The novel association of circulating tumor cells and circulating megakaryocytes with prostate cancer prognosis

Lei Xu,[1,2]* Xueying Mao,[1]* Tianyu Guo,[1] Pui Ying Chan,[3] Greg Shaw,[4] John Hines,[4] Elzbieta Stankiewicz,[1] Yuqin Wang,[1] Tim Oliver,[1] Amar Ahmad,[5] Daniel Berney,[1] Jonathan Shamash,[3] Yong-Jie Lu[1]

[1]Centre for Molecular Oncology, Barts Cancer Institute, Queen Mary University of London, London, UK.

[2]Department of Urology, Zhongshan Hospital, Fudan University, Shanghai, China.

[3]Department of Medical Oncology, Barts Health NHS, London, UK.

[4]Department of Urology, Barts Health NHS, London, UK.

[5]Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, UK

*These authors contributed equally to this work.

Running title: Circulating tumor cells and megakaryocyte for prognosis

Key words: circulating tumor cell, circulating megakaryocyte, epithelial-mesenchymal transition, repeated FISH, prostate cancer prognosis

Abbreviations list

AUC: area under the ROC curve

CI: confidential interval

CMS: combined mesenchymal CTC and megakaryocyte score

CRPC: castration-resistant prostate cancer

CRS: combined risk score

CTC: circulating tumor cell

EMT: epithelial-mesenchymal transition

FISH: fluorescence *in situ* hybridization

GS: Gleason score

HR: hazard ratio

IQR: interquartile range

mCRPC: metastatic castration-resistant prostate cancer

NCCN: National Comprehensive Cancer Network

PSA: prostate specific antigen

ROC: receiver operating characteristic

VIM: vimentin

Corresponding Author: Yong-Jie Lu, Centre for Molecular Oncology, Barts Cancer Institute, Queen Mary University of London, John Vane Science Centre, Charterhouse Square, London, EC1M 6BQ, UK. Phone: +44 (0)20 7882 3597; Fax: +44 (0)20 7882 3884; Email: y.j.lu@qmul.ac.uk

Translational relevance: Here we firstly developed an optimized approach to robustly wash off immunofluorescence signals and perform multiple rounds of fluorescence *in situ* hybridization analysis on the same cells. This facilitates the genomic analysis and confirmation of circulating tumor cells (CTCs), valuable for cancer diagnosis, prognosis and progression monitoring. Prostate cancer patients with advanced clinical features showed an increasing number of epithelial/EMTing/mesenchymal CTCs. The presence of metastasis was efficiently predicted by a combined risk score based on both serum PSA level and EMTing CTC count. These analyses greatly enhance our ability to investigate metastasis process and to predict/monitor cancer progression. The most important novel finding was the association of circulating megakaryocytes with good prognosis. In combination with mesenchymal CTC count, they showed great cancer prognosis potential. The identification of circulating megakaryocytes and their association with cancer prognosis potentially opens a new revenue to investigate cancer progression and to develop novel prognostic biomarkers.

Abstract:

Purpose: To develop an approach for the investigation of different subtypes of circulating tumor cells (CTCs) and other cells to evaluate their potential prognostic value of prostate cancer.

Experimental Design: Malignancy of CTCs undergoing epithelial to mesenchymal transition (EMT) was confirmed by repeated Fluorescence *in situ* hybridization. Subgroups of CTCs in 81 patients with prostate cancer (43 castration resistant and 38 untreated localized) were correlated to disease aggressiveness parameters. Area under curve analysis was applied to compare the performance for metastasis prediction between serum PSA level alone and a combined risk score using both PSA and EMTing CTC count. Circulating megakaryocytes and cancer patient survival association was performed using Cox model.

Results: The majority of vimentin(VIM)+/CD45- cells were malignant with genomic alterations in several genomic regions. The number of cytokeratin(CK)-/VIM+/CD45- CTCs correlated with disease burden, tumor aggressiveness and poorer survival. Meanwhile, CK+/VIM+/CD45- CTCs were associated with metastases better than other subtypes of CTCs in these limited samples. Combination of PSA level and the number of CK+/VIM+/CD45- CTCs enhanced the prediction of cancer metastases (AUC 0.921, 95%CI: 0.858-0.985). The number of circulating megakaryocytes was potentially associated with good patient survival in advanced prostate cancer (HR:0.849, 95%CI:0.628-1.146, per cell increase) and the difference between the number of mesenchymal CTCs and megakaryocytes strongly correlated to poor survival (HR:10.17, 95%CI:2.164-47.789, if score ≥2.0).

Conclusions: This CTC analysis approach and the potential association of megakaryocytes with cancer prognosis may greatly enhance our ability to investigate the cancer metastasis process and to predict/monitor cancer progression.

Introduction

Cancer cells evolve during disease progression and in response to treatment (1), influencing their sensitivity to treatment. Circulating tumor cells (CTCs) are considered to be the seeds of metastases (2) and, at the same time, a source of information on tumor tissue that can be acquired through a simple blood sample. Even before a metastatic tumor is clinically evident, it has been reported that a large number of CTCs already exist in the circulation (3,4). This provides a more accessible 'liquid biopsy' than tumor tissue biopsies to predict/monitor disease progression and therapy response at both cellular and molecular level. However, current CTC analysis mainly focused on cells expressing epithelial phenotype-specific markers (2). In fact, activation of epithelial-mesenchymal transition (EMT) is a key process in cancer metastasis (5) and an important factor in promoting invasiveness of cancer cells and their resistance to therapy (6). Both down-regulation of epithelial markers, such as epithelial cell adhesion molecule (EpCAM), cytokeratin (CK), and E-cadherin, and up-regulation of mesenchymal markers, such as vimentin (VIM), and N-cadherin, have been described (7,8). Hence, CTCs undergoing EMT as part of the metastatic process may be missed when isolated based on their epithelial characteristics only. To address this challenge, we have optimized a novel epitope independent CTC isolation system, Parsortix™, to capture CTCs based on the much larger size and less deformability nature of tumor cells compared to normal blood cells (9). The efficiency of Parsortix in capturing CTCs has been independently validated by other research groups in comparison to CellSearch® in patients with small cell lung cancer (10), and other types of cancer including breast, colon and lung (11). Here we explored its clinical application in prostate cancer prognosis and disease state monitoring.

Prostate cancer is the most common cancer in Western men (12,13). Current criteria for risk stratification of newly diagnosed prostate cancer are mainly based on clinical features, including serum prostate specific antigen (PSA), clinical stage, and biopsy/surgical specimen

Gleason score (GS)(14). However, these factors are not sufficient to discriminate between patients with indolent and aggressive disease. Furthermore, when metastatic disease progresses to castration-resistant prostate cancer (CRPC), the established PSA test is not satisfactory for disease prognosis and monitoring. These current clinical unmet needs urge us to identify novel biomarkers that can monitor disease status precisely, accessibly and in real time. Therefore, we investigated the potential of analyzing CTCs with both epithelial and mesenchymal features for cancer prognosis using Parsortix™ system in prostate cancer as a disease model. We developed a novel technique to perform five rounds of fluorescence *in situ* hybridization (FISH) on the same slides after immunofluorescence staining. Using this technique, we simultaneously identified the epithelial and mesenchymal cell features and multiple genomic alterations, confirming the malignancy of circulating cells with mesenchymal phenotype. We also demonstrated the correlation of CTCs with advanced disease features. Unexpectedly, we discovered that an increase of circulating megakaryocytes had the trend to be correlated with good prognosis in patients with progressive disease and the combination of CTC and megakaryocyte count may effectively predict survival in advanced disease.

Materials and Methods

Patients and blood samples

A total of 81 patients with written consent were recruited from December 2014 in St Bartholomew's Hospital, Barts Health NHS, London, UK, comprising 38 with untreated localized prostate cancer, and 43 with progressive CRPC (40 with metastasis) ready to commence an alternative treatment. The clinical information is summarized in **Table 1** and details of sample collection and clinical information were shown in Supplementary Data and **Table S1**. Blood specimens from 24 healthy male donors were collected with signed Ethics

committee approved consent forms. 7.5mL of whole blood was donated from each participant for CTC enumeration. Use of blood samples from patients and healthy donors in this study was approved by National Research Ethics Service committee London City & East with a Research Ethics Committee reference of 09/H0704/4+5.

CTC isolation and enumeration

Isolation of CTCs from whole blood using a size- and deformability-based system Parsortix™ was performed as previously described (9). Sample harvest and the process of immunostaining are detailed in Supplementary Materials and Methods. Different DAPI+ populations of cells were recorded, including CK+/VIM-/CD45-, CK+/VIM+/CD45-, and CK-/VIM+/CD45-.

FISH after immunofluorescence analysis

Immunofluorescence signals were washed in a stripping buffer (containing 2% SDS, 0.0625M Tris-HCl pH6.8 and 0.8% β-mercaptoethanol) at 50°C for 20 minutes, followed by washing three times in PBS for 5 minutes. Details of FISH probe and multiple FISH procedure are indicated in Supplementary Materials and Methods. Ten FISH probes for nine commonly altered genes/genomic regions in prostate cancer, including 6q16, *NKX3.1*, *C-MYC*, *PTEN*, *CCND1, RB1*, 16q22.1, *ERG* and *AR* were used for CTC analysis. The copy number of each FISH probe was counted in CK+/VIM-/CD45-, CK+/VIM+/CD45-, CK-/VIM+/CD45- and CK-/VIM-/CD45- cells. The copy number of lymphocytes from the prostate cancer patients and healthy donors were used to calculate the baseline for each genetic change. FISH signals from >100 lymphocytes were counted in each sample.

Statistics

Unless specifically noted, Wilcoxon rank-sum test was applied to assess the equality of CTCs between subgroups based on CTC-score as well as different clinical features, such as

metastasis, primary GS, and risk classification in localized disease. Data shown was as median (interquartile range [IQR]). Spearman's rank correlation was used to assess the association between CTC counts and concurrent PSA level. Bivariate logistic regression was performed with PSA and EMTing CTC counts as predictors for imaging detected metastasis (yes, no). A combined risk score (CRS) was computed as the linear predictor of the fitted bivariate logistic model with PSA and EMTing CTC count as only predictors (as CRS = a* PSA + b * EMTing CTC count, where the values of 'a' and 'b' are the estimated log odds ratios). Receiver operating characteristic (ROC) curve analysis was used to test the ability of different subtypes of CTCs as well as CRS to distinguish patients with metastasis. Optimal cut-off point was calculated with an optimal corrected classified value to provide best available sensitivity and specificity. Rocgold function was used to independently test the equality of the ROC area of each method against the PSA as a standard curve. The combined mesenchymal CTC and megakaryocyte score (CMS) was calculated as:

$$\frac{mesenchymal\ CTC\ count - megakaryocte\ count}{megakaryocte\ count}$$

Due to possible absence of megakaryocytes detected in patients, megakaryocyte count equal to 0 was recorded as 1 for this score calculation. To determine the factors that predict survival time, variables considered as potential predictors were selected for univariate analyses using Cox model. The hazard ratio associated with each biomarker was derived from the Cox model as an increase in the hazard per unit increase in biomarker. Due to the small sample size, p value for the hazard ratio were further adjusted by false discovery rate using Benjamini & Hochberg's adjustment method (15). Survival curves were generated using the Kaplan-Meier method and compared using the log-rank test. All statistical tests were two sided and p-values less than 0.05 (including multiple test adjustment where relevant) were considered as statistically significant. Bonferroni correction test was performed to modify p

values for multiple tests through dividing the critical p value by the number of comparisons being made. Statistical analyses were performed using Stata 13.0.

Results

Detection of circulating cells with epithelial and mesenchymal features

Using our optimized CTC isolation and detection method (9), we analyzed blood samples from 81 prostate cancer patients (**Table 1 and S1**) to identify CK+/VIM-/CD45-, CK+/VIM+/CD45- and CK-/VIM+/CD45- circulating cells (**Fig. 1A**). All three types of cells have been detected at high frequencies both in the blood of CRPC and localized cancer patients and they are higher in CRPC patients than those in untreated patients (**Table 1**). Of 81 patients studied, 15 (19%) patients had no detectable CK+/CD45- cells but detectable CK-/VIM+/CD45- circulating cells.

Genetic evidence that CK-/VIM+/CD45- circulating cells are malignant cells with genomic alterations

To confirm the malignant feature of CK-/VIM+/CD45- circulating cells by detecting multiple genomic alterations, we developed a technology for repeated multiple rounds of FISH analysis after immunofluorescence staining, and applied it on the CTC samples. This post-immunofluorescence multiple FISH analysis technique was developed using slides with lymphocytes spiked with PC3 cells, in which we removed the immunofluorescence signals completely by the stripping buffer, but not by 2XSSC buffer, fix solution or proteinase K digestion. The length of poly-lysine slide coating time was optimized to 45 minutes to best preserve cells for downstream repeated FISH analysis after signal striping. Less coating time frequently resulted in damaged or lost cells after signal striping. Using these optimized conditions, we detected clear nuclear morphology and FISH signals in up to the fifth round of

FISH on cells after immunofluorescence (**Fig. S1**). In the leucocytes from the healthy donors, the average false positive rates for the probes range from 0.7%-7.1% (**Table S2**).

The multiple FISH technique was then successfully applied to 11 prostate cancer CTC cases for five rounds of FISH (**Fig. 1B**) and one case for two rounds due to strong florescence background, to investigate the genomic alterations of nine genomic regions, including chromosomal copy number changes and/or rearrangements. In the leucocytes from three randomly selected patients, the average false positive rates for the probes range from 1.5%-7.8% (**Table S2**). In CTCs from an individual, each of the genomic changes was only detected in a proportion of cells, indicating genomic heterogeneity. In the limited number of cases and limited number of CTCs in each case, no obvious genomic change patterns specific to a subgroup of epithelial or mesenchymal CTCs were observed. Changes of more than 30% of the genomic regions were detected in 68% of CK+/VIM-/CD45- cells, 57% of CK+/VIM+/CD45- cells and 54% of CK-/VIM+/CD45- cells, but only detected in 3.7% of CK-/VIM-/CD45+ leucocytes and 7.8% of CK-/VIM-/CD45- cells (**Table S3**). The similar rate of genetic changes in the CK+/VIM-/CD45-, CK+/VIM+/CD45- and CK-/VIM+/CD45- circulating cells indicates that the majority of CK-/VIM+/CD45- cells were CTCs. While it is possible that small proportions of the above three circulating cell categories are of non-malignant origin, we considered all CK+/VIM-/CD45-, CK+/VIM+/CD45- and CK-/VIM+/CD45- cells for the correlation analysis between CTC numbers and clinical features, and categorized them as epithelial, EMTing or mesenchymal CTCs, respectively. When classifying cases as positive or negative for CTCs, the number of CK+/VIM-/CD45-, CK+/VIM+/CD45- and CK-/VIM+/CD45- circulating cells found in non-cancer healthy control cases were considered.

Association of CTC positivity with advanced clinical features in localized and metastatic prostate cancer

Analyzing blood samples from 24 healthy male donors, we detected one, two and three CK-/VIM+/CD45- cells in three, two and two samples, respectively, and none in the remaining 17 samples (median 0, range 0 to 3 cells per 7.5 mL), and complete absence of CK+/VIM-/CD45- or CK+/VIM+/CD45- circulating cells. Consequently we defined positive CTC cases as those showing any CK+/VIM-/CD45-, any CK+/VIM+/CD45-, and/or >3 CK-/VIM+/CD45- cells to prevent potential false positive, although a better definition may be worked out from future large cohort studies. Based on these criteria, 24 of 38 patients with untreated localized disease (63%) and 41 of 43 CRPC patients (95%) scored positive for CTCs. Excluding mesenchymal CTCs reduced CTC positivity to 20 (53%) and 38 (88%) in these two groups of patients respectively. In all 81 patients, high serum PSA level, high GS and metastatic status were significantly correlated with CTC-score positive patients. In the 38 untreated localized diseases, CTC-score positive cases have a trend to be associated with high risk classification based on National Comprehensive Cancer Network (NCCN) guideline version 1, 2016 (16), which should be further validated in a large sample cohort (**Table S4**).

Association of subgroups of CTCs with serum PSA level, primary biopsy GS and the risk of localized tumor

We further analyzed the relationship of different sub-populations of CTCs to PSA level, primary GS and the risk of localized cancer. While all three sub-populations of CTC numbers were significantly correlated to serum PSA levels (Spearman's $\varrho = 0.28$, 0.26, and 0.36, and p = 0.01, 0.019 and 0.0009 respectively for CK+/VIM-/CD45-, CK+/VIM+/CD45- and CK-/VIM+/CD45- CTCs), the mesenchymal type had the most significant association (**Fig. 2A**). As expected, the total number of CTCs had an even more significant correlation (Spearman's $\varrho = 0.49$, $p < 0.0001$) due to the greater number of events included for analysis. Due to the multiple tests for analysis, Bonferroni correction test was performed to modify the significant p values. After adjustment, mesenchymal CTC alone but not epithelial and EMTing CTC

alone reached statistical significance. Traditional CTCs (combination of epithelial and EMTing CTC) had a significant correlation.

Higher CTC count was also detected in patients with higher GS of primary tumor regardless of therapies received and mesenchymal CTCs had the most significant association (p = 0.001). When total CTCs were included for analysis, a stronger association was observed (p = 0.0001) (**Fig. 2B, Table S5**). After multiple test correction, both epithelial and mesenchymal subtypes of CTCs alone were correlated with higher GS.

When the 38 untreated patients with localized disease were divided into low/intermediate and high risk groups based on NCCN guideline, the number of each subtype of CTCs was higher in high risk group compared to low/intermediate group (**Fig. 2C, Table S5**). After adjustment for multiple tests, EMTing CTCs and the traditional CTCs (epithelial and EMTing CTCs) still remained statistical significance.

EMTing CTC number was significantly associated with the presence of metastasis

When we correlated the numbers of different sub-populations of CTCs to cancer metastasis, the presence of metastases was significantly associated with higher number of any type of CTCs and the association with EMTing CTCs was most significant (p = 0.0001), which was similar in significance when considering total CTCs (**Fig. 2D, Table S5**). Each subtype of CTCs alone achieved a statistical significance after p value adjustment for multiple tests. When a ROC curve was applied to explore the role of CTCs in metastases prediction in comparison with PSA, EMTing CTC count (with an optimal cutoff point at ≥ 2 cells) had the highest area under the ROC curve (AUC) score (0.755, 95%CI: 0.654-0.856) of all types of CTCs. Although the efficiency was lower than PSA (AUC =0.823, 95%CI: 0.720-0.927, with an optimal cutoff point at ≥ 23 ng/mL)(**Table S6**), when EMTing CTC count and PSA were combined to create a CRS (as CRS = 0.012 * PSA + 0.115 * EMTing CTC count), the CRS

(AUC = 0.921, 95%CI: 0.858-0.985) was a significantly (p = 0.03) better predictor of metastasis than PSA score alone (**Fig. 2E**). The increased performance of the CRS for metastasis prediction compared to PSA at different sensitivity and specificity level are presented in **Table S7**. Box plots of CRS in patients with/without metastasis are shown in **Fig. S2A**.

Mesenchymal CTC count is potentially associated with poor survival

Survival analysis was performed in the 40 metastatic CRPC (mCRPC) patients with a median follow-up time of 11.8 months (range 0.7-19.7 months), of which 11 (27.5%) had died. Median follow-up time for patients still alive at endpoint was 13.3 months (range 4.1-19.7 months). **Table 2** shows the association between potential risk factors and survival derived from Cox model with adjusted p value using Benjamini & Hochberg's method (15). High PSA level (p = 0.015), high serum ALP level (p = 0.015), and high serum LDH level (p = 0.040) were found to be associated with increased risk of death in univariate analyses. None of the subtypes of CTCs had a significant association in this small cohort. However, when a cut-off was selected to optimally demonstrate the significance (**Fig. S3**), patients with mesenchymal CTCs (≥ 5 cells) had an HR of 8.458 (95%CI: 1.815-39.411, p = 0.001). Patients with epithelial (≥ 3 cells), EMTing (≥ 2 cells) and total CK+ (≥ 6 cells) CTCs had an HR of 2.765 (95%CI: 0.732-10.438), 1.858 (95%CI: 0.491-7.033), and 1.537 (95%CI: 0.449-5.261), respectively, none reached statistical significance.

The presence of circulating megakaryocytes was associated with better survival

Unexpectedly, we found a rare population of cells in the harvest samples, with big nuclei (larger than both lymphocytes and most CTCs) of strong DAPI staining, but negative for CK, VIM and CD45 (called BigNeg below)(**Fig. 3A**). Due to the strong DAPI staining and larger size of nuclei, such cells were easily identified from other cells with DAPI staining alone

(**Fig. 3B**). Through the absence of CK, VIM and CD45 staining, it was possible to exclude CTCs. FISH analysis for these cells showed 100% polyploidy and most (61.9%) of them had more than ten copies of the genome in each cell (**Fig. 3C**), suggesting a special type of cells, distinct to lymphocytes or CTCs. The mean nuclear diameter of BigNeg cells was 17.5 ± 4.9 μm (range: 10 – 32 μm), while the mean nuclear diameter for lymphocytes was 6.5 ± 0.8 μm (range: 4 – 8 μm). No overlapping in nuclear size was observed between these two categories (**Fig. 3D**). Therefore, BigNeg cells were distinguishable from other cells by means of immunostaining and size measurement using the following criteria: cells with a strong DAPI stained nucleus ≥ 10 μm in size and negative for CD45, CK, and VIM staining. Using these criteria, such cells were counted in all 81 blood samples. The numbers were variable and lower in mCRPC patients who died during follow-up (median 1 cell/7.5 mL, IQR: 2-1 cells/7.5mL) than those mCRPC patients still alive (median 2 cells/7.5 mL, IQR: 4-1 cells/7.5 mL) and also than patients with untreated localized disease (median 2 cells/7.5 mL, IQR: 4.25-0 cells/7.5 mL), though none of the differences reached significance (p value of 0.18 and 0.33, respectively). In the 40 mCRPC patients, high BigNeg cell count had a trend to be associated with better survival (HR: 0.849, 95%CI: 0.628-1.146, adjusted p = 0.31, **Table 2**). When a cut-off was selected, patients with BigNeg (≥ 3 cells) had an HR of 0.144 (95% CI: 0.018-1.129, p = 0.02). The estimated survival rates were significantly higher in those with ≥ 3 BigNeg cells by Kaplan-Meier curve (p = 0.03, **Fig. 4A**). Based on previous results that high mesenchymal CTC counts generally represent more aggressive tumors and are associated with a higher risk for death, we hypothesize that the difference between the number of mesenchymal CTCs and BigNeg cells may further enhance the ability to predict survival. Using the formula (Mesenchymal CTC count – BigNeg cell count)/BigNeg cell count as detailed in the method and a univariate Cox model for survival analysis, the HR for per unit increase of the score is 1.282 (95%CI: 1.097-1.499, adjusted p = 0.015) (**Table 2**).

When a cut-off was selected, patients with CMS ($\geq$ 2.0) had an HR of 10.170 (95% CI: 2.164-47.789, p = 0.0005) and by Kaplan-Meier survival analysis, a CMS 2.0 cut off most significantly separated different survival groups (p = 0.0003, **Fig. 4B**).

In blood samples from the 24 healthy donors, the median number of BigNeg cells was 1 cell/7.5 mL (range: 0-15, IQR: 2.5-0, **Fig. S2B**), which was slightly lower than that in the 38 patients with untreated localized disease (median: 2, range: 0-20, IQR: 4.25-0, **Fig. S2B**) (p = 0.19) and marginally significantly lower than that in 43 CRPC patients (median: 2, range: 0-23, IQR: 3-1, **Fig. S2B**) (p = 0.078).

Finally, we characterized the BigNeg cells. Their consistent hyperploidy revealed by FISH analysis highly suggested a potential megakaryocyte-like origin. We established the megakaryocytes immunofluorescence analysis method by detecting CD34 and CD41 expression in megakaryocytes induced by phorbol myristate acetate from K562 cells (**Fig. 4C**). All BigNeg cells were CD34 positive with 92% (101/110) CD41 positive, confirming them as a specific group of megakaryocytes (**Fig. 4D**). CD45+ lymphocytes, epithelial CTCs, EMTing CTCs, and mesenchymal CTCs were all negative for CD34 and CD41.

Discussion

We have previously shown that CD45 negative cells with both epithelial and mesenchymal features can be detected in prostate cancer patients using Parsortix™ (9). Here, we developed a repeated FISH analysis approach following immunofluorescence analysis on the same slide to investigate multiple genomic regions for alterations and, consequently, establish the malignancy of VIM+/CD45- cell populations, which were associated with disease burden and poor prognosis and have the potential to serve as an additional marker to traditional epithelial CTCs for clinical application. As EMT is a critical step for cancer metastasis, analyzing EMT CTCs has great potential for cancer prognosis and progression monitoring (7,17). Several non-EpCAM based CTC isolation platforms have been developed to analyze CTCs with

EMT, such as ISET (18,19) and Vitatex CAM platform (20). CTCs with ongoing EMT or completely changed into mesenchymal features have been identified and their biological and genetic difference from CellSearch captured CTCs have been revealed. To our knowledge, the present study is the first genetic investigation of EMT CTCs using repeated FISH and the first investigation of potential prognostic value of EMT CTCs in prostate cancer.

Repeated FISH analysis of CTCs generates genomic alteration information for multiple genomic regions in single cells, enabling the confirmation of the malignant genomic feature of suspicious CTCs, investigation of the heterogeneity of cancer cells by analyzing the differences in genomic alterations between individual cells and correlation of genomic alterations with cellular features and different types of CTCs to understand mechanisms of metastases. FISH analysis performed on cells analyzed by immunostaining (21-24) have been previously performed by using traditional fixation with ethanol or methanol mixed with acetic acid (24) and treatment with pepsin (21) to remove immunofluorescence signals. However, based on our experience, it was difficult to achieve complete removal of previous immunofluorescence signals using these pre-treatments and the leftover fluorescence signals often interfered with FISH signal interpretation. Our stripping buffer method, robustly removing immunofluorescence signals completely without damaging cell morphology, facilitates the analysis of multiple genomic alterations on the same cells after immunofluorescence, which increases the chance to detect genomic alterations in most CTCs to confirm their malignancy nature. A panel of genes is better than a single gene or genetic change for cancer prognosis (25). Multiple genomic region analysis also makes it possible to combine a number of genomic changes for the development of CTC-based genetic prognostic biomarkers. However, in this study, FISH was primarily used to confirm the malignancy of EMT CTCs. We observed a high frequency of alterations in CTCs in the genomic regions investigated, suggesting that extensive genomic alterations occur in CTCs at the stage of

cancer metastasis. This is consistent with a previous report that metastatic prostate cancer has more genomic alterations than primary tumor (26). Our results also showed that genomic heterogeneity exists in prostate cancer CTCs.

In this study, we demonstrated the association of CTC number with GS and the aggressiveness of localized disease, which have not been reported previously. Association of detection of $\geq$ 5 traditional CellSearch[®] isolated CTCs/7.5mL with poor prognosis in patients with advanced prostate cancer was initially reported around a decade ago (27-29) and have been recently further validated in several large-scaled clinical trials (30-33). In our small patient cohort, we observed the associations of $\geq$ 3 epithelial CTCs (CK+/VIM-/CD45-), $\geq$ 2 EMTing CTCs (CK+/VIM+/CD45-), and $\geq$ 6 traditional CTCs (combination of epithelial and EMTing type) with advanced clinical features, CRPC development and poorer survival, which is consistent with these previous reports analyzing epithelial CTCs. By analyzing mesenchymal CTCs, we detected an increase of CTC positive cases in both CRPC and localized diseases. Of interest, high baseline mesenchymal CTC count ($\geq$ 5 cells/7.5mL) before an alternative treatment in CRPC patients was better associated with poorer survival than the traditional CK+ CTCs in our cohort, which warrants further investigation. Moreover, mesenchymal CTCs associated with PSA level and GS better than epithelial CTCs and EMTing CTCs were most significantly correlated to cancer metastasis. All these data indicate a potential better prognostic value of mesenchymal than epithelial CTCs in prostate cancer. As the number of patients in this study is still small, these associations should be validated in large cohorts of samples. EMT has been increasingly recognized for its role in tumor cell invasion, metastatic dissemination and acquisition of therapeutic resistance (6). Loss of epithelial markers, such as E-cadherin, and gain of mesenchymal markers, such as VIM and N-cadherin, have been proven to be associated with more invasive phenotype or higher GS in prostate cancer cells (34,35). EMTing CTCs were most significantly associated with both

high-risk localized disease and metastasis, suggesting that these cells under active EMT process are potential indicators for cancer aggressive invasion and metastasis. Taken together, compared to traditional CK+ CTCs, CTCs undergoing or undergone EMT provided more information about disease burden and intrinsic tumor biology.

Radionuclide bone-scan and computed tomography, the current gold standard procedures to detect metastatic sites, are costly, time-consuming and expose patients to radiation. Patients who are unlikely to have metastasis are better off avoiding these costly and potentially harmful procedures. PSA level has been used to predict bone metastases and select patients for bone scans (16). While our data also showed a strong association between PSA and metastasis and none of the CTC analyses taken alone outperformed PSA for metastatic correlation, a CRS based on both PSA level and EMTing CTC count, significantly improved the metastasis prediction accuracy compared to PSA alone. This suggests that although PSA level correlates with CTC count, they are independent factors contributing to cancer metastasis, potentially by representing different aspects of tumor biology. Large-scale trials are warranted to confirm this superior metastasis prediction model.

The most striking finding is the unexpected discovery of increased circulating megakaryocytes in cancer patients and its association with the survival of CRPC patients. As the nuclei of megakaryocytes are big and with strong DAPI staining, they can be easily identified during CTC analysis. A group of large hyperploid cells with no detectable biomarkers (including epithelial and hepatocellular cancer specific markers) have also been reported recently (36). As the authors did not include any mesenchymal markers, they considered all these cells as CTCs with EMT. Based on our study, these cells might be a mixture of mesenchymal CTCs and circulating megakaryocytes. Megakaryocytes have also been identified in a previous CTC study of prostate cancer, but the investigators exclude them for further analysis (23). Taking advantage of the size and/or deformability based CTC

isolation platform, for example Parsortix™ here, such cells can be captured with high efficiency for cancer prognosis analysis. With the development of multiple treatment approaches for CRPC patients, real-time prognostic and therapeutic response predictive biomarkers are critical for stratified patient management. However, currently only CellSearch based CTC analysis, which mainly detects cells with epithelial features, is approved by FDA for CRPC patient survival prediction (2). Here we not only showed that mesenchymal CTCs has the potential to be better associated with patient survival than CTCs with epithelial features, but also found that the combination of mesenchymal CTC and megakaryocyte counts has a great power to predict CRPC patient survival, with 1.28-fold risk of death per unit increase of the CMS and 10-fold risk for patients with a score ≥ 2.0.

In addition to the bone marrow and the spleen, megakaryocytes have also been detected at very low numbers in circulation (37,38). CD34 expresses in hematopoietic stem cells and megakaryocyte progenitor cells, which are all diploidy cells (39). Its expression can also be seen on mature-appearing megakaryocytes in reactive and disorders of bone marrows (40). CD41 is a megakaryocytic lineage specific marker, expressing neither in endothelial cells nor monocytes (39,41,42). The circulating megakaryocytes detected in our study with high DNA content and consistent expression of CD34, represent a specific and potentially abnormal type of megakaryocytes, in which endonuclear DNA replication is active but the CD34 expression is not switched off, hence premature high DNA content megakaryocytes without apparent lobulated nuclei. This is consistent to the finding of CD34 expression in mature-appearing megakaryocytes in reactive and disorders of bone marrows (40) and increased ploidization of megakaryocytes have been reported in patients with metastatic tumors (43). Platelets, which are released from megakaryocytes in bone marrow has recently been recognized to play an important role in cancer metastasis (44,45) and the gene expression profile of platelets has been shown to efficiently distinguish individuals with or without

cancers (46). Platelet counts have also been investigated for cancer prognosis and the data are still conflict (47). In this study, there was no correlation of the number of platelets with either circulating megakaryocytes or cancer prognosis.

The association of high number of megakaryocytes with a good survival of CRPC patients suggests an anti-tumor effect of megakaryocytes. To support this, megakaryocytes have been reported to increase their number in response to cancer bone metastasis and inhibit prostate and breast cancer cell growth both *in vitro* and *in vivo* (48-50). While the prognostic value and biological functions of megakaryocytes in cancer development and progression need to be further investigated, models that analyze the number difference between CTCs and megakaryocytes should be developed to make an efficient biomarker for survival prediction for patients with mCRPCs and potentially for other cancers as well. With further validation and cut off optimization in large sample cohorts, this study has potential to be translated into clinical use, for patients not only with prostate cancer but also a range of other human malignancies.

References:

1.	Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, *et al.* Tumour evolution inferred by single-cell sequencing. Nature **2011**;472(7341):90-4.

2.	Xu L, Shamash J, Lu YJ. Circulating Tumor Cells: a window to understand cancer metastasis, monitor and fight against cancers. Journal of Cancer Research Updates **2015**;4(1):13-29.

3.	Husemann Y, Geigl JB, Schubert F, Musiani P, Meyer M, Burghart E, *et al.* Systemic spread is an early step in breast cancer. Cancer Cell **2008**;13(1):58-68.

4.	Lucci A, Hall CS, Lodhi AK, Bhattacharyya A, Anderson AE, Xiao L, *et al.* Circulating tumour cells in non-metastatic breast cancer: a prospective study. Lancet Oncol **2012**;13(7):688-95.

5.	Ruscetti M, Quach B, Dadashian EL, Mulholland DJ, Wu H. Tracking and Functional Characterization of Epithelial-Mesenchymal Transition and Mesenchymal Tumor Cells during Prostate Cancer Metastasis. Cancer Res **2015**;75(13):2749-59.

6.	Polyak K, Weinberg RA. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. Nat Rev Cancer **2009**;9(4):265-73.

7.	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(6119):580-4.

8.	Armstrong AJ, Marengo MS, Oltean S, Kemeny G, Bitting RL, Turnbull JD, *et al.* Circulating tumor cells from patients with advanced prostate and breast cancer display both epithelial and mesenchymal markers. Mol Cancer Res **2011**;9(8):997-1007.

9.	Xu L, Mao X, Imrali A, Syed F, Mutsvangwa K, Berney D, *et al.* Optimization and Evaluation of a Novel Size Based Circulating Tumor Cell Isolation System. PLoS One **2015**;10(9):e0138032.

10.	Chudziak J, Burt DJ, Mohan S, Rothwell DG, Mesquita B, Antonello J, *et al.* Clinical evaluation of a novel microfluidic device for epitope-independent enrichment of circulating tumour cells in patients with small cell lung cancer. Analyst **2016**;141(2):669-78.

11.	Hvichia GE, Parveen Z, Wagner C, Janning M, Quidde J, Stein A, *et al.* A novel microfluidic platform for size and deformability based separation and the subsequent molecular characterization of viable circulating tumor cells. Int J Cancer **2016**;138(12):2894-904.

12.	Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, *et al.* Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer **2013**;49(6):1374-403.

13.   Society AC. Cancer Facts & Figures 2016. Atlanta: American Cancer Society **2016**.

14.   D'Amico AV, Whittington R, Malkowicz SB, Schultz D, Blank K, Broderick GA*, et al.* Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA **1998**;280(11):969-74.

15.   Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical andpowerful approach to multiple testing. J Roy Statist Soc Ser B (Methodological) **1995**;57:289-300.

16.   Mohler JL, Armstrong AJ, Bahnson RR, D'Amico AV, Davis BJ, Eastham JA*, et al.* Prostate Cancer, Version 1.2016. J Natl Compr Canc Netw **2016**;14(1):19-30.

17.   Satelli A, Mitra A, Brownlee Z, Xia X, Bellister S, Overman MJ*, et al.* Epithelial-mesenchymal transitioned circulating tumor cells capture for detecting tumor progression. Clin Cancer Res **2015**;21(4):899-906.

18.   Lecharpentier A, Vielh P, Perez-Moreno P, Planchard D, Soria JC, Farace F. Detection of circulating tumour cells with a hybrid (epithelial/mesenchymal) phenotype in patients with metastatic non-small cell lung cancer. Br J Cancer **2011**;105(9):1338-41.

19.   Massard C, Oulhen M, Le Moulec S, Auger N, Foulon S, Abou-Lovergne A*, et al.* Phenotypic and genetic heterogeneity of tumor tissue and circulating tumor cells in patients with metastatic castration-resistant prostate cancer: A report from the PETRUS prospective study. Oncotarget **2016**;7(34):55069-82.

20.   Friedlander TW, Ngo VT, Dong H, Premasekharan G, Weinberg V, Doty S*, et al.* Detection and characterization of invasive circulating tumor cells derived from men with metastatic castration-resistant prostate cancer. Int J Cancer **2014**;134(10):2284-93.

21.   Fehm T, Sagalowsky A, Clifford E, Beitsch P, Saboorian H, Euhus D*, et al.* Cytogenetic evidence that circulating epithelial cells in patients with carcinoma are malignant. Clin Cancer Res **2002**;8(7):2073-84.

22.   Shaffer DR, Leversha MA, Danila DC, Lin O, Gonzalez-Espinoza R, Gu B*, et al.* Circulating tumor cell analysis in patients with progressive castration-resistant prostate cancer. Clin Cancer Res **2007**;13(7):2023-9.

23.   Leversha MA, Han J, Asgari Z, Danila DC, Lin O, Gonzalez-Espinoza R*, et al.* Fluorescence in situ hybridization analysis of circulating tumor cells in metastatic prostate cancer. Clin Cancer Res **2009**;15(6):2091-7.

24. Attard G, Swennenhuis JF, Olmos D, Reid AH, Vickers E, A'Hern R, *et al.* Characterization of ERG, AR and PTEN gene status in circulating tumor cells from patients with castration-resistant prostate cancer. Cancer Res **2009**;69(7):2912-8.

25. Bostrom PJ, Bjartell AS, Catto JW, Eggener SE, Lilja H, Loeb S, *et al.* Genomic Predictors of Outcome in Prostate Cancer. Eur Urol **2015**;68(6):1033-44.

26. Hong MKH, Macintyre G, Wedge DC, Van Loo P, Patel K, Lunke S, *et al.* Tracking the origins and drivers of subclonal metastatic expansion in prostate cancer. Nat Commun **2015**;6:6605.

27. Danila DC, Heller G, Gignac GA, Gonzalez-Espinoza R, Anand A, Tanaka E, *et al.* Circulating tumor cell number and prognosis in progressive castration-resistant prostate cancer. Clin Cancer Res **2007**;13(23):7053-8.

28. de Bono JS, Scher HI, Montgomery RB, Parker C, Miller MC, Tissing H, *et al.* Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. Clin Cancer Res **2008**;14(19):6302-9.

29. Scher HI, Jia X, de Bono JS, Fleisher M, Pienta KJ, Raghavan D, *et al.* Circulating tumour cells as prognostic markers in progressive, castration-resistant prostate cancer: a reanalysis of IMMC38 trial data. Lancet Oncol **2009**;10(3):233-9.

30. Vogelzang NJ, Fizazi K, Burke JM, De Wit R, Bellmunt J, Hutson TE, *et al.* Circulating Tumor Cells in a Phase 3 Study of Docetaxel and Prednisone with or without Lenalidomide in Metastatic Castration-resistant Prostate Cancer. Eur Urol **2017**;71(2):168-71.

31. Scher HI, Heller G, Molina A, Attard G, Danila DC, Jia X, *et al.* Circulating tumor cell biomarker panel as an individual-level surrogate for survival in metastatic castration-resistant prostate cancer. J Clin Oncol **2015**;33(12):1348-55.

32. Goldkorn A, Ely B, Quinn DI, Tangen CM, Fink LM, Xu T, *et al.* Circulating tumor cell counts are prognostic of overall survival in SWOG S0421: a phase III trial of docetaxel with or without atrasentan for metastatic castration-resistant prostate cancer. J Clin Oncol **2014**;32(11):1136-42.

33. Lorente D, Olmos D, Mateo J, Bianchini D, Seed G, Fleisher M, *et al.* Decline in Circulating Tumor Cell Count and Treatment Outcome in Advanced Prostate Cancer. Eur Urol **2016**;70(6):985-92.

34. Singh S, Sadacharan S, Su S, Belldegrun A, Persad S, Singh G. Overexpression of vimentin: role in the invasive phenotype in an androgen-independent model of prostate cancer. Cancer Res **2003**;63(9):2306-11.

35. Tran NL, Nagle RB, Cress AE, Heimark RL. N-Cadherin expression in human prostate carcinoma cell lines. An epithelial-mesenchymal transformation mediating adhesion withStromal cells. Am J Pathol **1999**;155(3):787-98.

36. Ogle LF, Orr JG, Willoughby CE, Hutton C, McPherson S, Plummer R*, et al.* Imagestream detection and characterisation of circulating tumour cells - A liquid biopsy for hepatocellular carcinoma? J Hepatol **2016**;65(2):305-13.

37. Psaila B, Lyden D, Roberts I. Megakaryocytes, malignancy and bone marrow vascular niches. J Thromb Haemost **2012**;10(2):177-88.

38. Bojko P, Hester JP, Durett AG, Maadani F, Korbling M, Champlin RE. Identification of megakaryocyte precursors in peripheral blood stem cell collections from normal donors. J Clin Apher **1998**;13(1):7-15.

39. Chen L, Kostadima M, Martens JH, Canu G, Garcia SP, Turro E*, et al.* Transcriptional diversity during lineage commitment of human blood progenitors. Science **2014**;345(6204):1251033.

40. Tang G, Wang SA, Menon M, Dresser K, Woda BA, Hao S. High-level CD34 expression on megakaryocytes independently predicts an adverse outcome in patients with myelodysplastic syndromes. Leuk Res **2011**;35(6):766-70.

41. Ishibashi T, Ruggeri ZM, Harker LA, Burstein SA. Separation of human megakaryocytes by state of differentiation on continuous gradients of Percoll: size and ploidy analysis of cells identified by monoclonal antibody to glycoprotein IIb/IIIa. Blood **1986**;67(5):1286-92.

42. Tomer A, Friese P, Conklin R, Bales W, Archer L, Harker LA*, et al.* Flow cytometric analysis of megakaryocytes from patients with abnormal platelet counts. Blood **1989**;74(2):594-601.

43. Winkelmann M, Stockler J, Grassmuck J, Pfitzer P, Schneider W. Ploidy pattern of megakaryocytes in patients with metastatic tumors with and without paraneoplastic thrombosis and in controls. Haemostasis **1984**;14(6):501-7.

44. Gay LJ, Felding-Habermann B. Contribution of platelets to tumour metastasis. Nat Rev Cancer **2011**;11(2):123-34.

45. Menter DG, Tucker SC, Kopetz S, Sood AK, Crissman JD, Honn KV. Platelets and cancer: a casual or causal relationship: revisited. Cancer Metastasis Rev **2014**;33(1):231-69.

46.   Best MG, Sol N, Kooi I, Tannous J, Westerman BA, Rustenburg F*, et al.* RNA-Seq of Tumor-Educated Platelets Enables Blood-Based Pan-Cancer, Multiclass, and Molecular Pathway Cancer Diagnostics. Cancer Cell **2015**;28(5):666-76.

47.   Pang Q, Liu C, Qu K, Liu S, Berasain C. Conflicting relationship between platelets and prognosis of hepatocellular carcinoma: is platelet-derived serotonin involved in? Liver Int **2015**;35(11):2484.

48.   Zaslavsky A, Baek KH, Lynch RC, Short S, Grillo J, Folkman J*, et al.* Platelet-derived thrombospondin-1 is a critical negative regulator and potential biomarker of angiogenesis. Blood **2010**;115(22):4605-13.

49.   Li X, Koh AJ, Wang Z, Soki FN, Park SI, Pienta KJ*, et al.* Inhibitory effects of megakaryocytic cells in prostate cancer skeletal metastasis. J Bone Miner Res **2011**;26(1):125-34.

50.   Jackson W, 3rd, Sosnoski DM, Ohanessian SE, Chandler P, Mobley A, Meisel KD*, et al.* Role of Megakaryocytes in Breast Cancer Metastasis to Bone. Cancer Res **2017**;77(8):1942-54.

Table 1. Summary of clinical characteristics and CTC count for all patients

	All patients, n = 81	Patient with untreated localized disease, n = 38	Patient with CRPC, n = 43
Age, y			
Median (IQR)	69 (76-62)	66.5 (72-57)	73 (81-67)
PSA, ng/mL			
Median (IQR)	15 (71-7.65)	8.91 (12.25-5.40)	61 (367-23)
Primary Gleason Score, n (%)			
6	13 (16)	11 (29)	2 (5)
3+4	17 (21)	14 (37)	3 (7)
4+3	14 (17)	9 (24)	5 (12)
8-10	24 (30)	4 (10)	20 (46)
n/a	13 (16)	0 (0)	13 (30)
Prior therapy, n (%)			
No treatment	38 (47)	38 (100)	0 (0)
Systemic therapy	43 (53)	0 (0)	43 (100)
Metastasis, n (%)			
No	41 (51)	38 (100)	3 (7)
Yes	40 (49)	0 (0)	40 (93)
CK+/VIM-/CD45- cell			
Detectable patient number (%)	49 (60)	18 (47)	31 (72)
Cell number, median (IQR)	1 (4-0)	0 (2-0)	2 (9-0)
CK+/VIM+/CD45- cell			
Detectable patient number (%)	37 (46)	7 (18)	30 (70)
Cell number, median (IQR)	0 (2-0)	0 (0-0)	2 (4-0)
CK-/VIM+/CD45- cell			
Detectable patient number (%)	59 (73)	21 (55)	38 (88)
Cell number, median (IQR)	2 (6-0)	1 (4-0)	4 (8-2)
Total CTC			
Detectable patient number (%)	73 (90)	30 (79)	43 (100)
Cell number, median (IQR)	6 (14-3)	3 (6-1)	11 (19-7)

IQR: interquartile range (Q75-Q25%); CRPC: castration-resistant prostate cancer; CK: cytokeratin; VIM: Vimentin.

Table 2. Predictors for survival in progressive mCRPC patients at baseline

	HR (95% CI)	LR χ^2 test	adjusted p*
Age (y)	0.994 (0.923-1.070)	0.02	0.875
PSA (ng/mL)	1.0006 (1.0003-1.0009)	9.84	0.015
GS	1.145 (0.598-2.193)	0.17	0.804
ALP (U/L)	1.003 (1.001-1.004)	8.77	0.015
LDH (U/L)	1.003 (1.001-1.006)	6.28	0.040
Epithelial CTCs (n)	1.005 (0.972-1.039)	0.08	0.847
EMTing CTCs (n)	1.066 (0.948-1.198)	0.95	0.477
Mesenchymal CTCs (n)	1.046 (1.008-1.086)	3.78	0.135
Total CK+ CTCs (n)	1.008 (0.979-1.038)	0.24	0.804
Total VIM+ CTCs (n)	1.032 (1.003-1.062)	3.21	0.158
Total CTCs (n)	1.014 (0.995-1.035)	1.58	0.338
BigNeg cells (n)	0.849 (0.628-1.146)	1.93	0.306
CMS	1.282 (1.097-1.499)	8.59	0.015

* Benjamini & Hochberg' adjustment (15)

HR: hazard ratio; GS: Gleason Score of the primary tumor; CI: confidential interval; ALP: alkaline phosphatase; LDH: lactate dehydrogenase; BigNeg cells: CK-/VIM-/CD45- cells with big nuclei; CMS: Combined mesenchymal CTC and Megakaryocyte score.

Figure Legends

Fig. 1. Representative images for different populations of detected cells in prostate cancer patients and five rounds of FISH in a CTC. (A) Immunofluorescence image for three types of circulating cells. Top row: One CK+/VIM-/CD45- cell (arrowed) adjacent to one CD45+ lymphocyte. Middle row: Two CK+/VIM+/CD45- cells (arrowed) adjacent to one CD45+ lymphocyte. Bottom row: Two CK-/VIM+/CD45- cells (arrowed) adjacent to one CD45+ lymphocyte. (B) Five rounds of FISH on one CK-/VIM+/CD45- CTC post-immunostaining. First round of FISH: *AR* (red) and 6q16 (green). Second round of FISH for *ERG* rearrangement: RP11-476D17 (red) and RP11-95I21 (green). Third round of FISH: *C-MYC* (red) and *NKX3.1* (green). Fourth round of FISH: *RB1* (red) and *PTEN* (green). Fifth round of FISH: *CCND1* (red) and 16q22.1 (green). FISH signals are indicated by arrows.

Fig. 2. The correlation of CTC count with clinical features. (A) Association of PSA with epithelial CTCs, EMTing CTCs, mesenchymal CTCs, total CK+ CTCs, total VIM+ CTCs and total CTCs. The numbers of all types of CTCs were higher in patients with higher GS of primary tumor (B), patients with untreated high-risk localized disease (C), and patients with metastases (D). (E) ROC analysis of the efficiencies of serum PSA level (AUC = 0.823) and CRS (AUC = 0.921) in discriminating metastatic prostate cancer patients from those without metastasis are shown in crosses and black dots respectively. An AUC equal to 1 suggests perfect discrimination.

Fig. 3. Representative immunofluorescence/FISH images and nuclear size of BigNeg cells. (A) An example of a BigNeg cell (arrowed) with a big and bright Nucleus and negative signals for CD45, CK and VIM and two adjacent CD45+/VIM+ lymphocytes. (B) BigNeg cells (arrowed) could be easily screened and identified under low resolution image based on nucleus size alone. (C) FISH analysis by probes of *AR* (red) and 6q16 (green) showed polyploidy of the BigNeg cell (arrowed) and an adjacent diploid lymphocyte (arrow-head) which was CD45 positive on previous immunostaining. (D) Comparison of nucleus diameter between BigNeg cells and lymphocytes where no size overlap was observed.

Fig. 4. BigNeg cell count was associated with survival and their megakaryocyte nature was confirmed by immunofluorescence. (A) Kaplan-Meier curve for overall survival showed progressive prostate cancer patients with less than three BigNeg cells had significantly shorter survival rates (p = 0.032). (B) Kaplan-Meier curve for overall survival showed progressive prostate cancer patients with CMS $\geq$ 2.0 had even poorer survival (p = 0.0003). (C) PMA-treated K562 cell lines were positive for CD34 and CD41 staining with various patterns: strong CD41 but weak CD34 (top row), similar signal strength of CD41 and CD34 (middle row), and strong CD34 but negative CD41 (bottom row). Cells with obvious CD41 signals were larger than those without or with very weak CD41 staining. (D) Top and middle row: BigNeg cells were positive for both CD41 and CD34, and signals of CD41 were relatively stronger. Bottom row: one BigNeg cell was positive for CD34 but not CD41. The adjacent CD45+ (image not shown) lymphocyte (with arrow-head) was negative for both CD34 and CD41. BigNeg cells positive for CD41 were relatively larger than those with no CD41 expression.
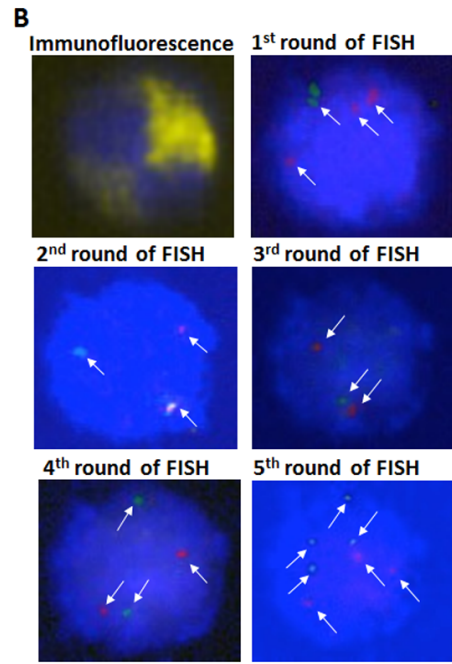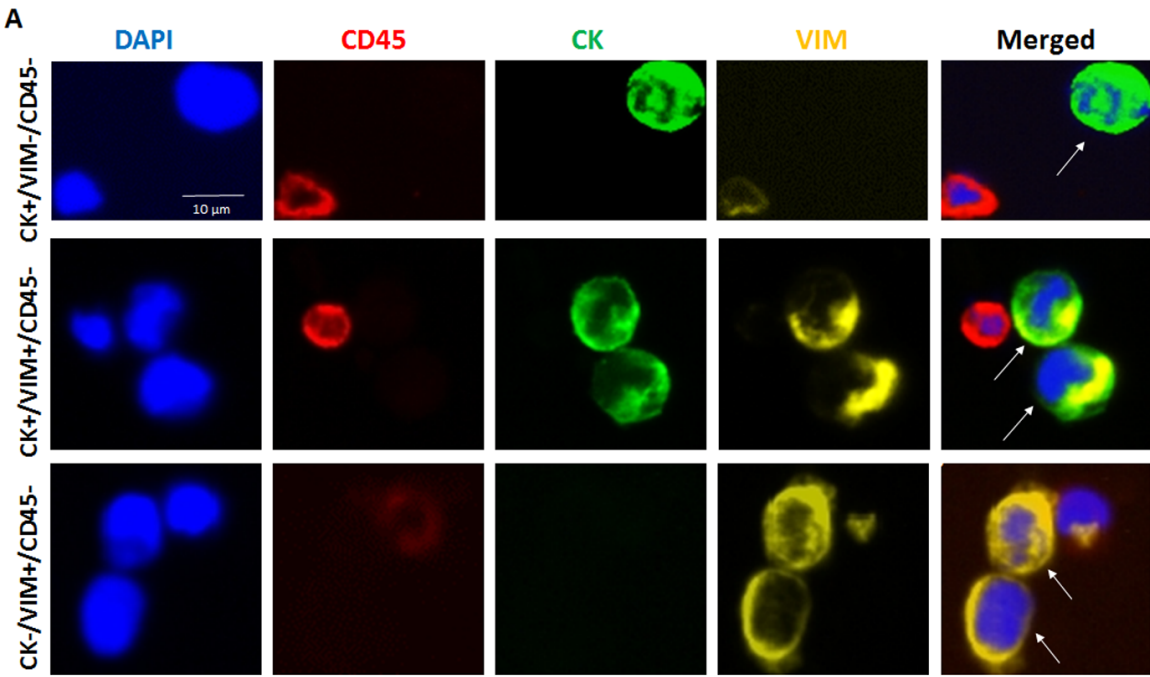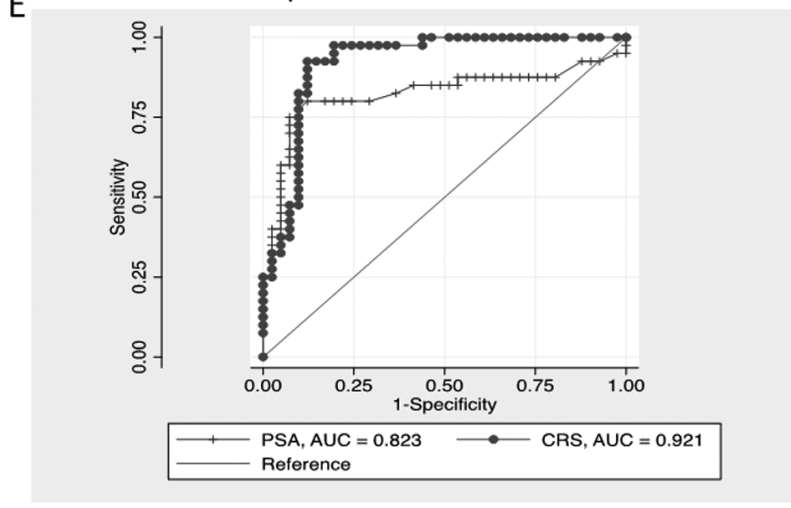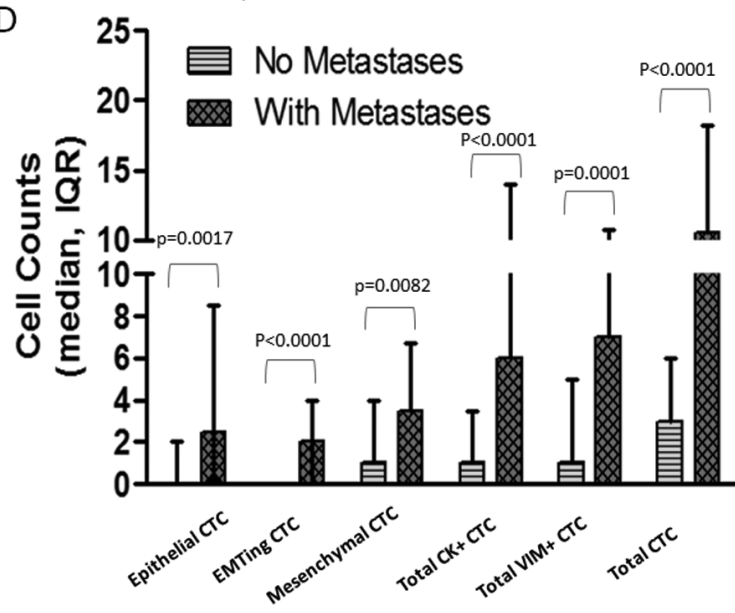
Figure 1

Figure 2

Figure 3

Figure 4