcinoma Susceptibility in the Eastern Chinese Han Population: A Case–Control Study Involving 2351 Subjects. *DNA Cell Biol* 2020; 39: 1521–1531.
54. Baslan T, Kendall J, Volyanskyy K, McNamara K, Cox H, D'Italia S *et al.* Novel insights into breast cancer copy number genetic heterogeneity revealed by single-cell genome sequencing. *Elife* 2020; 9: 1–21.

55. Yasui K, Imoto I, Fukuda Y, Pimkhaokham A, Yang Z-Q, Naruto T *et al.* Identification of target genes within an amplicon at 14q12-q13 in esophageal squamous cell carcinoma. *Genes, Chromosom Cancer* 2001; 32: 112–118.
56. Liu Z, Zhang A, Zheng L, Johnathan A-F, Zhang J, Zhang G. The Biological Significance and Regulatory Mechanism of c-Myc Binding Protein 1 (MBP-1). *Int J Mol Sci* 2018; 19: 3868.
57. SAWADA G, UEO H, MATSUMURA T, UCHI R, ISHIBASHI M, MIMA K *et al.* CHD8 is an independent prognostic indicator that regulates Wnt/ β -catenin signaling and the cell cycle in gastric cancer. *Oncol Rep* 2013; 30: 1137–1142.
58. Huang Y, Lin L, Shen Z, Li Y, Cao H, Peng L *et al.* CEBPG promotes esophageal squamous cell carcinoma progression by enhancing PI3K-AKT signaling. *Am J Cancer Res* 2020; 10: 3328–3344.

Figures

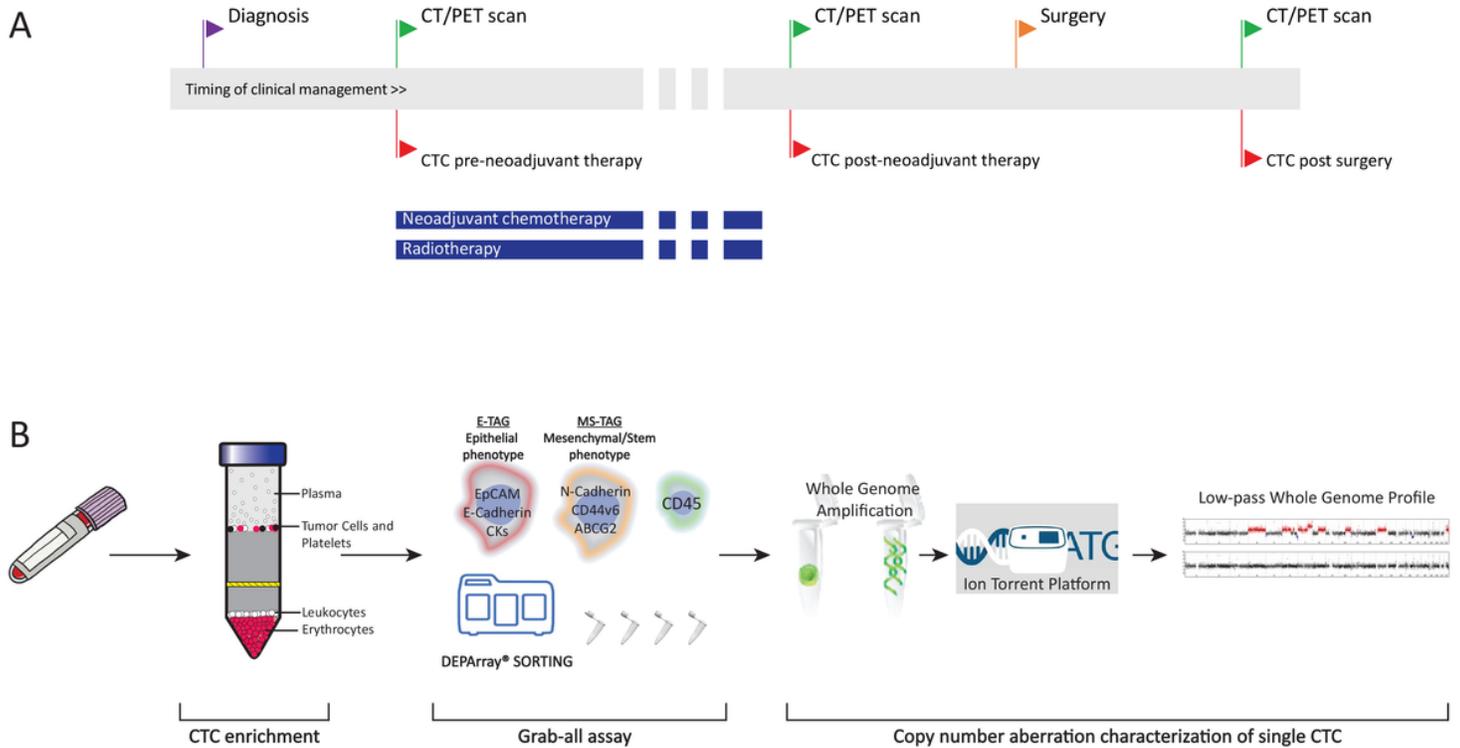


Figure 1

Clinical history of esophageal cancer (EC) patients and Experimental Workflow. [A] Outline of the clinical therapeutic history of patients with non-metastatic EC. Along the timeline, patient's clinical management milestones are paired with blood biopsies (red flags). Treatment regimens are shown as blue bars below the timeline. [B] Experimental workflow adopted for CTC analysis. CTC enrichment from peripheral blood was performed by density gradient based Oncoquick, followed by Grab-all assay to identify CTCs with different phenotypes using DEPArray platform. Once identified, CTCs were sorted as single cells and subjected to whole genome amplification. Low-pass whole genome profiles were obtained from amplified DNA of single CTCs using the Ion Torrent S5 platform.

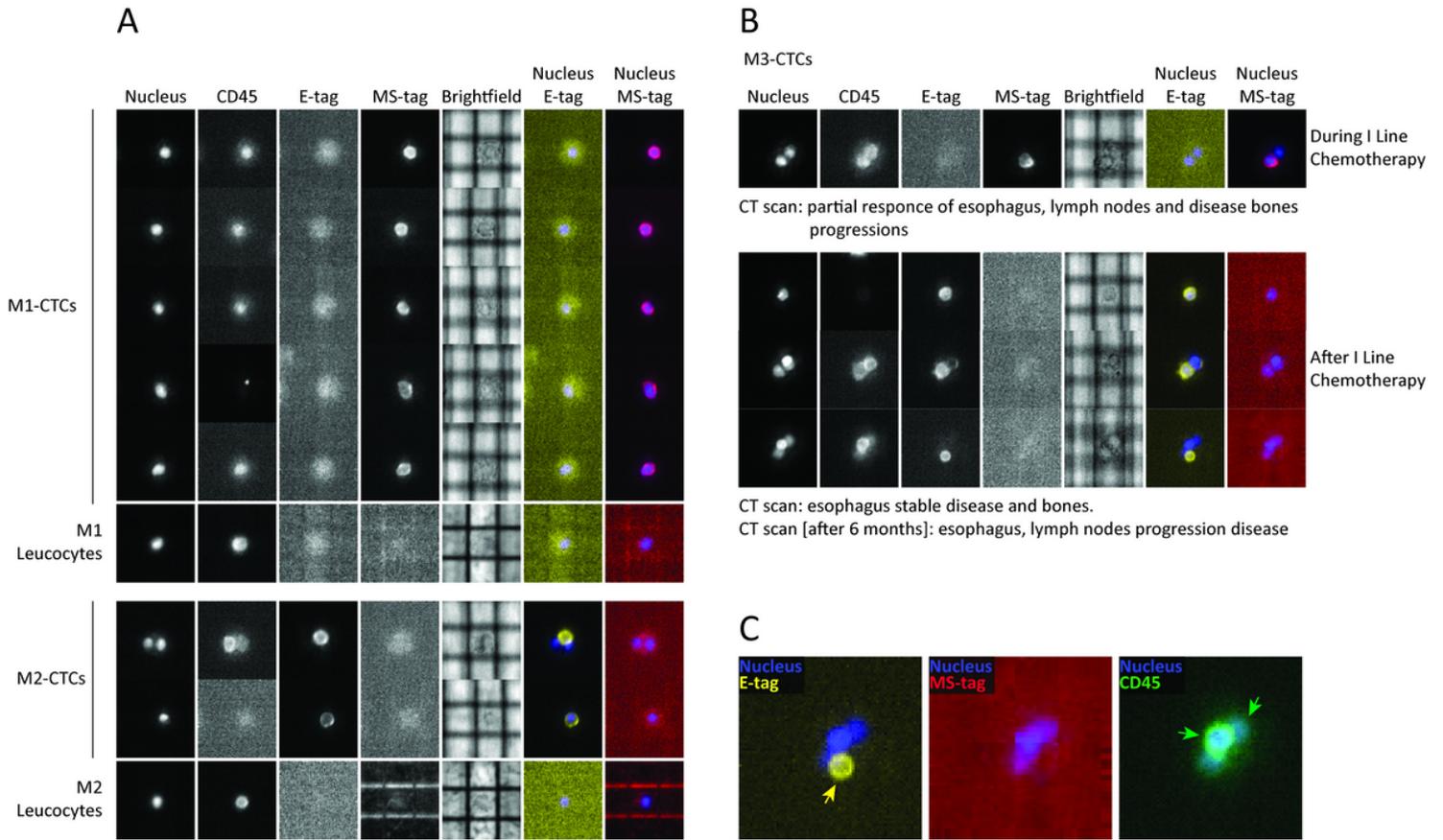


Figure 2

Identification of different CTC phenotypes by Grab-all assay on samples from metastatic esophageal cancer (EC) patients. [A] Panel of representative CTCs with different phenotypes, frankly mesenchymal/stem (patient M1) (red pseudocolour), frankly epithelial (patient M2) (yellow pseudocolour); CD45+ matched leucocytes as controls. [B] Phenotypic evolution of CTCs of patient M3: single mesenchymal/stem CTC during first line chemotherapy and epithelial CTCs after treatment; CT scan outcomes performed at the time of blood draws or after six months are listed under the panels. [C] Zoomed DEPArray image of a cell cluster composed of an epithelial CTC (yellow pseudocolour) and leucocytes (green pseudocolour).

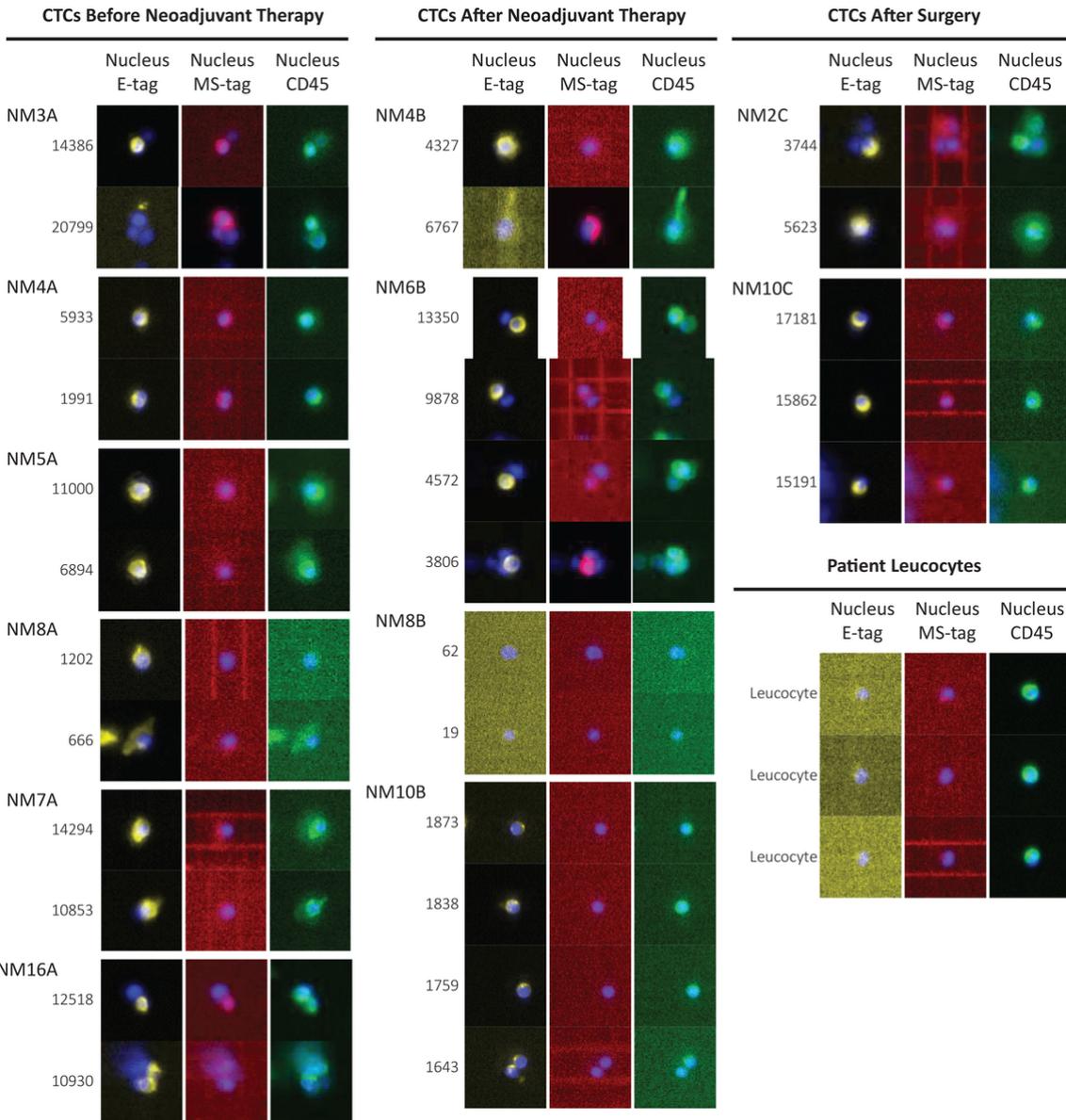


Figure 3

DEPArray panels of representative CTCs identified in non-metastatic esophageal cancer (EC) patients. Images of CTCs obtained before neoadjuvant treatment, after neoadjuvant treatment and after surgery are grouped in separate columns (each including three different imaging channels). A fourth group lists images of patients' leukocytes. Every line represents an identified cell with its specific ID. Autofluorescence signals on FITC channel (green) of CTCs are well distinct from specific CD45 positivity of leukocytes. We identified frankly epithelial CTCs (yellow pseudocolour) or mesenchymal/stem CTCs (red pseudocolour) or hybrid phenotypes (ID 3806). In patient NM8's blood drawn after neoadjuvant therapy, we identified cells that were both negative for CD45 and other Grab-all assay tags. We decided to sort these single cells to perform CNA analyses in order to classify them as tumoral or normal cells.

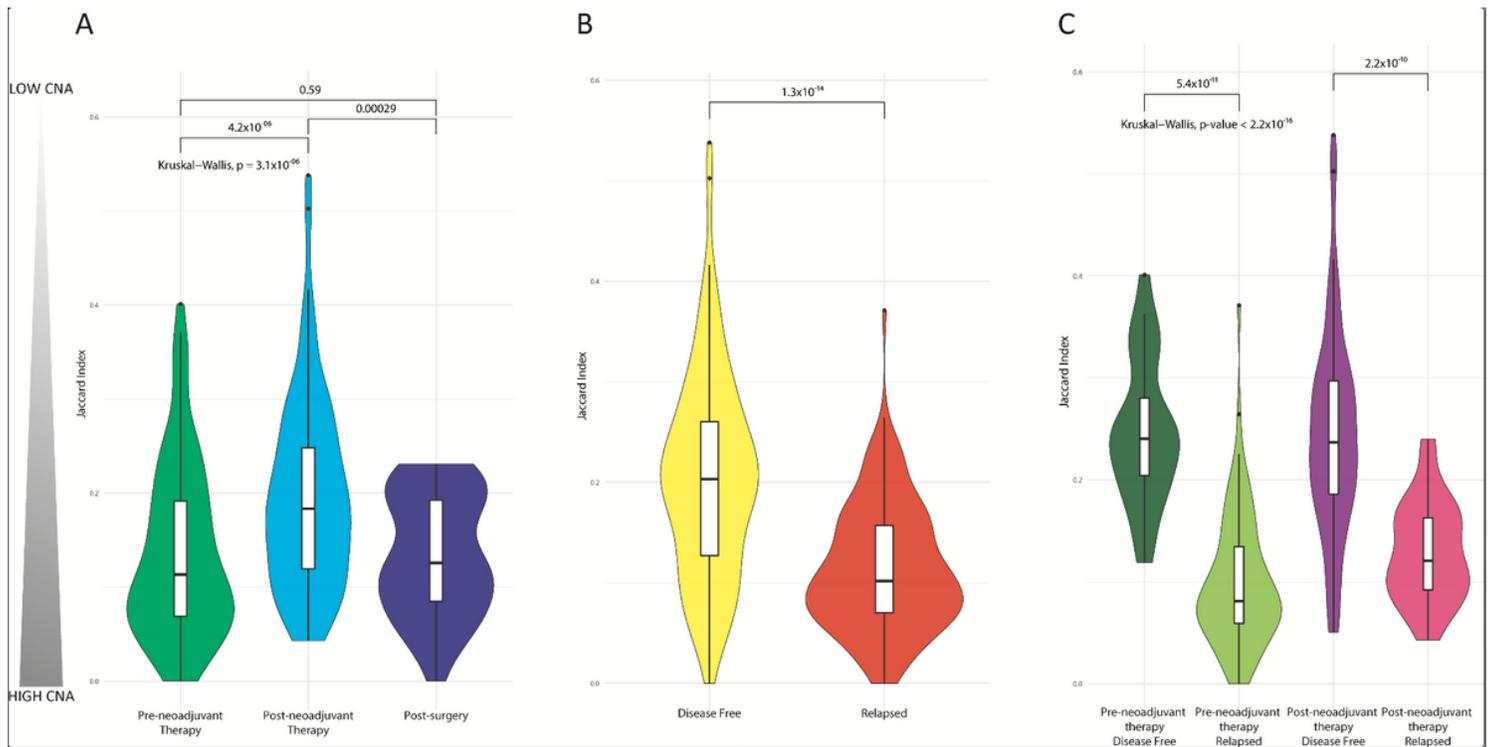


Figure 5

CTC genomic imbalance is measured relative to a matched diploid genome using a Jaccard similarity index (JI) between the set of CNAs present in the CTC and those in the patient's leukocytes. A high JI value (max 1) indicates that the CTC's genome has a similar CNA profile to that of matched leukocytes, a small JI value (min 0) indicate a CTC's genome with a CNA profile that differs from that of the matched leukocytes. JI values close to zero indicate increasing genomic imbalance. The panel of violin represent the distribution of CTC's JIs for cells grouped based on: time point of blood draws [A], patients clinical outcomes [B] or a combination of time point and patient's clinical outcomes [C]. At the top, we report the p-values of Kruskal-Wallis and Wilcoxon tests.

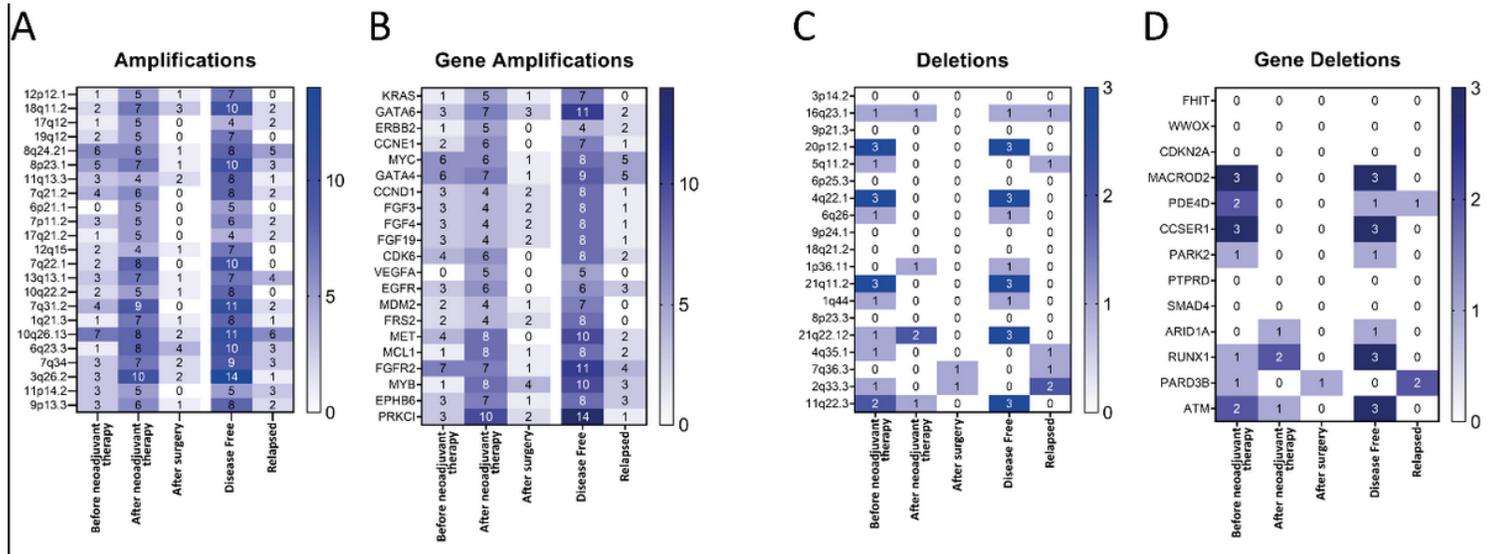


Figure 6

Panel of known esophageal cancer (EC) specific focal copy number changes which we found in identified CTCs. Frequently amplified regions [A], genes included within amplified regions [B], frequently deleted regions [C] and genes included in deleted regions [D]. CTCs are grouped based on the time point of blood draws and patients' clinical outcomes. The absolute frequency of CTCs bearing a specified focal copy number change is represented by a number and on a color scale.

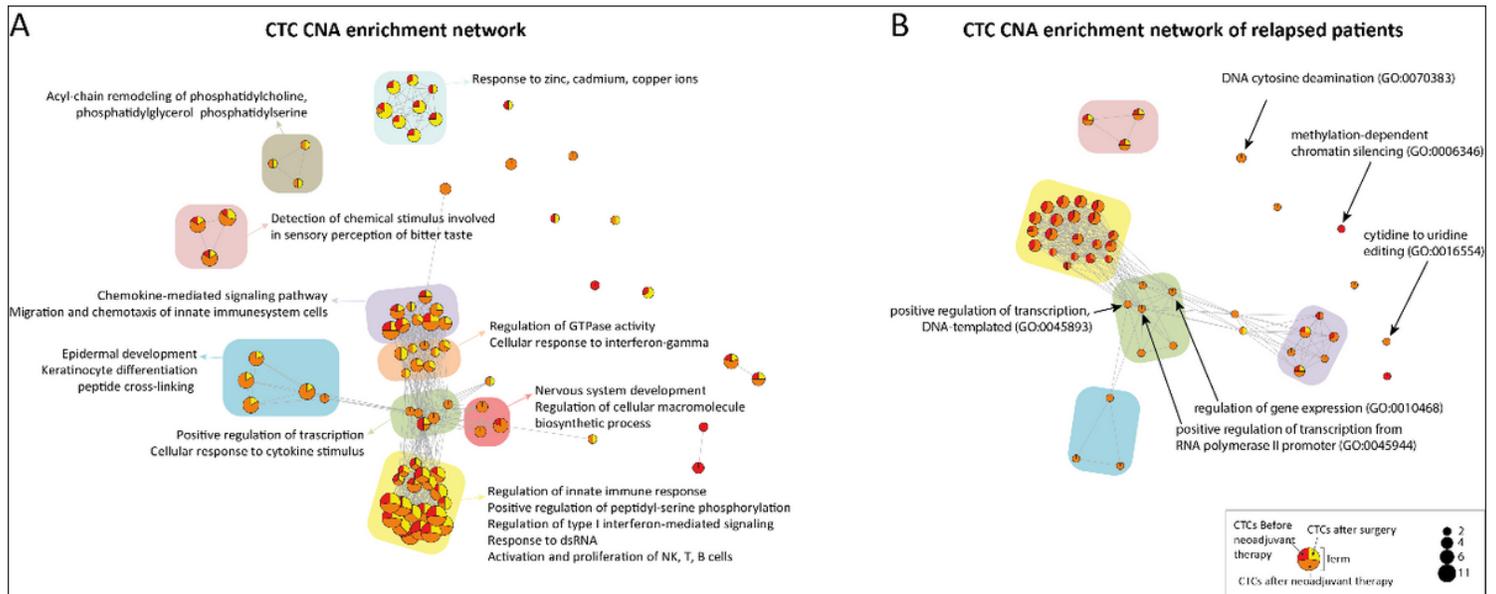


Figure 7

Networks on enriched GO (Biological Process) terms of CNAs of identified CTCs. GO terms are represented as nodes (circles). Sizing of the nodes is based on the number of CTCs with enriched term. Every circle is divided into sectors representing a proportion of the clinical step where CTCs were been identified: red, orange sectors before and after neoadjuvant therapy respectively, yellow after surgery; edges (grey lines) represent at least ten common genes overlapping the two terms. Network on enriched GO terms of all identified CTCs (A) and only of CTCs identified on relapsed patients (B). GO terms in the larger communities were color-coded accordingly.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SUPPLEMENTARYTABLE1.xlsx](#)
- [SUPPLEMENTARYFIGURE1.tif](#)
- [SUPPLEMENTARYFIGURE2.tif](#)
- [SUPPLEMENTARYFIGURE3.tif](#)
- [SUPPLEMENTARYFIGURE4.tif](#)