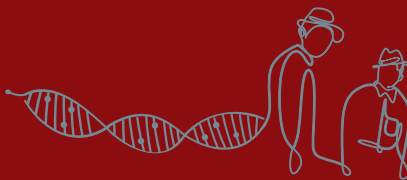# Unravelling the Genomic Landscape of Metastatic Prostate Cancer

a prospect on patient stratification
using blood-based biomarkers

Lisanne Francisca van Dessel

Unravelling the Genomic Landscape of Metastatic Prostate Cancer

a prospect on patient stratification using blood-based biomarkers

Lisanne Francisca van Dessel

Unravelling the Genomic Landscape of Metastatic Prostate Cancer

a prospect on patient stratification
using blood-based biomarkers

Het ontrafelen van het genomische landschap van gemetastaseerd prostaatkanker

een vooruitzicht op patiënt stratificatie door gebruik van bloed-gebaseerde biomarkers

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. F.A. van der Duijn Schouten

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op

dinsdag 1 juni 2021 om 10:30 uur

door

Lisanne Francisca van Dessel

geboren te Rotterdam.

Erasmus University Rotterdam

*"Be brave, be curious,
be determined, overcome the odds.
It can be done."*

- Stephen Hawking

Promotiecommissie

Promotoren
prof. dr. R. de Wit
prof. dr. ir. G.W. Jenster
prof. dr. J.W.M. Martens

Overige leden
prof. dr. C.H. Bangma
prof. dr. A.C. Dingemans
prof. dr. J.A. Schalken

Copromotor
dr. M.P.J.K. Lolkema

Paranimfen
drs. S.C. Eindhoven
drs. L. Angus

Table of Contents
Unravelling the Genomic Landscape of Metastatic Prostate Cancer
a prospect on patient stratification using blood-based biomarkers

Chapter 1

General introduction

Metastatic prostate cancer

Prostate cancer is the most prevalent malignancy in men worldwide and in the Netherlands over 12,000 men were diagnosed with this disease in 2018[1,2]. In general, prostate cancer has an indolent course, but men with de novo or those that progress to metastatic disease represent a subgroup with poor prognosis. Metastatic prostate cancer can be divided into two disease phases: hormone-sensitive prostate cancer (HSPC) that will be followed by castration-resistant prostate cancer (CRPC). The mainstay of therapy for metastatic prostate cancer is targeting the androgen receptor (AR) pathway by reducing circulating testosterone levels (androgen deprivation therapy, ADT). Most patients will initially respond to ADT often for several years (HSPC phase), however, prostate tumor cells will inevitably become resistant to ADT and progress (CRPC phase), leading to considerable morbidity and ultimately to death (Figure 1). For CRPC patients the available treatment options have rapidly evolved and include chemotherapy, second-line AR-targeted therapy, radionucleotides, immunotherapy and molecular-targeted therapy (Table 1). Recently, some of these treatment options have shifted to being applied in the HSPC phase as well. Although the treatment landscape for metastatic prostate cancer has dramatically changed the last decade, selecting the right treatment at the right time in the right order, sequential or in combination for an individual patient remains challenging as we lack predictive and monitoring biomarkers.
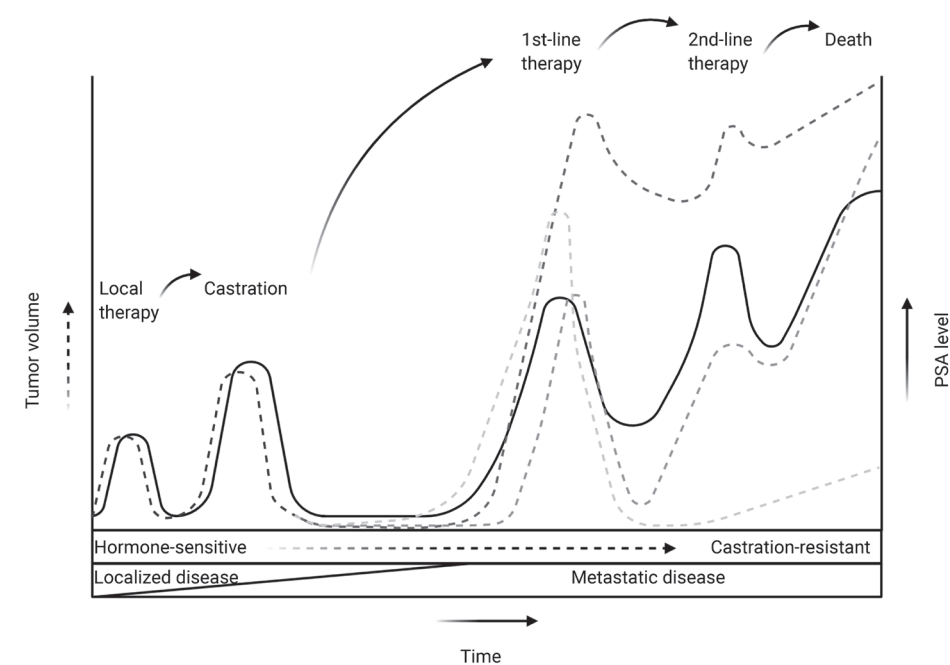

Figure 1. Disease progression of prostate cancer

Table 1. Available treatment options for castration-resistant prostate cancer patients

Chemotherapy	Second-line AR-targeted therapy	Radionucleotides	Immunotherapy	Molecular-targeted therapy	Bone-modifying agents
Docetaxel	Abiraterone acetate	Radium-223	Sipuleucel-T*	Olaparib	Zoledronic acid
Cabazitaxel	Enzalutamide	Strontium-89	*Pembrolizumab*		Denosumab
Mitoxantrone	Apalutamide	Samarium-153			
		Rhenium-188			

*Not available in The Netherlands
Treatments in italics indicate drugs currently investigated in clinical trials

Current biomarkers and clinical trial endpoints

Biomarkers are characteristics that are measured as an indicator of (patho)physiological processes or of responses to an intervention[3]. They comprise clinical, laboratory and imaging measurements, but also genetic and molecular parameters. Furthermore, based on their application several subtypes exist including diagnostic, prognostic, predictive and monitoring biomarkers. Biomarkers can be used as endpoint in clinical trials if they are precisely defined and reflect an outcome of interest that can be analyzed to address a particular research question[3].

Currently, the recommended biomarkers to use in clinical trials for metastatic prostate cancer are clinical measures (symptoms/performance status), blood-based markers (prostate-specific antigen (PSA), alkaline phosphatase, lactate dehydrogenase, complete blood count and circulating tumor cells), imaging modalities (computer tomography (CT) scan/magnetic resonance imaging (MRI) scan and bone scintigraphy scan), and patient-reported outcome measures[4].
Measuring symptoms, performance status and patient-reported outcomes is of importance to assess if a patient is fit for therapy, to assess therapy toxicity and to evaluate therapy response. Importantly, pain is associated with worse survival and adverse quality of life[5,6].
PSA is the most commonly used biomarker in prostate cancer and is used as a diagnostic, prognostic and monitoring marker. However, PSA is a prostate-specific marker, rather than a prostate cancer-specific marker and is in itself an endocrine-dependent enzyme, regulated by an androgen-dependent promoter[7,8]. Thus, PSA levels will be directly affected by ADT and second-line AR-targeted therapies and does not simply reflect suppression of tumor burden. Importantly and seemingly contradictory, in CRPC PSA levels do not always correlate with disease burden and clinical benefit[9]. Therefore, PSA dynamics should be interpreted with caution and it is not recommended to base treatment decisions on PSA dynamics alone[3]. Other blood-based markers are measured at baseline to determine prognosis and are often not used to assess therapy response. Circulating tumor cells will be discussed in more detail below.

Image modalities like CT, MRI and bone scans are useful to evaluate disease burden and to assess (late) therapy response using well-defined criteria. Unfortunately, this can lead to prolonged treatment in unresponsive patients with a potentially toxic and/or expensive agent. Furthermore, bone scans measure bone metabolism using ^{99}mTc-MDP that is incorporated into hydroxyapatite and thus do not directly assess prostate tumor cells themselves but rather their derived effects on bone[10].

Using these currently recommended biomarkers in clinical trials it is evident that they are insufficient to monitor early therapy response and fail to capture the highly dynamic tumor biology. In addition, clinical trial outcome discriminates clear responders from non-responders to tested drugs, suggesting presence of inherent tumor features that determine why some patients fail treatment. Thus, stratification based on specific tumor features that predict treatment response could improve the clinical outcome of patients.

Genomic landscape of metastatic prostate cancer

Over the recent years comprehensive genomic analyses have shown that metastatic prostate cancer harbors a very complex genomic landscape consisting of multiple single nucleotide variants and structural rearrangements[11,12]. Well-known prostate cancer driver genes are *AR*, *PTEN, TP53* and *RB1,* and the *TMPRSS2-ERG* gene fusion[11,13,14]. Still, *AR* is the most widely used actionable genomic target in prostate cancer, although tumors eventually become resistant. Resistance mechanisms roughly encompass two mechanisms: (1) increasing androgen sensitivity by activating mutation, overexpression and alternative splicing of AR; or (2) increasing androgen availability by changes in androgen biosynthesis and metabolism[15-19]. However, sequencing efforts showed that actionable targets are also present in non-AR related pathways, including PI3K, Wnt and DNA repair[12]. In addition, non-coding alterations appear to be of prognostic value and might even have functional implications[11]. Whether these alterations represent driving events or actionable targets is still not clear.

Liquid biopsies

The vastly growing amount of genomic data advances our understanding of the biology of metastatic prostate cancer. It has become clear that metastatic prostate cancer is a highly heterogeneous disease. The genomic makeup of a tumor evolves over time and differs between patients and even between tumor sites within one patient. This urges us to reconsider our current 'gold standard', a tumor biopsy, to acquire genomic information from the tumor. Obtaining tumor biopsies is a burdensome procedure with risk of complications for the patient. Moreover,

metastatic prostate cancer is frequently confined to the bone, making biopsies not always successful (success rate: 25-75%) and not readily performed sequentially during a treatment course[20,21]. In clinical practice, this often leads to a long time interval between the obtained tumor biopsy and the given treatment.

With the emergence of liquid biopsies a new approach to obtain tumor-derived DNA has become available with diagnostic, predictive, prognostic and disease monitoring applications. Liquid biopsies include circulating cell-free DNA (cfDNA) and RNA (cfRNA), circulating tumor cells (CTCs), and exosomes (Figure 2). The main advantage is that liquid biopsies can be obtained in a minimally-invasive and safe way by sampling body fluids like blood, urine and ascites, and thus can be easily repeated at different stages during the disease. Since all tumor sites can potentially release such molecules, it might better reflect the intrapatient heterogeneity. Since this thesis focusses on cfDNA and CTCs, these two biomarkers will be discussed in more detail.
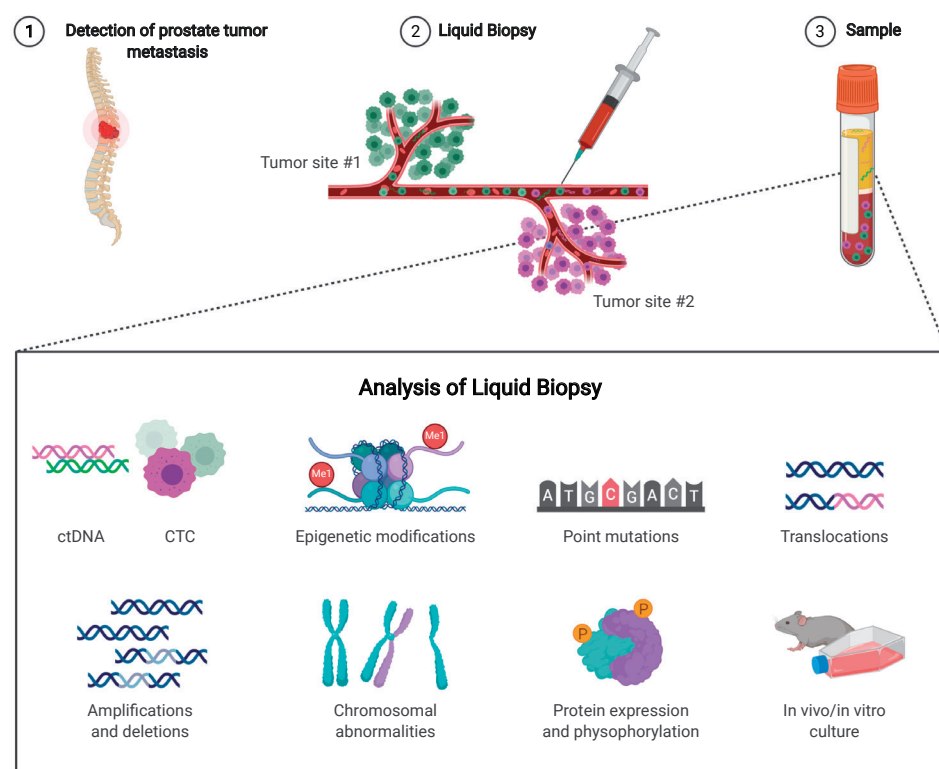
Figure 2. Obtaining a liquid biopsy

Cell-free DNA

Cell-free DNA are small fragments of nucleic acid (180-200 bp lengths) that are released in the peripheral blood circulation by apoptotic or necrotic cells that mostly originate from tissue cells and hematological cells[22-25]. When tumor cells shed tumor DNA into the blood stream it is called circulating tumor DNA (ctDNA). The amount of cell-free DNA measured depends on multiple factors, like (patho)physiological conditions (exercise, disease) and preanalytical conditions (storage temperature, type of blood collection tube) but is usually low, around 10 ng per mL plasma[26]. The fraction of ctDNA depends on tumor type and disease stage and can vary from extremely low (< 0.01%) to very high (60%) but is often less than 1% of the total cfDNA pool[27,28]. Thus, the biggest challenge in clinical practice might be the detection limit. Yet, by the quantification and characterization of ctDNA, detection and follow up of tumor burden, and even stratification for specific therapies is possible[28-35].

Circulating tumor cells

Circulating tumor cells are tumor cells derived from the primary tumor or from its metastases that have been released in the peripheral blood circulation. Since CTCs can be derived from all tumor sites in the body, it is conceived that they capture the genetic and phenotypic heterogeneity of the cancer[36]. As for ctDNA, the detection limit of CTCs is a big challenge since they have a low abundance in blood of generally 1 CTC/mL of blood[37,38]. For CTC detection and quantification, the FDA has approved the CellSearch system, which magnetically separates CTCs from other blood cells using EpCAM-coated ferrofluids. CTCs are the most studied type of liquid biopsy in metastatic prostate cancer and are of prognostic and predictive value. Metastatic patients with ≥ 5 CTCs/7.5 mL of blood have a worse outcome and increasing numbers of CTCs over time also have prognostic significance[39-41]. Expression of AR splice variant 7 (AR-V7), yielding a constitutively active form of AR, in CTCs has been strongly associated with poor response to second-line AR-targeted therapy, like enzalutamide and abiraterone. On the other hand, good response to chemotherapy is observed in patients with AR-V7 expression in CTCs[42-44].

Steroidomics

As all metastatic prostate cancer patients are treated with ADT, circulating androgen levels are below castrate level (50 ng/mL). However, in CRPC patients disease progression occurs at suppressed testosterone levels. Growing evidence support the concept of continued importance of the AR pathway in CRPC that is not only dependent of genomic aberrations, but also involves changes in steroid biosynthesis and metabolism[15-19]. Steroid hormones are small signaling molecules that regulate gene expression through binding and relocation of nuclear receptors[45-47]. Steroid hormones are derived from cholesterol and include five groups, including glucocorticoids

and androgens[48,49]. Steroid measurements are predominantly performed in serum samples and mass spectrometry (MS)-based techniques are the gold standard in clinical diagnostics[50,51].

In metastatic prostate cancer patients, steroid analysis mostly focusses on measuring testosterone levels to evaluate efficacy of ADT. Nevertheless, several reports show that the level of testosterone may be a prognostic and predictive marker as well. Patients with high testosterone levels (> 0.05 ng/mL) before start of therapy seem to have a better outcome on second-line AR-targeted therapy, whereas patients with a low testosterone level do better on taxane-based chemotherapy[52-55]. Thus, a renewed interest has grown for studying steroid metabolites, i.e. steroidomics, in metastatic prostate cancer.

Focus of this thesis

In this thesis, I aim to unravel the genomic landscape of metastatic prostate cancer to understand disease progression (Chapter 2) to expedite the implementation of liquid biopsies in routine cancer care. For the latter as well as for steroid hormone measurements, it is vital to standardize preanalytical conditions of sample collection. My research on optimization of preanalytical conditions can contribute to drafting standard-operating-procedures and guidelines (Chapter 3, 4, 7). To show the potential applications of liquid biopsies in metastatic prostate cancer, ctDNA was analyzed during treatment (Chapter 5), and the ability to derive tumor organoids from CTCs was explored (Chapter 6). Finally, to advance our understanding of the androgen status in metastatic prostate cancer the importance of 11-ketotestosterone is highlighted (Chapter 8).

Acknowledgements

The figures in this chapter were created with Biorender.com.

References

1. Ferlay J EM, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F. Global Cancer Observatory: Cancer Today. International Agency for Research on Cancer. (https://gco.iarc.fr/today).
2. Kankerregistratie N. Cijfers over Kanker. Integraal Kankercentrum Nederland. (www.cijfersoverkanker.nl).
3. Group F-NBW. BEST (Biomarkers, EndpointS, and other Tools) Resource. 2016/03/25 ed: Silver Spring (MD): Food and Drug Administration (US); Bethesda (MD): National Institutes of Health (US), 2016.
4. Scher HI, Morris MJ, Stadler WM, et al. Trial Design and Objectives for Castration-Resistant Prostate Cancer: Updated Recommendations From the Prostate Cancer Clinical Trials Working Group 3. J Clin Oncol 2016;34(12):1402-18. (In eng). DOI: JCO.2015.64.2702 [pii]
10.1200/JCO.2015.64.2702.
5. Halabi S, Vogelzang NJ, Kornblith AB, et al. Pain predicts overall survival in men with metastatic castration-refractory prostate cancer. Journal of Clinical Oncology 2008;26(15):2544-9.
6. Autio KA, Bennett AV, Jia X, et al. Prevalence of pain and analgesic use in men with metastatic prostate cancer using a patient-reported outcome measure. J Oncol Pract 2013;9(5):223-9. (In eng). DOI: JOP.2013.000876 [pii]
10.1200/JOP.2013.000876.
7. Riegman PHJ, Vlietstra RJ, Vanderkorput JAGM, Brinkmann AO, Trapman J. The Promoter of the Prostate-Specific Antigen Gene Contains a Functional Androgen Responsive Element. Mol Endocrinol 1991;5(12):1921-1930. (In English) (<Go to ISI>://A1991HH37200018).
8. Young CYF, Andrews PE, Montgomery BT, Tindall DJ. Tissue-Specific and Hormonal-Regulation of Human Prostate-Specific Glandular Kallikrein. Biochemistry-Us 1992;31(3):818-824. (In English). DOI: Doi 10.1021/Bi00118a026.
9. Armstrong AJ, Eisenberger MA, Halabi S, et al. Biomarkers in the management and treatment of men with metastatic castration-resistant prostate cancer. Eur Urol 2012;61(3):549-59. (In eng). DOI: S0302-2838(11)01242-5 [pii]
10.1016/j.eururo.2011.11.009.
10. Kanishi D. Tc-99m-Mdp Accumulation Mechanisms in Bone. Oral Surg Oral Med O 1993;75(2):239-246. (In English). DOI: Doi 10.1016/0030-4220(93)90100-I.
11. Fraser M, Sabelnykova VY, Yamaguchi TN, et al. Genomic hallmarks of localized, non-indolent prostate cancer. Nature 2017;541(7637):359-364. (In eng). DOI: nature20788 [pii]
10.1038/nature20788.
12. Robinson D. Integrative Clinical Genomics of Advanced Prostate Cancer. Cell 2015;161:1215-1228.
13. Cancer Genome Atlas Research N. The Molecular Taxonomy of Primary Prostate Cancer. Cell 2015;163(4):1011-25. (In eng). DOI: S0092-8674(15)01339-2 [pii]
10.1016/j.cell.2015.10.025.
14. Schoenborn JR, Nelson P, Fang M. Genomic Profiling Defines Subtypes of Prostate Cancer with the Potential for Therapeutic Stratification. Clin Cancer Res 2013;19(15):4058-4066. (In English). DOI: 10.1158/1078-0432.Ccr-12-3606.
15. Taylor BS, Schultz N, Hieronymus H, et al. Integrative Genomic Profiling of Human Prostate Cancer. Cancer Cell 2010;18(1):11-22. (In English). DOI: DOI 10.1016/j.ccr.2010.05.026.
16. Stanbrough M, Bubley GJ, Ross K, et al. Increased expression of genes converting adrenal androgens to testosterone in androgen-independent prostate cancer. Cancer Res 2006;66(5):2815-2825. (In English). DOI: Doi 10.1158/0008-5472.Can-05-4000.
17. Mohler JL, Gregory CW, Ford OH, 3rd, et al. The androgen axis in recurrent prostate cancer. Clin Cancer Res 2004;10(2):440-8. (In eng). DOI: 10.1158/1078-0432.ccr-1146-03.
18. Montgomery RB, Mostaghel EA, Vessella R, et al. Maintenance of intratumoral androgens in metastatic prostate cancer: a mechanism for castration-resistant tumor growth. Cancer Res 2008;68(11):4447-54. (In eng). DOI: 68/11/4447 [pii]
10.1158/0008-5472.CAN-08-0249.
19. Linja MJ, Savinainen KJ, Saramaki OR, Tammela TL, Vessella RL, Visakorpi T. Amplification and overexpression of androgen receptor gene in hormone-refractory prostate cancer. Cancer Res 2001;61(9):3550-5. (In eng) (https://www.ncbi.nlm.nih.gov/pubmed/11325816).
20. Lorente D, Omlin A, Zafeiriou Z, et al. Castration-Resistant Prostate Cancer Tissue Acquisition From Bone Metastases for Molecular Analyses. Clin Genitourin Cancer 2016;14(6):485-493. (In eng). DOI: S1558-7673(16)30106-9 [pii]
10.1016/j.clgc.2016.04.016.

21. McKay RR, Zukotynski KA, Werner L, et al. Imaging, procedural and clinical variables associated with tumor yield on bone biopsy in metastatic castration-resistant prostate cancer. Prostate Cancer Prostatic Dis 2014;17(4):325-31. (In eng). DOI: pcan201428 [pii]
10.1038/pcan.2014.28.

22. Diehl F, Li M, Dressman D, et al. Detection and quantification of mutations in the plasma of patients with colorectal tumors. P Natl Acad Sci USA 2005;102(45):16368-16373. (In English). DOI: DOI 10.1073/pnas.0507904102.

23. Heitzer E, Auer M, Hoffmann EM, et al. Establishment of tumor-specific copy-number alterations from plasma DNA of patients with cancer. Int J Cancer 2013;133(2):346-356. (In English). DOI: Doi 10.1002/Ijc.28030.

24. Jahr S, Hentze H, Englisch S, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. Cancer Res 2001;61(4):1659-65. (In eng) (http://www.ncbi.nlm.nih.gov/pubmed/11245480).

25. Bortner CD, Oldenburg NB, Cidlowski JA. The role of DNA fragmentation in apoptosis. Trends Cell Biol 1995;5(1):21-6. (In eng). DOI: S0962892400889321 [pii].

26. Fleischhacker M, Schmidt B. Circulating nucleic acids (CNAs) and cancer--a survey. Biochim Biophys Acta 2007;1775(1):181-232. (In eng). DOI: S0304-419X(06)00059-X [pii]
10.1016/j.bbcan.2006.10.001.

27. Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 2014;6(224):224ra24. (In eng). DOI: 6/224/224ra24 [pii]
10.1126/scitranslmed.3007094.

28. Diehl F, Schmidt K, Choti MA, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med 2008;14(9):985-90. (In eng). DOI: nm.1789 [pii]
10.1038/nm.1789.

29. Bidard FC, Madic J, Mariani P, et al. Detection rate and prognostic value of circulating tumor cells and circulating tumor DNA in metastatic uveal melanoma. Int J Cancer 2014;134(5):1207-13. (In eng). DOI: 10.1002/ijc.28436.

30. Dawson SJ, Tsui DW, Murtaza M, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med 2013;368(13):1199-209. (In eng). DOI: 10.1056/NEJMoa1213261.

31. Diaz LA, Jr., Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. J Clin Oncol 2014;32(6):579-86. (In eng). DOI: JCO.2012.45.2011 [pii]
10.1200/JCO.2012.45.2011.

32. Forshew T, Murtaza M, Parkinson C, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med 2012;4(136):136ra68. (In eng). DOI: 4/136/136ra68 [pii]
10.1126/scitranslmed.3003726.

33. Murtaza M, Dawson SJ, Tsui DW, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature 2013;497(7447):108-12. (In eng). DOI: nature12065 [pii]
10.1038/nature12065.

34. Shinozaki M, O'Day SJ, Kitago M, et al. Utility of circulating B-RAF DNA mutation in serum for monitoring melanoma patients receiving biochemotherapy. Clin Cancer Res 2007;13(7):2068-74. (In eng). DOI: 13/7/2068 [pii]
10.1158/1078-0432.CCR-06-2120.

35. Elshimali YI, Khaddour H, Sarkissyan M, Wu Y, Vadgama JV. The clinical utilization of circulating cell free DNA (CCFDNA) in blood of cancer patients. Int J Mol Sci 2013;14(9):18925-58. (In eng). DOI: ijms140918925 [pii]
10.3390/ijms140918925.

36. Miyamoto DT, Zheng Y, Wittner BS, et al. RNA-Seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance. Science 2015;349(6254):1351-6. (In eng). DOI: 349/6254/1351 [pii]
10.1126/science.aab0917.

37. Allard WJ, Matera J, Miller MC, et al. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. Clin Cancer Res 2004;10(20):6897-904. (In eng). DOI: 10/20/6897 [pii]
10.1158/1078-0432.CCR-04-0378.

38. Thalgott M, Rack B, Maurer T, et al. Detection of circulating tumor cells in different stages of prostate cancer. J Cancer Res Clin Oncol 2013;139(5):755-63. (In eng). DOI: 10.1007/s00432-013-1377-5.

39. Moreno JG, Miller MC, Gross S, Allard WJ, Gomella LG, Terstappen LW. Circulating tumor cells predict survival in patients with metastatic prostate cancer. Urology 2005;65(4):713-8. DOI: 10.1016/j.urology.2004.11.006.

40. Danila DC, Heller G, Gignac GA, et al. Circulating tumor cell number and prognosis in progressive castration-resistant prostate cancer. Clin Cancer Res 2007;13(23):7053-8. DOI: 10.1158/1078-0432.CCR-07-1506.

41. de Bono JS, Scher HI, Montgomery RB, et al. Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. Clin Cancer Res 2008;14(19):6302-9. (In eng). DOI: 14/19/6302 [pii]
10.1158/1078-0432.CCR-08-0872.

42. Scher HI, Graf RP, Schreiber NA, et al. Assessment of the validity of nuclear-localized androgen receptor splice variant 7 in circulating tumor cells as a predictive biomarker for castration-resistant prostate cancer. JAMA Oncol 2018;4(9):1179-1186. (Article) (In English). DOI: 10.1001/jamaoncol.2018.1621.

43. Onstenk W, Sieuwerts AM, Kraan J, et al. Efficacy of Cabazitaxel in Castration-resistant Prostate Cancer Is Independent of the Presence of AR-V7 in Circulating Tumor Cells. Eur Urol 2015;68(6):939-45. DOI: 10.1016/j.eururo.2015.07.007.

44. Antonarakis ES, Lu C, Wang H, et al. AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. N Engl J Med 2014;371(11):1028-38. DOI: 10.1056/NEJMoa1315815.

45. Norman AW, Mizwicki MT, Norman DP. Steroid-hormone rapid actions, membrane receptors and a conformational ensemble model. Nat Rev Drug Discov 2004;3(1):27-41. (In eng). DOI: 10.1038/nrd1283
nrd1283 [pii].

46. Evans RM. The steroid and thyroid hormone receptor superfamily. Science 1988;240(4854):889-95. (In eng) (https://www.ncbi.nlm.nih.gov/pubmed/3283939).

47. Mangelsdorf DJ, Thummel C, Beato M, et al. The nuclear receptor superfamily: the second decade. Cell 1995;83(6):835-9. (In eng). DOI: 0092-8674(95)90199-X [pii].

48. Miller WL, Auchus RJ. The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. Endocr Rev 2011;32(1):81-151. (In eng). DOI: er.2010-0013 [pii]
10.1210/er.2010-0013.

49. Chien Y, Rosal K, Chung BC. Function of CYP11A1 in the mitochondria. Mol Cell Endocrinol 2017;441:55-61. (In eng). DOI: S0303-7207(16)30441-5 [pii]
10.1016/j.mce.2016.10.030.

50. Taylor AE, Keevil B, Huhtaniemi IT. Mass spectrometry and immunoassay: how to measure steroid hormones today and tomorrow. Eur J Endocrinol 2015;173(2):D1-12. (In eng). DOI: EJE-15-0338 [pii]
10.1530/EJE-15-0338.

51. Monaghan PJ, Keevil BG, Trainer PJ. The use of mass spectrometry to improve the diagnosis and the management of the HPA axis. Rev Endocr Metab Disord 2013;14(2):143-57. (In eng). DOI: 10.1007/s11154-013-9240-1.

52. Hashimoto K, Tabata H, Shindo T, et al. Serum testosterone level is a useful biomarker for determining the optimal treatment for castration-resistant prostate cancer. Urol Oncol Semin Orig Invest 2019;37(7):485-491. (Article) (In English). DOI: 10.1016/j.urolonc.2019.04.026.

53. Shiota M, Kashiwagi E, Murakami T, et al. Serum testosterone level as possible predictive marker in androgen receptor axis-targeting agents and taxane chemotherapies for castration-resistant prostate cancer. Urol Oncol Semin Orig Invest 2019;37(3):180.e19-180.e24. (Article) (In English). DOI: 10.1016/j.urolonc.2018.10.020.

54. Sakamoto S, Maimaiti M, Xu M, et al. Higher serum testosterone levels associated with favorable prognosis in enzalutamide- and abiraterone-treated castration-resistant prostate cancer. J Clin Med 2019;8(4) (Article) (In English). DOI: 10.3390/jcm8040489.

55. Ando K, Sakamoto S, Takeshita N, et al. Higher serum testosterone levels predict poor prognosis in castration-resistant prostate cancer patients treated with docetaxel. Prostate 2020;80(3):247-255. (Article) (In English). DOI: 10.1002/pros.23938.

Part 1

Genomic Landscape of Metastatic Prostate Cancer
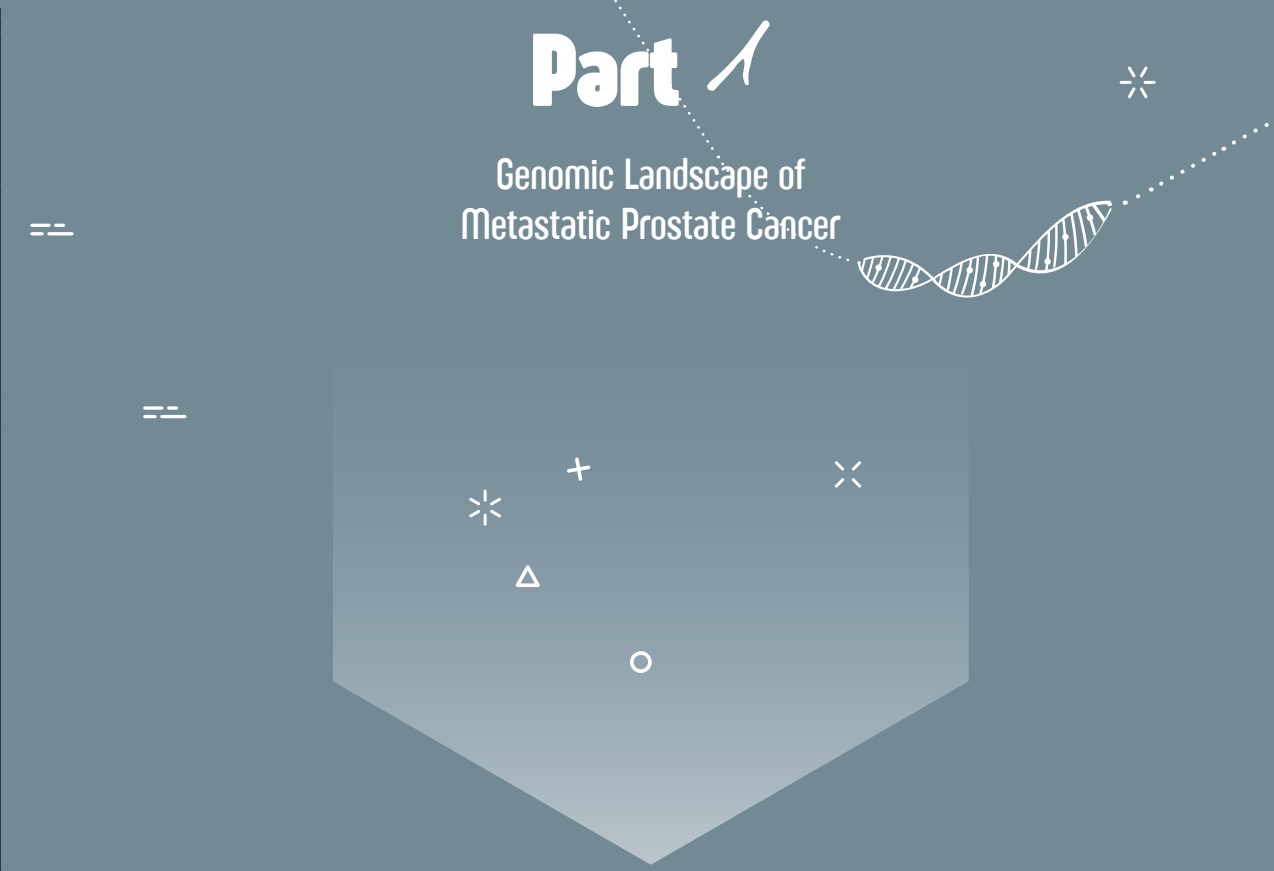
Chapter 2

The genomic landscape of metastatic castration-resistant prostate cancers reveals multiple distinct genotypes with potential clinical impact

Chapter 2

Lisanne F. van Dessel*, Job van Riet*, Minke Smits, Yanyun Zhu, Paul Hamberg, Michiel S. van der Heijden, Andries M. Bergman, Inge M. van Oort, Ronald de Wit, Emile E. Voest, Neeltje Steeghs, Takafumi N. Yamaguchi, Julie Livingstone, Paul C. Boutros, John W.M. Martens, Stefan Sleijfer, Edwin Cuppen, Wilbert Zwart, Harmen J.G van de Werken, Niven Mehra, Martijn P. Lolkema
* These authors contributed equally

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration–resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

Abstract

Metastatic castration-resistant prostate cancer (mCRPC) has a highly complex genomic landscape. With the recent development of novel treatments, accurate stratification strategies are needed. Here we present the whole-genome sequencing (WGS) analysis of freshfrozen metastatic biopsies from 197 mCRPC patients. Using unsupervised clustering based on genomic features, we define eight distinct genomic clusters. We observe potentially clinically relevant genotypes, including microsatellite instability (MSI), homologous recombination deficiency (HRD) enriched with genomic deletions and *BRCA2* aberrations, a tandem duplication genotype associated with *CDK12*[-/-] and a chromothripsis-enriched subgroup. Our data suggests that stratification on WGS characteristics may improve identification of MSI, *CDK12*[-/-] and HRD patients. From WGS and ChIP-seq data, we show the potential relevance of recurrent alterations in non-coding regions identified with WGS and highlight the central role of AR signaling in tumor progression. These data underline the potential value of using WGS to accurately stratify mCRPC patients into clinically actionable subgroups.

Introduction

Prostate cancer is known to be a notoriously heterogeneous disease and the genetic basis for this interpatient heterogeneity is poorly understood[1,2]. The ongoing development of new therapies for metastatic prostate cancer that target molecularly defined subgroups further increases the need for accurate patient classification and stratification[3-5]. Analysis of whole-exome sequencing data of metastatic prostate cancer tumors revealed that 65% of patients had actionable targets in non-androgen receptor related pathways, including PI3K, Wnt, and DNA repair[6]. Several targeted agents involved in these pathways, including mTOR/AKT pathway inhibitors[7] and PARP inhibitors[8], are currently in various phases of development and the first clinical trials show promising results. Therefore, patients with metastatic prostate cancer could benefit from better stratification to select the most appropriate therapeutic option. More extensive analysis using whole-genome sequencing (WGS)-based classification of tumors may be useful to improve selection of patients for different targeted therapies. The comprehensive nature of WGS has many advantages, including the detection of mutational patterns, as proven by the successful treatment of patients with high-tumor mutational burden with immune checkpoint blockade therapy[9-12]. Moreover, WGS unlike exome sequencing, can detect structural variants and aberrations in non-coding regions, both important features of prostate cancer.

The stratification of prostate cancer patients, based on differences in the mutational landscape of their tumors, has mainly focused on mutually-exclusive mutations, copy-number alterations, or distinct patterns in RNA-sequencing caused by the abundant *TMPRSS2-ERG* fusion, which is recurrent in 50% of primary prostate tumors[6,13-18]. More recently, WGS of metastatic prostate cancer tumors demonstrated that structural variants arise from specific alterations such as *CDK12*[-/-] and *BRCA2*[-/-] genotypes, and are strongly associated with genome-wide events such as large tandem duplications or small genomic deletions, respectively[19-23]. Advances in WGS analysis and interpretation have revealed rearrangement signatures in breast cancer relating to disease stage, homologous recombination deficiency (HRD) and *BRCA1/BRCA2* defects based on size and type of structural variant[22,24]. Thus, WGS enables the identification of patterns of DNA aberrations (i.e. genomic scars) that may profoundly improve classification of tumors that share a common etiology, if performed in a sufficiently powered dataset.

In this study, we analyzed the WGS data obtained from 197 metastatic castration-resistant prostate cancer (mCRPC) patients. We describe the complete genomic landscape of mCRPC, including tumor specific single- and multi-nucleotide variants (SNVs and MNVs), small insertions and deletions (InDels), copy-number alterations (CNAs), mutational signatures, kataegis, chromothripsis, and structural variants (SVs). Next, we compared the mutational frequency of the detected driver genes and genomic subgroups with an unmatched WGS cohort of primary prostate cancer (n = 210), consisting of exclusively of Gleason score 6–7 tumors[15,25]. We investigated the presence of
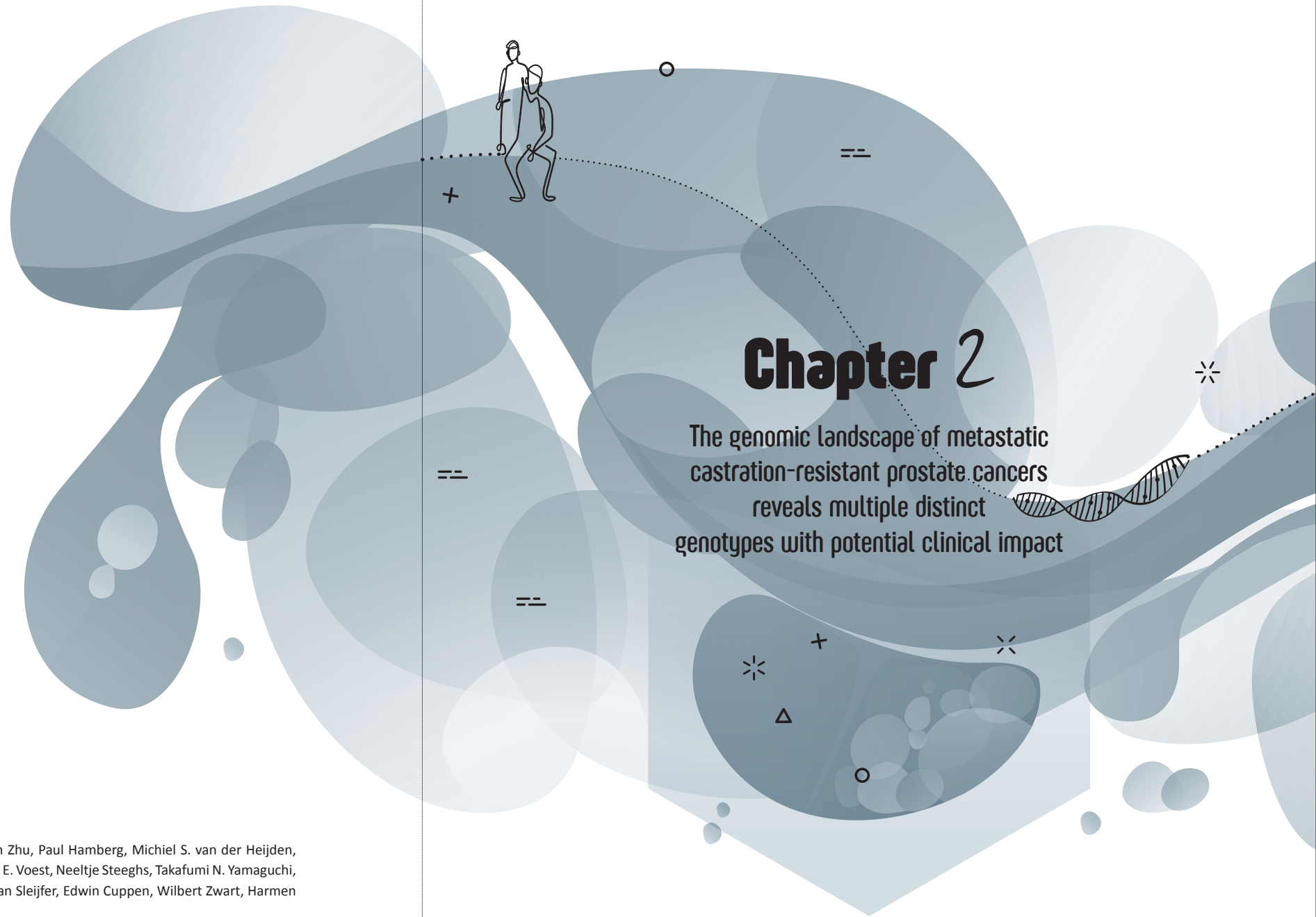
Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration–resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

possible driver genes by analyzing genes with enriched (non-synonymous) mutational burdens and recurrent or high-level copy-number alterations[26,27]. By utilizing various basic genomic features reflecting genomic instability and employing unsupervised clustering, we were able to define eight distinct genomic subgroups of mCRPC patients. We combined our genomic findings with AR, FOXA1, and H3K27me ChIP-seq data, and confirmed that important regulators of AR-mediated signaling are located in non-coding regions with open chromatin and highlight the central role of AR signaling in tumor progression.

Materials and Methods

Patient cohort and study procedures

Patients with metastatic prostate cancer were recruited under the study protocol (NCT01855477) of the Center for Personalized Cancer Treatment (CPCT). This consortium consists of 41 hospitals in The Netherlands (Tabel S1). This CPCT-02 protocol was approved by the medical ethical committee (METC) of the University Medical Center Utrecht and was conducted in accordance with the Declaration of Helsinki. Patients were eligible for inclusion if the following criteria were met: 1) age ≥ 18 years; 2) locally advanced or metastatic solid tumor; 3) indication for new line of systemic treatment with registered anti-cancer agents; 4) safe biopsy according to the intervening physician. For the current study, patients were included for biopsy between 03 May 2016 and 28 May 2018. Data were excluded of patients with the following characteristics: 1) hormone-sensitive prostate cancer; 2) neuro-endocrine prostate cancer (as assessed by routine diagnostics); 3) unknown disease status; 4) prostate biopsy (Figure 1A). All patients provided written informed consent before any study procedure. The study procedures consisted of the collection of matched peripheral blood samples for reference DNA and image-guided percutaneous biopsy of a single metastatic lesion. Soft tissue lesions were biopsied preferentially over bone lesions. The clinical data provided by CPCT have been locked at 1st of July 2018.

Collection and sequencing of samples[28]

Blood samples were collected in CellSave preservative tubes (Menarini-Silicon Biosystems, Huntington Valley, PA, USA) and shipped by room temperature to the central sequencing facility at the Hartwig Medical Foundation. Tumor samples were fresh-frozen in liquid nitrogen directly after the procedure and send to a central pathology tissue facility. Tumor cellularity was estimated by assessing a hematoxylin-eosin (HE) stained 6 micron section. Subsequently, 25 sections of 20 micron were collected for DNA isolation. DNA was isolated with an automated workflow (QIAsymphony) using the DSP DNA Midi kit for blood and QIAsymphony DSP DNA Mini kit for tumor samples according to the manufacturer's protocol (Qiagen). DNA concentration was measured by Qubit™ fluorometric quantitation (Invitrogen, Life Technologies, Carlsbad, CA, USA). DNA libraries for Illumina sequencing were generated from 50-100 ng of genomic DNA using

standard protocols (Illumina, San Diego, CA, USA) and subsequently whole-genome sequenced in a HiSeq X Ten system using the paired-end sequencing protocol (2x 150 bp). Whole-genome alignment (GRCh37), somatic variants (SNV, InDel (max. 50 bp), MNV), structural variant and copy-number calling and in silico tumor cell percentage estimation were performed in a uniform manner as detailed by Priestley et al. (2019)[28]. Mean read coverages of reference and tumor BAM were calculated using Picard Tools (v1.141; CollectWgsMetrics) based on GRCh37[29].

Additional annotation of somatic variants and heuristic filtering

In addition, heuristic filtering removed somatic SNV, InDel and MNV variants based on the following criteria: 1) minimal alternative reads observations ≤ 3; 2) gnomAD exome (ALL) allele frequency ≥ 0.001 (corresponding to ~ 62 gnomAD individuals); and 3) gnomAD genome (ALL) ≥ 0.005 (~ 75 gnomAD individuals)[30]. gnomAD database v2.0.2 was used. Per gene overlapping a genomic variant, the most deleterious mutation was used to annotate the overlapping gene. Structural variants, with BAF ≥ 0.1, were further annotated by retrieving overlapping and nearest up- and downstream annotations using custom R scripts based on GRCh37 canonical UCSC promoter and gene annotations with respect to their respective up- or downstream orientation (if known)[31]. Only potential fusions with only two different gene-partners were considered (e.g. *TMPRSS2-ERG*); structural variants with both breakpoints falling within the same gene were simply annotated as structural variant mutations. Fusion annotation from the COSMIC (v85), CGI and CIVIC databases were used to assess known fusions[32-34]. The COSMIC (v85), OncoKB (July 12, 2018), CIVIC (July 26, 2018), CGI (July 26, 2018) and the list from Martincorena et al. (dN/dS) were used to classify known oncogenic or cancer-associated genes[26,32-34].

Ploidy and copy-number analysis

Ploidy and copy-number (CN) analysis was performed by a custom pipeline as detailed by Priestley et al. (2019)[28]. Briefly, this pipeline combines B-allele frequency (BAF), read depth and structural variants to estimate the purity and CN profile of a tumor sample. Recurrent focal and broad CN alterations were identified by GISTIC2.0 (v2.0.23)[27]. GISTIC2.0 was run with the following parameters: a) genegistic 1; b) gcm extreme; c) maxseg 4000; d) broad 1; e) brlen 0.98; f) conf 0.95; g) rx 0; h) cap 3; i) saveseg 0; j) armpeel 1; k) smallmem 0; l) res 0.01; m) ta 0.1; n) td 0.1; o) savedata 0; p) savegene 1; q) gvt 0.1. Categorization of shallow and deep CN aberration per gene was based on thresholded GISTIC2 calls. Focal peaks detected by GISTIC2 were re-annotated, based on overlapping genomic coordinates, using custom R scripts and UCSC gene annotations. GISTIC2 peaks were annotated with all overlapping canonical UCSC genes within the wide peak limits. If a GISTIC2 peak overlapped with ≤ 3 genes, the most-likely targeted gene was selected based on oncogenic or tumor suppressor annotation in the COSMIC (v85), OncoKB (July 12, 2018), CIVIC (July 26, 2018) and CGI (July 26, 2018) lists[26,32-34]. Peaks in gene deserts were annotated with their nearest gene.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

Estimation of tumor mutational burden (TMB)

The mutation rate per megabase (Mbp) of genomic DNA was calculated as the total genome-wide amount of SNV, MNV and InDels divided over the total amount of callable nucleotides (ACTG) in the human reference genome (hg19) FASTA sequence file:

$$TMB_{genomic} = \frac{(SNV_g + MNV_g + InDels_g)}{\left(\frac{2858674662}{10^6}\right)} \tag{1}$$

The mutation rate per Mbp of coding mutations was calculated as the amount of coding SNV, MNV and InDels divided over the summed lengths of distinct non-overlapping coding regions, as determined on the subset of protein-coding and fully supported (TSL = 1) transcripts in GenCode v28 (hg19)[35]:

$$TMB_{coding} = \frac{(SNV_c + MNV_c + InDels_c)}{\left(\frac{28711682}{10^6}\right)} \tag{2}$$

MSI and HR-deficiency prediction

HR-deficiency/BRCAness was estimated using the CHORD classifier (Nguyen, van Hoeck and Cuppen, manuscript in preparation). This classifier was based on the HRDetect[36] algorithm, however, redesigned to improve its performance beyond primary breast cancer. The binary prediction score (ranging from zero to one) was used to indicate BRCAness level within a sample. To elucidate the potential target gene(s) in the HR-deficient samples (Figure 4), we used the list of BRCAness genes from Lord et al. (2016)[37].

MSI status was determined based on the following criteria; if a sample contained more than 11,436 genomic InDels (max. 50 bp, with repeat-stretches of ≥ 4 bases, repeat length sequence between 2 and 4, or if these consists of a single repeat sequence which repeats ≥ 5 times), the sample was designated as MSI[28].

Detection of (onco-)genes under selective pressure

To detect (onco-)genes under tumor-evolutionary mutational selection, we employed a Poisson-based dN/dS model (192 rate parameters; under the full trinucleotide model) by the R package dndscv (v0.0.0.9)[26]. Briefly, this model tests the normalized ratio of non-synonymous (missense, nonsense and splicing) over background (synonymous) mutations whilst correcting for sequence composition and mutational signatures. A global q-value ≤ 0.1 (with and without taking InDels into consideration) was used to identify statistically-significant (novel) driver genes.

Identification of hypermutated foci (kataegis)

Putative kataegis events were detected using a dynamic programming algorithm which determines a globally optimal fit of a piecewise constant expression profile along genomic coordinates as described by Huber et al. and implemented in the tilingarray R package (v1.56.0)[38]. Only SNVs were used in detecting kataegis. Each chromosome was assessed separately and the maximum number of segmental breakpoints was based on a maximum of five consecutive SNVs (max. 5000 segments per chromosome). Fitting was performed on $\log_{10}$-transformed intermutational distances. Per segment, it was assessed if the mean intermutational distance was ≤ 2000 bp and at least five SNVs were used in the generation of the segment. A single sample with > 200 distinct observed events was set to zero observed events as this sample was found to be hypermutated throughout the entire genome rather than locally. Kataegis was visualized using the R package karyoploteR (v1.4.1)[39].

Mutational signatures analysis

Mutational signatures analysis was performed using the MutationalPatterns R package (v1.4.2)[40]. The thirty consensus mutational signatures, as established by Alexandrov et. al, (matrix Sij; i = 96; number of trinucleotide motifs; j = number of signatures) were downloaded from COSMIC (as visited on 23-05-2018)[41]. Mutations (SNVs) were categorized according to their respective tri-nucleotide context (hg19) into a mutational spectrum matrix Mij (i = 96; number of trinucleotide contexts; j = number of samples) and subsequently, per sample a constrained linear combination of the thirty consensus mutational signatures was constructed using non-negative least squares regression implemented in the R package pracma (v1.9.3).

Between two and fifteen custom signatures were assessed using the NMF package (v0.21.0) with 1000 iterations[42]. By comparing the cophenetic correlation coefficient, residual sum of squares and silhouette, we opted to generate five custom signatures. Custom signatures were correlated to existing (COSMIC) signatures using cosine similarity.

Detection of chromothripsis-like events

Rounded absolute copy-number (excluded Y chromosome) and structural variants (BAF ≥ 0.1) were used in the detection of chromothripsis-like events by the Shatterseek software (v0.4) using default parameters[43]. As a precise standardized definition of chromothripsis has not yet been fully established, and as per the author's instruction, we performed visual inspection of reported chromothripsis-like events after dynamically adapting criteria thresholds (taking the recommended thresholds into consideration). We opted to use the following criteria: a) Total number of intrachromosomal structural variants involved in the event ≥ 25; b) max. number of oscillating CN segments (2 states) ≥ 7 or max. number of oscillating CN segments (3 states) ≥ 14; c) Total size of chromothripsis event ≥ 20 Mbp; d) Satisfying the test of equal distribution of SV types (P > 0.05); and e) Satisfying the test of non-random SV distribution within the cluster region or chromosome (P ≤ 0.05).

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration–resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

Unsupervised clustering of mCRPC WGS characteristics

Samples were clustered using the Euclidian distance of the Pearson correlation coefficient (1-r) and Ward.D hierarchical clustering based on five basic whole genome characteristics; number of mutations per genomic Mbp (SNV, InDel and MNV), mean genome-wide ploidy, number of structural variants and the relative frequencies of structural variant categories (inversions, tandem duplications (larger and smaller than 100 kbp), deletions (larger and smaller than 100 kbp), insertions and interchromosomal translocations). Data was scaled but not centered (root mean square) prior to calculating Pearson correlation coefficients. After clustering, optimal leaf ordering (OLO) was performed using the seriation package (v1.2.3)[44]. The elbow method was employed to determine optimal number of discriminating clusters (Figure S10) using the factoextra package (v1.0.5). Bootstrapping was performed using the pvclust package (v2.0) with 5000 iterations.

Cluster-specific enrichment of aberrant genes (either through SV, deep copy-number alteration or coding SNV/InDel/MNV), kataegis, chromothripsis, GISTIC2 peaks and predicted fusions) between clusters was tested using a two-sided Fisher's Exact Test and Benjamini-Hochberg correction.

A principal component analysis (with scaling and centering) using the prcomp R package[45] was performed on the chosen genomic features and cos2 values for each feature per principal component were retrieved to determine the importance of each feature per respective principal component.

To test the robustness of our clustering, we performed unsupervised clustering, and also other techniques, using various combinations of structural variants and clustering mechanisms as a surrogate for different genome-instability metrics but this analysis did not reveal any striking new clusters.

Supervised clustering based on mutually-exclusive aberrations

Samples were sorted on mutual-exclusivity of *SPOP*, *FOXA1*, and *IDH1* coding mutations and copy-number aberrations and *ETS* family gene fusions (and overexpression) per promiscuous partner (*ERG, ETV1, ETV4,* and *FLI1*) as defined in primary prostate cancer[13]. Tabel S1A of the article 'The Molecular Taxonomy of Primary Prostate Cancer' [13] was used to determine the relative frequency and mutational types of each of the respective primary prostate cancer within the TCGA cohort. In addition, as the TCGA cohort did not denote high-level/deep amplifications, we did not incorporate these either in this analysis.

Correlation of the detection rate of genomic aberrations versus tumor cell percentages

Absolute counts of SNV, InDels, MNV and SV were correlated to the in silico estimated tumor cell percentage using Spearman's correlation coefficient.

Correlation of pretreatment history with detected aberrations and WGS characteristics

Pretreatment history of patients were summarized into 10 groups:
1. Only chemo-treatment (with radio-nucleotides)
2. Only chemo-treatment (without radio-nucleotides)
3. Only radio-nucleotides
4. Only secondary anti-hormonal therapy (with radio-nucleotides)
5. Only secondary anti-hormonal therapy (without radio-nucleotides)
6. Secondary anti-hormonal therapy + one chemo-treatment (with radio-nucleotides)
7. Secondary anti-hormonal therapy + two chemo-treatments (with radio-nucleotides)
8. Secondary anti-hormonal therapy + one chemo-treatment (without radio-nucleotides)
9. Secondary anti-hormonal therapy + two chemo-treatments (without radio-nucleotides)
10. No additional treatment after androgen deprivation therapy

Association with mutated genes, presence of chromothripsis, presence of kataegis, MSI-status and genomic subtypes was tested with a two-sided Fisher's exact test with Benjamini-Hochberg correction.

ChIP-seq experimental set-up and analysis

ChIP-seq cell culturing
VCaP cells were incubated in RPMI medium supplemented with 10% fetal bovine serum (FBS). Bicalutamide resistant VCaP cells (VCaP-Bic) were cultured in RPMI medium supplemented with 10% dextran charcoal-stripped bovine serum (DCC) and 10^6M bicalutamide. VCaP cells were hormone deprived in RPMI medium supplemented with 10% DCC for 3 days before the ChIP-seq experiment.

ChIP-seq and peak calling analysis
For both cell and tissue ChIPs, 5 µg of antibody and 50 µL of magnetic protein A or G beads (10008D or 10009D, ThermoFisher Scientific) were used per IP. The following antibodies were used: Foxa1/2 (M-20, sc-6554 Santa Cruz Biotechnology), AR (N-20, sc-816 Santa Cruz Biotechnology), and H3K27ac (39133, Active Motif). ChIP-seq was performed as described previously[46]. In brief, fresh frozen tissue was cryosectioned into 30-µm-thick slices and stored at −80°C till processing. Samples were fixed using 2 mM DSG (20593; Thermo Fisher Scientific) in solution A (50 mM

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

Hepes-KOH, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) while rotating for 25 min at room temperature, followed by the addition of 1% formaldehyde and another 20 min incubation at room temperature. The reaction was quenched by adding a surplus of glycine. Subsequently, tissue sections were pelleted and washed with cold PBS. Tissue was disrupted using a motorized pellet pestle (Sigma-Aldrich) to disrupt the tissue in cold PBS and to obtain a cell suspension, after which the nuclei were isolated and the chromatin was sheared. During immunoprecipitation, human control RNA (4307281; Thermo Fisher Scientific) and recombinant Histone 2B (M2505S; New England Biolabs) were added as carriers, as described previously[47].

Immunoprecipitated DNA was processed for sequencing using standard protocols and sequenced on an Illumina HiSeq 2500 with 65 bp single end reads. Sequenced samples were aligned to the reference human genome (Ensembl release 55: Homo sapiens GRCh 37.55) using Burrows-Wheeler Aligner (BWA, v0.5.10)[48], reads with a mapping quality > 20 were used for further downstream analysis.

For the tissues, peak calling was performed using MACS2[49] with option --nomodel. In addition, peaks were called against matched input using DFilter[50] in the refine setting with a bandwidth of 50 and a kernel size of 30. Only peaks that were shared between the two algorithms were considered.
    For the cell lines, peaks were obtained with MACS (v1.4; p ≤ 10-7).

The AR and FOXA1 ChIP-seq data for LNCAP with/-out R1881 was obtained from GSE94682[51]. The H3K27ac ChIP-seq data for LNCAP was obtained from GSE114737[46]

Determining enrichment of enhancer to gene ratios
Absolute copy-numbers segments overlapping the gene loci and putative enhancer region (as detected by GISTIC2; focal amplification peaks with a width < 5000 bp) were retrieved per sample. If regions overlapped multiple distinct copy-number segments, the maximum copy-number value of the overlapping segments was used to represent the region. Samples with gene-to-enhancer ratios deviating > 1 studentized residual from equal 1:1 gene-to-enhancer ratios (linear model: $\log_2$(copy-number of enhancer) - $\log_2$(copy-number of gene locus) ~ 0) were categorized as gene or enhancer enriched. Based on the direction of the ratio, samples were either denoted as enhancer (if positive ratio) or gene (if negative ratio) enriched.

Comparison of unmatched primary prostate cancer and mCRPC
Mutational frequencies of the drivers (dN/dS and or GISTIC2) and subtype-specific genes were compared to a separate (unmatched) cohort of primary prostate cancer (n = 210) focusing on Gleason score of 3+3, 3+4 or 4+3, as described by Fraser et al. (2017) and Espiritu et al. (2018)[15,25]. Briefly, whole-genome sequencing reads were mapped to the human reference genome (hg19)

using BWA[48] (v0.5.7) and downstream analysis was performed using Strelka[52] (v.1.0.12) for mutational calling using a matched-normal design (SNVs and InDels), copy-number alterations were estimated with TITAN[53] (v1.11.0) and SNP array data as described in Espiritu et al. (2018)[25] with Delly[54] (v0.5.5 and v0.7.8) was used for detecting structural variants (translocations, inversions, tandem-duplications and deletions). Large insertion calls and overall ploidy was not available for the primary prostate cancer cohort.

TMB was calculated by dividing the number of SNVs and InDels by the total amount of callable bases in the human reference genome (hg19), identical to equation 1. MNV calls were not available for the primary prostate cancer cohort.

Multiple aberrations per gene within a sample were summarized as a single mutational event, e.g. a deletion and mutation in *PTEN* would only count for a single mutation sample in the sample. Only non-synonymous mutations and gains / deletions overlapping with coding regions were used. Statistically significant differences in mutational frequencies were calculated using a two-sided Fisher's Exact test with Benjamini-Hochberg correction.

The primary prostate cancer dataset was clustered together with the mCRPC cohort using the Euclidian distance of the Pearson correlation coefficient (1-r) and Ward.D hierarchical clustering based on three basic whole genome characteristics which were available for all samples; number of mutations per genomic Mbp (SNVs and InDels), number of structural variants, and the relative frequencies of structural variant categories (inversions, tandem duplications (larger and smaller than 100 kbp), deletions (larger and smaller than 100 kbp), and interchromosomal translocations).

Results

Characteristics of the mCRPC cohort and sequencing approach
We analyzed fresh-frozen metastatic tumor samples and matched blood samples from 197 castration-resistant prostate cancer (CRPC) patients using WGS generating to date the largest WGS dataset for mCRPC (Figure 1A). Clinical details on biopsy site, age, and previous treatments of the included patients are described in Figure 1B and 1C and Tabel S2. WGS data was sequenced to a mean coverage of 104X in tumor tissues and 38X in peripheral blood (Figure S1A). The median estimated tumor cell purity using in silico analysis of our WGS data was 62% (range: 16-96%; Figure S1B). Tumor cell purity correlated weakly with the frequency of called SNVs (Spearman correlation; rho = 0.2; P = 0.005), InDels (Spearman correlation; rho = 0.35; P < 0.001), MNVs (Spearman correlation; rho = 0.25; P < 0.001) and structural variants (Spearman correlation; rho = 0.22; P = 0.002; Figure S1C).

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration−resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact
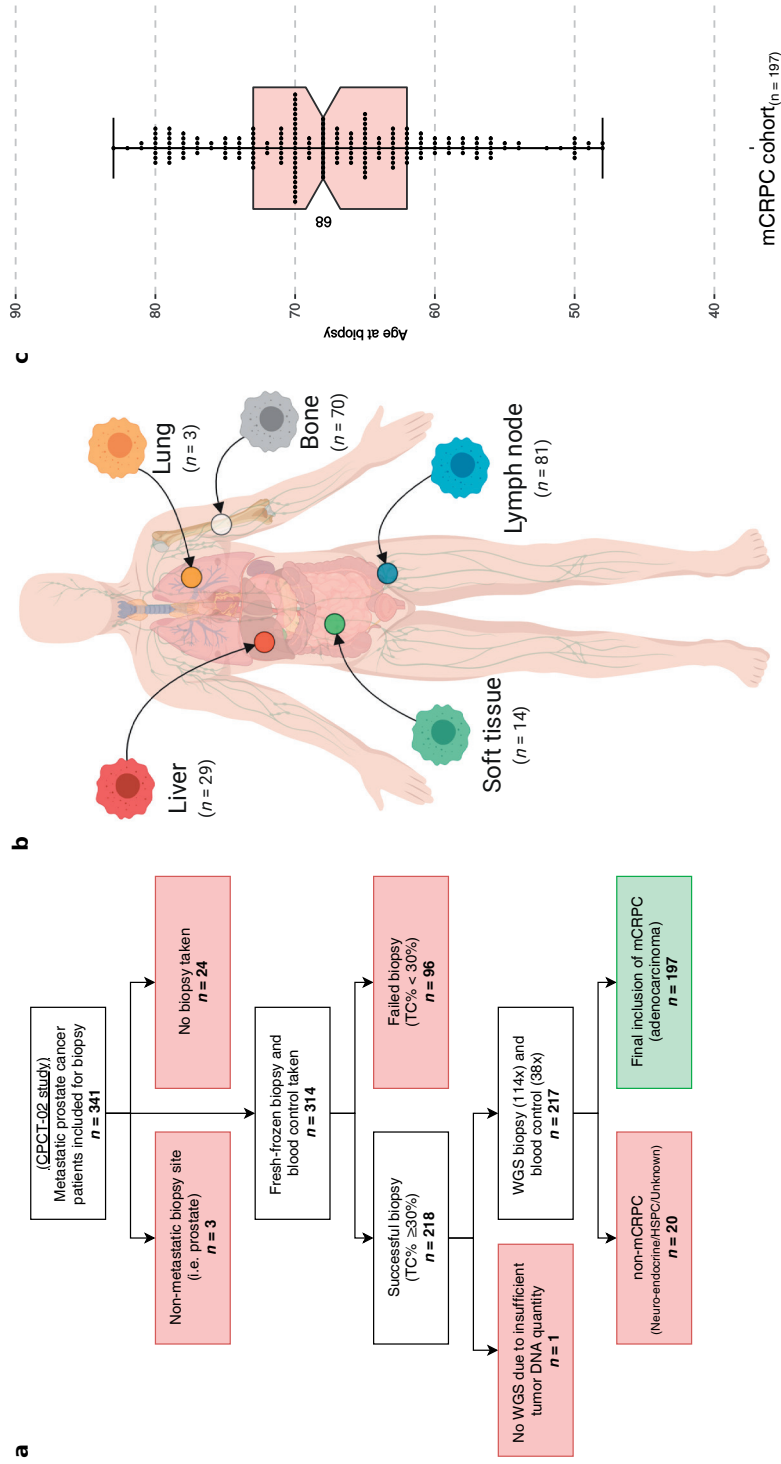
Figure 1. Overview of study design and patient cohort (n = 197)

a) Flowchart of patient inclusion. From the CPCT-02 cohort, patients with metastatic prostate cancer were selected. Patients were excluded if data from metastatic samples were not available and if clinical data indicated that patients had hormone-sensitive or neuro-endocrine prostate cancer or unknown disease status at the time of analysis.
b) Overview of the biopsy sites. Number of biopsies per metastatic site analyzed with WGS.
c) Age of patients at biopsy. Bee-swarm boxplot with notch of the patient age distribution. Boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Data points outside the IQR are shown.

Landscape of mutational and structural variants in mCRPC

The median tumor mutational burden (TMB) at the genomic level (SNVs and InDels per Mbp) was 2.7 in our mCRPC cohort, including 14 patients with high TMB (> 10). We found a median of 6,621 SNVs (IQR: 5,048−9,109), 1,008 small InDels insertions and deletions (IQR: 739−1,364), 55 MNVs (IQR: 34−86) and 224 SVs (IQR: 149−370) per patient (Figure S2A−C). We observed a highly complex genomic landscape consisting of multiple driver mutations and structural variants in our cohort.

We confirmed that known driver genes of prostate cancer were enriched for non-synonymous mutations (Figure 2 and Figure S2E)[13,15,55]. In total, we detected 11 genes enriched with non-synonymous mutations: *TP53, AR, FOXA1, SPOP, PTEN, ZMYM3, CDK12, ZFP36L2, PIK3CA,* and *APC*. *ATM* was mutated in 11 samples, but after multiple-testing correction appeared not to be enriched.

Our copy-number analysis revealed distinct amplified genomic regions including 8q and Xq and deleted regions including 8p, 10q, 13q, and 17p (Figure S2D). Well-known prostate cancer driver genes[8,16], such as *AR, PTEN, TP53,* and *RB1*, are located in these regions. In addition to large-scale chromosomal copy-number alterations, we could identify narrow genomic regions with recurrent copy-number alterations across samples, which could reveal important prostate cancer driver genes (Data file S1).

TMPRSS2-ERG gene fusions were the most common fusions in our cohort (n = 84 out of 197; 42.6%) and were the majority of *ETS* family fusions (n = 84 out of 95; 88.4%; Figure 2 and Figure S3). This is comparable to primary prostate cancer, where *ETS* fusions are found in approximately 50% of tumors[13,15]. The predominant break point was located upstream of the second exon of *ERG*, which preserves its *ETS*-domain in the resulting fusion gene.

In 42 patients (21.3%), we observed regional hypermutation (kataegis; Figure 2 and Figure S4). In addition, we did not observe novel mutational signatures specific for metastatic disease or possible pretreatment histories (Figure S5)[41].

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact
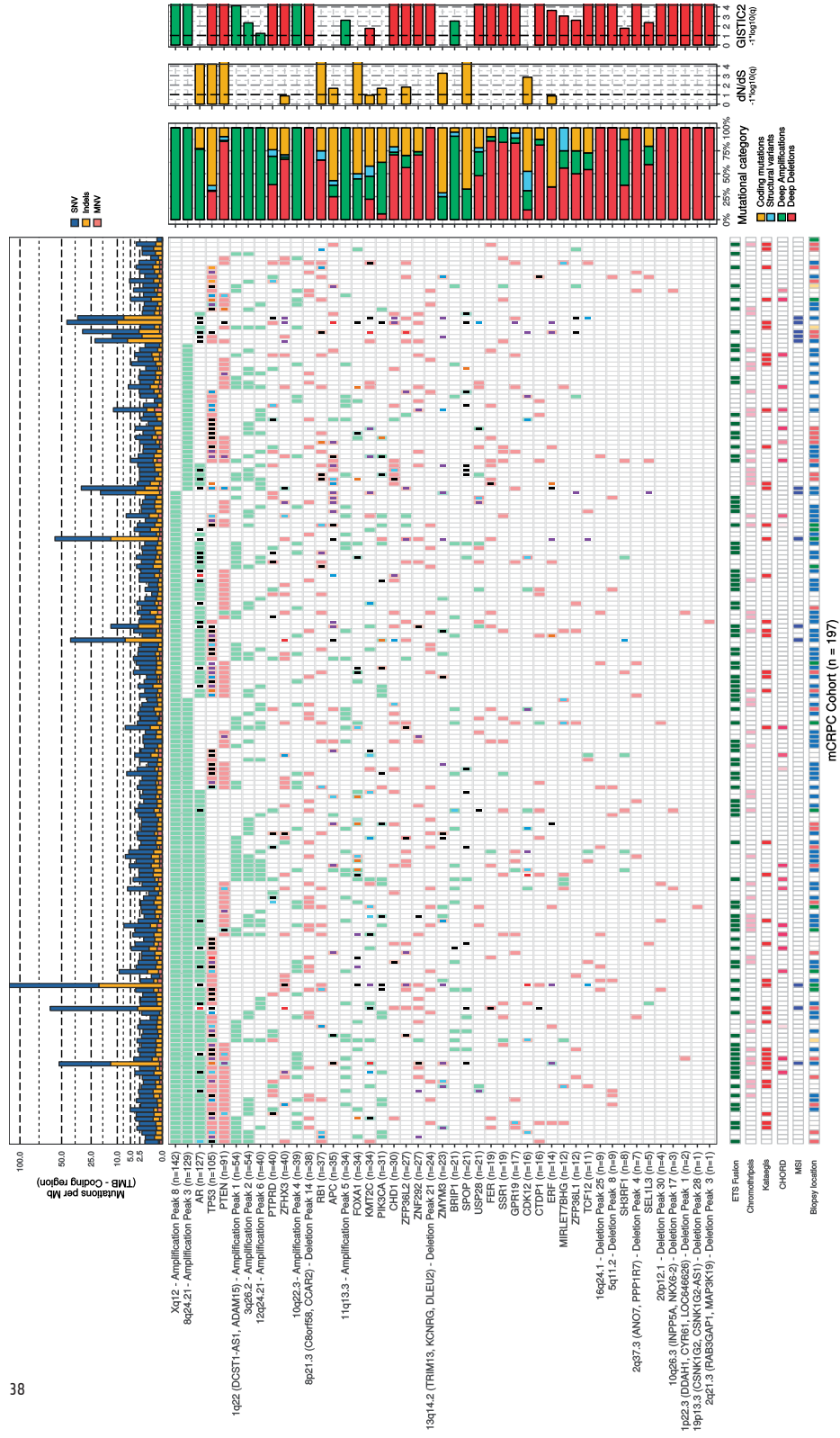
Figure 2. mCRPC shows multiple recurrent somatic alterations affecting several oncogenic pathways

Based on dN/dS (q ≤ 0.1) and GISTIC2 focal peak (q ≤ 0.1) criteria, we show the genes and focal genomic foci that are recurrently mutated, amplified, or deleted in our mCRPC cohort of 197 patients. The upper track (top bar plot) displays the number of genomic mutations per Mbp (TMB) per SNV (blue), InDel (yellow), and MNV (orange) category in coding regions (square-root scale). Samples are sorted based on mutual-exclusivity of the depicted genes and foci. The heatmap displays the type of mutation(s) per sample; (light-)green or (light-)red backgrounds depict copy-number aberrations while the inner square depicts the type of (coding) mutation(s). Relative proportions of mutational categories (coding mutations [SNV, InDels, and MNV] (yellow), SV (blue), deep amplifications [high-level amplifications resulting in many additional copies] (green), and deep deletions [high-level losses resulting in (near) homozygous losses] (red)) per gene and foci are shown in the bar plot next to the heatmap. Narrow GISTIC2 peaks covering ≤ 3 genes were reduced to gene-level rows if one of these genes is present in the dN/dS (q ≤ 0.1) analysis or is a known oncogene or tumor-suppressor. For GISTIC2 peaks covering multiple genes, only deep amplifications and deep deletions are shown. Recurrent aberrant focal genomic foci in gene deserts are annotated with their nearest gene. Significance scores (−1*log₁₀ (q)) of the dN/dS and GISTIC2 analysis are shown on the outer-right bar plots; bars in the GISTIC2 significance plot are colored red if these foci were detected as a recurrent focal deletion and green if detected as a recurrent focal gain. Per sample, the presence of (predicted) ETS fusions (green), chromothripsis (light pink), kataegis (red), CHORD prediction score (HR-deficiency) (pink gradient), MSI status (dark blue), and biopsy location are shown as bottom tracks.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration–resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

To further investigate whether our description of the genome-wide mutational burden and observed alterations in drivers and/or subtype-specific genes in mCRPC were metastatic specific, we compared our data against an unmatched WGS cohort of primary prostate cancer (n = 210)[15,25], consisting of Gleason score 6–7 disease. Comparison of the median genome-wide TMB (SNVs and InDels per Mbp) revealed that the TMB was roughly 3.8 times higher in mCRPC (Figure 3A) and the frequency of structural variants was also higher (Figure 3B) between disease stages, increasing as disease progresses. Analysis on selected driver and subtype-specific genes showed that the mutational frequency of several genes (*AR, TP53, MYC, ZMYM3, PTEN, PTPRD, ZFP36L2, ADAM15, MARCOD2, BRIP1, APC, KMT2C, CCAR2, NKX3-1, C8orf58,* and *RYBP*) was significantly altered (q ≤ 0.05) between the primary and metastatic cohorts (Figure 3C–E). All genes for which we observed significant differences in mutational frequency, based on coding mutations, were enriched in mCRPC (Figure 3D). We did not identify genomic features that were specific for the metastatic setting, beyond androgen deprivation therapy-specific aberrations revolving *AR* (no aberrations in hormone-sensitive setting versus 137 aberrations in castration-resistant setting). We cannot exclude from these data that matched sample analysis or larger scale analysis could reveal such aberrations.

We next determined whether previous treatments affect the mutational landscape. Using treatment history information, we grouped prior secondary anti-hormonal therapy, taxane-based chemotherapy and systemic radionucleotide therapy into different groups (Figure S6). This analysis did not reveal systematic biases due to pretreatment in aberrations, such as TMB, kataegis, chromothripsis, *ETS* fusions, or somatically altered genes (Data file S1).
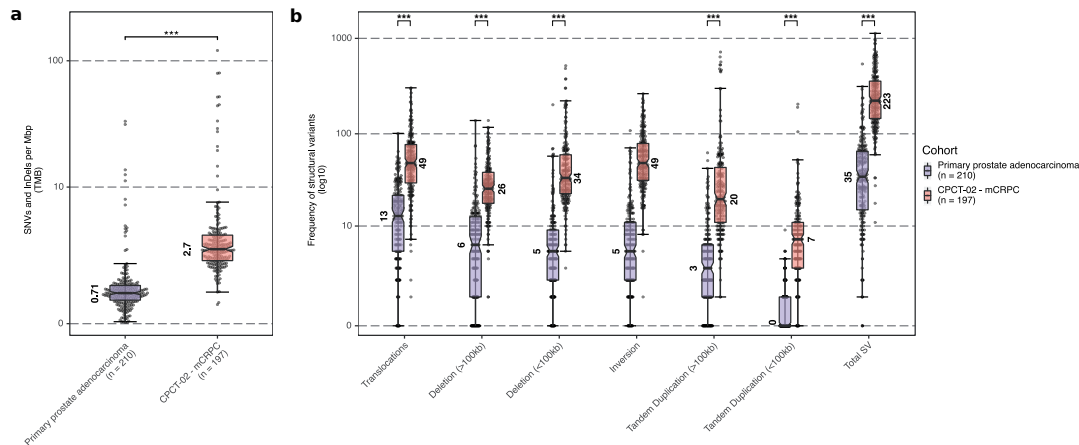
Figure 3. Comparison of the mutational landscape between primary prostate cancer and mCRPC
a) Tumor mutational burden (SNVs and InDels per Mbp) from a primary prostate cancer (n = 210) and the CPTC-02 mCRPC cohort (n = 197). Bee-swarm boxplot with notch of the tumor mutational burden. Boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Data points outside the IQR are shown. Statistical significance was tested with Wilcoxon rank-sum test and P ≤ 0.001 is indicated as ***.
b) Frequency of structural variant events from an unmatched cohort of primary prostate cancer (n = 210) and the CPTC-02 mCRPC cohort (n = 197). Boxplot depicts the upper and lower quartiles, with the median
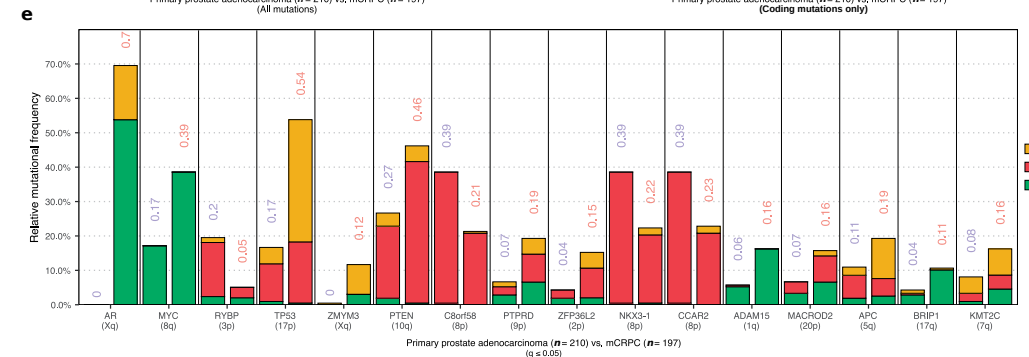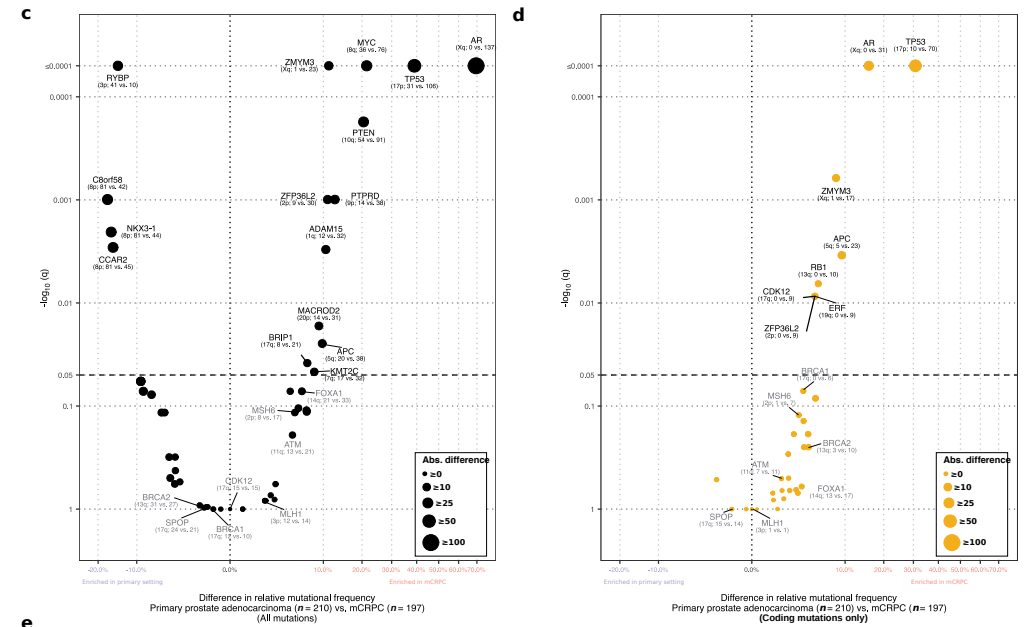
shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Data points outside the IQR are shown. Statistical significance was tested with Wilcoxon rank-sum test and P ≤ 0.001 is indicated as ***.
c) Comparison of the mutational frequencies for driver genes detected by dN/dS and/or GISTIC2, or subtype-specific genes, enriched in mCRPC relative to primary prostate cancer or vice-versa. The difference in relative mutational frequency is shown on the x-axis and the adjusted P-value (two-sided Fisher's Exact Test with BH correction) is shown on the y-axis. Size of the dot is proportional to the absolute difference in mutational frequency between both the cohorts. Symbols of genes with P-values below 0.05 are depicted in black and additional genes-of-interests are highlighted in gray. The general genomic foci of the gene and absolute number of samples with an aberration per cohort in primary prostate cancer and mCRPC, respectively, is shown below the gene symbol. This analysis was performed on coding mutations, gains and deletions per gene.
d) Same as in c but using only coding mutations.
e) Overview of the mutational categories (coding mutations [yellow], deletions [red] and amplifications [green]) of the driver genes detected by dN/dS and/or GISTIC2, or subtype-specific genes, enriched in mCRPC relative to primary prostate cancer (q ≤ 0.05). For each gene the frequency in primary prostate cancer is displayed followed by the frequency in mCRPC.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration–resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

The role of the AR-pathway in mCRPC

Focusing on the AR-pathway revealed that aberrant AR signaling occurred in 80% of our patients. In 57.3% of patients both *AR* and the *AR*-enhancer (~ 66.13 Mb on chromosome X; located about 629 kbp upstream of the *AR* gene[20]) were affected (Figure 4A). In an additional 6.6% and 14.7% of tumors only *AR* gene alterations or *AR*-enhancer amplification occurred, respectively. The percentage of mCRPC patients with the exclusive *AR*-enhancer amplification (29 out of 197; 14.7%) versus exclusively *AR*-locus amplification (13 out of 197; 6.6%) is similar to previous observations which showed 21 out of 94 CRPC patients (10.3%) with exclusively *AR*-enhancer amplification versus 4 out of 94 CRPC patients (4.3%) with exclusively AR-locus amplification[20]. Concurrent amplification of the *AR* gene and the *AR*-enhancer was not necessarily of equal magnitude, which resulted in differences in copy-number enrichment of these loci (Figure 4B).

To date, no AR ChIP-seq data has been reported in human mCRPC samples and evidence of increased functional activity of the amplified enhancer thus far is based on cell line models[56]. To resolve this, we performed AR ChIP-seq on two selected mCRPC patient samples with *AR*-enhancer amplification based on WGS data. As controls we used two prostate cancer cell-lines (LNCaP and VCaP) and three independent primary prostate cancer samples that did not harbor copy-number alterations at this locus (Figure S7)[57]. We observed active enhancer regions (H3K27ac) in the castration-resistant setting, co-occupied by AR and FOXA1, at the amplified *AR*-enhancer. This is substantially stronger when compared to the hormone-sensitive primary prostate cancer samples without somatic amplifications (Figure 4C and Figure S7). Furthermore, a recurrent focal amplification in a non-coding region was observed at 8q24.21 near *PCAT1*. This locus bears similar epigenetic characteristics to the *AR*-enhancer with regard to H3K27ac and, to a lesser extent, binding of AR and/or FOXA1 in the mCRPC setting (Figure 4D and Figure S7).

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

>>

b)
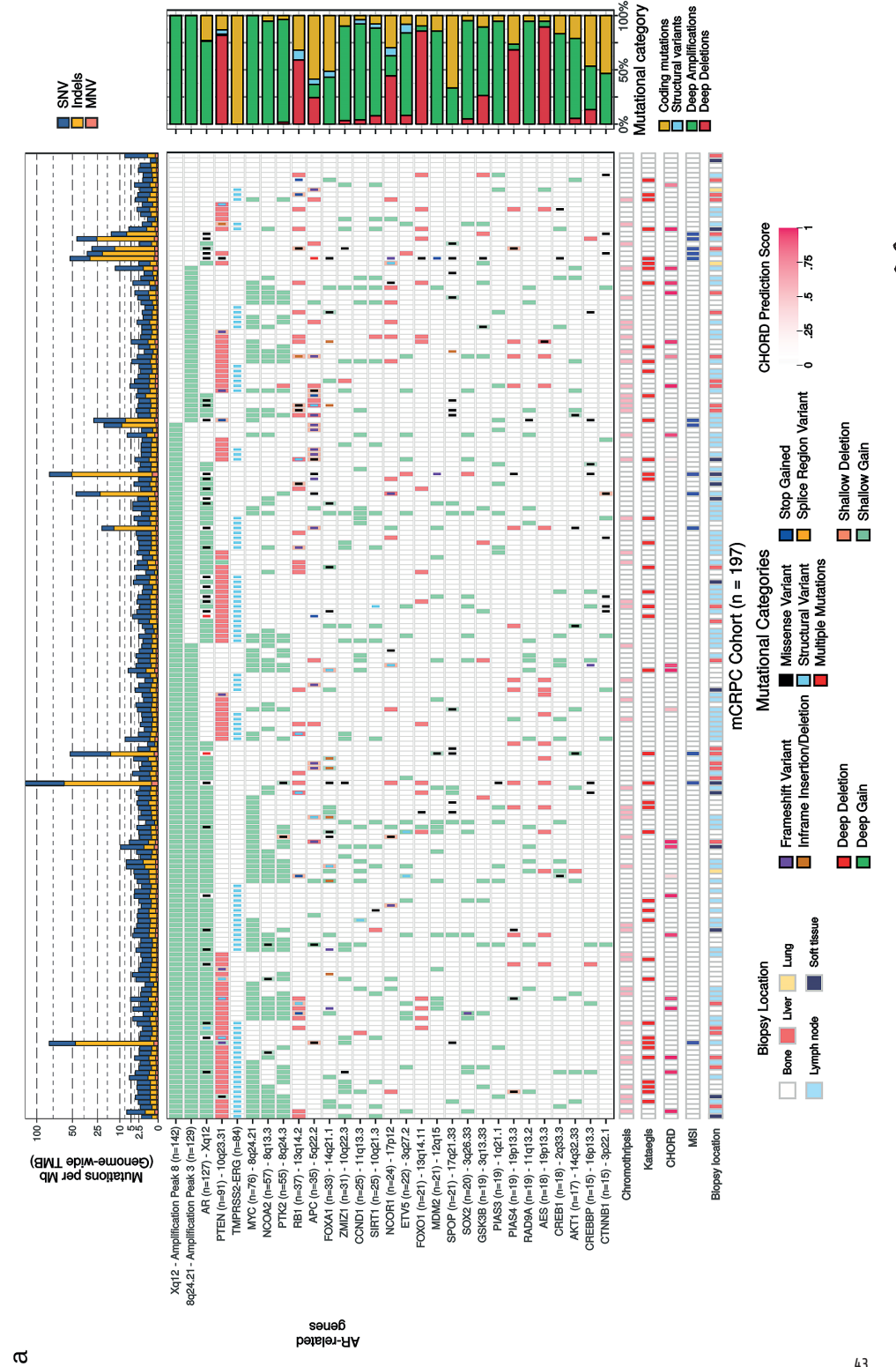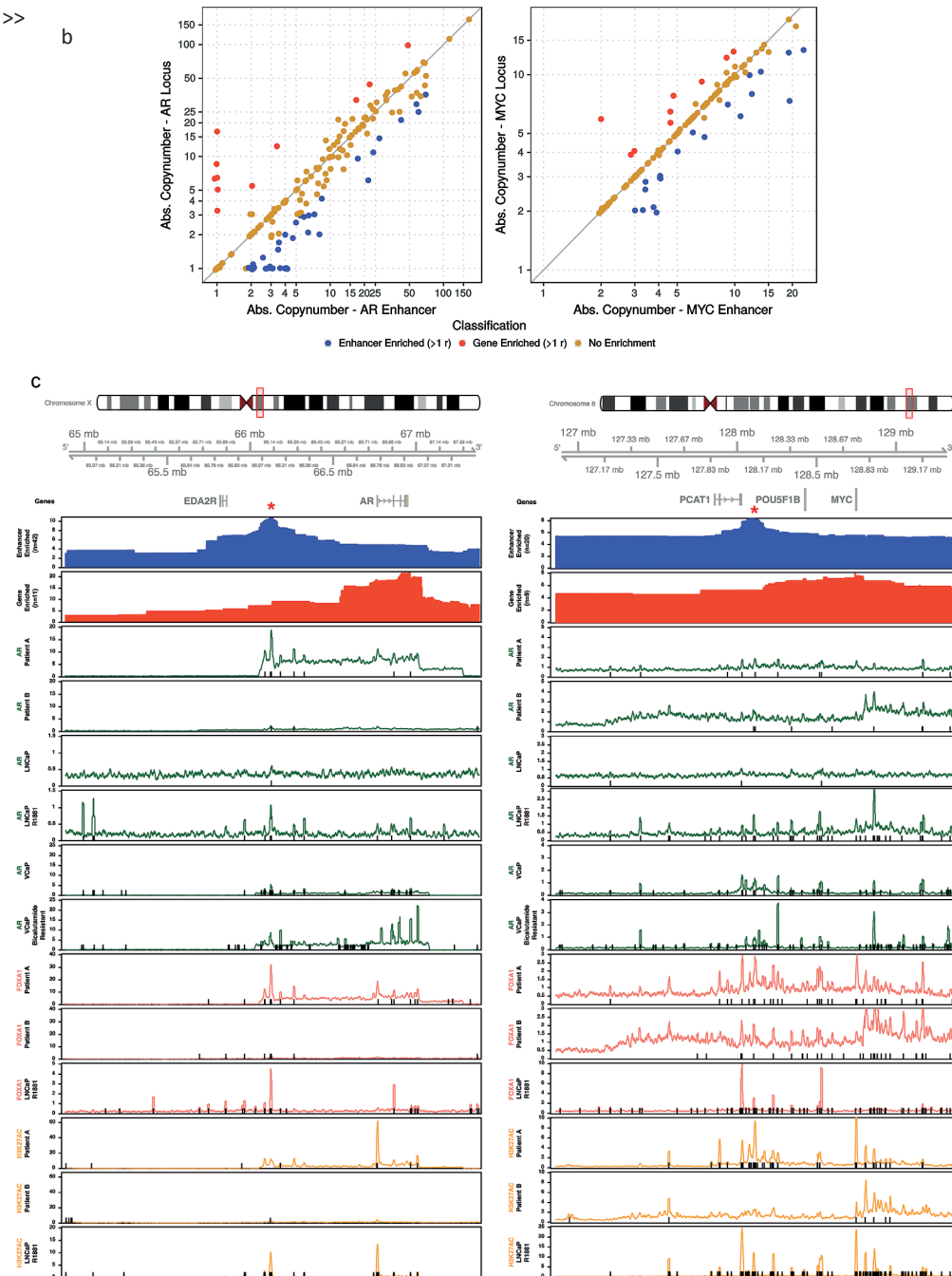
Figure 4. WGS reveals novel insight into the various (non-coding) aberrations affecting AR regulation

a) Mutational overview of top recurrently mutated genes affecting AR regulation and their putative enhancer foci (as detected by GISTIC2). The first track represents the number of genomic mutations per Mbp (TMB) per SNV (blue), InDels (yellow), and MNV (orange) category genome-wide (square-root scale). Samples are sorted

based on mutual-exclusivity of the depicted genes and foci. The heatmap displays the type of mutation(s) per sample, (light-)green or (light-)red backgrounds depict copy-number aberrations while the inner square depicts the type of (coding) mutation(s). Relative proportions of mutational categories (coding mutations [SNV, InDels and MNV] (yellow), SV (blue), deep amplifications (green), and deep deletions (red)) per gene and foci are shown in the bar plot next to the heatmap. The presence of chromothripsis (light pink), kataegis (red), CHORD prediction score (HR-deficiency) (pink gradient), MSI status (dark blue), and biopsy location are shown as bottom tracks.

b) Overview of the copy-number deviations between putative enhancer and gene regions for AR and MYC. Samples were categorized as enhancer- (blue) or gene- (red) enriched if enhancer-to-gene ratio deviated > 1 studentized residual (residual in standard deviation units) from a 1:1 ratio.

c) Copy-number and ChIP-seq profiles surrounding the AR and PCAT1/MYC gene loci (with 1.25 additional Mbp up-/downstream). The upper panel displays the selected genomic window and the overlapping genes. The first and second track display the aggregated mean copy-number (per 0.1 Mbp window) of the enhancer- and gene-enriched samples, respectively. These profiles identify distinct amplified regions (indicated by red asterisk) in proximity to the respective gene bodies. The 3rd to 8th tracks represent AR ChIP-seq profiles (median read-coverage per 0.1 Mbp windows) in two mCRPC patients (# 3 and 4), LNCaP (# 5) and LNCaP with R1881 treatment (# 6), VCaP (# 7) and bicalutamide-resistant VCaP (# 8). The 9th to 11th tracks represent FOXA1 ChIP-seq profiles (median read-coverage per 0.1 Mbp windows) in two mCRPC patients (#9 and 10) and LNCaP with R1881 treatment (# 11). The 12th to 14th tracks represent H3K27ac ChIP-seq profiles (median read-coverage per 0.1 Mbp windows) in two mCRPC patients (# 12 and 13) and LNCaP with R1881 treatment (# 14) reflecting active enhancer regions. ChIP-seq peaks (MACS/MACS2; q < 0.01) are shown as black lines per respective sample.

WGS-based stratification defines genomic subgroups in mCRPC

Our comprehensive WGS data and large sample size enabled us to perform unsupervised clustering on several WGS characteristics to identify genomic scars that can define subgroups of mCRPC patients. We clustered our genomic data using the total number of SVs, relative frequency of SV category (translocations, inversions, insertions, tandem duplications, and deletions), genome-wide TMB encompassing SNV, InDels and MNV, and tumor ploidy. Prior to clustering, we subdivided tandem duplications and deletions into two major categories based on the respective genomic size of the aberration (smaller and larger than 100 kbp) since previous studies revealed distinctions based on similar thresholds for these structural variants in relation to specific mutated genes[19-21,58]. Similarly, we observed a difference in genomic size and number in our subgroups of mCRPC patients (Figure S8).

This analysis defined eight distinct subgroups (Figure 5–6 and Figure S8–11): (A) microsatellite instability (MSI) signature with high TMB and association with mismatch repair deficiency; (B) tandem duplication (> 100 kbp) phenotype associated with biallelic *CDK12* inactivation; (D) homologous recombination deficiency (HRD) features with many deletions (> 100 kbp) and association with (somatic) mutations in BRCAness-associated genes; this was supported by high HR-deficiency scores (CHORD; Figure S8–9); (F) chromothripsis; (C, E, G, H); non-significant genomic signature without any currently known biological association. Table 1 summarizes the key features of each subgroup.
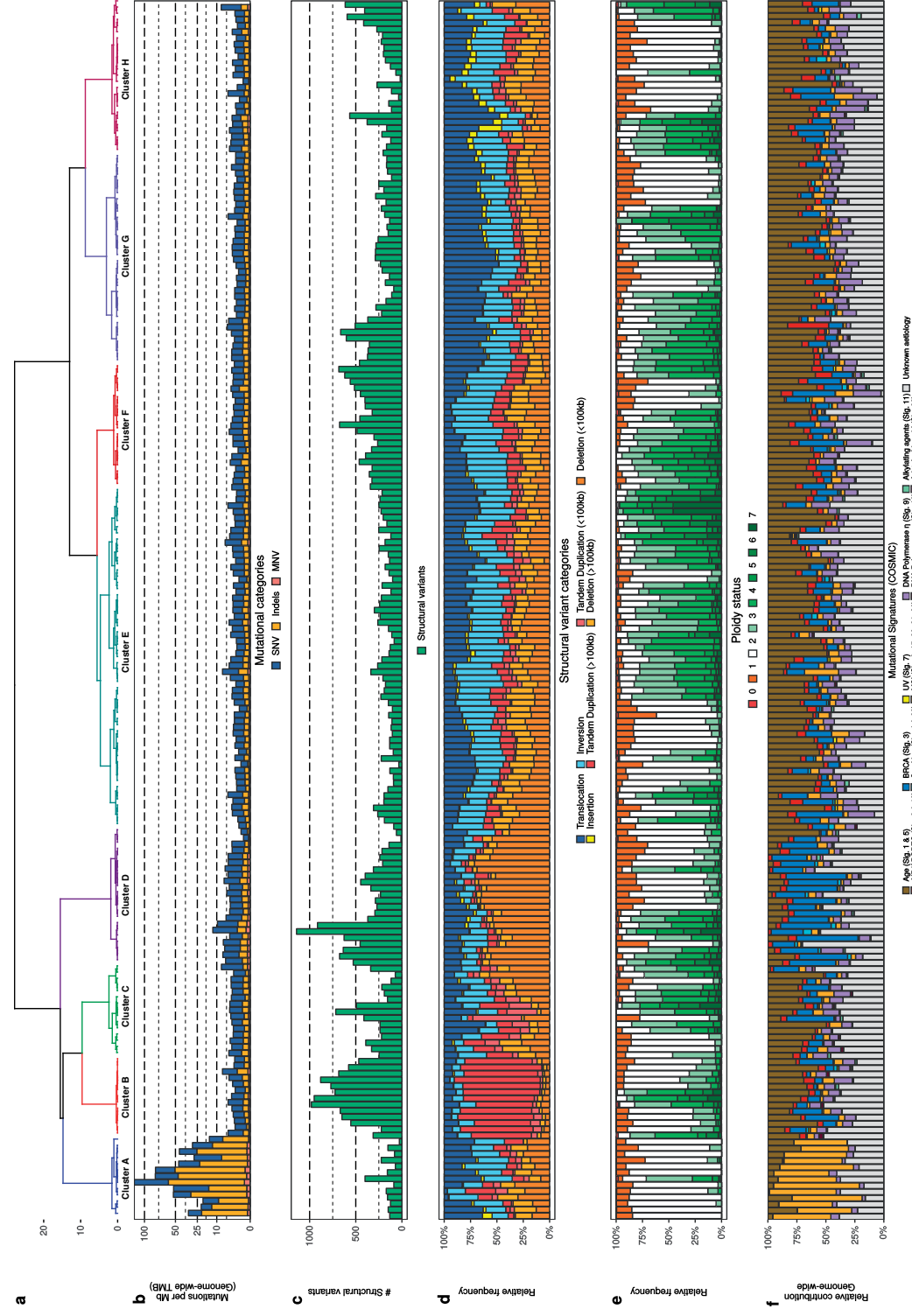
Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

Figure 5. Unsupervised clustering of mCRPC reveals distinct genomic phenotypes

a) Dendrogram of unsupervised clustering with optimal leaf ordering. Top eight clusters are highlighted and denoted based on order of appearance (left to right); A to H. y-axis displays clustering distance (Pearson correlation; ward.D).

b) Number of genomic mutations per Mbp (TMB) per SNV (blue), InDels (yellow), and MNV (orange) category. All genome-wide somatic mutations were taken into consideration (square-root scale).

c) Absolute number of unique structural variants per sample.

d) Relative frequency per structural variant category (translocations, inversions, insertions, tandem duplications, and deletions). Tandem duplications and deletions are subdivided into > 100 kbp and < 100 kbp categories. This track shows if an enrichment for particular category of (somatic) structural variant can be detected, which in turn, can be indicative for a specific mutational aberration.

e) Relative genome-wide ploidy status, ranging from 0 to ≥ 7 copies. This track shows the relative percentage of the entire genome, which is (partially) deleted (ploidy < 2 per diploid genome) or amplified (ploidy > 2 per diploid genome).

f) Relative contribution to mutational signatures (COSMIC) summarized per proposed etiology. This track displays the proposed etiology of each SNV based on their mutational contexts.

g) Relative frequency of different SNV mutational changes.

h) HR-deficient prediction score as assessed by CHORD. The binary prediction score of CHORD (ranging from 0 to 1) is shown, in which higher scores reflect more evidence for HR-deficiency in a given sample.

i) MSI status as determined using a stringent threshold of MSI characteristics[28].

j) Presence of a fusion with a member of the ETS transcription factor family. Green color indicates a possible fusion.

k) Presence of chromothripsis. Pink color indicates presence of chromothripsis as estimated by ShatterSeek.

l) Presence of kataegis. Red color indicates presence of one or more regions showing kataegis.

m) General biopsy location.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

Clusters A and B represent previously identified genomic subgroups (MSI and *CDK12*[-/-])[6,19,21,59]. In cluster B, only two patients were allocated to this subgroup without a specific somatic mutation in the identifying gene. The well-known mismatch repair genes: *MLH1*, *MSH2*, and *MSH6* are among the cluster-specific mutated genes in cluster A (Figure 6A). Twelve out of these thirteen patients had at least one inactivating alteration in one of these genes (Figure 6B). Interestingly, cluster B (*CDK12*[-/-]) harbors two patients without non-synonymous *CDK12* mutation or copy-number alteration; the cause of their tandem duplication phenotype is currently unknown (Figure 6B). Cluster D shows significant features of HRD, specifically biallelic *BRCA2* inactivation (Figure S12), mainly mutational signature 3, enrichment of deletions (< 100 kbp) and is supported by high HR-deficiency scores (CHORD) (Figure S8 and 9)[22,60]. Remarkably, seven out of twenty-two patients did not have a biallelic *BRCA2* inactivation. However, four of these patients showed at least one (deleterious) aberration in other BRCAness-related genes (Figure 6B)[37]. Cluster F was enriched for chromothripsis events, however we could not reproduce a previous finding, suggesting chromothripsis was associated with inversions and p53 inactivation in prostate cancer[21]. Apart from the chromothripsis events, no clear gene aberration was associated with this cluster (Figure 6B). In the remaining patients, there were no distinct genomic signatures or biologic rationale for patient clustering (cluster C, E, G, H). In cluster C, conjoint aberrations of *BRCA1* and *TP53* were observed in one patient with a high HR-deficiency prediction score (CHORD), which is known to lead to a small tandem duplication phenotype (< 100 kbp)[58]. Two other patients within cluster C displayed a weak CHORD scoring associated with HR-deficiency, however no additional definitive evidence was found for a *BRCA1* loss-of-function mutation within these patients.

Table 1. Overview of the distinctive characteristics for each cluster (A-H)

	Number of patients (n; % of cohort)	Tumor mutational burden (CDS)	SNV/InDel ratio	Number of structural variants	Main structural variant category or differentiating category	Ploidy status	Main mutational signature	Top 3 cluster-specific aberrations (% of cluster)	ETS-fusions (n)	Chromothripsis (n)	Kataegis (n)
Cluster A	13 (6.6)	36.88	0.99	149	None	1.92	MSI	MSH6 (69.2) JAK1 (69.2) CIC (58.3)	3	1	6
Cluster B	13 (6.6)	2.44	7.07	669	Tandem duplications (> 100 kb)	2.39	N/A	CDK12 (84.6) FGF3 (69.2) FGF4 (69.2)	2	0	1
Cluster C	15 (7.6)	3.00	6.73	237	Tandem duplications (< 100 kb)	3.19	N/A	None	7	1	2
Cluster D	22 (11.2)	4.39	7.28	323	Deletions (> 100kb)	2.16	BRCA	BRCA2 (68.2)	7	5	5
Cluster E	55 (27.9)	2.12	7.13	178	None	3.24	N/A	None	25	8	13
Cluster F	20 (10.2)	2.51	6.15	400	None	3.35	N/A	Chromothripsis (80)	10	16	7
Cluster G	34 (17.3)	2.12	6.13	222	None	2.98	N/A	None	23	8	5
Cluster H	25 (12.7)	2.30	5.81	201	Insertions	1.97	N/A	None	16	7	3

All numbers are median of the cluster, unless otherwise indicated.
CDS, coding sequence
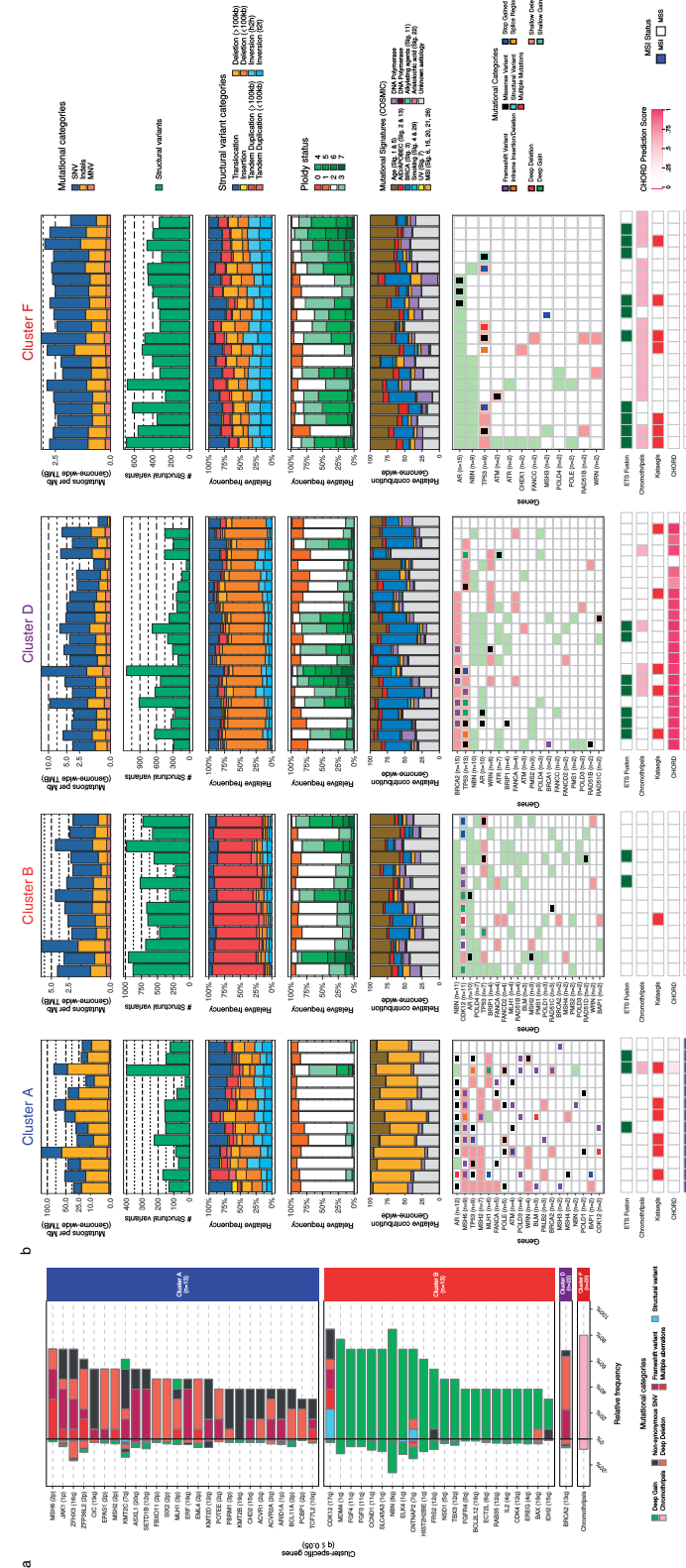
Figure 6. Distinct genomic phenotypes in mCRPC are enriched by mutually exclusive aberrations in key pathways

a) Cluster-specific enrichment of mutated genes (multiple colors), chromothripsis (light pink), and structural variants (light blue) (Fisher's Exact Test with BH correction; q ≤ 0.05). Percentages to the left of the black line represent the relative mutational frequency in mCRPC samples, which are not present in the respective cluster, while the percentages to the right of the black line represent the relative mutational frequency present in the samples from the tested cluster.
b) Genomic overview with biologically relevant genes in the clusters A, B, D, and F with mutational enrichment of genes or large-scale events. The first track represents the number of genomic mutations per Mbp (TMB) per SNV (blue), InDels (yellow), and MNV (orange) category genome-wide (square-root scale). The second track represents the absolute number of unique structural variants (green) per sample. The third track represents the relative frequency per structural variant category. Tandem duplications and deletions are subdivided into > 100 kbp and < 100 kbp categories. The fourth track represents relative genome-wide ploidy status, ranging from 0 to ≥ 7 copies. The fifth track represents the relative contribution to mutational signatures (COSMIC) summarized per proposed etiology. The sixth track displays somatic mutations in the relevant genes found in at least one cluster. The lower tracks represent presence of ETS fusions (green), chromothripsis (pink), kataegis (red), CHORD prediction scores (HR-deficiency) (pink gradient), and MSI status (blue) based on a threshold of MSI characteristics.

49

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

In addition to our unsupervised clustering approach, we clustered our samples using the clustering scheme proposed by TCGA (Figure S13A), which defines seven clusters based on coding mutations and copy-number aberrations in *SPOP, FOXA1, IDH1,* and *ETS* family gene fusions (and overexpression) per promiscuous partner (*ERG, ETV1, ETV4,* and *FLI1*)[13]. Unfortunately, we currently lack matched mRNA-sequencing data in our cohort and therefore cannot observe overexpression of fused *ETS* family members, which restricted us to only characterize the genomic breaks of these promiscuous partners. Without incorporation of *ETS* family overexpression, this proposed clustering scheme categorizes 61% of mCRPC into these seven groups versus 68% of the original cohort containing primary prostate cancer described by TCGA (Figure S13B)[13]. There was no significant correlation between the TCGA clustering scheme and our defined genomic subtypes such as MSI, BRCAness or $CDK12^{-/-}$. In addition, we did not detect statistical enrichment or depletion (q ≤ 0.05) between these supervised clusters and additional mutated genes, kataegis and chromothripsis, only the known enrichment of homozygous *CHD1* deletions in the *SPOP*-cluster was observed[13].

Performing unsupervised clustering and principal component analysis on the primary prostate cancer and metastatic cohorts revealed no striking primary-only genomic subgroup nor did we detect the presence of the mCRPC-derived genomic subgroups in the primary prostate cancer cohort (Figure S14). This could reflect the absence of *CDK12* mutations and the presence of only three sporadic *BRCA2*-mutated samples (1%) in the primary prostate cancer cohort. Furthermore, only one sample (1%) with MSI-like and high TMB (> 10), respectively, was observed in the primary cancer cohort. Indeed, there is a striking difference in the mutational load between both disease settings.

Discussion

We performed WGS of metastatic tumor biopsies and matched-normal blood obtained from 197 patients with mCRPC to provide an overview of the genomic landscape of mCRPC. The size of our cohort enables classification of patients into distinct disease subgroups using unsupervised clustering. Our data suggest that classification of patients using genomic events, as detected by WGS, improves patient stratification, specifically for clinically actionable subgroups such as BRCA-deficient and MSI patients. Furthermore, we confirm the central role of AR signaling in mCRPC that mediates its effect through regulators located in non-coding regions and the apparent difference in primary versus metastatic prostate cancers.

The classification of patients using WGS has the advantage of being, in theory, more precise in determining genomically defined subgroups in prostate cancer compared to analyses using targeted panels consisting of a limited number of genes, or exome sequencing. The identification of subgroups based on predominant phenotypic characteristics encompassing genomic

signatures may be clinically relevant and our clustering analysis refines patient classification. In cluster A, we observed a high TMB, which has been associated in other tumor types with a high sensitivity to immune check-point inhibitors[9,11,12]. Clinical trials using pembrolizumab in selected mCRPC patients are underway (KEYNOTE-028, KEYNOTE-199)[61,62]. Interestingly, in both cluster B and cluster D, we identified patients that did not have the defining biallelic *CDK12* or *BRCA2* (somatic) mutation. Such patients might be deemed false-negatives or false-positives when using FDA-approved assays (BRCAnalysis™ and FoundationFocus™), currently used in breast cancer diagnosis and based on the presence of *BRCA1/2* mutations, to predict response to poly(ADP-ribose) polymerase (PARP) inhibitors and/or platinum compounds based on the presence of *BRCA1/2* mutations. The first clinical trials combining PARP inhibitors with AR-targeted therapies in mCRPC show promising results[8]. Thus, WGS-based stratification may improve the patient classification of DNA repair-deficient tumors as it uses the genome-wide scars caused by defective DNA repair to identify tumors that have these deficiencies.

The use of WGS also allowed us to gain more insight into the role of non-coding regions of the genome in prostate cancer. We confirmed the amplification of a recently reported *AR*-enhancer[20,21,56]. In line with the cell line-based observations, we show AR binding at these mCRPC-specific enhancer regions, providing the first clinical indication that *AR*-enhancer amplification also increases AR signaling in mCRPC tumors. These findings are supported by previous studies demonstrating that this amplification ultimately resulted in significantly elevated expression of *AR* itself[20,21,56]. Furthermore, we confirm a recurrent focal amplification near *PCAT1*, which shows robust chromatin binding for AR in mCRPC samples, providing clinical proof-of-concept of a functional enhancer that is also active and AR-bound in cell line models. Recent research elucidated to the functional importance of this region in regulating *MYC* expression in prostate cancer, which could highlight a putative role of this somatically acquired amplification[57]. However, the WGS and ChIP-seq data presented here are not conclusive in elucidating the definitive role of this amplified region in regulating *MYC* expression and further mechanistic studies are needed to establish a potential link to *MYC* regulation.

In addition, *PCAT1* is a long non-coding RNA, which is known to be upregulated in prostate cancer and negatively regulates *BRCA2* expression while positively affecting *MYC* expression[63,64]. Combining our WGS approach with AR, FOXA1, and H3K27ac ChIP-seq data, we identify non-coding regions affecting both *AR* itself, and possibly *MYC*, through *AR*-enhancer amplification as a potential mechanism contributing to castration resistance.

A potential pitfall of our clustering analysis is the selection of features used; for this we made a number of assumptions based on the literature and distribution of the structural variants within our cohort[19-21,58]. As the input of features and weights for clustering analysis are inherent to the clustering outcome, we performed additional clustering analyses using various combinations of these features and applied alternative approaches but did not detect striking differences

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration−resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

compared to the current approach. Another potential pitfall of the employed hierarchical clustering scheme is that patients are only attributed to a single cluster. An example of this can be seen in cluster A where a patient is grouped based on its predominant genotype (MSI) and associated mutations in mismatch repair-related genes (*MLH1, POLE, POLD3,* and *BLM*), but this sample also displays an increased number of structural variants and increased ploidy status and harbors a pathogenic *BRCA2* mutation. However, it is missing the characteristic number of genomic deletions (< 100 kbp) and BRCA mutational signature associated with *BRCA2*[-/-] samples that define cluster D. Despite these pitfalls we conclude that unbiased clustering contributes towards improved classification of patients.

The CPCT-02 study was designed to examine the correlation of genomic data with treatment outcome after biopsy at varying stages of disease. Our cohort contains patients with highly variable pretreatment history and since the treatments for mCRPC patients nowadays significantly impacts overall survival, the prognosis of patients differs greatly. Therefore, correlation between genomic data and clinical endpoints, such as survival is inherently flawed due to the very heterogeneous nature of the patient population. Moreover, our analysis comparing primary and metastatic samples shows a significant increase in the number of genomic aberrations with advancing disease, meaning that the difference in timing of the biopsies may bias the prognostic value of the data. In future studies, we plan to gather all known clinically defined prognostic information and determine whether the genomic subtypes increase the ability to predict outcome. Unfortunately, some clinical parameters with prognostic importance such as ethnicity will not be available due to ethical regulations. Moreover, we will increase the sample size, in order to correlate genomic features to clinical parameters to better determine whether the subtypes we identified are stable over time. Therefore, we are currently unable to present meaningful correlations between clinical endpoints and the clusters we identified.

Overall, we show the added value of WGS-based unsupervised clustering in identifying patients with genomic scars who are eligible for specific therapies. Since our clustering method does not rely on one specific genetic mutation we are able to classify patients even when WGS (or our methodology) does not find conclusive evidence for (bi-allelic) mutations in the proposed gene-of-interest. Further research should validate clinical response and outcome on specific therapies in matched subgroups. This study also shows that a large population of mCRPC patients do not fall into an as-of-yet clinically relevant or biologically clear genotype and further research can help elucidate the oncogenic driver events and provide new therapeutic options.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

References

1. Boyd LK, Mao X, Lu YJ. The complexity of prostate cancer: genomic alterations and heterogeneity. Nat Rev Urol 2012;9(11):652-64. DOI: 10.1038/nrurol.2012.185.
2. Wei L, Wang J, Lampert E, et al. Intratumoral and Intertumoral Genomic Heterogeneity of Multifocal Localized Prostate Cancer Impacts Molecular Classifications and Genomic Prognosticators. Eur Urol 2017;71(2):183-192. DOI: 10.1016/j.eururo.2016.07.008.
3. Mullane SA, Van Allen EM. Precision medicine for advanced prostate cancer. Curr Opin Urol 2016;26(3):231-9. DOI: 10.1097/MOU.0000000000000278.
4. Ciccarese C, Massari F, Iacovelli R, et al. Prostate cancer heterogeneity: Discovering novel molecular targets for therapy. Cancer Treat Rev 2017;54:68-73. DOI: 10.1016/j.ctrv.2017.02.001.
5. Shtivelman E, Beer TM, Evans CP. Molecular pathways and targets in prostate cancer. Oncotarget 2014;5(17):7217-59. DOI: 10.18632/oncotarget.2406.
6. Robinson D, Van Allen EM, Wu YM, et al. Integrative Clinical Genomics of Advanced Prostate Cancer. Cell 2015;161(5):1215-1228. (In English). DOI: 10.1016/j.cell.2015.05.001.
7. Chow H, Ghosh PM, deVere White R, et al. A phase 2 clinical trial of everolimus plus bicalutamide for castration-resistant prostate cancer. Cancer 2016;122(12):1897-904. DOI: 10.1002/cncr.29927.
8. Clarke N, Wiechno P, Alekseev B, et al. Olaparib combined with abiraterone in patients with metastatic castration-resistant prostate cancer: a randomised, double-blind, placebo-controlled, phase 2 trial. Lancet Oncol 2018;19(7):975-986. DOI: 10.1016/S1470-2045(18)30365-6.
9. Yarchoan M, Hopkins A, Jaffee EM. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. N Engl J Med 2017;377(25):2500-2501. (In eng). DOI: 10.1056/NEJMc1713444.
10. Chan TA, Yarchoan M, Jaffee E, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. Ann Oncol 2019;30(1):44-56. DOI: 10.1093/annonc/mdy495.
11. Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science 2015;348(6230):124-8. (In eng). DOI: science.aaa1348 [pii] 10.1126/science.aaa1348.
12. Samstein RM, Lee CH, Shoushtari AN, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet 2019;51(2):202-206. (In eng). DOI: 10.1038/s41588-018-0312-8 10.1038/s41588-018-0312-8 [pii].
13. Cancer Genome Atlas Research N. The Molecular Taxonomy of Primary Prostate Cancer. Cell 2015;163(4):1011-25. (In eng). DOI: S0092-8674(15)01339-2 [pii] 10.1016/j.cell.2015.10.025.
14. Angeles AK, Bauer S, Ratz L, Klauck SM, Sultmann H. Genome-Based Classification and Therapy of Prostate Cancer. Diagnostics (Basel) 2018;8(3). DOI: 10.3390/diagnostics8030062.
15. Fraser M, Sabelnykova VY, Yamaguchi TN, et al. Genomic hallmarks of localized, non-indolent prostate cancer. Nature 2017;541(7637):359-364. (In eng). DOI: nature20788 [pii] 10.1038/nature20788.
16. Schoenborn JR, Nelson P, Fang M. Genomic Profiling Defines Subtypes of Prostate Cancer with the Potential for Therapeutic Stratification. Clin Cancer Res 2013;19(15):4058-4066. (In English). DOI: 10.1158/1078-0432.Ccr-12-3606.
17. Nam RK, Sugar L, Wang Z, et al. Expression of TMPRSS2:ERG gene fusion in prostate cancer cells is an important prognostic factor for cancer progression. Cancer Biol Ther 2007;6(1):40-5. DOI: 10.4161/cbt.6.1.3489.
18. Tomlins SA, Laxman B, Varambally S, et al. Role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia 2008;10(2):177-88. DOI: 10.1593/neo.07822.
19. Wu YM, Cieslik M, Lonigro RJ, et al. Inactivation of CDK12 Delineates a Distinct Immunogenic Class of Advanced Prostate Cancer. Cell 2018;173(7):1770-1782 e14. (In eng). DOI: S0092-8674(18)30565-8 [pii] 10.1016/j.cell.2018.04.034.
20. Viswanathan SR, Ha G, Hoff AM, et al. Structural Alterations Driving Castration-Resistant Prostate Cancer Revealed by Linked-Read Genome Sequencing. Cell 2018;174(2):433-447.e19. (Article) (In English). DOI: 10.1016/j.cell.2018.05.036.
21. Quigley DA, Dang HX, Zhao SG, et al. Genomic Hallmarks and Structural Variation in Metastatic Prostate Cancer. Cell 2018;175(3):889. (In eng). DOI: S0092-8674(18)31328-X [pii] 10.1016/j.cell.2018.10.019.
22. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 2016;534(7605):47-54. DOI: 10.1038/nature17676.
23. Taylor RA, Fraser M, Livingstone J, et al. Germline BRCA2 mutations drive prostate cancers with distinct evolutionary trajectories. Nat Commun 2017;8:13671. DOI: 10.1038/ncomms13671.
24. Davies H, Morganella S, Purdie CA, et al. Whole-Genome Sequencing Reveals Breast Cancers with Mismatch Repair Deficiency. Cancer Res 2017;77(18):4755-4762. DOI: 10.1158/0008-5472.CAN-17-1083.
25. Espiritu SMG, Liu LY, Rubanova Y, et al. The Evolutionary Landscape of Localized Prostate Cancers Drives Clinical Aggression. Cell 2018;173(4):1003-1013 e15. DOI: 10.1016/j.cell.2018.03.029.
26. Martincorena I, Raine KM, Gerstung M, et al. Universal Patterns of Selection in Cancer and Somatic Tissues. Cell 2017;171(5):1029-1041 e21. DOI: 10.1016/j.cell.2017.09.042.
27. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol 2011;12(4):R41. DOI: 10.1186/gb-2011-12-4-r41.
28. Priestley P, Baber J, Lolkema MP, et al. Pan-cancer whole-genome analyses of metastatic solid tumors. Nature 2019;575(7781):210-216. (In eng). DOI: 10.1038/s41586-019-1689-y 10.1038/s41586-019-1689-y [pii].
29. Institute B. Picard tools. 2016.
30. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536(7616):285-91. DOI: 10.1038/nature19057.
31. Casper J, Zweig AS, Villarreal C, et al. The UCSC Genome Browser database: 2018 update. Nucleic Acids Res 2018;46(D1):D762-D769. DOI: 10.1093/nar/gkx1020.
32. Forbes SA, Beare D, Boutselakis H, et al. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res 2017;45(D1):D777-D783. DOI: 10.1093/nar/gkw1121.
33. Tamborero D, Rubio-Perez C, Deu-Pons J, et al. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. Genome Med 2018;10(1):25. DOI: 10.1186/s13073-018-0531-8.
34. Griffith M, Spies NC, Krysiak K, et al. CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. Nat Genet 2017;49(2):170-174. DOI: 10.1038/ng.3774.
35. Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 2012;22(9):1760-74. (In eng). DOI: 22/9/1760 [pii] 10.1101/gr.135350.111.
36. Davies H, Glodzik D, Morganella S, et al. HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. Nat Med 2017;23(4):517-525. DOI: 10.1038/nm.4292.
37. Lord CJ, Ashworth A. BRCAness revisited. Nat Rev Cancer 2016;16(2):110-20. DOI: 10.1038/nrc.2015.21.
38. Huber W, Toedling J, Steinmetz LM. Transcript mapping with high-density oligonucleotide tiling arrays. Bioinformatics 2006;22(16):1963-70. DOI: 10.1093/bioinformatics/btl289.
39. Gel B, Serra E. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics 2017;33(19):3088-3090. DOI: 10.1093/bioinformatics/btx346.
40. Blokzijl F, Janssen R, van Boxtel R, Cuppen E. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. Genome Med 2018;10(1):33. DOI: 10.1186/s13073-018-0539-0.
41. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. Nature 2013;500(7463):415-21. DOI: 10.1038/nature12477.
42. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. BMC Bioinformatics 2010;11:367. DOI: 10.1186/1471-2105-11-367.
43. Cortes-Ciriano I, Lee JJ, Xi R, et al. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. Nat Genet 2020;52(3):331-341. DOI: 10.1038/s41588-019-0576-7.
44. Hahsler M, Hornik K, Buchta C. Getting Things in Order: An Introduction to the R Package seriation. 2008 2008;25(3):34. DOI: 10.18637/jss.v025.i03.
45. Ripley WNVBD. Modern Applied Statistics with S. 4 ed: Springer-Verlag New York, 2002.
46. Singh AA, Schuurman K, Nevedomskaya E, et al. Optimized ChIP-seq method facilitates transcription factor profiling in human tumors. Life Sci Alliance 2019;2(1):e201800115. DOI: 10.26508/lsa.201800115.
47. Zwart W, Koornstra R, Wesseling J, Rutgers E, Linn S, Carroll JS. A carrier-assisted ChIP-seq method for estrogen receptor-chromatin interactions from breast cancer core needle biopsy samples. BMC Genomics 2013;14:232. DOI: 10.1186/1471-2164-14-232.
48. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25(14):1754-60. DOI: 10.1093/bioinformatics/btp324.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

49.   Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008;9(9):R137. DOI: 10.1186/gb-2008-9-9-r137.

50.   Kumar V, Muratani M, Rayan NA, et al. Uniform, optimal signal processing of mapped deep-sequencing data. Nat Biotechnol 2013;31(7):615-22. DOI: 10.1038/nbt.2596.

51.   Stelloo S, Nevedomskaya E, Kim Y, et al. Endogenous androgen receptor proteomic profiling reveals genomic subcomplex involved in prostate tumorigenesis. Oncogene 2018;37(3):313-322. DOI: 10.1038/onc.2017.330.

52.   Kim S, Scheffler K, Halpern AL, et al. Strelka2: fast and accurate calling of germline and somatic variants. Nat Methods 2018;15(8):591-594. DOI: 10.1038/s41592-018-0051-x.

53.   Ha G, Roth A, Khattra J, et al. TITAN: inference of copy-number architectures in clonal cell populations from tumor whole-genome sequence data. Genome Res 2014;24(11):1881-93. DOI: 10.1101/gr.180281.114.

54.   Rausch T, Zichner T, Schlattl A, Stutz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics 2012;28(18):i333-i339. DOI: 10.1093/bioinformatics/bts378.

55.   Armenia J, Wankowicz SAM, Liu D, et al. The long tail of oncogenic drivers in prostate cancer. Nat Genet 2018;50(5):645-651. (In eng). DOI: 10.1038/s41588-018-0078-z
10.1038/s41588-018-0078-z [pii].

56.   Takeda DY, Spisak S, Seo JH, et al. A Somatically Acquired Enhancer of the Androgen Receptor Is a Noncoding Driver in Advanced Prostate Cancer. Cell 2018;174(2):422-432 e13. DOI: 10.1016/j.cell.2018.05.037.

57.   Mazrooei Pea. Somatic Mutations and Risk-Variants Converge on Cis-Regulatory Elements to Reveal the Cancer Driver Transcription Regulators in Primary Prostate Tumors. SSRN Electron 2018. DOI: doi:10.2139/ssrn.3245213.

58.   Menghi F, Barthel FP, Yadav V, et al. The Tandem Duplicator Phenotype Is a Prevalent Genome-Wide Cancer Configuration Driven by Distinct Gene Mutations. Cancer Cell 2018;34(2):197-210 e5. DOI: 10.1016/j.ccell.2018.06.008.

59.   Pritchard CC, Morrissey C, Kumar A, et al. Complex MSH2 and MSH6 mutations in hypermutated microsatellite unstable advanced prostate cancer. Nat Commun 2014;5:4988. DOI: 10.1038/ncomms5988.

60.   Polak P, Kim J, Braunstein LZ, et al. A mutational signature reveals alterations underlying deficient homologous recombination repair in breast cancer. Nat Genet 2017;49(10):1476-1486. DOI: 10.1038/ng.3934.

61.   Hansen AR, Massard C, Ott PA, et al. Pembrolizumab for advanced prostate adenocarcinoma: findings of the KEYNOTE-028 study. Ann Oncol 2018;29(8):1807-1813. (In eng). DOI: S0923-7534(19)34146-8 [pii]
10.1093/annonc/mdy232.

62.   Bono JSD, Goh JC, Ojamaa K, et al. KEYNOTE-199: Pembrolizumab (pembro) for docetaxel-refractory metastatic castration-resistant prostate cancer (mCRPC). Journal of Clinical Oncology 2018;36(15_suppl):5007-5007. DOI: 10.1200/JCO.2018.36.15_suppl.5007.

63.   Prensner JR, Chen W, Iyer MK, et al. PCAT-1, a long noncoding RNA, regulates BRCA2 and controls homologous recombination in cancer. Cancer Res 2014;74(6):1651-60. DOI: 10.1158/0008-5472.CAN-13-3159.

64.   Prensner JR, Chen W, Han S, et al. The long non-coding RNA PCAT-1 promotes prostate cancer cell proliferation through cMyc. Neoplasia 2014;16(11):900-8. DOI: 10.1016/j.neo.2014.09.001.

Supplementary information
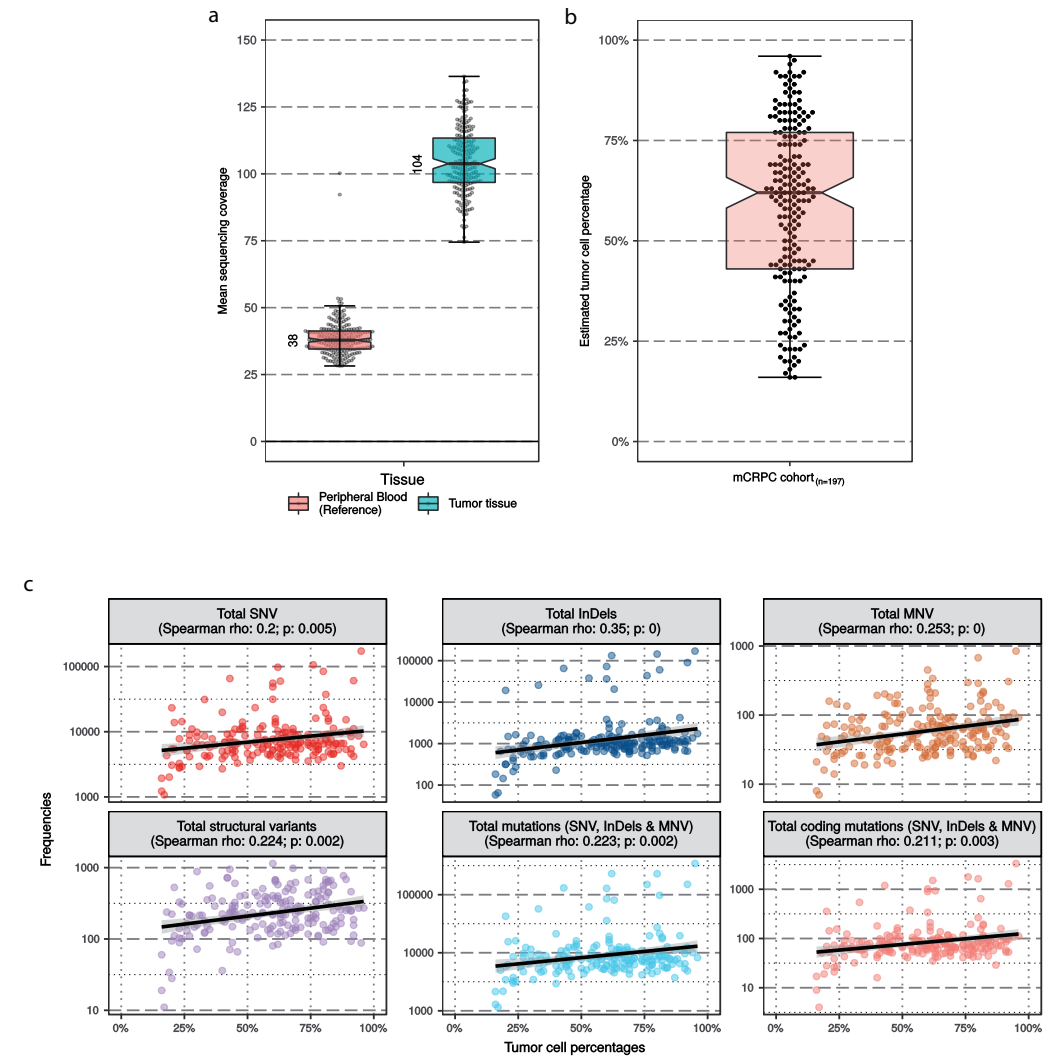
Figure S1. Overview of sequencing quality metrics
(a) Bee-swarm boxplot with notch of the mean read coverage per sample of reference and tumor tissues. Boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Data points outside the IQR are shown.
(b) Bee-swarm boxplot with notch of the estimated (in silico) cohort-wide tumor cell percentages. Boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Data points outside the IQR are shown.
(c) Correlation (Spearman) of estimated tumor cell percentages with observed aberrations per mutational category. Based on the low rank correlation coefficients (Spearman rho) we did not find high correlation with tumor cell percentages and detected events, however a minor correlation could indeed be seen.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration–resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

>>
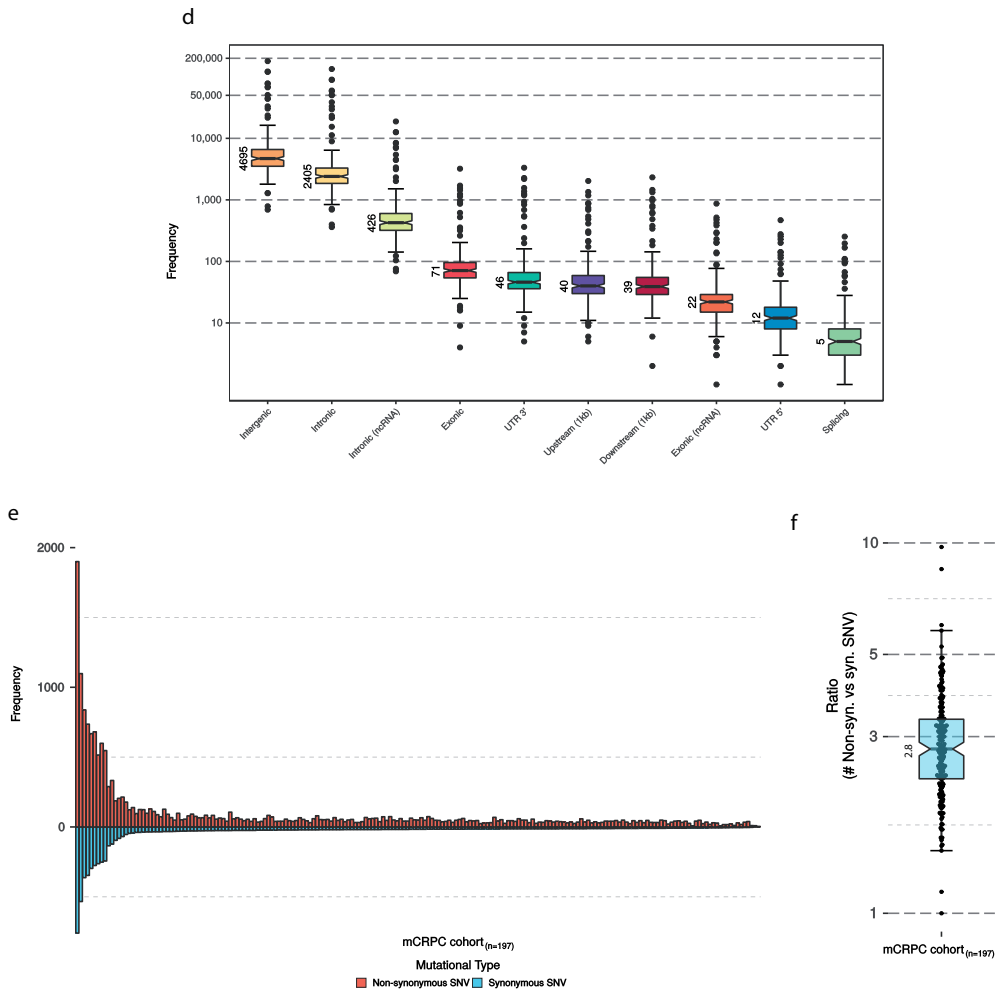
(d) Overview of the locations of variants (SNV / InDels / MNV) in respect to UCSC gene-models. Boxplot with notch depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Data points outside the IQR are shown.
(e) Frequency of non-synonymous (red) and synonymous (blue) SNV per mCRPC sample.
(f) Ratio of non-synonymous over synonymous SNV for the entire mCRPC cohort. Bee-swarm boxplot with notch of the ratio. Boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Data points outside the IQR are shown.
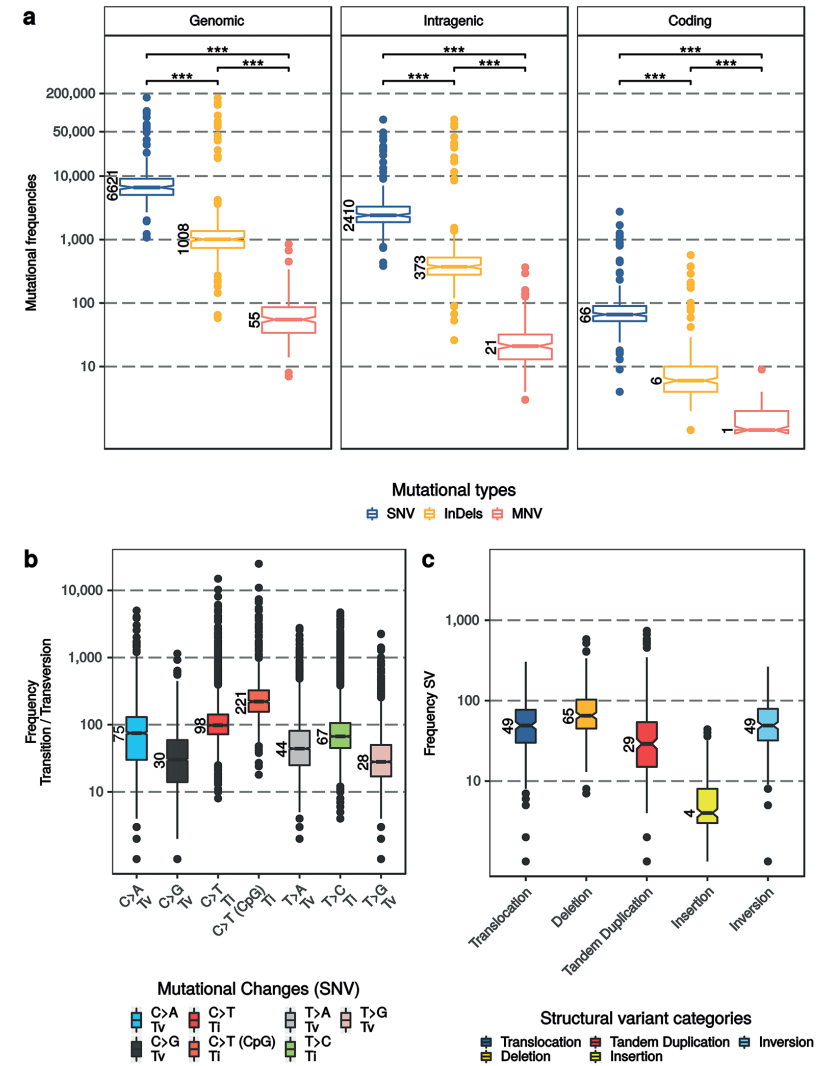
Figure S2. Overview of cohort-wide mCRPC somatic characteristics
(a) Number of SNV (blue), InDels (yellow) and MNV (orange) per whole-genome sequenced sample over three resolutions; genome-wide, within intragenic regions and within coding regions. Boxplot with notch depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Data points outside the IQR are shown. Statistical significance (Wilcoxon rank-sum test) is denoted per comparison.
(b) Type of genome-wide SNVs. Transition (Ti) and transversion (Tv), with a special attention for C to T Ti in CpG context, are indicated per sample. Boxplot with notch depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Data points outside the IQR are shown.
(c) Frequency of Tandem Duplications (DUP), Insertions (INS), Inversions (INV), Deletions (DEL) and interchromosomal translocations (BND) are indicated per sample. Boxplot with notch depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Data points outside the IQR are shown.

>>

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

e

(d) Overview of recurrent copy-number aberrations as detected by GISTIC2. G-scores are depicted on the y-axis ranging from 0 to ≥ 2. Regions with amplifications (G-score > 0) are depicted in green and deletions (G-score < 0) in blue. Regions with significant (and recurring) copy-number aberrations (q ≤ 0.1) are denoted with a darker shade of green or blue, respective of amplification or deletion. Per region, the foci of maximal amplification or deletion (focal peaks; q ≤ 0.1) are denoted in the inner track; the peak identifier is also denoted as presented in Table S3.

(e) Overview of genes detected by the dN/dS algorithm and corresponding mutational categories. Genes not present in one of our lists of known (onco)genes are colored red; (COSMIC v85, CGI, CIVIC and the list from Martincorena et al.[26]). The upper figure displays absolute frequencies per mutational category in the detected genes and the lower figure displays the respective q-value (-1*log$_{10}$(q)). The red line in the bottom figure indicates the threshold for statistical significance (q = 0.01).

a
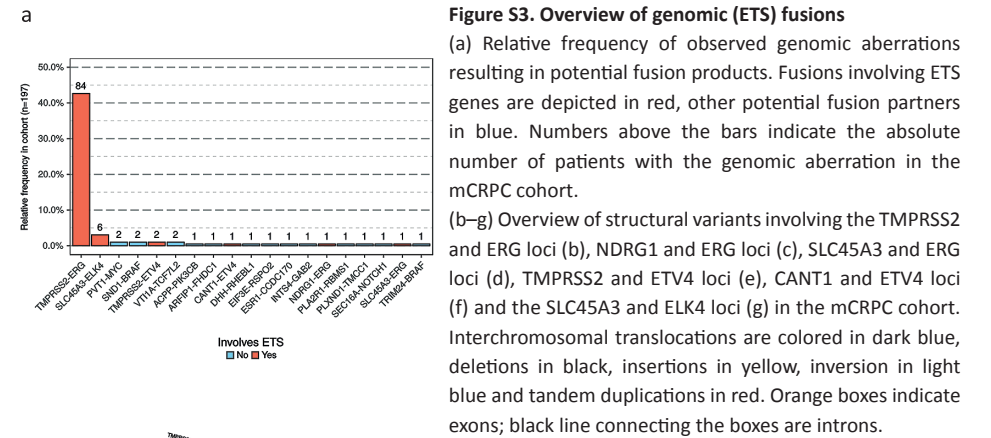
Figure S3. Overview of genomic (ETS) fusions

(a) Relative frequency of observed genomic aberrations resulting in potential fusion products. Fusions involving ETS genes are depicted in red, other potential fusion partners in blue. Numbers above the bars indicate the absolute number of patients with the genomic aberration in the mCRPC cohort.

(b–g) Overview of structural variants involving the TMPRSS2 and ERG loci (b), NDRG1 and ERG loci (c), SLC45A3 and ERG loci (d), TMPRSS2 and ETV4 loci (e), CANT1 and ETV4 loci (f) and the SLC4A3 and ELK4 loci (g) in the mCRPC cohort. Interchromosomal translocations are colored in dark blue, deletions in black, insertions in yellow, inversion in light blue and tandem duplications in red. Orange boxes indicate exons; black line connecting the boxes are introns.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration−resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact
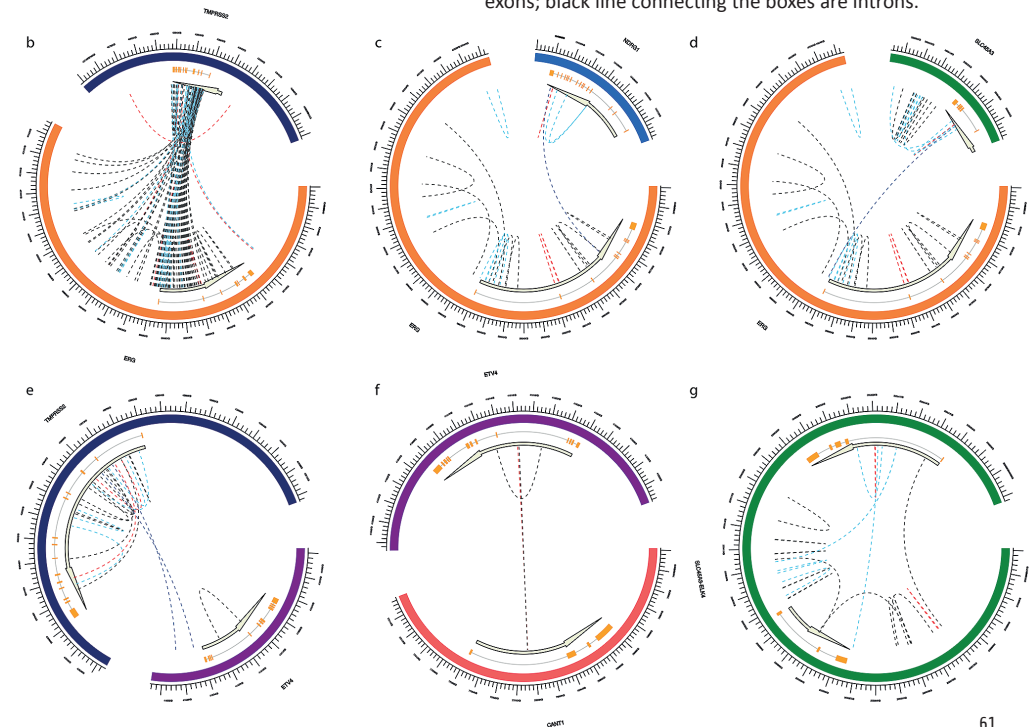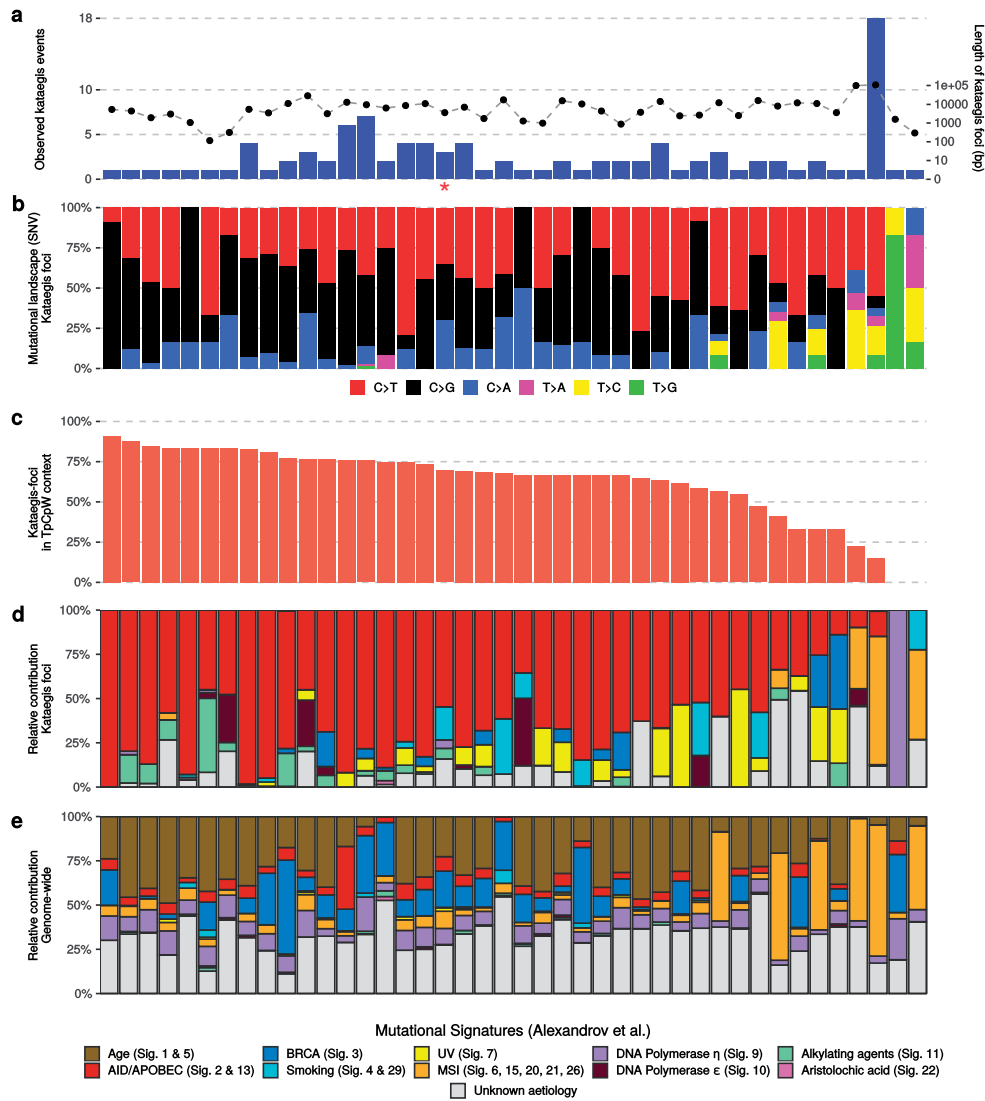
Figure S4. Kataegis prevalence in mCRPC
(a) Number of observed kataegis events in mCRPC cohort samples (n = 42, blue bars) and the respective genomic width of all observed kataegis foci per sample (right y-axis; black points).
(b) Relative frequency of mutational contexts (of SNV) found in all observed kataegis foci per sample.
(c) Relative frequency of SNV in observed kataegis foci in APOBEC-related TpCpW mutational context. W stands for T or A.
(d) Relative contribution to mutational signatures (COSMIC) within the kataegis foci.
(e) Relative contribution to mutational signatures (COSMIC) of all genome-wide events of the sample.
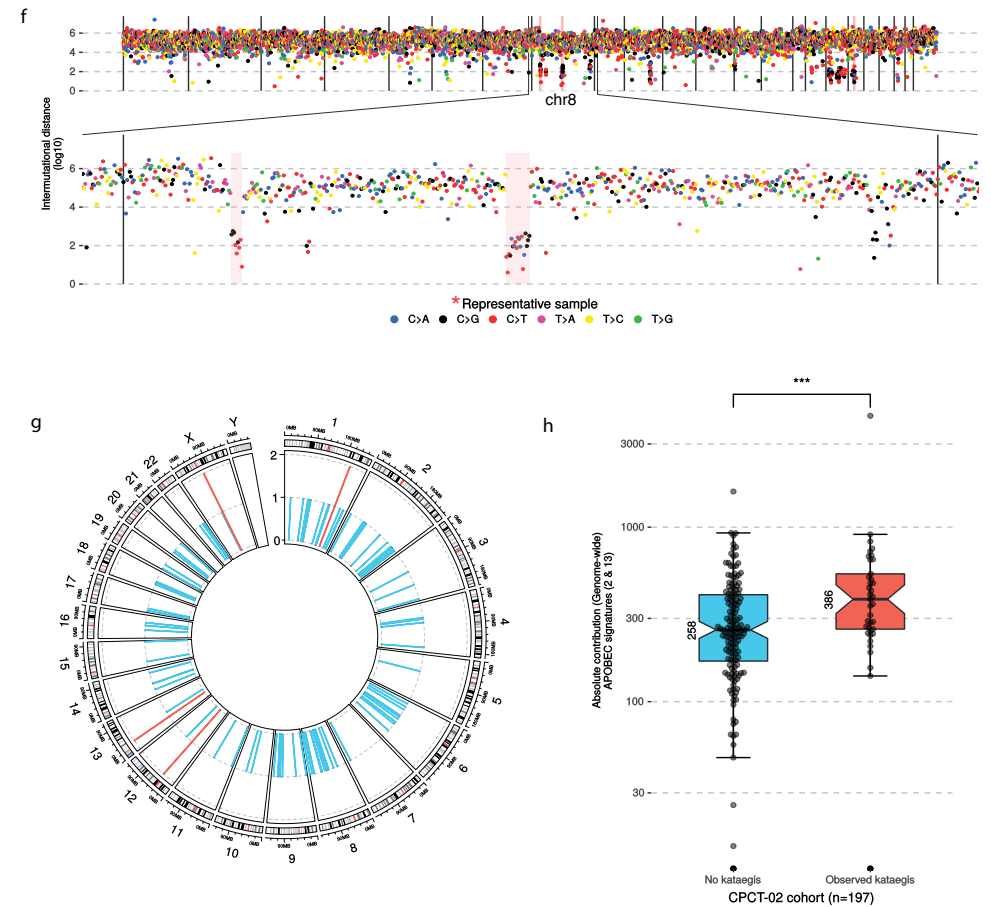
(f) Representation of two distinct kataegis foci on chromosome 8 within a single respective sample (highlighted with * in (a). SNV (colored on Ti/Tv type) are shown with relative genomic distances (in $\log_{10}$) to neighboring SNV. Observed kataegis foci are highlighted with a transparent red background.
(g) Frequency and locations of cohort-wide observed kataegis foci, binned per 1 Mbp. Bins with 2 kataegis events in distinct samples are colored red, else blue.
(h) Absolute contribution of APOBEC signatures (2 & 13) in samples without (n = 155) and with (n = 42) observed kataegis. Bee-swarm boxplot with notch of the mean absolute contribution of APOBEC signatures (2 & 13). Boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Data points outside the IQR are shown. Statistical significance was tested with Wilcoxon rank-sum test.
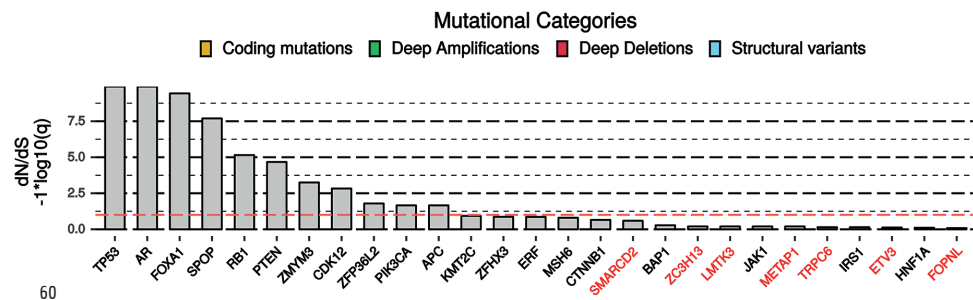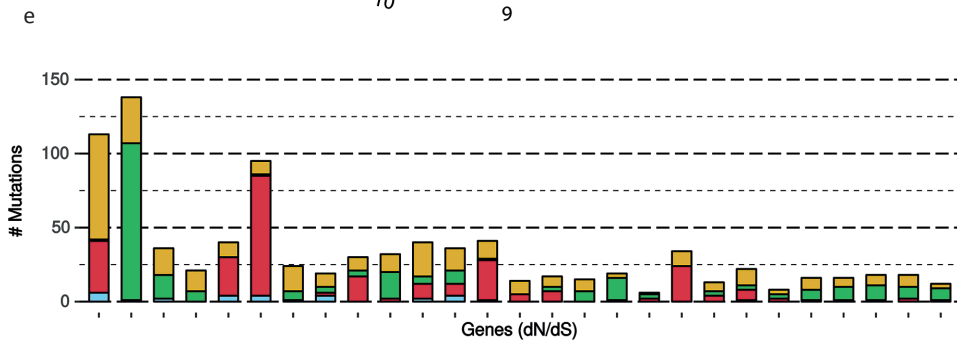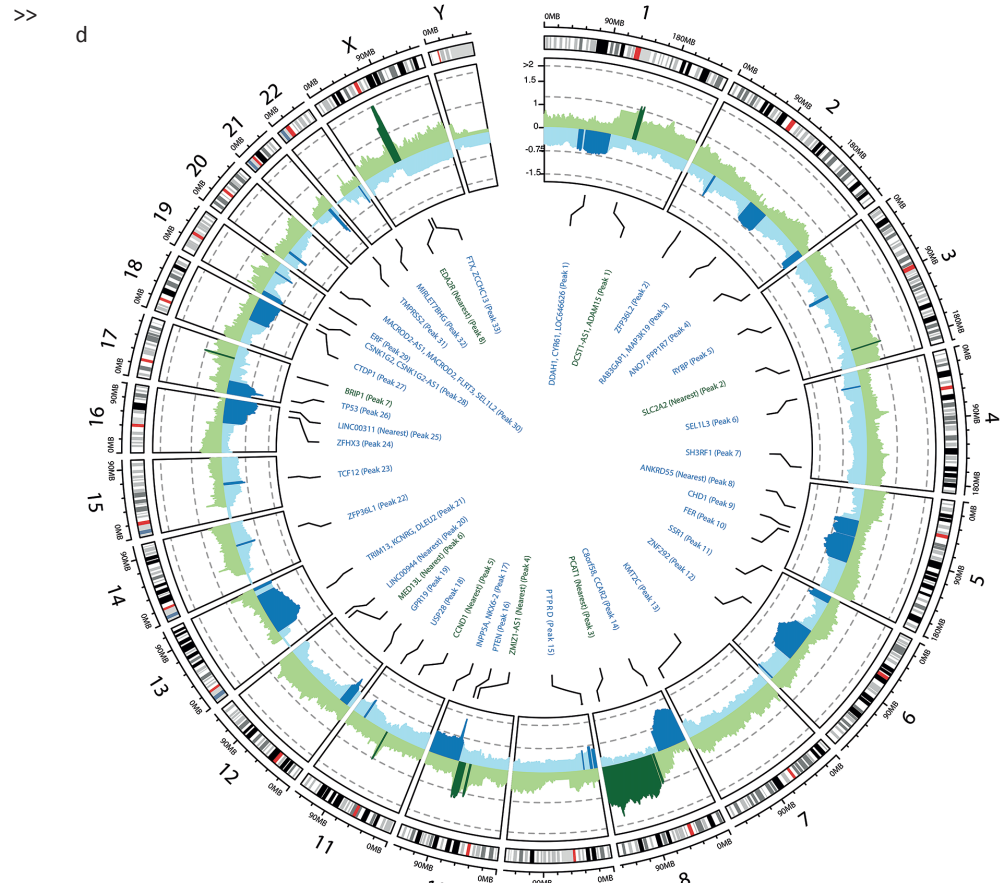
Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact
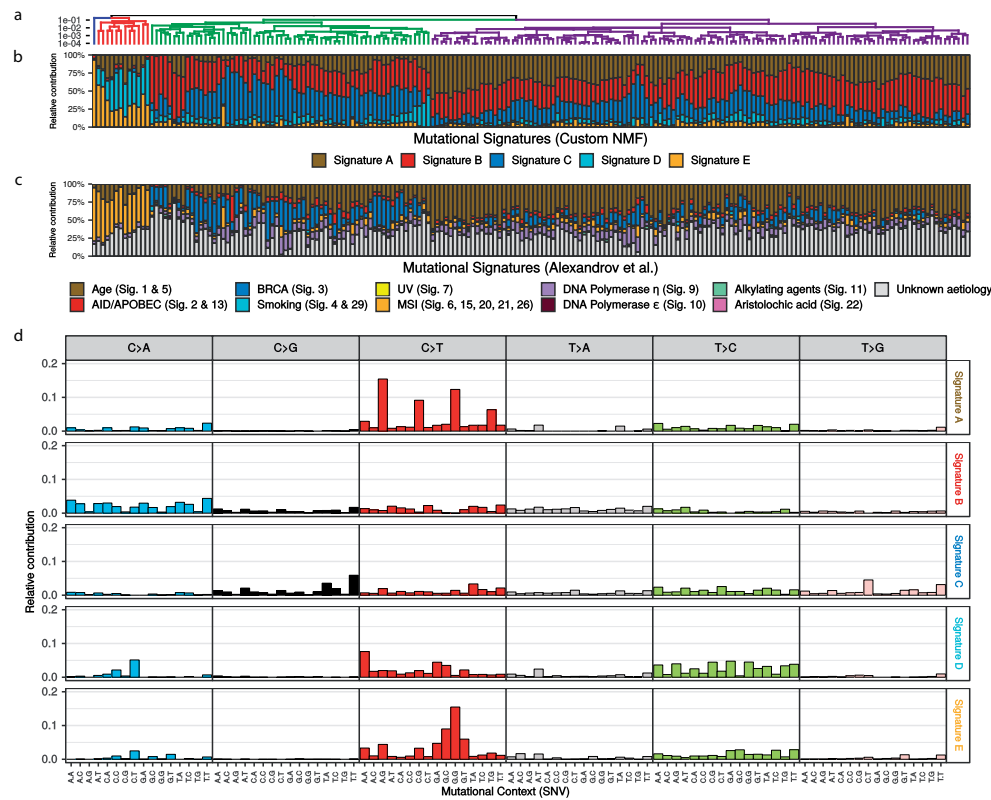
Figure S5. Mutational signature analysis

(a) Dendrogram of unsupervised clustering (Euclidean distance; Ward.D method) on absolute contributions
of SNVs in custom signatures A-E.

(b) Relative contribution to the five custom mutational signatures A-E.

(c) Relative contribution to COSMIC mutational signatures.

(d) Relative distribution of the 96 mutational contexts present in the custom signatures.
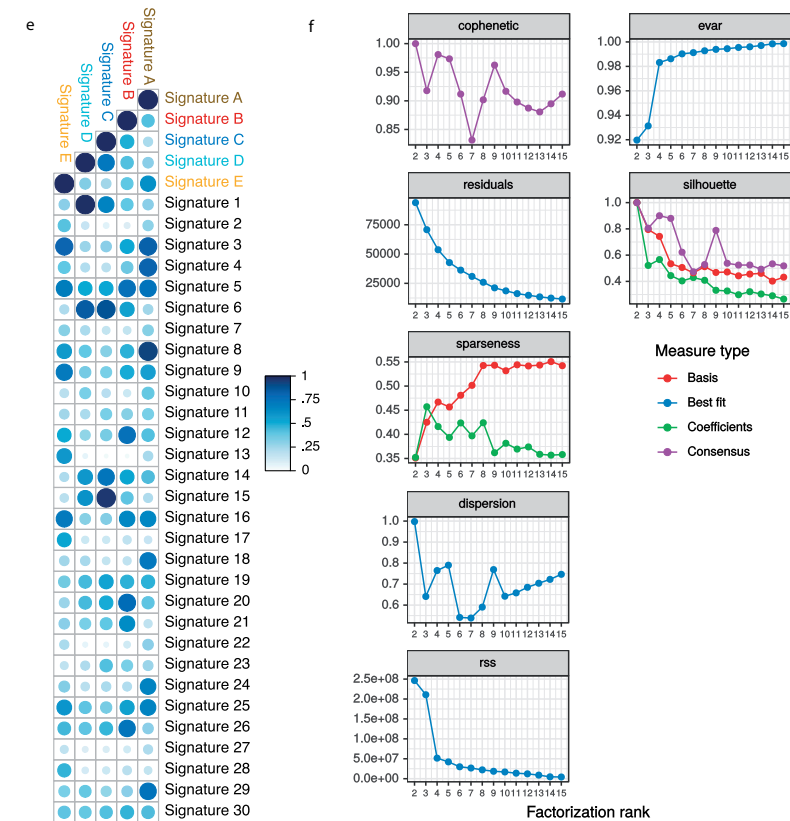
(e) Correlation (cosine similarity) of novel signatures with COSMIC signatures. The size of the dot reflects the
cosine similarity, with higher cosine similarity values shown as larger dots. The color gradient indicates the
level of cosine similarity.

(f) Quality metrics of NMF between two to fifteen ranks.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

Figure S6. The mutational landscape of mCRPC seems unrelated to treatment history

The upper track displays the number of genomic mutations per Mbp (TMB) of SNV (blue), InDels (yellow) and MNV (orange) categories. The heatmap displays the type of mutation(s) per sample; (light-)green or (light-)red backgrounds depict copy-number aberrations whilst the inner square depicts the type of (coding) mutation(s). Relative proportions of mutational categories (coding mutations [SNV, InDels and MNV] (yellow), SV (blue), deep amplifications [high-level amplifications resulting in many additional copies] (green) and deep deletions [high-level losses resulting in (near) homozygous losses] (red)) per gene and foci are shown in the bar plot next to the heatmap. Narrow GISTIC2 peaks covering ≤ 3 genes were reduced to gene-level rows if one of these genes is present in the dN/dS (q ≤ 0.1) analysis or is a known oncogene or tumor-suppressor. For GISTIC2 peaks covering multiple genes, only deep amplifications and deep deletions are shown. Recurrent aberrant focal genomic foci in gene deserts are annotated with their nearest gene. Significance scores (-1*log₁₀(q)) of the dN/dS and GISTIC2 analysis are shown on the outer-right bar plots; bars in the GISTIC2 significance plot are colored red if these foci were detected as a recurrent focal deletion and green if detected as a recurrent focal gain. Per sample, the presence of (predicted) ETS fusions (green), chromothripsis (light pink), kataegis (red), CHORD prediction score (HR-deficiency) (pink gradient), MSI status (dark blue), biopsy location and treatment history are shown as bottom tracks.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact
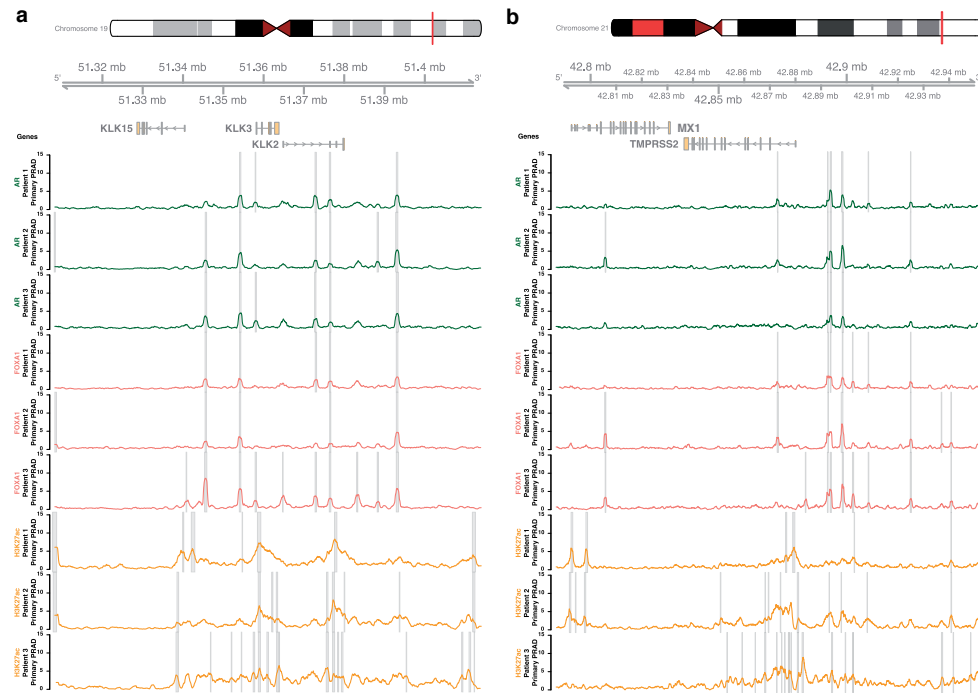
Figure S7. ChIP-seq profiles in primary prostate cancer for known driver genes

ChIP-seq profiles from three independent primary prostate cancer patients surrounding the AR and PCAT1/MYC gene loci (with 1.25 additional Mbp up-/downstream) and two known AR-regulated positive controls (KLK3 and TMPRSS2 with additional 0.5 Mbp up-/downstream). Per subplot, the upper panel displays the selected genomic window and the overlapping genes. The 1th to 3th tracks represent AR ChIP-seq profiles (median read-coverage per 1000 bp windows) in the three primary prostate cancer patients. The 4th to 6th tracks represent FOXA1 ChIP-seq profiles (median read-coverage per 1000 bp windows) in the three primary prostate cancer patients. Finally, the 7th to 9th track represent H3K27ac ChIP-seq profiles (median read-coverage per 1000 bp windows) in the three primary prostate cancer patients. ChIP-seq peaks (MACS/MACS2; q < 0.01) are shown as grey transparent lines per respective sample.

(a) ChIP-seq profiles surrounding the positive control KLK3 region.

(b) ChIP-seq profiles surrounding the positive control TMPRSS2 region.

(c) ChIP-seq profiles surrounding the AR region. The red asterisk denotes the location of the amplified region within the mCRPC setting.

(d) ChIP-seq profiles surrounding the PCAT1/MYC region. The red asterisk denotes the location of the amplified region within the mCRPC setting.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration–resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

a

b

c

d

e

f

g

h

i

j

k

l

m

n
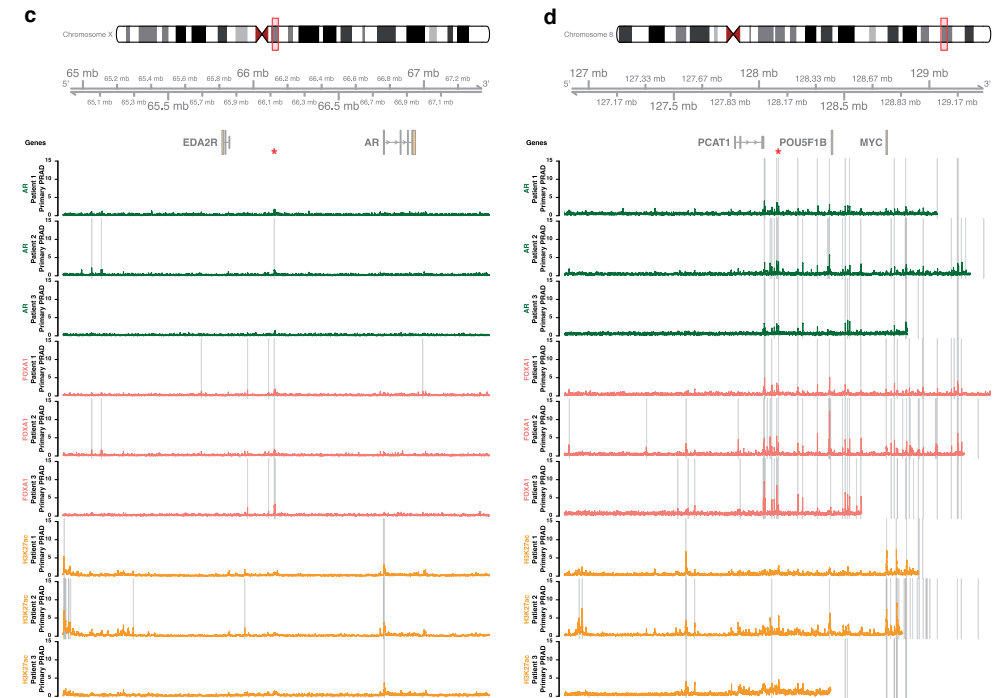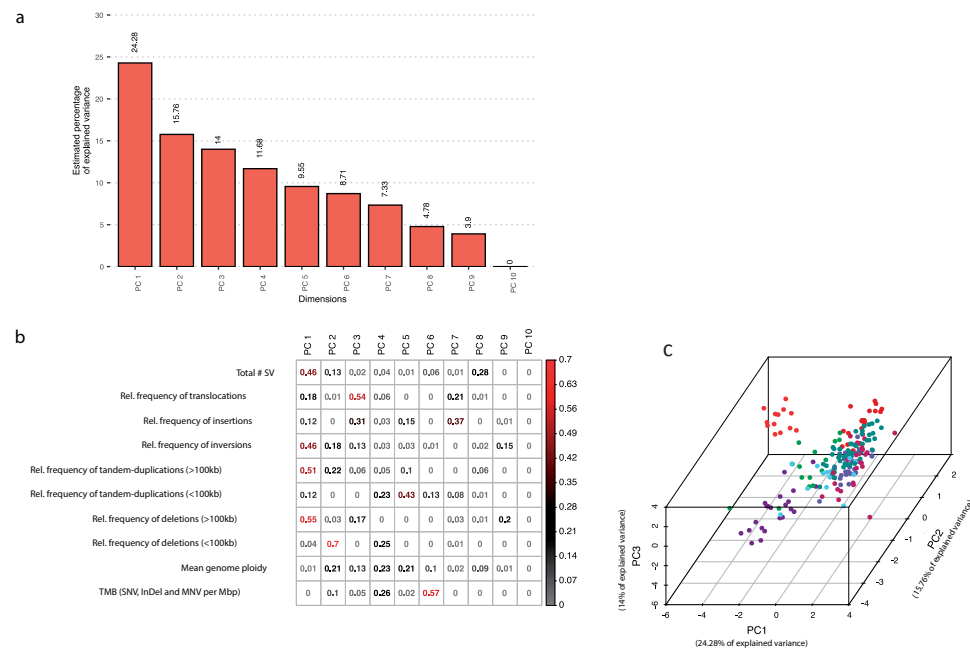
Figure S8. Rationale of the chosen genomic features for unsupervised clustering
Principal component analysis (PCA) and overview of the genomic features included in the unsupervised clustering analysis highlighting the chosen size cut-offs and striking differences between samples.
(a) Overview of the explained variance per principal component (PC) in PCA.
(b) The quality of representation for each feature per principal component (cos2), this ranges from 0 (no importance / representation in PC) to 1 (absolute importance / representation in PC). Color gradient (0 to 0.7) denotes cos2, red values denote important / representation of feature within PC. Numbers shown are the cos2 values.
(c) Visualization of the first three principal components of PCA, each sample is colored based on their assigned cluster (depicted in n) after unsupervised clustering on their genomic features.
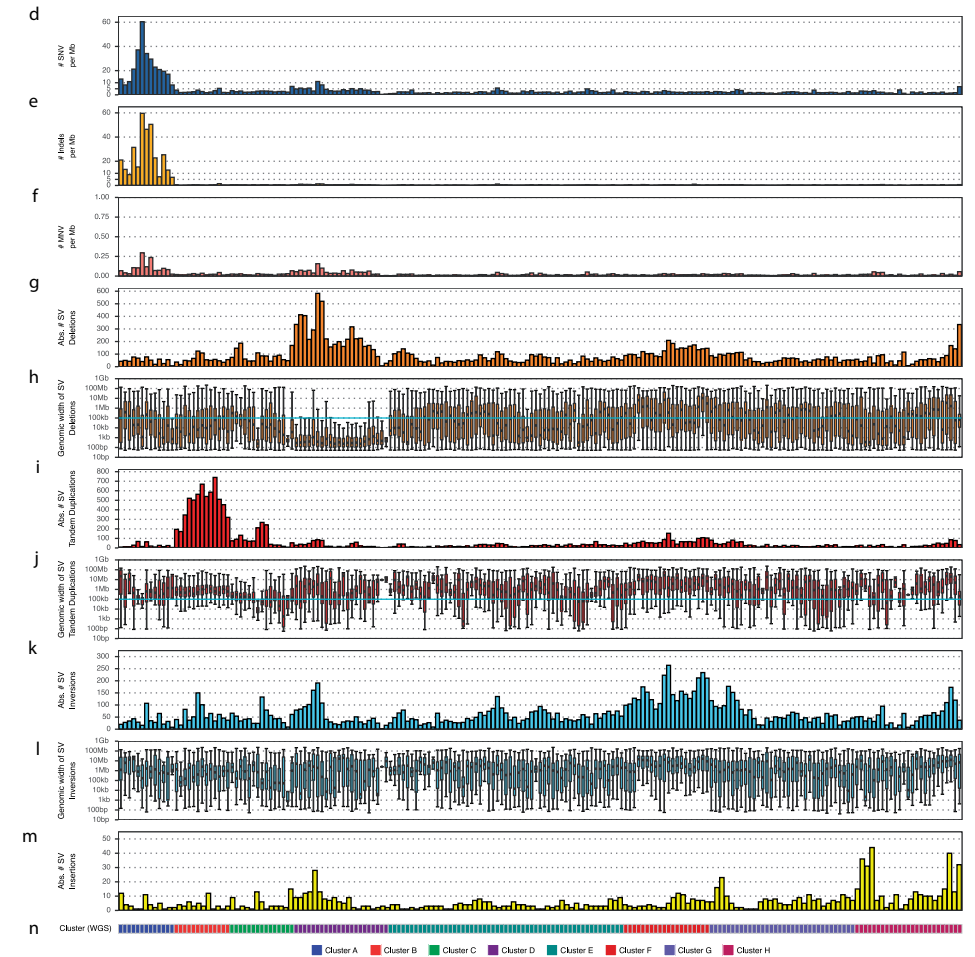
(d) All genome-wide somatic SNVs per Mbp (square root scale).
(e) All genome-wide somatic InDels per Mbp (square root scale).
(f) All genome-wide somatic MNV per Mbp (square root scale).
(g) Absolute number of deletions (SV) per sample.
(h) Distribution of the genomic width of deletions (SV) per sample. Cyan line indicates the chosen size cut-offs (< 100 kbp and ≥ 100 kb).
(i) Absolute number of tandem duplications (SV) per sample.
(j) Distribution of the genomic width of tandem duplications (SV) per sample. Cyan line indicates the chosen size cut-offs (< 100 kbp and ≥ 100 kb).
(k) Absolute number of inversions (SV) per sample.
(l) Distribution of the genomic width of inversions (SV) per sample.
(m) Absolute number of insertions (SV) per sample. Genomic width of insertions could not be estimated accurately due to repeat-like sequences.
(n) Assigned clusters based on unsupervised clustering of genomic features.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration–resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact
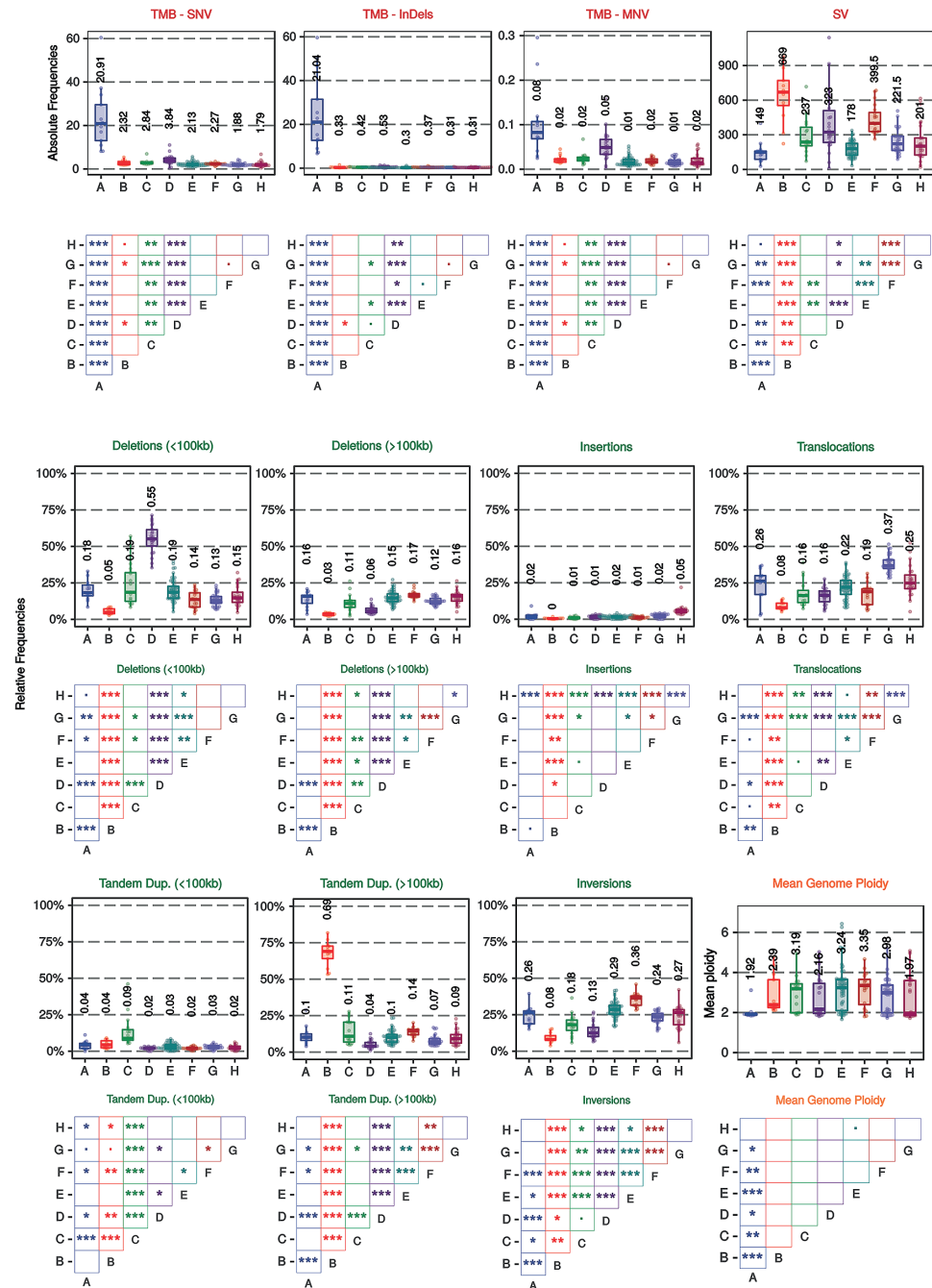
Figure S9. Cluster characteristics
Overview of genomic characteristics and COSMIC mutational signatures per cluster (A-H) derived from unsupervised clustering of the mCRPC cohort using basic WGS characteristics. Bee-swarm boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Data points outside the IQR are shown. A pairwise Wilcoxon rank-sum test (BH correction) was performed to detect statistically significant differences between clusters; * denotes P ≤ 0.05, ** denotes P ≤ 0.01, *** denotes P ≤ 0.001. Significant differences of events were also found in clusters without a clear biological association (C, E and G-H), such as increased numbers of translocations in cluster G and insertions in cluster H.
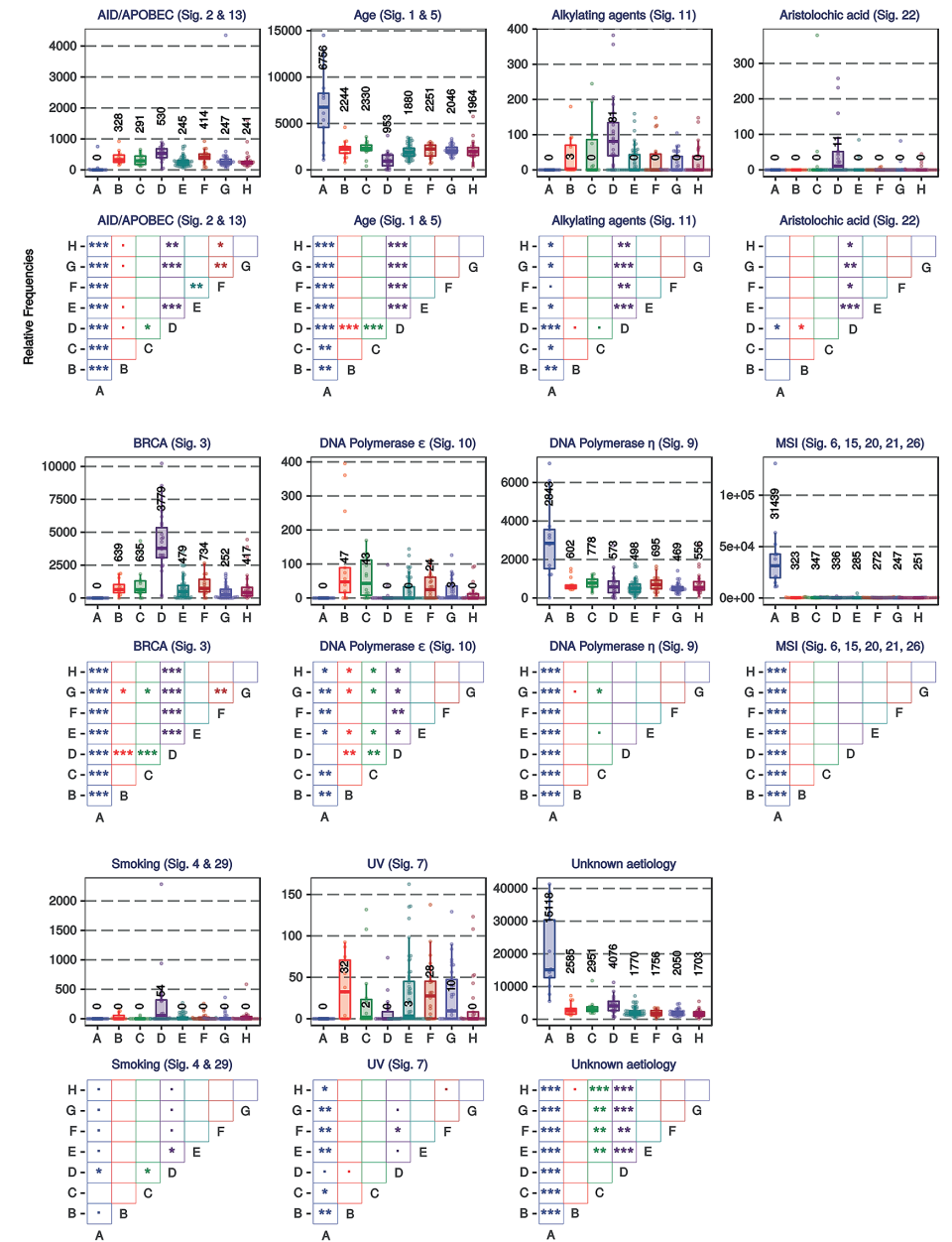
Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration–resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

a

Optimal number of clusters
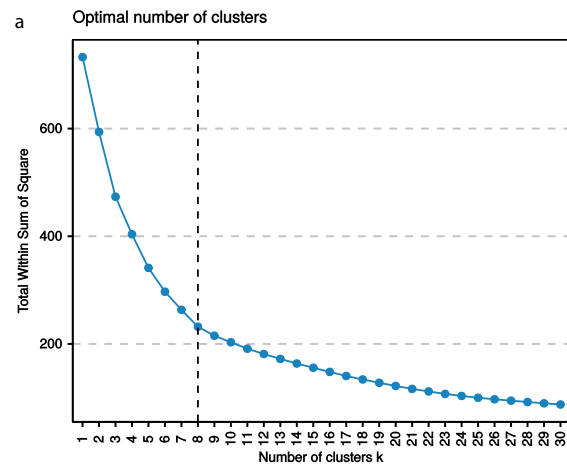
Figure S10. Clustering QC
(a) Clustering estimation using optimum total within-cluster sum of square (wss). The final selection of the most optimal number of clusters was based on the knee in the blue line, e.g. the moment when increasing the number of clusters does not dramatically decrease wss.
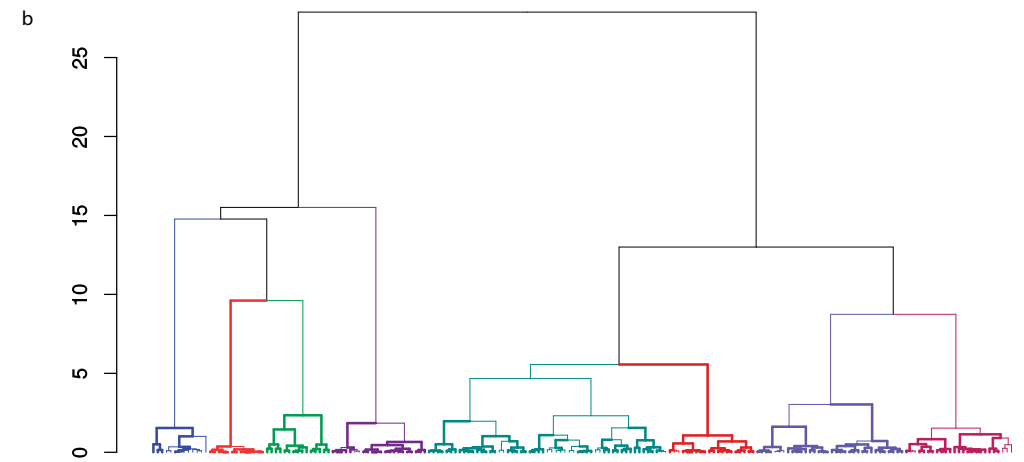
(b) Bootstrapping results (5000 iterations) of the unsupervised hierarchical clustering with additional coloring of the eight defined clusters as used in this manuscript. Branches with Approximately Unbiased P-values (AU) ≤ 0.05 are highlighted with bolder lines and reflect significantly robust groups of samples based on similar characteristics derived from WGS. Y-axis displays clustering distance (Pearson correlation; ward.D).

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

Cluster A
n = 13

Cluster B
n = 13

Cluster C
n = 15

Cluster D
n = 22

Cluster E
n = 55

Cluster F
n = 20

Cluster G
n = 34
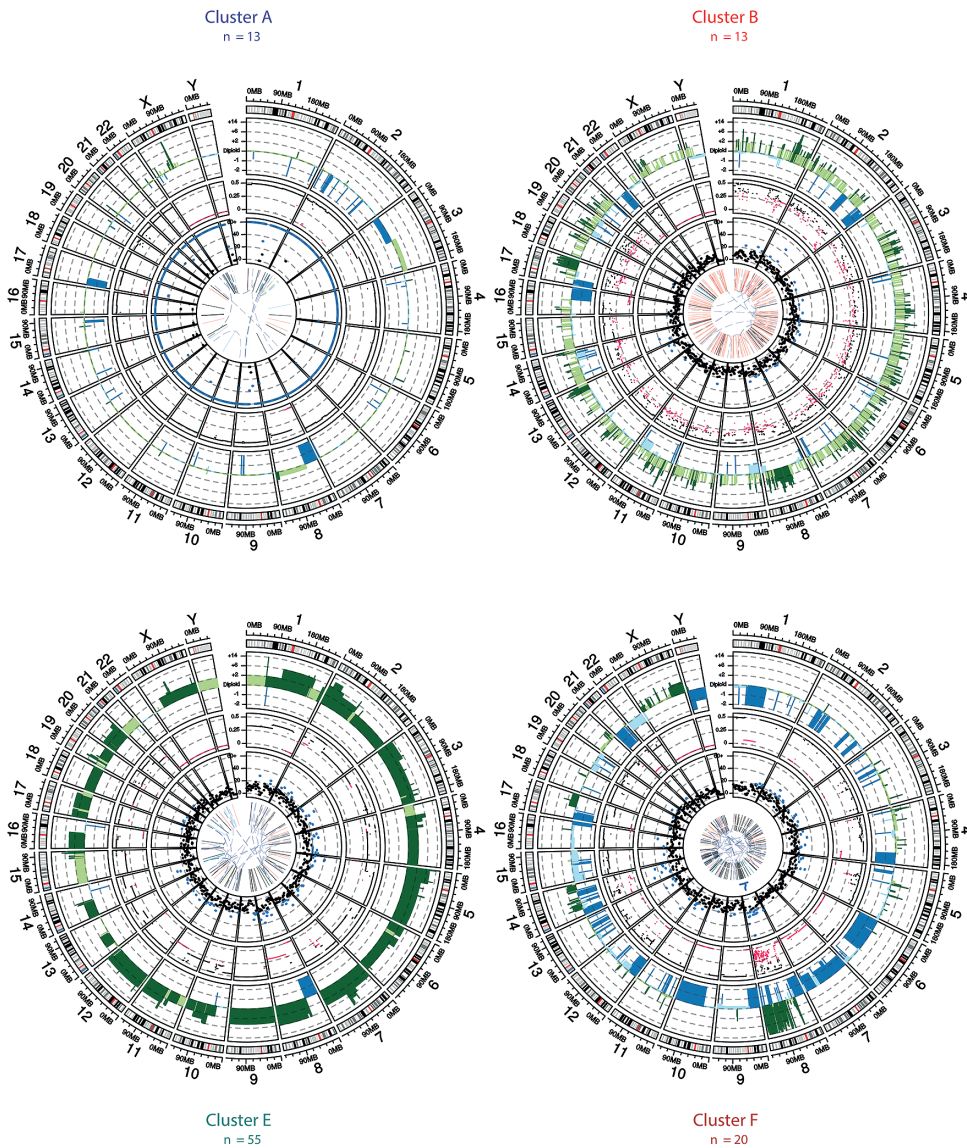
Cluster H
n = 25

Figure S11. Representative mCRPC sample per cluster

Genomic overviews of a single representative sample per (unsupervised) cluster. The outer track displays the genomic ideogram, the second-outer track displays copy-number profiles (amplification in light green; deep amplification beyond sample-specific threshold (GISTIC2) in dark green, deletions in blue; deep deletions beyond sample-specific threshold (GISTIC2) in dark blue). The third track displays tumor cell percentage-corrected minor allele-frequency (MAF) values of individual copy-number segments (MAF ≤ 0.33 in pink; MAF ≥ 0.33 in black). The fourth track displays the number of mutations per 5 Mbp, ranging from 0 to 60+; bins with ≥ 20 mutations are highlighted in blue. The innermost track displays structural variants; interchromosomal translocations in dark blue, deletions in grey, insertions in yellow, inversion in light blue and tandem duplications in red.

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

Mutations
BRCA2

32900000

Cluster G - chr13:32900730T>G (c.838 - Missense)

Cluster D - chr13:32907202TA>T (c.1815 - Frameshift)
Cluster E - chr13:32912615G>T (c.4350 - Truncating)
Cluster D - chr13:32913166T>A (c.4901 - Missense)
Cluster D - chr13:32914157A>AT (c.5893 - Frameshift)
Cluster H - chr13:32915094C>G (c.6829 - Missense)
Cluster D - chr13:32915120G>C (c.6855 - Missense)
Cluster D - chr13:32915237GCCACACATTCT>G (c.6973 - Frameshift)

32920000

Cluster D - chr13:32915254TACATGTCCCG>T (c.6990 - Frameshift)

32940000

Cluster E - chr13:32945129C>T (c.8751*† - Missense)

Cluster D - chr13:32954022C>CA (c.9317† - Frameshift) Pathogenic
Cluster A - chr13:32954022CA>C (c.9317† - Frameshift) Pathogenic

Cluster D - chr13:32954022C>CA (c.9317† - Frameshift) Pathogenic

32960000

Exon Frequency of variant allele
UTR   Frequency of reference allele

Figure S12. Overview of BRCA2 mutations
Genomic distribution of non-synonymous BRCA2 mutations found within our mCRPC cohort. Mutations are displayed as pie-charts depicting the variant allele frequency (red portion of the pie-chart) and reference allele frequency (white portion of the pie-chart). Samples are colored based on their respective cluster after unsupervised clustering (figure 5). Known COSMIC mutations are annotated with * and/or known dbSNP variants with a †. Alleles with known pathogenicity within ClinVar are highlighted, mutations without ClinVar annotation could be considered as variants with as-of-yet uncertain significance.

Mutational categories
SNV Indels MNV

Structural variants

Structural variant categories
Translocation Inversion
Deletion (>100kb) Tandem Duplication (>100kb)
Tandem Duplication (<100kb) Deletion (<100kb)

Ploidy status
0 1 2 3 4 5 6 7

Mutational Signatures (COSMIC)
Age (Sig. 1, 5) MSI (Sig. 6, 15, 20, 21, 26)
APOBEC (Sig. 2 & 13) DNA Polymerase ε (Sig. 9)
BRCA (Sig. 3) DNA Polymerase ε (Sig. 10)
Smoking (Sig. 4 & 29) Alkylating agents (Sig. 11)
UV (Sig. 7) Aristolochic acid (Sig. 22)
Unknown aetiology

Mutational Categories
Frameshift Variant Shallow Deletion
Inframe Insertion/Deletion Shallow Gain
Missense Variant Deep Deletion
Structural Variant Deep Gain
Stop Gained
Splice Region Variant

MSI Status
MSI MSS

CHORD Prediction Score
0 .25 .5 .75 1

Mutations per Mb
(Genome-wide TMB)

Structural variants

Relative Frequency

Relative frequency

Relative contribution
Genome-wide

Genes
ERG (n=85)
FOXA1 (n=20)
SPOP (n=14)
ETV4 (n=3)
IDH1 (n=2)
FLI1 (n=1)
ETV1 (n=1)
TP53 (n=104)
PTEN (n=90)
MXD3-1 (n=42)
RB1 (n=37)
KMT2C (n=29)
CHD1 (n=26)
BRCA2 (n=25)
CDK12B (n=19)
KMT2D (n=18)
ZMYM3 (n=13)
ATM (n=13)
PIK3CA (n=14)
CDK12 (n=13)
BRAF (n=8)
CTNNB1 (n=8)
MED12 (n=5)
AKT1 (n=5)
HRAS (n=3)
ZNF770 (n=3)
CHORD
MSI
Chromothripsis
Kataegis
Cluster (WGS)

a

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

>>

Figure S13. Supervised clustering of mCRPC based on TCGA criteria

(a) Samples are sorted based on mutual-exclusivity of the same genes and aberrations as the TCGA clustering, depicted in red colors in the heatmap. In addition, all genes which the TCGA defined as recurrent alterations in primary prostate cancer are shown, genes also discovered in the mCRPC cohort as enriched in non-synonymous mutations or copy-number alterations (dN/dS and/or GISTIC2) are depicted with a red asterisk. The upper track displays the number of genomic mutations per Mbp (TMB) of SNV (blue), InDels (yellow) and MNV (orange) categories. Second track displays the absolute number of unique structural variants per sample. Third track displays the relative frequency per structural variant category, Tandem Duplications and Deletions are subdivided into > 100 kbp and < 100 kbp categories. The fourth track displays the relative genome-wide ploidy status, ranging from 0 to ≥ 7 copies and the fifth track displays the relative contribution to mutational signatures (COSMIC) summarized per proposed etiology. The heatmap displays the type of mutation(s) per sample; (light-)green or (light-)red backgrounds depict copy-number aberrations whilst the inner square depicts the type of (coding) mutation(s). In addition, the lower tracks display CHORD prediction score (HR-deficiency) (pink gradient), MSI status (blue), chromothripsis (pink), presence of kataegis (red) and in which of the eight genomic cluster, as defined by this manuscript, each sample falls.

(b) Overview of the relative frequency of samples captured per mutually-exclusive group for both the TCGA and mCRPC cohort. Promiscuous ETS family fusions (ETV1, ETV4 and FLI1) which were captured in the TCGA cohort using mRNA overexpression were split as the mCRPC cohort did not have accompanying mRNA sequencing data to perform a similar capturing.

>>

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration-resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact
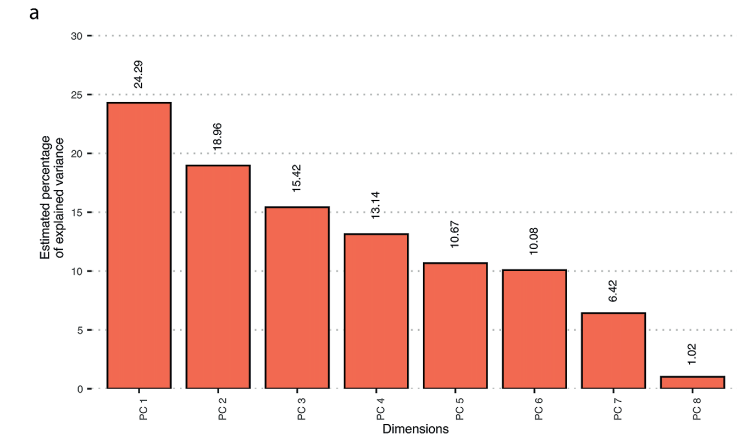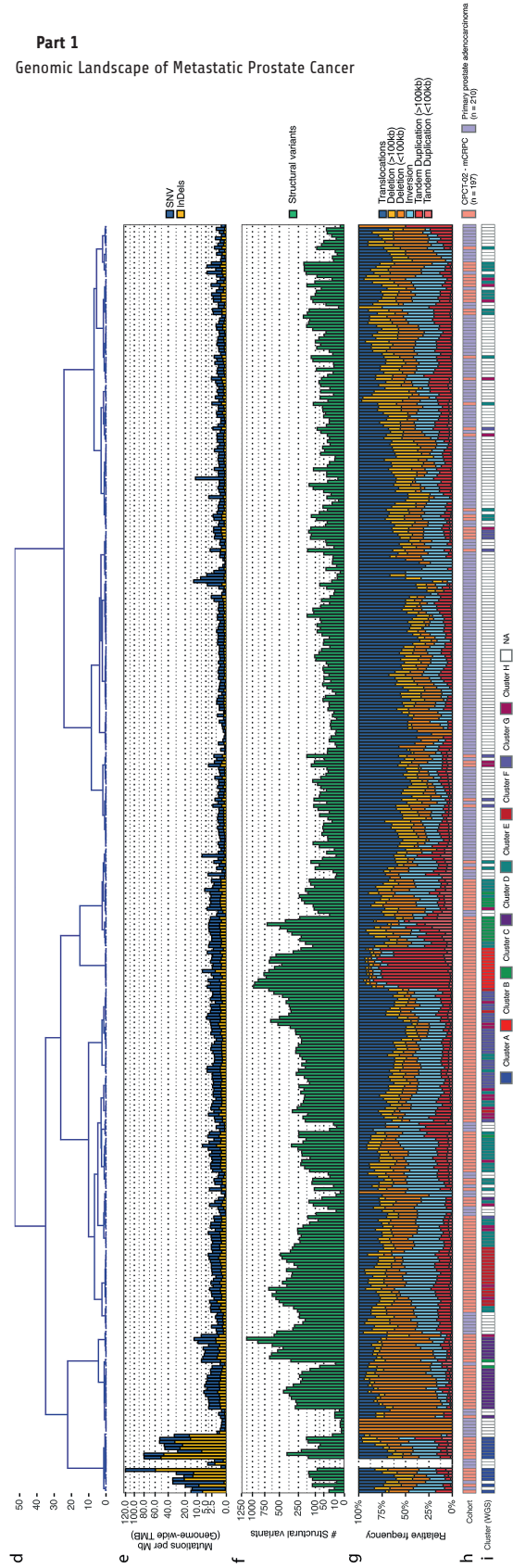
Figure S14. Overview of clustering scheme on primary prostate cancer and mCRPC
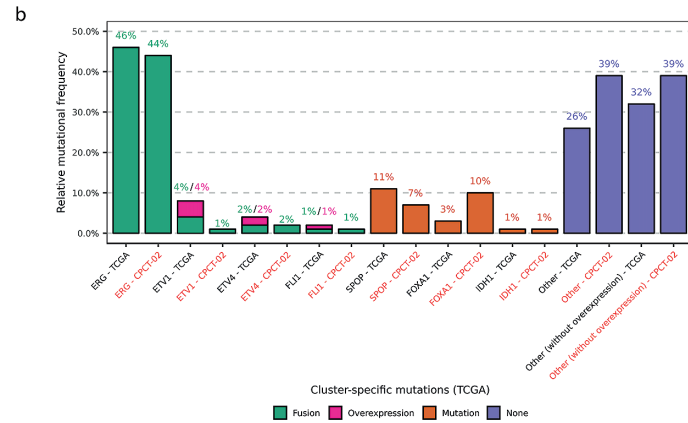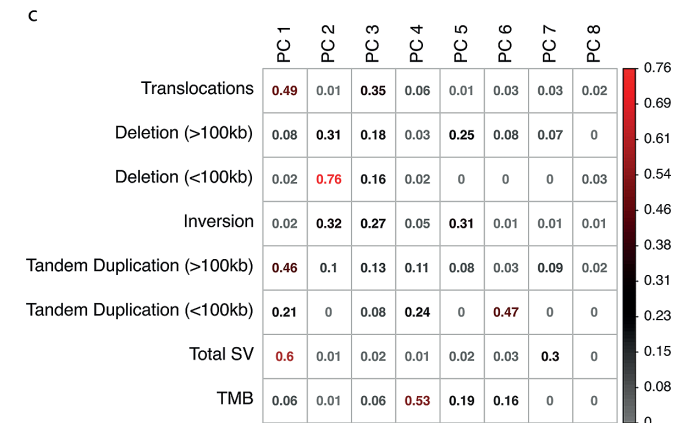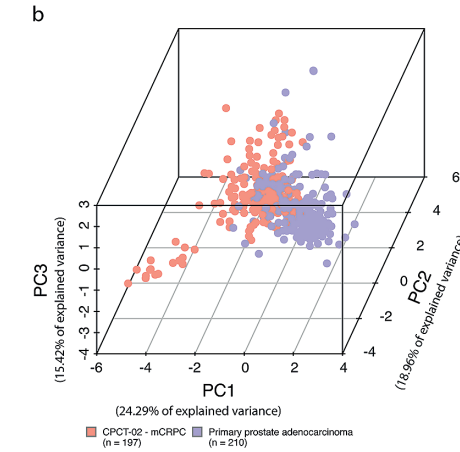
(a) Overview of the explained variance per principal component (PC) in PCA of the combined dataset of primary prostate cancer (n = 210) and mCRPC (n = 197). PCA was performed in the following features: Total number of SV, genome-wide TMB (SNV and InDels) and relative frequency of structural variants (except insertions).

(b) Visualization of the first three principal components of PCA, each sample is colored based on their respective disease-setting, primary prostate cancer is colored as violet whilst mCRPC is colored as light-red.

(c) The quality of representation for each feature per principal component (cos2), this ranges from 0 (no importance / representation in PC) to 1 (absolute importance / representation in PC). Color gradient (0 to 0.7) denotes cos2, red values denote important / representation of feature within PC. Numbers shows are the cos2 values.

(d) Dendrogram of unsupervised clustering with optimal leaf ordering on genomic features. Y-axis displays clustering distance (Pearson correlation; ward.D).

(e) All genome-wide somatic SNVs (blue) and InDels (yellow) per Mbp (square root scale).

(f) Absolute frequency of structural variants per sample (square root scale).

(g) Relative frequency per structural variant category, Tandem Duplications and Deletions are subdivided into > 100 kbp and < 100 kbp categories.

(h) Respective cohort of the samples, primary prostate cancer is colored as violet whilst mCRPC is colored as light-red.

(i) Assigned clusters of mCRPC samples based on unsupervised clustering of genomic features as described in figure 5.

Table S1. Participating centers

Organization	Local principal investigator	Included patients for this study (n)
Radboud UMC, Nijmegen	Carla van Herpen	91
Erasmus MC, Rotterdam	Martijn Lolkema	38
Franciscus Gasthuis & Vlietland, Rotterdam	Paul Hamberg	20
NKI-AVL, Amsterdam	Neeltje Steeghs	15
Isala, Zwolle	Jan Willen de Groot	3
Martini Ziekenhuis, Groningen	Johan van Rooijen	3
Medisch Centrum Leeuwarden	Hiltje de Graaf	3
Maastricht UMC, Maastricht	Vivianne Tjan-Heijnen	3
Noordwest Ziekenhuisgroep, Alkmaar	Mathijs Hendriks	3
UMC Utrecht, Utrecht	Els Witteveen	3
Amphia Ziekenhuis, Breda	Bert Jan ten Tije	2
Reinier de Graaf Gasthuis, Delft	Annelie Vulink	2
Treant Zorggroep, Hoogeveen	Sophia van den Boogerd	2
Zuyderland Medisch Centrum, Geleen	Frans Erdkamp	2
ETZ Elisabeth, Tilburg	Laurens Beerepoot	1
Leids Universitair Medisch Centrum, Leiden	Hans Gelderblom	1
Maasstad Ziekenhuis, Rotterdam	Rineke Leys	1
Meander Medisch Centrum, Amersfoort	Haiko Bloemendal	1
St. Antonius Ziekenhuis, Utrecht	Maartje Los	1
VUmc, Amsterdam	Henk Verheul	1
ZGT, Almelo	Esther Siemerink	1

Part 1
Genomic Landscape of Metastatic Prostate Cancer

Chapter 2
The genomic landscape of metastatic castration–resistant prostate cancers
reveals multiple distinct genotypes with potential clinical impact

Table S2. Patient characteristics

Patients (n = 197)		
	n	**%**
Age at biopsy		
Median	68	
Range (min-max)	48-83	
Prior ADT		
Yes	197	100.0
Drug-based	181	91.9
Surgery-based (orchiectomy)	3	1.5
With Docetaxel	6	3.0
No clear documentation of ADT type	7	3.6
Prior systemic therapy (other than ADT)		
0 previous treatments	27	13.7
≥ 1 previous treatments	170	86.3
1 previous treatment	45	22.8
2 previous treatments	69	35.0
3 previous treatments	31	15.7
4 previous treatments	19	9.6
5 previous treatments	6	3.0
Type of prior systemic therapy (other than ADT)		
Hormonal therapy only	20	10.2
Chemotherapy only	37	18.8
Radionucleotide therapy only	4	2.0
Immunotherapy only (Dendritic cell therapy)	4	2.0
Targeted therapy only	0	0.0
Hormonal and chemotherapy	68	34.5
Hormonal and radionucleotide therapy	3	1.5
Chemotherapy and radionucleotide therapy	3	1.5
Hormonal and immunotherapy	3	1.5
Chemotherapy and immunotherapy	3	1.5
Hormonal, chemotherapy and radionucleotide therapy	15	7.6
Hormonal, chemotherapy and immunotherapy	4	2.0
Hormonal, radionucleotide and immunotherapy	2	1.0
Hormonal, chemotherapy and targeted therapy (Olaparib)	2	1.0
Hormonal, chemotherapy, radionucleotide and immunotherapy	1	0.5
Unknown at time of analysis	1	0.5
Prior radiotherapy		
Yes (curative radiotherapy of the prostate and/or palliative radiotherapy of metastases)	117	59.4
No	77	39.1
Unknown at time of analysis	3	1.5
Started therapy after biopsy for whole-genome sequencing		
Yes	138	70.1
Hormonal therapy	53	26.9
Chemotherapy	56	28.4
Radionucleotide therapy	12	6.1
Immunotherapy (Pembrolizumab)	6	3.0
Targeted therapy	3	1.5
Combinational therapy	6	3.0
Other*	2	1.0
No	19	9.6
Unknown at time of analysis	40	20.3
Biopsy site		
Liver	29	14.7
Lymph node	81	41.1
Bone	70	35.5
Lung	3	1.5
Soft tissue/Other**	14	7.1

*Boneregulating agent
**Soft tissue/other: (sub)cutis, muscle, peritoneum, kidney, bladder, adrenal gland*

Data file S1.

Can be accessed through https://www.nature.com/articles/s41467-019-13084-7
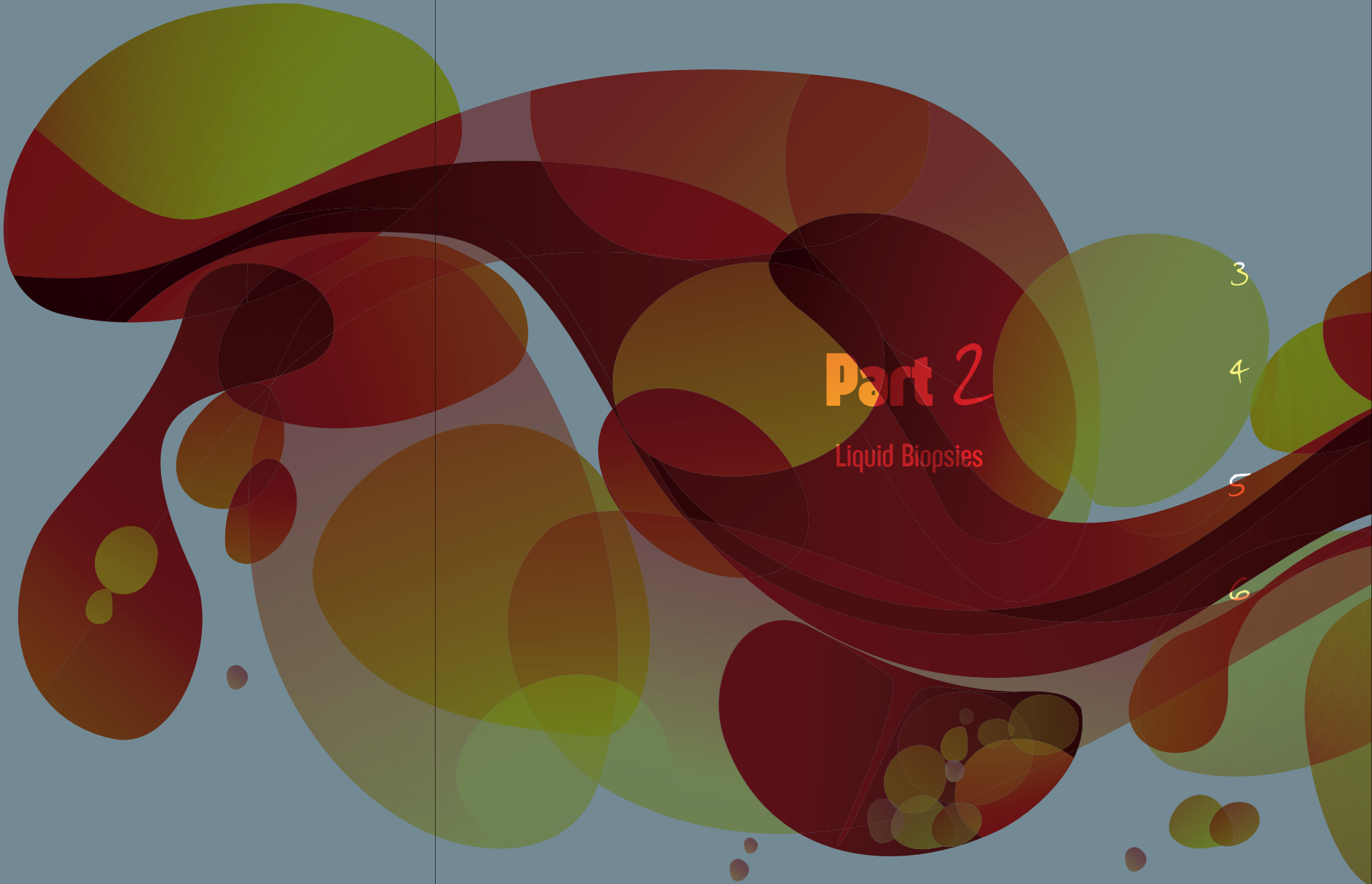
Part 2

Liquid Biopsies

Chapter 3

Application of circulating tumor DNA
in prospective clinical oncology trials – standardization
of preanalytical conditions

Chapter 3

Lisanne F. van Dessel, Nick Beije, Jean C. Helmijr, Silvia R. Vitale, Jaco Kraan, Maxime P. Look, Ronald de Wit,
Stefan Sleijfer, Maurice P.H.M. Jansen, John W.M. Martens, Martijn P. Lolkema

Part 2
Liquid Biopsies

Chapter 3
Application of circulating tumor DNA in prospective clinical oncology trials – standardization
of preanalytical conditions

Abstract

Circulating tumor DNA (ctDNA) has emerged as a potential new biomarker with diagnostic, predictive, and prognostic applications for various solid tumor types. Before beginning large prospective clinical trials to prove the added value of utilizing ctDNA in clinical practice, it is essential to investigate the effects of various preanalytical conditions on the quality of cell-free DNA (cfDNA) in general and of ctDNA in particular in order to optimize and standardize these conditions. Whole blood samples were collected from patients with metastatic cancer bearing a known somatic variant. The following preanalytical conditions were investigated: (a) different time intervals to plasma isolation (1, 24, and 96 h) and (b) different preservatives in blood collection tubes (EDTA, CellSave, and BCT). The quality of cfDNA/ctDNA was assessed by DNA quantification, digital polymerase chain reaction (dPCR) for somatic variant detection and a β-actin fragmentation assay for DNA contamination from lysed leukocytes. In 11 (69%) of our 16 patients, we were able to detect the known somatic variant in ctDNA. We observed a time-dependent increase in cfDNA concentrations in EDTA tubes, which was positively correlated with an increase in wild-type copy-numbers and large DNA fragments (> 420 bp). Using different preservatives did not affect somatic variant detection ability, but did stabilize cfDNA concentrations over time. Variant allele frequency was affected by fluctuations in cfDNA concentration only in EDTA tubes at 96 h. Both CellSave and BCT tubes ensured optimal ctDNA quality in plasma processed within 96 h after blood collection for downstream somatic variant detection by dPCR.

Introduction

Circulating tumor DNA (ctDNA) has emerged as a potential new biomarker in the field of oncology. The quantification and characterization of ctDNA in plasma creates numerous potential applications, including detection of minimal residual disease, early evaluation of treatment response and stratification for targeted therapy according to specific genetic changes[1-7].

The application of ctDNA-based diagnostic tests into the clinic still faces several technical difficulties. The biggest hurdle might be the detection limit: ctDNA may comprise < 1.0% of the total cell-free DNA (cfDNA), making detection of the tumor-specific fraction challenging[1,8,9]. The majority of cfDNA is derived from apoptotic tissue and hematological cells which release their DNA in the circulation[10,11]. Thus, the absolute quantity of cfDNA ('the background') determines our ability to detect ctDNA, and quantification of the tumor-specific variant frequency depends both on the abundance of ctDNA molecules and on the total amount of cfDNA. One of the most important factors impacting the total amount of cfDNA is the time to plasma processing after blood collection, which increases the release of wild-type DNA from lysed hematological cells present in the blood collection tube[12,13]. To avoid this, plasma needs to be separated from the blood sample within hours after the blood withdrawal, but the maximum time frame to do so, remains to be revealed.

Due to logistical and practical reasons, it is often not possible to process and store blood samples immediately after blood withdrawal to ensure optimal ctDNA quality; especially in the context of large multi-center prospective clinical trials, which are essential to definitely establish ctDNA as a clinically relevant new biomarker, there is a need for standardization of preanalytical conditions that allow longer processing time of blood samples. To overcome this problem, specialized 'cell-stabilizing' blood collection tubes have been developed. These tubes should not only minimize contamination by wild-type DNA from lysed hematological cells in the blood tube, but also preserve the quality of ctDNA for reliable downstream analyses.

Until today, a number of studies have tested the different available blood collection tubes to optimally preserve cfDNA/ctDNA[12,14-16]. They all demonstrate a time-dependent increase in cfDNA concentrations in ethylenediaminetetraacetic acid (EDTA) tubes, while cfDNA concentrations remained stable in both BCT and CellSave tubes. Toro et al.[14] included the PAXgene blood DNA tube in their study, but this tube did not improve the results obtained with EDTA tubes. Yet, even though preservation methods have been compared[17], thorough direct comparisons between BCT and CellSave tubes at clinically relevant time frames are missing. We set out to compare the available preservatives for their ability to allow easier implementation of ctDNA-based tests into larger clinical trials where processing of samples within 1 h presents a major logistical challenge. The purpose of this study was to investigate the effect on the quality of cfDNA in general and

Part 2
Liquid Biopsies

Chapter 3
Application of circulating tumor DNA in prospective clinical oncology trials – standardization
of preanalytical conditions

of ctDNA in particular in patients with metastatic cancer under the following preanalytical conditions: (a) different time intervals to plasma isolation (1, 24, and 96 h); and (b) different types of preservative in the blood collection tubes (EDTA, CellSave, and BCT tubes). To this purpose, the amount of cfDNA isolated from plasma was quantitated, its' size determined, and the fraction of ctDNA determined.

Materials and Methods

Patient characteristics and somatic variant status of tumor

Between October 2015 and January 2016 cancer patients within the Erasmus MC Cancer Institute in Rotterdam, the Netherlands were invited to contribute blood samples for this study by their treating physician. Patients were included if they had metastatic disease, were not currently receiving systemic treatment and if a validated dPCR assay (TaqMan® SNP genotyping assays, ThermoFisher Scientific, Waltham, MA USA; see also section on 'Digital PCR TaqMan® SNP genotyping and β-actin fragmentation assay') was available for the known somatic variant in their primary and/or metastatic lesion. Somatic variant status and variant allele frequency (VAF) in tissue had been assessed as part of the standard of care by the molecular diagnostics laboratory of the department of pathology in the Rotterdam region by either Sanger sequencing (patient #10 and #16), SNaPshot analysis (patient #05) or NGS analysis (all other patients). The DNA input for these analyses ranged from 0.48 to 10 ng. The calculation of VAF was performed through NGS analysis by calculating the coverage of the variant nucleotide relative to the total coverage on that position. For tissue samples analyzed by Sanger sequencing the VAF was calculated by determining the ratio between the variant peak and the wild-type peak. All patients provided written informed consent, and the institutional review board approved the protocols (Erasmus MC ID MEC-15-616).

Preanalytical conditions

After obtaining written informed consent, 9x 10 mL of blood samples were collected within a single blood withdrawal (Figure S1). Matched blood samples were collected in sterile 3x 10 mL K$_2$EDTA vacutainer® (BD, Franklin Lakes, NJ, USA), 3x 10 mL Cell-Free DNA BCT® (Streck, Omaha, NE, USA) and 3x 10 mL CellSave Preservative (Janssen Diagnostics, Raritan, NJ, USA) blood collection tubes according to manufacturer instructions. The blood samples from one of each type of tube (EDTA, BCT and CellSave) were processed for plasma isolation at 3 different time points: within 1 h after blood withdrawal, at 24 h and at 96 h after blood withdrawal (Figure S1). Plasma was isolated using 2 sequential centrifugation steps: 1) 1711g for 10 minutes at room temperature; 2) 12,000g for 10 minutes at room temperature. Plasma was stored at -80°C in 1 mL aliquots immediately after centrifugation until further processing.

cfDNA isolation and quantification

For cfDNA isolation plasma samples were thawed at 4°C and 3 mL of plasma per sample was used. cfDNA was isolated using the QIAamp® Circulating Nucleic Acid kit (Qiagen, Venlo, Limburg, The Netherlands) according to manufacturer's instructions. cfDNA was eluted from the Qiagen® Mini column using 50 µL buffer AVE which was applied 3 times to the column to obtain the highest cfDNA concentration possible. cfDNA was stored at -20°C. cfDNA concentrations were quantified using the Quant-iT dsDNA high-sensitivity assay (Invitrogen, Life Technologies, Carlsbad, CA, USA) according to manufacturer's instructions and the Qubit fluorometer (Invitrogen) was used as readout.

Digital PCR TaqMan® SNP genotyping and β-actin fragmentation assay

cfDNA samples were thawed at room temperature. Validated TaqMan® SNP genotyping assays (ThermoFisher Scientific, Waltham, MA, USA) were used for somatic variant and wild-type detection according to manufacturer's instructions (Tabel S1). Accordingly, the limit of detection of this assay is 0.1% [18]. The maximum volume input of 7.8 µL of the final cfDNA eluate was used, unless the amount of cfDNA in this volume exceeded the maximal input of 30 ng cfDNA, then 30 ng cfDNA was used. Depending on the obtained cfDNA concentration after plasma isolation, at least 2.57 ng cfDNA was analyzed, leading to a detection rate of 0.78% at the most.

The TaqMan® β-actin fragmentation assay was based on the assay developed by Norton et al.[12] to detect small (136 bp) and long (420 bp) β-actin fragments. We adapted the assay so that both fragments were measured within a single experiment using the reported primers, but different probes for each fragment (Tabel S2). For the β-actin fragmentation assay a standardized input of 2 ng cfDNA was used to minimize the change of having multiple DNA fragments in one well.

All dPCR reactions were performed with the QuantStudio 3D Digital PCR System (ThermoFisher Scientific) according to the manufacturer's protocol. In short, dPCR reaction mix was prepared containing 8.7 µL QuantStudio 3D Digital PCR Master Mix v2, 0.44 µL Taqman primer/probe mix, up to 7.8 µL of cfDNA and the total volume was completed with PCR grade H$_2$O to a final volume of 17.4 µL. Using the QuantStudio 3D Digital PCR Chip Loader samples were partitioned on a 20,000 wells QuantStudio 3D Digital PCR Chip v2 followed by a PCR reaction on a ProFlex 2x Flat PCR System with the following program: 10 min at 96°C, 40x cycles of 2 min at 60°C, and followed by 30 sec at 98°C, 2 min at 60°C and pause at 10°C. The dPCR data were then acquired with the QuantStudio 3D Digital PCR Instrument and the data was analyzed with the QuantStudio 3D Analysis Suite by one technician (JH) to account for inter-observer variability.

Statistical analysis

The Wilcoxon signed rank test was used to compare the difference between matched 1 h and 24 h samples relative to the difference between matched 1 h and 96 h samples. The Friedman test was used to test the order of the three 1 h samples. To correct for multiple testing, we adjusted the P

Part 2
Liquid Biopsies

Chapter 3
Application of circulating tumor DNA in prospective clinical oncology trials – standardization
of preanalytical conditions

value for significance using the Bonferroni correction. Significance was thus defined as P < 0.008 (0.05/6). Correlations were tested by Spearman's rank correlation coefficient.

cfDNA concentrations determined by the Quant-iT dsDNA high-sensitivity assay were corrected for the plasma input and were converted from ng/mL plasma to copies/mL plasma by taking into consideration that 3.3 pg of human DNA contains 1 copy of a single gene. cfDNA concentrations were then log-transformed.

To correct for differences in plasma input used for cfDNA isolation and for differences in elution volume after cfDNA isolation, we expressed dPCR results as variant/wild-type copy-numbers per mL plasma. To calculate variant/wild-type copy-numbers per mL plasma the following equation as described by Lo et al.[19] was used:

$$C = Q * (V_{DNA})/V_{PCR}) * (1/V_{ext})$$

where C is variant/wild-type copy-numbers per mL plasma; Q is the total number of variant/wild-type copy-numbers determined by dPCR; V_{DNA} is the total volume of cfDNA obtained after cfDNA isolation; V_{PCR} is the volume of cfDNA solution used for the dPCR reaction; and V_{ext} is the volume of plasma used for cfDNA isolation.

To calculate VAF we divided the variant copy-numbers per mL plasma by the sum of variant and wild-type copy-numbers per mL plasma.

All statistical analyses were performed using STATA version 14.1. All figures were plotted using R version 3.2.3.

Results

Somatic variant detection rate in ctDNA of recruited patients

A total of 16 patients were included who all met the set criteria to investigate the effect of different preanalytical conditions on the quality of ctDNA. Somatic variant status of the primary and/or metastatic lesion had been previously assessed, either by targeted next generation sequencing (13 of 16 patients), by SNaPshot analysis (one of 16 patients), or by traditional Sanger sequencing (2 of 16 patients). Table 1 lists the origin of the primary tumor, the site and number of metastases, and the VAF in the tumor tissue. Using the specific TaqMan SNP genotyping assay (Table S1), we were able to detect in 11 (69%) of our 16 patients the known somatic variant in ctDNA isolated within 1 h from EDTA tubes. This corresponds to the detection of 13 (68%) of 19 of the total number of somatic variants tested as some patients had multiple known somatic variants.

Temporal effect of storage in EDTA tubes on cfDNA quality

To investigate the effect of different time intervals from the blood withdrawal to plasma isolation on cfDNA quality, we measured cfDNA concentration isolated from plasma collected in EDTA

Table 1. Tumor characteristics and somatic variant detection

Patient ID (#)	Primary tumor	Site and number of metastases (x)	Interval tumor tissue and plasma analysis (months)	Known somatic variant (nucleotide change)	Variant allele frequency in tissue (%)	Variant allele frequency in plasma EDTA 1 h (%)	Cell-free DNA concentration in plasma EDTA 1 h (copies/mL plasma)
01	Cholangio-carcinoma	Li (3), Lu (2), LN (1)	2	KRAS p.G12D (c.35G>A)	40	0.00	3655
02	Pancreatic cancer	Li (3), Lu (1) LN (6)	9	KRAS p.G12V (c.35G>T)	62	0.00	4055
	CRC			BRAF p.V600E (c.1799T>A)	39	0.97	
				PIK3CA p.H1047R (c.3140A>G)	38	1.86	
03	Breast cancer	LN (>2)	-1*	PIK3CA p.H1047L (c.3140A>T)	26	0.00	2788
04	Melanoma	Li (2), LN (5)	2	BRAF p.V600E (c.1799T>A)	3	1.44	1615
05	CRC	Li (6), LN (2)	6	KRAS p.G13D (c.38G>A)	Unknown	65.46	223,130
06	CRC	Li (3), Lu (4)	18	KRAS p.G12D (c.35G>A)	44	8.61	2215
07	Melanoma	Brain (2), Abd (7)	8	NRAS p.Q61R (c.182A>G)	68	17.22	4245
08	Melanoma	LN (3), Lu (6), Li (>15), Spleen (1), Bone (4), Peritonitis carcinomatosa, pleuritis carcinomatosa	1	BRAF p.V600E (c.1799T>A)	64	37.21	22,442
09	Melanoma	LN (5)	1	BRAF p.V600E (c.1799T>A)	70	6.42	2739
10	CRC	Brain (2), Li (1), Lu (8)	87	KRAS p.G13D (c.38G>A)	50	0.00	6030
11	CRC	Lu (2)	5	KRAS p.G13D (c.38G>A)	57	0.84	4670
12	CRC	Li (>20), LN (1)	3	KRAS p.Q61R (c.182A>G)	46	0.00	16,136
13	NSCLC	Brain (8), Adrenal gland (1)	7	EGFR p.T790M (c.2369C>T)	17	1.18	5358
				EGFR p.L858R (c.2573T>G)	17	2.62	
14	Melanoma	LN (7), Lu (5), adnexa	22	BRAF p.V600E (c.1799T>A)	56	5.37	3539
15	NSCLC	Li (unknown)	1	EGFR p.T790M (c.2369C>T)	65	27.60	14,085
16	Melanoma	Brain (1)	38	BRAF p.V600E (c.1799T>A)	>50	0.00	3012

CRC: colorectal cancer; NSCLC: non-small cell lung cancer; Li: liver; Lu: lung; LN: lymph node; Abd: abdomen
* A new biopsy was taken 2.5 weeks after the blood collection.

Part 2
Liquid Biopsies

Chapter 3
Application of circulating tumor DNA in prospective clinical oncology trials – standardization
of preanalytical conditions

tubes. We observed a significant increase in cfDNA concentrations in samples isolated after 96 h compared to samples isolated within 1 h (P < 0.001; Figure 1 and Figure S2). This increase in cfDNA concentration was significantly positively correlated with an increase in wild-type copy-numbers (rho = 0.85; P < 0.001; Figure 2A). If a somatic variant was detected in the 1 h sample, the somatic variant could also be detected in 24 h and 96 h samples. We also observed a significant positive correlation between variant copy-numbers and cfDNA concentration, although this was less strong (rho = 0.42; P < 0.001; Figure 2B).
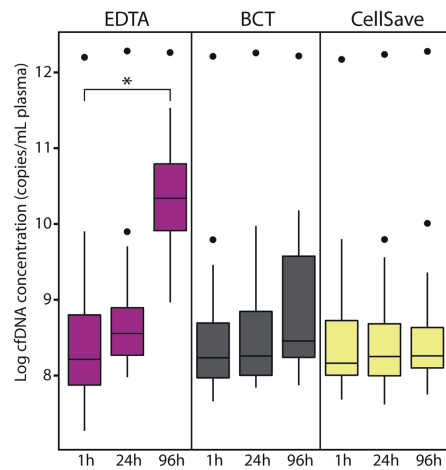
Figure 1. cfDNA concentrations for different preanalytical conditions
Boxes (interquartile ranges [IQR]) and whiskers (1.5x IQR) are shown together with the median (black horizontal line) of the log cfDNA concentrations in copies per mL plasma of 16 patients for the different preanalytical conditions. Outliers are displayed as black dots. The Wilcoxon signed rank test was used to compare the difference between matched 1 h and 24 h samples relative to the difference between matched 1 h and 96 h samples. *P < 0.001.

To investigate whether the increase in cfDNA concentration and wild-type copy-numbers was due to the release of intact DNA from lysed leukocytes, we used the β-actin fragmentation assay (Figure 3A). In all preanalytical conditions, we detected low amounts of large fragments. We observed significantly more large fragments in samples from 96 h than in samples from 1 h (420 bp, P < 0.001; 2000 bp, P < 0.001; Figure 3B). There was also a small but significant increase in fragmented DNA in samples from 96 h compared to samples from 1 h (136 bp, P = 0.002; Figure 3B).
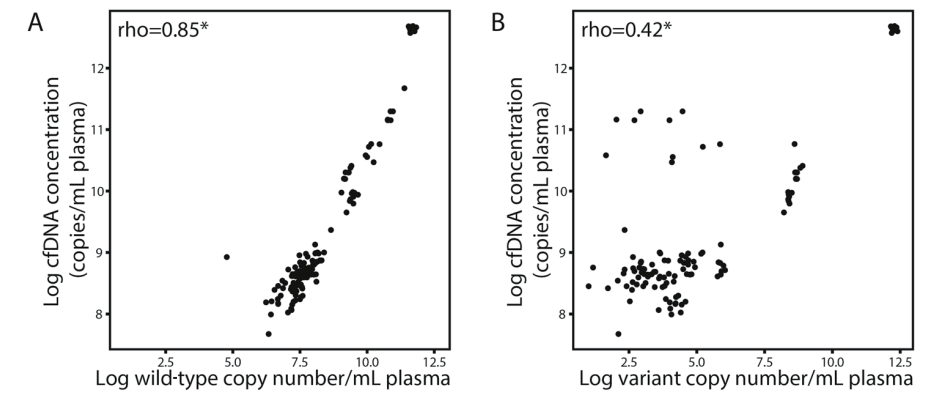
Figure 2. Correlation between wild-type or variant copy-numbers and cfDNA concentration
The log number of wild-type copies (A) or variant copies (B) in copies per mL plasma on the x-axis is plotted against the log cfDNA concentrations in copies per mL plasma on the y-axis. Data points correspond to single sample measurements from each time interval and each type of preservative. Correlations were tested by Spearman's rank correlation coefficient. *P < 0.001. Five patients with undetectable variant copy-numbers in ctDNA are removed from plot B.
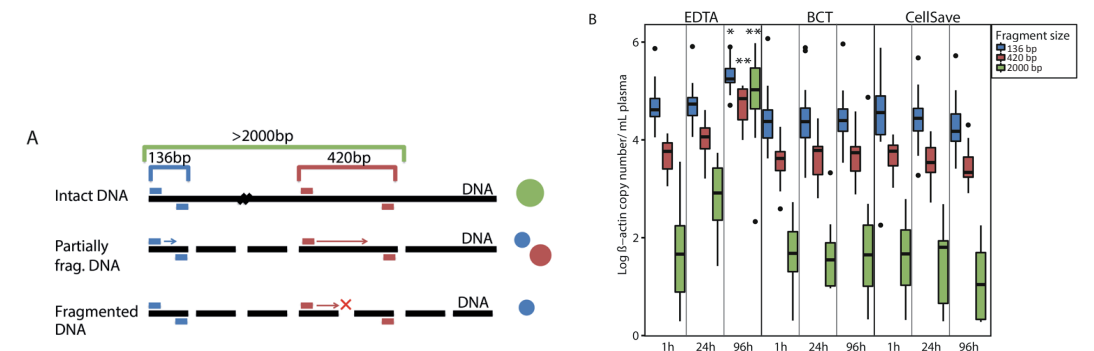
Figure 3. β-actin fragmentation assay for different preanalytical conditions
(A) Principle of β-actin fragmentation assay. dPCR wells containing only 136-bp signal are indicative of fragmented DNA (fragments < 200 bp), whereas the 420-bp primer set will only bind to intact DNA (> 420 bp). When a large intact DNA fragment (> 2000 bp) is present in one of the wells, both primer sets can bind, resulting in a mixed signal. In theory, this can also occur when a small (< 200 bp) and large (> 420 bp) DNA fragment is present together in one well.
(B) Results of β-actin fragmentation assay. Boxes (interquartile ranges [IQR]) and whiskers (1.5x IQR) are shown together with the median (black horizontal line) of the number of β-actin copies for the different preanalytical conditions. Outliers are displayed as black points. The Wilcoxon signed rank test was used to compare the difference between matched 1 h and 24 h samples relative to the difference between matched 1 h and 96 h samples for the different fragment sizes. *P = 0.002; **P < 0.001.

Part 2
Liquid Biopsies

Chapter 3
Application of circulating tumor DNA in prospective clinical oncology trials – standardization
of preanalytical conditions

The interaction between different preservatives and plasma isolation time intervals and cfDNA quality

Next, we studied the effect of different preservatives in blood collection tubes on cfDNA quality. We compared cfDNA concentrations isolated from plasma collected in EDTA, BCT, and CellSave tubes processed within 1 h. Cell-free DNA concentrations were similar in all blood collection tubes (Figure 1 and Figure S2). We also did not observe any differences in the DNA fragment size distribution with the β-actin fragmentation assay for the different tubes at 1 h (Figure 3B).

In order to investigate whether the used preservatives in BCT and CellSave tubes could prevent the time-dependent increase in cfDNA concentration observed in EDTA tubes, we measured cfDNA concentrations in samples isolated after 24 h and 96 h after blood withdrawal. We observed stable cfDNA concentrations in all 24 h and 96 h samples compared to their matched 1 h samples (Figure 1 and Figure S2). Also, we did not observe any differences in the DNA size distribution with the β-actin fragmentation assay for the matched time intervals for both tube types (Figure 3B).

The interaction between different preservatives and plasma isolation time intervals on somatic variant detection in ctDNA

To study the effect of time-dependent increase in cfDNA concentrations and wild-type copy-numbers on somatic variant detection, we analyzed VAF in the different preanalytical conditions compared to their matched 1 h sample. If a somatic variant was detected in the EDTA 1 h sample, the somatic variant could also be detected in all BCT and CellSave samples. There was no correlation between the VAF in tumor tissue and the VAF in plasma (Figure S3). There was a significant decrease in VAFs in samples from EDTA 96 h (P = 0.003; Figure 4), which was not observed for the other preanalytical conditions. Because all tubes were drawn within a single blood withdrawal, we expected, in contrast to VAF, that all tubes within each patient contains similar amounts of variant copy-numbers. Indeed, variant copy-numbers appeared largely similar between tubes and in all tubes compared to their matched 1 h sample (Figure 5 and Figure S4).
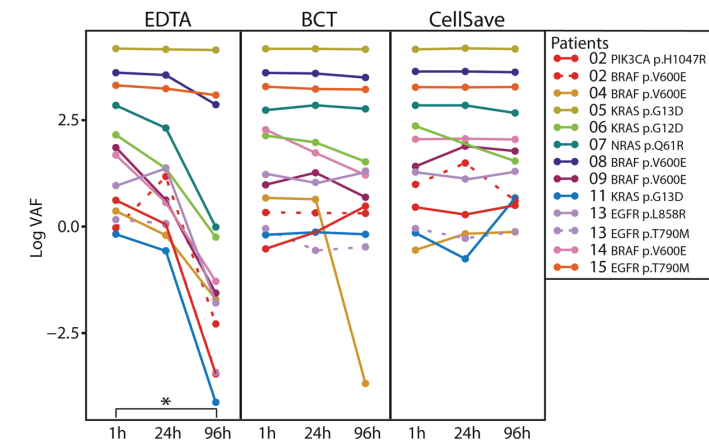
Figure 4. VAF of 11 patients for different preanalytical conditions
Data points correspond to VAF for each individual patient and assay. The Wilcoxon signed rank test was used to compare the difference between matched 1 h and 24 h samples relative to the difference between matched 1 h and 96 h samples. *P = 0.003.
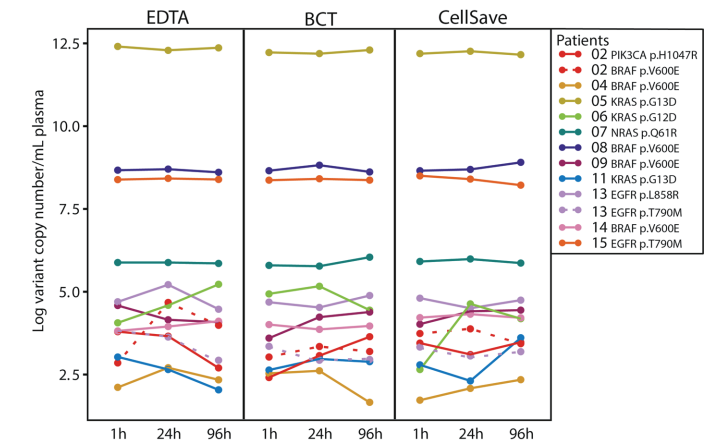
Figure 5. Variant copy-numbers of 11 patients for different preanalytical conditions
Data points correspond to variant copy-numbers for each individual patient and assay. The Wilcoxon signed rank test was used to compare the difference between matched 1 h and 24 h samples relative to the difference between matched 1 h and 96 h samples.

Part 2
Liquid Biopsies

Chapter 3
Application of circulating tumor DNA in prospective clinical oncology trials – standardization
of preanalytical conditions

Discussion and Conclusions

The purpose of this study was to investigate the effects of various preanalytical conditions on the quality of cfDNA in general and of ctDNA in particular. The main aim was to investigate whether BCT and CellSave tubes processed within 96 h after blood withdrawal into plasma were suitable for downstream analyses of ctDNA.

Patients were recruited with a high prior probability to harbor ctDNA in their plasma, that is, patients with metastatic disease without current anti-cancer treatment. In 69% of our patients, we were able to detect the known somatic variant from tissue in ctDNA and this corresponds to the detection of 68% of all tested somatic variants. In two of six missed somatic variants, the somatic variant status in tissue was assessed > 3 years ago. It may be possible that other cancer subclones have emerged, resulting in undetectable somatic variants in ctDNA. Unfortunately, in these cases, more recent information on somatic variant status was not available. Detection of somatic variants in plasma may also be influenced by the site and extent of metastases, which is exemplified by patient #05. This patient had a widespread pattern of metastases with corresponding high levels of cfDNA and high levels of variant copy-numbers in plasma. However, due to our heterogeneous cohort, this relationship could not be tested statistically for the other patients.

The clinical utility and potential importance of our methods is evidenced by our findings in patient #02, who was thought to have metastases from his pancreatic carcinoma (first primary cancer) harboring a *KRAS* mutation. However, we could only detect *BRAF* and *PIK3CA* mutations in his ctDNA, highly suggestive that the metastases were originating from the patients' colorectal cancer (second primary cancer), which can have important implications for his disease management.

The formation of small DNA fragments (180–200 bp lengths) is a biochemical hallmark of apoptosis, whereas during cell lysis or necrosis intact genomic DNA and thus much larger DNA fragments (50–300 kbp) remain[20]. Through an increase in wild-type copy-numbers and mainly intact DNA fragments, we were able to demonstrate that the time-dependent increase in cfDNA concentration in EDTA tubes indeed originates from leukocyte lysis. In addition, we observed low levels of intact DNA fragments in all preanalytical conditions, indicating a background level of leukocyte lysis here. Both Norton et al.[12] and Rothwell et al.[15] observed a similar increase in cfDNA concentrations in samples collected in EDTA tubes. In both BCT and CellSave tubes, cfDNA concentrations, wild-type copy-numbers and β-actin fragment sizes remained stable up to 96 h, indicating that the preservative in these tubes does not adversely affect cfDNA quality. Interestingly, there was also a significant increase in fragmented DNA in samples from EDTA 96 h, which might be attributed to nucleases remaining active.

As we only used dPCR for downstream analysis of ctDNA, we cannot rule out the possibility that the used preservatives in BCT and CellSave tubes could potentially damage the cfDNA and thus affect other downstream analyses. Rothwell et al. assessed the number of single nucleotide variants through whole-genome sequencing of cfDNA isolated from plasma collected in CellSave tubes[15]. They did not observe introduction of DNA errors. Thus, the preservative used in CellSave tubes does not seem to influence cfDNA downstream analysis using NGS.

Despite the contamination with intact cfDNA, we were still able to detect all somatic variants in ctDNA from EDTA 96 h samples, in those samples where we were able to detect a somatic variant in the EDTA 1 h samples. These data suggest that stored samples which have not been processed optimally for ctDNA analysis can still be used to determine the presence of somatic variants in ctDNA. As a consequence of increased cfDNA concentrations and correlated wild-type copy-numbers, we did observe a significant decrease in VAF in the EDTA 96 h samples. With respect to ctDNA applications for treatment response evaluation, this could result in serious misinterpretations of VAFs. However, variant copy-numbers remained stable in all tubes and might thus be a more accurate outcome measure to evaluate treatment response in patients with cancer. Further investigation is needed to determine the inter-assay variability regarding the range of variant copy-numbers and VAFs we observed among the different tubes.

The results in this study indicate that EDTA tubes processed at 96 h after blood withdrawal are not suitable for blood collection for subsequent cfDNA/ctDNA analysis as the time-dependent increase in cfDNA concentration, resulting from leukocyte lysis, significantly affects VAF. In patient samples with low variant copy-numbers, this increase in cfDNA concentration may cause variant copies to fall below the limit of detection of the dPCR assay and thus may lead to false-negative results. Both BCT and CellSave tubes preserve cfDNA/ctDNA quality equally well up to 96 h and the used preservatives did not affect downstream cfDNA/ctDNA analyses by dPCR. Variant copy-numbers and VAFs also remained stable in these tubes.

Therefore, we recommend for all future clinical studies, in which flexibility regarding the processing of blood samples is needed, to isolate plasma from blood collected in either BCT or CellSave tubes within 96 h. This will make large multi-center trials using a central processing facility feasible, and will lead to optimal quality of ctDNA for research and diagnostics.

Acknowledgements

Part 2
Liquid Biopsies

Chapter 3
Application of circulating tumor DNA in prospective clinical oncology trials – standardization
of preanalytical conditions

References

1. Diehl F, Schmidt K, Choti MA, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med 2008;14(9):985-90. (In eng). DOI: nm.1789 [pii] 10.1038/nm.1789.

2. Shinozaki M, O'Day SJ, Kitago M, et al. Utility of circulating B-RAF DNA mutation in serum for monitoring melanoma patients receiving biochemotherapy. Clin Cancer Res 2007;13(7):2068-74. (In eng). DOI: 13/7/2068 [pii] 10.1158/1078-0432.CCR-06-2120.

3. Forshew T, Murtaza M, Parkinson C, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med 2012;4(136):136ra68. (In eng). DOI: 4/136/136ra68 [pii] 10.1126/scitranslmed.3003726.

4. Dawson SJ, Tsui DW, Murtaza M, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med 2013;368(13):1199-209. (In eng). DOI: 10.1056/NEJMoa1213261.

5. Murtaza M, Dawson SJ, Tsui DW, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature 2013;497(7447):108-12. (In eng). DOI: nature12065 [pii] 10.1038/nature12065.

6. Bidard FC, Madic J, Mariani P, et al. Detection rate and prognostic value of circulating tumor cells and circulating tumor DNA in metastatic uveal melanoma. Int J Cancer 2014;134(5):1207-13. (In eng). DOI: 10.1002/ijc.28436.

7. Diaz LA, Jr., Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. J Clin Oncol 2014;32(6):579-86. (In eng). DOI: JCO.2012.45.2011 [pii] 10.1200/JCO.2012.45.2011.

8. Diehl F, Li M, Dressman D, et al. Detection and quantification of mutations in the plasma of patients with colorectal tumors. P Natl Acad Sci USA 2005;102(45):16368-16373. (In English). DOI: DOI 10.1073/pnas.0507904102.

9. Holdhoff M, Schmidt K, Donehower R, Diaz LA, Jr. Analysis of circulating tumor DNA to confirm somatic KRAS mutations. J Natl Cancer Inst 2009;101(18):1284-5. (In eng). DOI: djp240 [pii] 10.1093/jnci/djp240.

10. Jahr S, Hentze H, Englisch S, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. Cancer Res 2001;61(4):1659-65. (In eng) (http://www.ncbi.nlm.nih.gov/pubmed/11245480).

11. Elshimali YI, Khaddour H, Sarkissyan M, Wu Y, Vadgama JV. The clinical utilization of circulating cell free DNA (CCFDNA) in blood of cancer patients. Int J Mol Sci 2013;14(9):18925-58. (In eng). DOI: ijms140918925 [pii] 10.3390/ijms140918925.

12. Norton SE, Lechner JM, Williams T, Fernando MR. A stabilizing reagent prevents cell-free DNA contamination by cellular DNA in plasma during blood sample storage and shipping as determined by digital PCR. Clin Biochem 2013;46(15):1561-5. (In eng). DOI: S0009-9120(13)00279-8 [pii] 10.1016/j.clinbiochem.2013.06.002.

13. Xue X, Teare MD, Holen I, Zhu YM, Woll PJ. Optimizing the yield and utility of circulating cell-free DNA from plasma and serum. Clin Chim Acta 2009;404(2):100-4. (In eng). DOI: S0009-8981(09)00109-0 [pii] 10.1016/j.cca.2009.02.018.

14. Toro PV, Erlanger B, Beaver JA, et al. Comparison of cell stabilizing blood collection tubes for circulating plasma tumor DNA. Clin Biochem 2015;48(15):993-8. (In eng). DOI: S0009-9120(15)00377-X [pii] 10.1016/j.clinbiochem.2015.07.097.

15. Rothwell DG, Smith N, Morris D, et al. Genetic profiling of tumors using both circulating free DNA and circulating tumor cells isolated from the same preserved whole blood sample. Mol Oncol 2015 (In Eng). DOI: S1574-7891(15)00205-7 [pii] 10.1016/j.molonc.2015.11.006.

16. Sherwood JL, Corcoran C, Brown H, Sharpe AD, Musilova M, Kohlmann A. Optimised Preanalytical Methods Improve KRAS Mutation Detection in Circulating Tumor DNA (ctDNA) from Patients with Non-Small Cell Lung Cancer (NSCLC). PLoS One 2016;11(2):e0150197. (In eng). DOI: 10.1371/journal.pone.0150197 PONE-D-15-55686 [pii].

17. Kang Q, Henry NL, Paoletti C, et al. Comparative analysis of circulating tumor DNA stability In KEDTA, Streck, and CellSave blood collection tubes. Clin Biochem 2016 (In Eng). DOI: S0009-9120(16)30040-6 [pii] 10.1016/j.clinbiochem.2016.03.012.

18. QuantStudio 3D rare mutation analysis solution. Applied Biosystems ThermoFisher Scientific. 16-06-2016 (https://www.thermofisher.com/content/dam/LifeTech/Documents/PDFs/PG1502-PJ8965-CO126182-Revise-QuantStudio-3D-Rare-Mutation-Product-Overview-Branding-Changes-New-SKUs-New-Assays-Global-FLR-2016-06-16.pdf).

19. Lo YM, Tein MS, Lau TK, et al. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. American Journal of Human Genetics 1998;62(4):768-75.

20. Bortner CD, Oldenburg NB, Cidlowski JA. The role of DNA fragmentation in apoptosis. Trends Cell Biol 1995;5(1):21-6. (In eng). DOI: S0962892400889321 [pii].

Part 2
Liquid Biopsies

Chapter 3
Application of circulating tumor DNA in prospective clinical oncology trials – standardization
of preanalytical conditions

Supplementary information
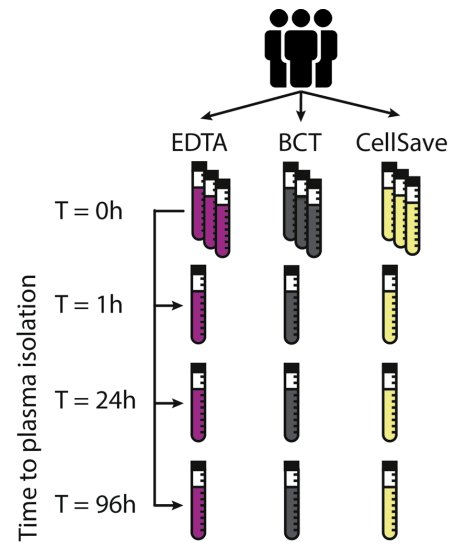
Figure S1. Overview of study design
At time point 0 h 9x 10 mL blood tubes (3x EDTA; 3x BCT; 3x CellSave) were collected within a single blood withdrawal from each recruited patient (n = 16). From each type of tube plasma was isolated within 1 h, after 24 h and after 96 h. Plasma was directly stored at -80°C after processing.
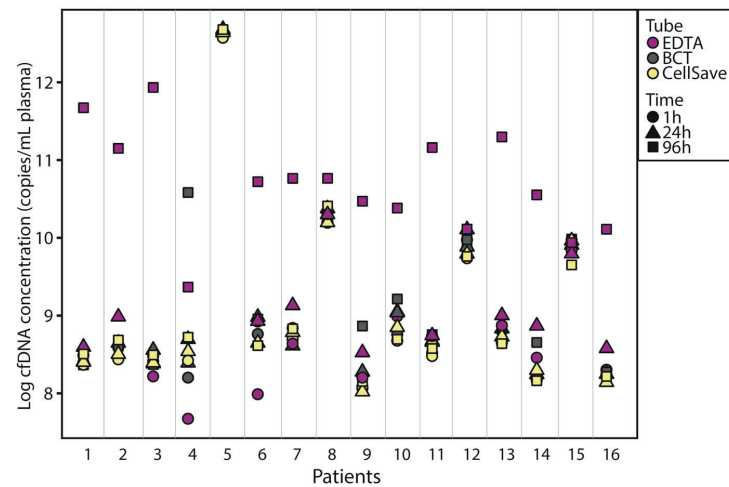
Figure S2. Cell-free DNA concentrations for each individual patient for different preanalytical conditions
Data points correspond to the log cfDNA concentrations in copies/mL plasma for each individual patient (n = 16) and for the different preanalytical conditions. Different tube types are indicated by different colors. Different time points are indicated by different symbols.
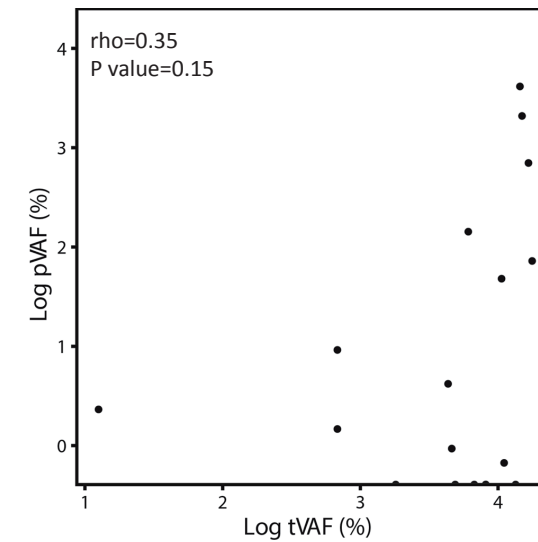
Figure S3. Correlation between variant allele frequency in tumor tissue and in ctDNA in plasma
The log percentage of tumor tissue variant allele frequency (tVAF) on the x-axis is plotted against the log precentage of ctDNA in plasma variant allele frequency (pVAF) on the y-axis. Data points correspond to single somatic variants. Correlation was tested by Spearman's rank correlation coefficient.
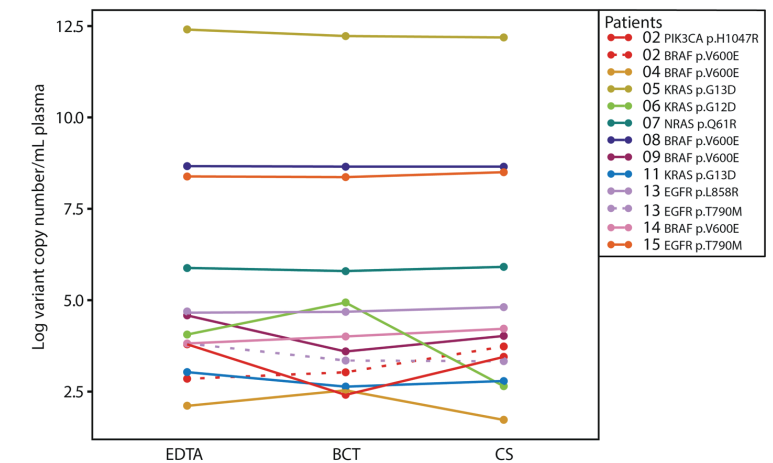
Figure S4. Variant copy-numbers for 1 h samples
The log variant copy-numbers from 11 patients for the 1 h samples. Data points correspond to variant copy-numbers for each individual patient and assay.

Part 2
Liquid Biopsies

Chapter 3
Application of circulating tumor DNA in prospective clinical oncology trials – standardization
of preanalytical conditions

Table S1. Used SNP genotyping assays

Assay ID	Assay name	Gene	Cosmic ID	Amino acid change	Nucleotide change
AH6R5PH	BRAF_476	*BRAF*	476	p.V600E	c.1799T>A
AHRSROS	EGFR_6240	*EGFR*	6240	p.T790M	c.2369C>T
AHRSRSV	EGFR_6224	*EGFR*	6224	p.L858R	c.2573T>G
AH6R5PI	KRAS_521	*KRAS*	521	p.G12D	c.35G>A
AHX1IHY	KRAS_520	*KRAS*	520	p.G12V	c.35G>T
AHD2BW0	KRAS_532	*KRAS*	532	p.G13D	c.38G>A
AHQJTKH	KRAS_552	*KRAS*	552	p.Q61R	c.182A>G
AHS1P6Q	NRAS_584	*NRAS*	584	p.Q61R	c.182A>G
AHLJ0TP	PIK3CA_776	*PIK3CA*	776	p.H1047L	c.3140 A>T
AHPAVCD	PIK3CA_775	*PIK3CA*	775	p.H1047R	c.3140 A>G

Table S2. Primer and probe designs for digital PCR

Gene	Forward primer	Reverse primer	Probe
β-actin 136 bp	5'-GCG CCG TTC CGA AAG TT-3'	5'- CGG CGG ATC GGC AAA -3'	6FAM-ACC GCC GAG ACC GCG TC-MGBNFQ
β-actin 420 bp	5'-CCG CTA CCT CTT CTG GTG-3'	5'-GAT GCA CCA TGT CAC ACT G-3'	VIC-CCT CCC TCC TTC CTG GCC TC-BHQ

Chapter 4

High-throughput isolation of circulating tumor DNA:
a comparison of automated platforms

Chapter 4

Lisanne F. van Dessel*, Silvia R. Vitale*, Jean C. Helmijr, Saskia M. Wilting, Michelle van der Vlugt-Daane,
Esther Oomen-de Hoop, Stefan Sleijfer, John W.M. Martens, Maurice P.H.M. Jansen, Martijn P. Lolkema
* These authors contributed equally

Chapter 4
High-throughput isolation of circulating tumor DNA:
a comparison of automated platforms

Abstract

The emerging interest in circulating tumor DNA (ctDNA) analyses for clinical trials has necessitated the development of a high-throughput method for fast, reproducible, and efficient isolation of ctDNA. Currently, the majority of ctDNA studies use the manual QIAamp (QA) platform to isolate DNA from blood. The purpose of this study was to compare two competing automated DNA isolation platforms (Maxwell [MX] and QIAsymphony [QS]) to the current 'gold standard' QA to facilitate high-throughput processing of samples in prospective trials. We obtained blood samples from healthy blood donors and metastatic cancer patients for plasma isolation. Total cell-free DNA (cfDNA) quantity was assessed by TERT quantitative PCR. Recovery efficiency was investigated by quantitative PCR analysis of spiked-in synthetic plant DNA. In addition, a β-actin fragmentation assay was performed to determine the amount of contamination by genomic DNA from lysed leukocytes. ctDNA quality was assessed by digital PCR for somatic variant detection. cfDNA quantity and recovery efficiency were lowest using the MX platform, whereas QA and QS showed a comparable performance. All platforms preferentially isolated small (136 bp) DNA fragments over large (420 and 2000 bp) DNA fragments. Detection of the number of variant and wild-type molecules was most comparable between QA and QS. However, there was no significant difference in variant allele frequency comparing QS and MX to QA. In summary, we show that the QS platform has comparable performance to QA, the 'gold standard', and outperformed the MX platform depending on the readout used. We conclude that the QS can replace the more laborious QA platform, especially when high-throughput cfDNA isolation is needed.

Introduction

With the discovery of cell-free DNA (cfDNA), first described in 1948 by Mandel and Metais[1], and subsequently circulating tumor DNA (ctDNA)[2], a novel biomarker in cancer research became available. Since then, many studies have shown its great potential for detecting minimal residual disease and evaluating treatment response[3-11]. However, to enable high-throughput ctDNA analyses a fast, accurate, and efficient cfDNA isolation method is highly needed.

Currently, the majority of ctDNA studies uses Qiagen's QIAamp (QA) platform for cfDNA isolation[12-14]. However, this manual platform is laborious and can only process up to 24 samples at a time rendering this method less suitable for large-scale studies. Automation of cfDNA isolation represents a potential solution provided that it is able to: (a) reduce hands-on time; (b) simultaneously process large numbers of samples; (c) accurately and reproducibly isolate cfDNA with a reasonable recovery; and (d) preserve the quality of ctDNA for downstream analyses.

Cell-free DNA is naturally fragmented (140–175 bp) and only present at low concentrations in the blood circulation (usually around 10 ng per mL plasma[15]). In addition, the fraction of ctDNA relative to cfDNA can vary from extremely low (< 0.01%) to very high (60%), as it is dependent on tumor type and stage[6,16]. Together these features make it imperative to carefully determine the efficacy of DNA isolation instead of merely investigating isolation yields. Furthermore, isolation of cfDNA and ctDNA therein is highly susceptible to genomic DNA contamination from lysed leukocytes[17,18], resulting in a potential underestimation of the ctDNA fraction and decreasing the detection sensitivity. As potential differences in cfDNA recovery efficiency between isolation methods might affect downstream analysis results of ctDNA by decreasing its detection sensitivity, standardized comparison of the different methods for cfDNA isolation is important and highly needed.

The purpose of this study was to compare two automated cfDNA isolation platforms, Maxwell (MX) and QIAsymphony (QS), to the current 'gold standard' QA isolation kit to determine whether these automated platforms can facilitate high-throughput processing of samples in prospective trials. Our analyses focused on both qualitative and quantitative parameters, including cfDNA yield, recovery efficiency, cfDNA fragmentation patterns, and ctDNA fraction retrieved, using optimally processed plasma samples of healthy blood donors (HBDs) and patients with metastatic cancer.

Part 2
Liquid Biopsies

Chapter 4
High-throughput isolation of circulating tumor DNA:
a comparison of automated platforms

Materials and Methods

Subjects

Blood samples were obtained from a total of 10 HBDs and 10 metastatic cancer patients. HBDs were either laboratory volunteers or blood donors of the Sanquin Blood Bank South-West Region, The Netherlands. Patients were enrolled in this study between September 2016 and September 2017 within the Erasmus MC Cancer Institute in Rotterdam, the Netherlands. Eligibility criteria for patients have been described previously[19]. All patients provided written informed consent, and the institutional review board approved the protocols (Erasmus MC ID MEC-15-616). The study methodologies conformed to the standards set by the Declaration of Helsinki. Patient and tumor characteristics are summarized in Table 2.

Blood collection

HBDs donated 20 mL of blood, collected either in 2x 10 mL CellSave preservative tubes (Janssen Diagnostics, Raritan, New Jersey, USA) or in 1x 10 mL EDTA tube (Becton, Dickinson and Company, Franklin Lakes, New Jersey, USA) and 1x 10 mL CellSave preservative tube. Patients donated 3x 10 mL of blood collected in CellSave preservative tubes. Blood samples were stored at room temperature until further processing. After blood withdrawal, samples in EDTA tubes were processed within 24 h, whereas samples in CellSave tubes were processed within 96 h for plasma isolation as previously described[19].

Cell-free DNA isolation

Cell-free DNA was isolated from 2 mL of plasma and eluted in 60 µL of the provided elution buffer. Three isolation platforms were evaluated (Table 1):
- QIAamp® (QA) Circulating Nucleic Acid kit (Qiagen, Hilden, North Rhine-Westphalia, Germany);
- QIAsymphony® (QS) SP Circulating DNA Kit (Qiagen);
- Maxwell® (MX) RSC LV ccfDNA Plasma Custom Kit (Promega, Madison, Wisconsin, USA).

All cfDNA isolations were performed according to the manufacturer's protocol, with some minor modifications. In more detail, cfDNA was isolated with QA as previously described[19]. The QS isolation was adapted by adding 1 µg of carrier RNA (cRNA, Qiagen) to the plasma sample preceding isolation. Using the MX platform, a third plasma centrifugation step at 2,000*g* for 10 minutes at room temperature was performed after thawing to eliminate residual leukocytes, as recommended by the manufacturer. The custom Maxwell® RSC ccfDNA Plasma Kit for large plasma volume protocol was used. In brief, 2 mL of plasma was added to an equal amount of binding buffer and 140 µL of magnetic beads. This mixture was incubated under rotation for 45 minutes at room temperature and subsequently centrifuged at 2,000*g* for 1 minute at room temperature. The pelleted mix of beads and cfDNA was then transferred to the cartridge and run on the MX instrument (Promega) according to the manufacturer's protocol.

Table 1. Specifications of cell-free DNA isolation platforms

Platform	Manufacturer	Protocol	cfDNA isolation kit	Plasma input (mL)	Number of samples per run	Handling-time per run (min)	Technique	Cost (€) per sample
QIAamp (QA)	Qiagen	Manual	QIAamp® Circulating Nucleic Acid Kit	1.0 – 5.0	24	180-240	Vacuum-column based	20
QIAsymphony (QS)	Qiagen	Automatic	QIAsymphony® Circulating DNA Kit	2.0 – 8.0*	96	30	Magnetic-bead based	24
Maxwell (MX)	Promega	Automatic	Maxwell® RSC LV ccfDNA Plasma Custom Kit	2.0 – 4.0*	16**	30	Magnetic-bead based	20

Upon request the manufacturer is able to adjust system settings and protocols for lower/higher plasma input volumes.
*** The Maxwell RSC 48 Instrument can process up to 48 samples per run.*

Testing of cRNA addition to the automated platforms

Plasma samples from several HBDs were pooled and divided into aliquots of 2 mL each. To each aliquot we added different amounts of cRNA, ranging from 0.25 µg up to 4 µg. As a control, plasma samples without cRNA were included. To allow determination of the recovery efficiency, synthetic plant DNA was added to plasma samples (see below).

Cell-free DNA quantification

All cfDNA samples were quantified by both Qubit™ fluorometric quantitation (Invitrogen, Life Technologies, Carlsbad, California, USA) and human TaqMan® copy-number reference assay TERT (Applied Biosystems, Life Technologies, Foster City, California, USA) by quantitative PCR (qPCR). The Qubit™ measurement was performed on 2 µL of each cfDNA sample using the Quant-iT dsDNA high-sensitivity assay (Invitrogen), according to the manufacturer's protocol. TERT qPCR reactions contained 5 µL cfDNA, 3.13 µL SensiFAST™ SYBR® Lo-Rox mix (Bioline, London, United Kingdom), 0.62 µL TERT assay in a total reaction volume of 12.5 µL. The qPCR reaction was performed on an Mx3000P Real-Time PCR System (Agilent, Santa Clara, California, USA) with a pre-incubation at 95°C for 10 minutes, followed by 45 cycles of 95°C for 10 seconds and 60°C for 22 seconds. cfDNA was quantified using a standard curve of human genomic DNA.

Synthetic plant DNA and plant DNA qPCR assay

The synthetic plant DNA assay developed by Kang et al.[20] was used as an exogenous control to calculate the recovery efficiency of each cfDNA isolation method. In short, 250 ng of a 150 bp gBlocks® gene fragment (Integrated DNA Technologies Incorporation [IDT], Coralville, Iowa, USA) was resuspended in LoTE buffer to a final concentration of 1.64×10^{0} ng/µL. The stock sample was serially diluted to a final concentration of 1.64×10^{-6} ng/µL of which 5 µL was spiked into plasma

Part 2
Liquid Biopsies

Chapter 4
High-throughput isolation of circulating tumor DNA:
a comparison of automated platforms

preceding cfDNA isolation. Plant DNA qPCR reactions were essentially performed as described above, using 900 nM of both forward and reverse primer and 250 nM of a FAM-labeled probe (Table S1). Recovery efficiency was determined using a standard curve including the amount of spiked-in plant DNA. Samples with a recovery efficiency < 5% or > 100% were excluded from further analysis as this strongly suggested an operator failure. This was further supported by the fact that recovery efficiency was not strongly correlated ($\rho = 0.45$) with cfDNA concentration (Figure S1).

Digital PCR TaqMan® SNP genotyping and β-actin fragmentation assay

The presence of somatic tumor-specific variants and wild-type DNA molecules was determined using standard and custom-made TaqMan® SNP genotyping assays (ThermoFisher Scientific, Waltham, Massachusetts, USA) according to the manufacturer's instructions (Table S2 and S3). The TaqMan® β-actin assay was used to investigate the fragment size distribution as an indication of leukocyte DNA contamination of the cfDNA, as previously reported[19]. In short, a standard amount of 2 ng of cfDNA was used to detect one small (136 bp) and two long (420 bp and 2000 bp) β-actin fragments within a single reaction. The used primers and probes are indicated in Table S1. The dPCR reaction was performed as previously described[19]. In short, a maximum volume input of 7.8 µL of the final cfDNA eluate was added to the dPCR reaction; the dPCR run was performed on the chip-based QuantStudio 3D Digital PCR System (ThermoFisher Scientific) according to the manufacturer's protocol. SNP genotyping assays were run at 56°C; the β-actin assay was run at 60°C. A negative (H_2O) and positive (cell genomic DNA with known variant) control was added to every experiment.

Sample size

To test whether QS and MX were comparable to QA we assumed a Cohen's effect size of 0.8, to be able to detect relevant differences. With a two-sided type I error probability (α) of 0.025 and a type II error probability (β) of 0.2, a power calculation determined that 18 subjects were needed for paired comparisons. Based on the foregoing 20 subjects were included (10 HBDs and 10 patients).

Calculations and statistical analysis

All assay results were corrected for variations in plasma input and eluate volume, as previously described[19], and expressed as either ng/mL plasma or as mutant/wild-type /β-actin copy-number per mL of plasma. The variant allele frequency (VAF) was calculated as follows: VAF = total variant copy-number / (total variant copy-number + total wild-type copy-number)

The statistical analyses and figure plotting were performed in R version 3.2.3. The Friedman test was used to test the difference between matched QA, MX and QS samples. Significant differences were post hoc analyzed using the Wilcoxon signed rank test. To correct for multiple testing, we adjusted the P value for significance by subsequently applying the Bonferroni correction. The Wilcoxon signed rank test was used to test the difference between matched EDTA and CellSave samples. Correlations were determined by Spearman's rank correlation coefficient.

Results

Optimization of cfDNA isolation using automated isolation platforms

In a small pilot study, we had previously observed a beneficial effect of cRNA addition to HBD plasma during isolation with the QS protocol on the cfDNA yield as determined by Qubit (Figure S2). Therefore, cRNA addition was implemented in our standard QS protocol. However, it has been reported that cRNA might interfere with Qubit-based DNA quantification and might not be a reliable read-out[21]. Therefore, we tested whether cfDNA isolation on the automated platforms (QS/MX) was beneficially or adversely affected by the addition of cRNA using multiple read-outs. We added varying amounts of cRNA to the plasma samples and measured the resulting cfDNA concentration by Qubit and TERT qPCR for both automated platforms. Using Qubit as read-out, the addition of cRNA increased the total amount of cfDNA extracted on both platforms (MX P < 0.001; QS P < 0.001; Figure 1A). However, using TERT qPCR as read-out, this increase could not be reproduced (Figure 1B). Next, we assessed the impact of cRNA on the recovery of spiked-in synthetic plant DNA. Addition of cRNA affected the recovery efficiency of plant DNA (MX P = 0.02; QS P = 0.04; Figure 1C). Independent of cRNA input, recovery of plant DNA was ~ 30% higher with QS (58.37 ± 9.52) than with MX (28.22 ± 6.67; P < 0.001). To assess whether the addition of cRNA biased the isolation of particular cfDNA fragment sizes, we performed the β-actin fragmentation assay (Figure 1D). For both methods, increasing amounts of cRNA reduced the number of small fragments (136 bp; MX P = 0.001; QS P < 0.001), while no effect on larger fragments was observed. For all post hoc analyses, paired testing of samples with and without addition of cRNA (0 µg) did not reveal any significant differences.

Part 2
Liquid Biopsies

Chapter 4
High-throughput isolation of circulating tumor DNA:
a comparison of automated platforms
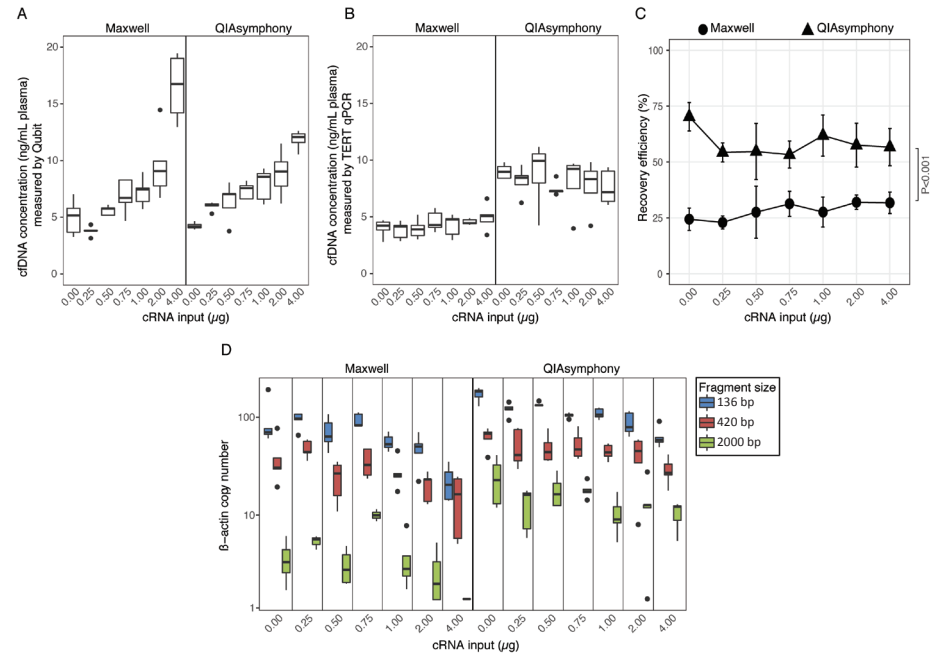
Figure 1. Effect of increasing cRNA input (0–4 μg) on cfDNA quantity and quality using the Maxwell and QIAsymphony platforms

The effect on cfDNA concentration (ng/mL plasma) was measured by Qubit (A) and TERT qPCR (B). The recovery efficiency of each platform was analyzed by qPCR using spiked-in synthetic plant DNA (C). Differences in cfDNA fragment size, expressed as number of β-actin fragments for each fragment size (136, 420 and 2000 bp), were analyzed by dPCR (D). Boxes (interquartile ranges [IQR]) and whiskers (1.5x IQR) are shown together with the median (black horizontal line). Outliers are indicated as single black points. Symbols ● and ▲ are mean values shown with whiskers (standard deviation). The Friedman test was used to test the group difference between Maxwell and QIAsymphony samples. Significant differences were post hoc analyzed using the Wilcoxon signed-rank test. n = 5.

Compatibility of CellSave preservative tubes with different isolation platforms

Previously, we have demonstrated the good performance of CellSave preservative tubes for ctDNA analysis[19]. However, the manufacturers of both automated platforms recommend to use plasma isolated from blood collected in EDTA tubes. To allow for a fair comparison with our CellSave QA results, we therefore first determined whether the automated platforms (QS/MX) were compatible with CellSave tubes by assessing the cfDNA quantity and quality.

Figure 2A shows cfDNA concentrations as measured by TERT qPCR analysis. For the MX platform, the median cfDNA concentration was 5.59 ng/mL plasma from EDTA tubes and was 2.19 ng/mL plasma from CellSave tubes (IQR: 5.06–6.21 and 2.07–3.37 ng/mL plasma, respectively; P = 0.008). For the QS platform, the median cfDNA concentration was 17.17 ng/mL plasma from EDTA tubes and 11.13 ng/mL plasma from CellSave tubes (IQR: 7.81–22.12 and 9.02–14.14 ng/

mL plasma, respectively). Although this was comparable, EDTA samples displayed a larger range in yielded cfDNA concentration. The potential effect of CellSave tubes on the recovery of synthetic plant DNA was determined as well. Comparable recovery efficiencies were observed in plasma collected in EDTA and CellSave tubes for both platforms (39.92% vs. 44.27% in MX and 67.92% vs. 66.19% in QS; Figure 2B). Finally, we used the β-actin fragmentation assay to evaluate cfDNA fragmentation patterns as a read out for general sample quality (Figure 2C). EDTA tubes yielded a higher number of large cfDNA fragments (2000 bp) irrespective of the platform used (median number of β-actin fragments and IQR MX: 33.08 (14.28–44.59); QS: 32.46 (25.53–55.44)) than in CellSave tubes (median number of β-actin fragments and IQR MX: 5.15 (2.42–9.17); QS: 13.80 (7.01–18.18); P = 0.008). The number of small DNA fragments (136 bp) did not differ between EDTA and CellSave tubes for MX, but was slightly higher for EDTA tubes on the QS platform (median number of β-actin fragments and IQR EDTA: 142.71 (110.28–198.18); CellSave: 89.71 (80.22–102.64); P = 0.04). Based on these results, we deemed CellSave tubes are compatible with both automated platforms and used them for all further experiments.
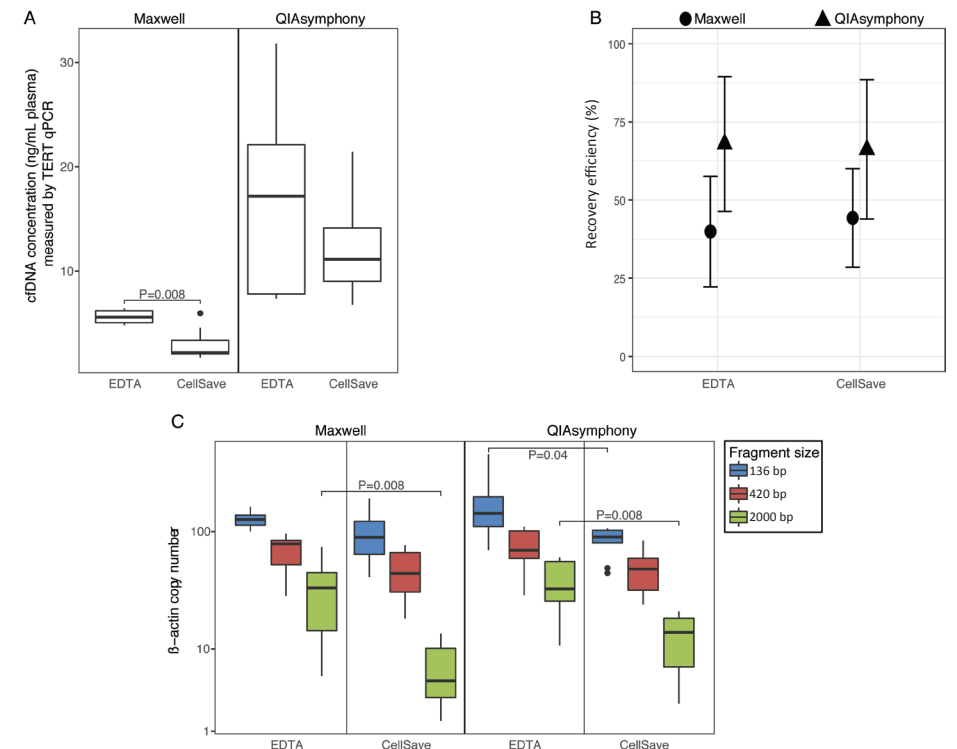
Figure 2. Compatibility of EDTA and CellSave blood collection tubes with the Maxwell and QIAsymphony platforms

The effects on cfDNA concentration (ng/mL plasma) measured by TERT qPCR (A), recovery efficiency measured by plant DNA qPCR (B), and β-actin fragmentation assay analyzed with dPCR are shown (C). Boxes (interquartile ranges [IQR]) and whiskers (1.5x IQR) are shown together with the median (black horizontal

Part 2
Liquid Biopsies

Chapter 4
High-throughput isolation of circulating tumor DNA:
a comparison of automated platforms

>>

line). Outliers are indicated as single black points. Symbols ● and ▲ are mean values shown with whiskers (standard deviation). The Wilcoxon signed-rank test was used to test the difference between blood collection tubes for each platform. n = 9.

Comparison of the performance of automated platforms on downstream cfDNA and ctDNA analyses

Next, we compared the quantity and quality of the obtained cfDNA using the current 'gold standard' manual QA platform to the automated QS and MX platforms using samples from 10 HBDs and 10 metastatic cancer patients. In HBDs, cfDNA concentrations measured by TERT qPCR analysis were comparable for all three isolation platforms (Figure 3A). In patients, the MX retrieved significantly less cfDNA compared to both QA (P = 0.002) and QS (P = 0.002; median cfDNA concentration and IQR QA: 15.84 (12.64–65.11); MX: 6.00 (3.80–20.43); QS: 14.50 (11.99–57.65) ng/mL plasma; Figure 3A). To determine the recovery efficiency of the three different platforms, 5 µL of synthetic plant DNA was added to each plasma sample preceding cfDNA isolation. The average recovery efficiency using QA (51.95 ± 12.02%) was similar to QS (43.45 ± 8.21%). However, MX performed worse (18.61 ± 5.81%; P < 0.001; Figure 3B). In HBDs, we did not observe cfDNA fragment size differences between either of the evaluated platforms (Figure 3C). In patients, MX isolated fewer small β-actin fragments (136 bp) than QA (median number of β-actin fragments and IQR for MX: 57.45 (53.17–66.72); and for QA: 83.18 (70.36–101.63); P < 0.01) and fewer large fragments (2000 bp) than QS (median number of β-actin fragments and IQR for MX: 2.08 (0.00–5.21); and for QS: 10.06 (6.70–13.72); P = 0.002).
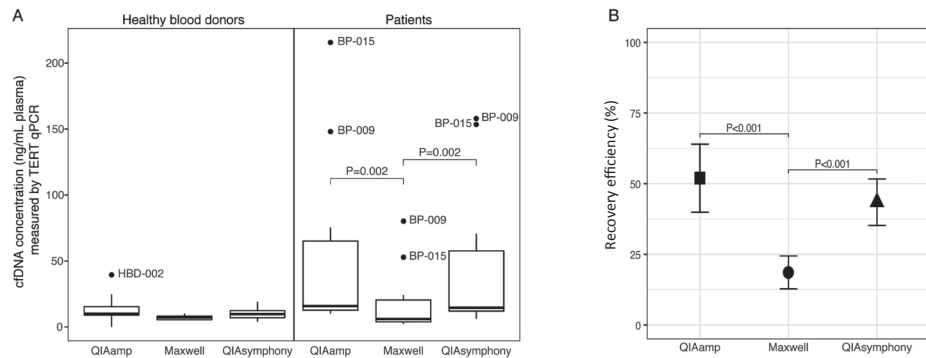
Figure 3. Effect of the different isolation platforms (QIAamp, Maxwell, and QIAsymphony) on downstream cfDNA analysis
Cell-free DNA was isolated from 2 mL matched plasma samples of HBDs (n = 10) and patients with metastatic cancer (n = 10) and analyzed by TERT qPCR assay for cfDNA concentration (ng/mL plasma) (A), plant DNA qPCR assay to determine recovery efficiency (B), and dPCR β-actin fragmentation assay to evaluate cfDNA fragment sizes (C).
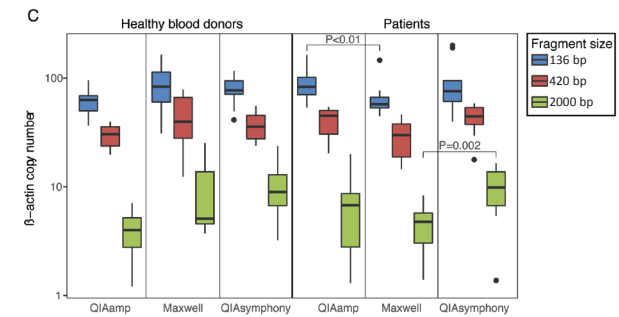
Boxes (interquartile ranges [IQR]) and whiskers (1.5x IQR) are shown together with the median (black horizontal line). Outliers are indicated as single black points. Symbols ■, ●, and ▲ are mean values shown with whiskers (standard deviation). The Friedman test was used to test the group difference between matched samples processed by the three platforms. Significant differences were post hoc analyzed using the Wilcoxon signed-rank test.

Finally, we compared somatic variant detection in ctDNA isolated by the different platforms. For this purpose, we used previously generated diagnostic sequencing results on the somatic variant status in the primary and/or metastatic lesions of the corresponding patients (Table 2). We detected the expected somatic variants in all patients for all isolation methods. QS results were most comparable to QA (Figure 4). In MX, fewer mutant molecules, though not significant, and significantly fewer wild-type molecules were isolated (Figure 4A, B). However, this did not result in a significantly different VAF (Figure 4C).
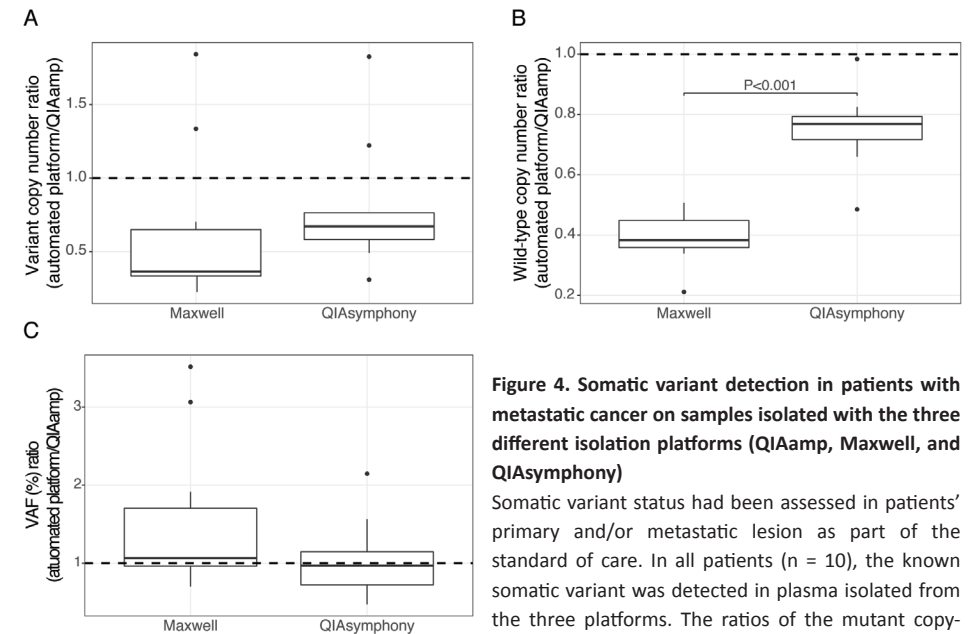
Figure 4. Somatic variant detection in patients with metastatic cancer on samples isolated with the three different isolation platforms (QIAamp, Maxwell, and QIAsymphony)
Somatic variant status had been assessed in patients' primary and/or metastatic lesion as part of the standard of care. In all patients (n = 10), the known somatic variant was detected in plasma isolated from the three platforms. The ratios of the mutant copy-

>>

Part 2
Liquid Biopsies

Chapter 4
High-throughput isolation of circulating tumor DNA:
a comparison of automated platforms

>>

number (A), wild-type copy-number (B), and variant allele frequency (VAF; C) measured in the Maxwell and QIAsymphony vs. QIAamp are shown. The dashed line (ratio of 1) resembles the situation when platforms have similar results. The Wilcoxon signed-rank test was used to test the difference between the platforms.

Table 2. Patient and tumor characteristics

Patient ID (#)	Primary tumor	Known somatic variant (nucleotide change)	Variant allele frequency in tissue (%)
BP-001	NSCLC	KRAS p.G12C (c.34G>T)	32
BP-003	Melanoma	NRAS p.Q61R (c.182A>G)	88
BP-004	Melanoma	BRAF p.V600E (c.1799_1800delinsAA)	50
BP-007	Melanoma	BRAF p.V600K (c.1798_1799delGTinsAA)	38
BP-008	CRC	KRAS p.G12D (c.35G>A)	45
BP-009	CRC	PIK3CA p.E545K (c.1633G>A)	45
BP-015	CRC	KRAS p.G13D (c.38G>A)	40
BP-016	CRC	KRAS p.G12V (c.35G>T)	Unknown
BP-023	CRC	KRAS p.G13D (c.38G>A)	Unknown
BP-028	Melanoma	BRAF p.V600K (c.1798_1799delinsAA)	55

CRC: colorectal cancer; NSCLC: non-small cell lung cancer.

Discussion

Up to now, several studies have investigated the effect of manual and automated cfDNA isolation platforms on ctDNA quantity and quality[22-24]. However, differences in preanalytical conditions, including plasma processing time, type of blood collection tube used, and storage conditions, hamper direct comparisons and straightforward conclusions. Here, we presented a study in which we have systematically optimized and compared automated isolation of cfDNA using QS and MX with the 'gold standard' QA.

The addition of carrier molecules like cRNA to plasma preceding cfDNA isolation increases the amount of cfDNA recovered during isolation by precipitating and binding of small molecules[25,26]. The manual QA platform requires addition of cRNA for the standard protocol, whereas the manufacturer's protocol of both the QS and MX does not require this. In a small pilot study, we observed that the addition of cRNA to the QS protocol improved cfDNA yield, so cRNA was

implemented into our standard QS protocol. However, Invitrogen has reported that cRNA might interfere with Qubit-based DNA quantification. Indeed, our findings suggest that the increase in cfDNA concentration as measured by Qubit for QS and MX is, at least in part, affected by the presence of cRNA. Data obtained from the TERT and plant DNA qPCR did not reveal any added value of cRNA to either of the automated platforms. Moreover, our fragmentation assay suggests that increasing amounts of cRNA reduce the amount of small fragments. Together, our results demonstrate that addition of cRNA to plasma does not improve cfDNA yields using these automated bead-based platforms.

In our previous study using the manual QA platform, we demonstrated the superiority of CellSave tubes over EDTA tubes for collecting plasma for cfDNA/ctDNA analysis as it ensures optimal ctDNA quality when processed within 96 h after blood withdrawal compared to only 24 h for EDTA tubes, enabling its use in multicenter clinical studies[19]. Therefore, we investigated the compatibility of CellSave tubes with QS and MX. On both platforms, we observed an increase in the isolation of large cfDNA fragments (2000 bp) in EDTA samples. This relates to the release of intact DNA from lysed leukocytes and a subsequent increase in cfDNA concentration, which we also observed here. As the recovery efficiency was not affected in CellSave tubes and the plasma samples were not contaminated with additional DNA from leukocytes, we recommend the use of CellSave tubes in combination with the QS or MX platform.

Currently, QA is widely used for cfDNA/ctDNA isolations, but its manual laborious and time-consuming protocol renders this method unsuitable for high-throughput isolations. The competing automated platforms QS and MX both use magnetic-bead-based protocols and have comparable hands-on times. However, costs and number of samples that can be processed per run differ (Table 1). In HBDs, cfDNA quantity and quality were similar on all platforms. However, in patients we saw for all assays that QA and QS yielded more cfDNA than MX. As this might suggest that higher amounts of cfDNA are less efficiently isolated by the MX platform, we spiked high amounts of fragmented DNA in HBD plasma and isolated this with MX (Figure S3). However, these high DNA amounts were isolated efficiently by MX. Another potential explanation for the difference in performance might be the absence of proteinase K incubation step in the MX protocol. Proteinase K is used in both the QA and QS protocols and can improve cfDNA yield by inhibiting nucleases and the release of protein-bound cfDNA. Moreover, recovery efficiency of plant DNA was lowest in MX. Altogether, this explains the lower yield of mutant and wild-type molecules isolated by MX, which may be a concern in samples with low frequent somatic variants. However, importantly, this lower yield did not translate into a significant difference in detected VAF (Figure S4). These data underline the importance of taking the used isolation method and read-out (mutant molecules/mL plasma or VAF) into consideration when comparing results between studies as well as for the diagnostic use of ctDNA. QS and QA performed comparable in detection of absolute numbers of mutant and wild-type molecules. Of note, other publications

Part 2
Liquid Biopsies

Chapter 4
High-throughput isolation of circulating tumor DNA:
a comparison of automated platforms

have observed similar performances of QA and MX in a head-to-head comparison[23,24]. This could be related to differences in preanalytical conditions (e.g. type of blood collection tube, plasma volume used as input), as multiple publications have demonstrated its relation to cfDNA quantity and quality[19,27,28]. In addition, we have optimized our QA protocol by re-eluting three times and thereby improving our cfDNA quantity. For automated magnetic-bead-based systems, this is not possible.

Conclusion

The results of this study show that the QS automated platform has comparable performance to the 'gold standard' QA and outperformed the MX platform depending on the read-out used. The QS platform is congruent with all our predefined goals as it (a) reduces hands-on time from 180–240 minutes to 30 minutes per run; (b) is able to process larger numbers of samples (96 instead of 24 at a time); (c) isolates comparable cfDNA yield with similar efficiency; and (d) has comparable ctDNA quantity and quality to QA. Therefore, the QS can replace the more laborious QA platform, especially when high-throughput cfDNA isolation is needed.

Acknowledgements

References

1. Mandel P, Metais P. Les acides nucleiques du plasma sanguin chez l'homme. C R Seances Soc Biol Fil 1948;142(3-4):241-3. (In und) (http://www.ncbi.nlm.nih.gov/pubmed/18875018).
2. Stroun M, Anker P, Maurice P, Lyautey J, Lederrey C, Beljanski M. Neoplastic characteristics of the DNA found in the plasma of cancer patients. Oncology 1989;46(5):318-22. (In eng) (http://www.ncbi.nlm.nih.gov/pubmed/2779946).
3. Bidard FC, Madic J, Mariani P, et al. Detection rate and prognostic value of circulating tumor cells and circulating tumor DNA in metastatic uveal melanoma. Int J Cancer 2014;134(5):1207-13. (In eng). DOI: 10.1002/ijc.28436.
4. Dawson SJ, Tsui DW, Murtaza M, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med 2013;368(13):1199-209. (In eng). DOI: 10.1056/NEJMoa1213261.
5. Diaz LA, Jr., Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. J Clin Oncol 2014;32(6):579-86. (In eng). DOI: JCO.2012.45.2011 [pii]
   10.1200/JCO.2012.45.2011.
6. Diehl F, Schmidt K, Choti MA, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med 2008;14(9):985-90. (In eng). DOI: nm.1789 [pii]
   10.1038/nm.1789.
7. Forshew T, Murtaza M, Parkinson C, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med 2012;4(136):136ra68. (In eng). DOI: 4/136/136ra68 [pii]
   10.1126/scitranslmed.3003726.
8. Murtaza M, Dawson SJ, Tsui DW, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature 2013;497(7447):108-12. (In eng). DOI: nature12065 [pii]
   10.1038/nature12065.
9. Shinozaki M, O'Day SJ, Kitago M, et al. Utility of circulating B-RAF DNA mutation in serum for monitoring melanoma patients receiving biochemotherapy. Clin Cancer Res 2007;13(7):2068-74. (In eng). DOI: 13/7/2068 [pii]
   10.1158/1078-0432.CCR-06-2120.
10. Herbreteau G, Vallee A, Knol AC, et al. Quantitative monitoring of circulating tumor DNA predicts response of cutaneous metastatic melanoma to anti-PD1 immunotherapy. Oncotarget 2018;9(38):25265-25276. (In eng). DOI: 10.18632/oncotarget.25404
    25404 [pii].
11. Pugh TJ. Circulating Tumor DNA for Detecting Minimal Residual Disease in Multiple Myeloma. Semin Hematol 2018;55(1):38-40. (In eng). DOI: S0037-1963(18)30018-0 [pii]
    10.1053/j.seminhematol.2018.03.002.
12. Zill OA, Greene C, Sebisanovic D, et al. Cell-Free DNA Next-Generation Sequencing in Pancreatobiliary Carcinomas. Cancer Discov 2015;5(10):1040-8. (In eng). DOI: 2159-8290.CD-15-0274 [pii]
    10.1158/2159-8290.CD-15-0274.
13. Oxnard GR, Paweletz CP, Kuang Y, et al. Noninvasive detection of response and resistance in EGFR-mutant lung cancer using quantitative next-generation genotyping of cell-free plasma DNA. Clin Cancer Res 2014;20(6):1698-1705. (In eng). DOI: 10.1158/1078-0432.CCR-13-2482.
14. Sefrioui D, Sarafan-Vasseur N, Beaussire L, et al. Clinical value of chip-based digital-PCR platform for the detection of circulating DNA in metastatic colorectal cancer. Dig Liver Dis 2015;47(10):884-90. (In eng). DOI: S1590-8658(15)00352-7 [pii]
    10.1016/j.dld.2015.05.023.
15. Fleischhacker M, Schmidt B. Circulating nucleic acids (CNAs) and cancer--a survey. Biochim Biophys Acta 2007;1775(1):181-232. (In eng). DOI: S0304-419X(06)00059-X [pii]
    10.1016/j.bbcan.2006.10.001.
16. Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 2014;6(224):224ra24. (In eng). DOI: 6/224/224ra24 [pii]
    10.1126/scitranslmed.3007094.
17. Jahr S, Hentze H, Englisch S, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. Cancer Res 2001;61(4):1659-65. (In eng) (http://www.ncbi.nlm.nih.gov/pubmed/11245480).
18. Elshimali YI, Khaddour H, Sarkissyan M, Wu Y, Vadgama JV. The clinical utilization of circulating cell free DNA (CCFDNA) in blood of cancer patients. Int J Mol Sci 2013;14(9):18925-58. (In eng). DOI: ijms140918925 [pii]
    10.3390/ijms140918925.

Part 2
Liquid Biopsies

Chapter 4
High-throughput isolation of circulating tumor DNA:
a comparison of automated platforms

19.   van Dessel LF, Beije N, Helmijr JC, et al. Application of circulating tumor DNA in prospective clinical oncology trials - standardization of preanalytical conditions. Mol Oncol 2017;11(3):295-304. (In eng). DOI: 10.1002/1878-0261.12037.

20.   Kang Q, Henry NL, Paoletti C, et al. Comparative analysis of circulating tumor DNA stability In K3EDTA, Streck, and CellSave blood collection tubes. Clin Biochem 2016;49(18):1354-1360. (In eng). DOI: S0009-9120(16)30040-6 [pii] 10.1016/j.clinbiochem.2016.03.012.

21.   Invitrogen. Qubit dsDNA assay specificity in the presence of single-stranded DNA. Application note.

22.   Devonshire AS, Whale AS, Gutteridge A, et al. Towards standardisation of cell-free DNA measurement in plasma: controls for extraction efficiency, fragment size bias and quantification. Anal Bioanal Chem 2014;406(26):6499-512. (In eng). DOI: 10.1007/s00216-014-7835-3.

23.   Sorber L, Zwaenepoel K, Deschoolmeester V, et al. A Comparison of Cell-Free DNA Isolation Kits: Isolation and Quantification of Cell-Free DNA in Plasma. J Mol Diagn 2017;19(1):162-168. (In eng). DOI: S1525-1578(16)30221-5 [pii]
10.1016/j.jmoldx.2016.09.009.

24.   Perez-Barrios C, Nieto-Alcolado I, Torrente M, et al. Comparison of methods for circulating cell-free DNA isolation using blood from cancer patients: impact on biomarker testing. Transl Lung Cancer Res 2016;5(6):665-672. (In eng). DOI: 10.21037/tlcr.2016.12.03
tlcr-05-06-665 [pii].

25.   Kishore R, Reef Hardy W, Anderson VJ, Sanchez NA, Buoncristiani MR. Optimization of DNA extraction from low-yield and degraded samples using the BioRobot EZ1 and BioRobot M48. J Forensic Sci 2006;51(5):1055-61. (In eng). DOI: JFO204 [pii]
10.1111/j.1556-4029.2006.00204.x.

26.   Shaw KJ, Thain L, Docker PT, et al. The use of carrier RNA to enhance DNA extraction from microfluidic-based silica monoliths. Anal Chim Acta 2009;652(1-2):231-3. (In eng). DOI: S0003-2670(09)00445-0 [pii]
10.1016/j.aca.2009.03.038.

27.   Volckmar AL, Sultmann H, Riediger A, et al. A field guide for cancer diagnostics using cell-free DNA: From principles to practice and clinical applications. Genes Chromosomes Cancer 2018;57(3):123-139. (In eng). DOI: 10.1002/gcc.22517.

28.   Haselmann V, Ahmad-Nejad P, Geilenkeuser WJ, et al. Results of the first external quality assessment scheme (EQA) for isolation and analysis of circulating tumor DNA (ctDNA). Clin Chem Lab Med 2018;56(2):220-228. (In eng). DOI: 10.1515/cclm-2017-0283
/j/cclm.ahead-of-print/cclm-2017-0283/cclm-2017-0283.xml [pii].

Supplementary information
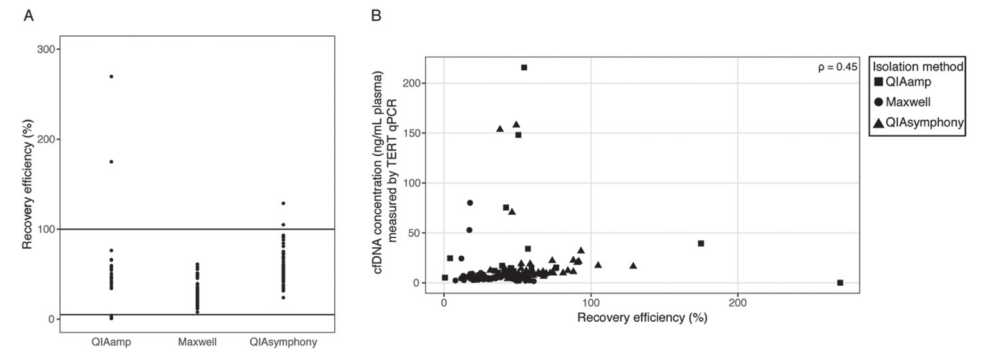
Figure S1. Overview of the recovery efficiency of synthetic plant DNA in all samples isolated with the different platforms (QA, MX, and QS)
(A) Dot plot of the recovery efficiency for each isolation platform, as analyzed by qPCR using spiked-in synthetic plant DNA. Samples with a recovery efficiency < 5% or > 100% (black horizontal lines) were excluded from the analyses.
(B) Correlation between recovery efficiency and cfDNA concentration (ng/mL plasma) measured by TERT qPCR assay. Correlations were tested by Spearman's rank correlation coefficient. *P < 0.001.
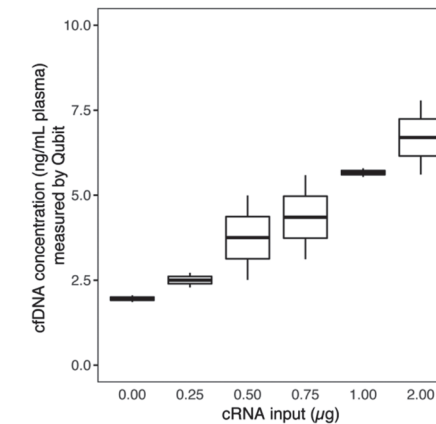
Figure S2. Effect of cRNA addition on cfDNA quantity using the QS platform
cfDNA concentration (ng/mL plasma) was determined by Qubit after adding increasing amounts of cRNA (0–4 µg) before start of the plasma isolation. Boxes (interquartile ranges [IQR]) and whiskers (1.5x IQR) are shown together with the median (black horizontal line).

Part 2
Liquid Biopsies

Chapter 4
High–throughput isolation of circulating tumor DNA:
a comparison of automated platforms

Figure S4. Representative data images of SNP genotyping dPCR assay isolated with the different platforms (QA, MX, and QS)
A subject with an intermediate (A), high (B) and low (C) VAF are shown. On the Y-axis, positive FAM signal represents mutant molecules (blue dots); on the X-axis, positive VIC signal represents wild-type molecules (red dots). Green dots reflect the presence of a mutant and a wild-type molecule in a single well.

Figure S3. Performance of the MX platform using increasing DNA input (0, 15, and 60 ng/mL fragmented cell line DNA has been spiked in HBD plasma)
The effects on (A) cfDNA concentration (ng/mL plasma) measured by TERT qPCR, (B) recovery efficiency measured by plant DNA qPCR, (C) total number of mutant molecules, and (D) VAF are shown. Boxes (interquartile ranges [IQR]) and whiskers (1.5x IQR) are shown together with the median (black horizontal line). Outliers are indicated as single black points. Symbol ● is mean value shown with whiskers (standard deviation). n = 5.

Table S1. Custom primer and probe sequences used for qPCR

Assay Name	Forward primer	Reverse primer	Probe	Amplicon size (bp)	Manufacturer
β-actin 136 bp	5'-GCG CCG TTC CGA AAG TT-3'	5'- CGG CGG ATC GGC AAA -3'	FAM-ACC GCC GAG ACC GCG TC-MGBNFQ	136	Invitrogen
β-actin 420 bp	5'-CCG CTA CCT CTT CTG GTG-3'	5'-GAT GCA CCA TGT CAC ACT G-3'	VIC-CCT CCC TCC TTC CTG GCC TC-BHQ	420	Invitrogen
β-actin 2000 bp	*The β-actin 2000 bp fragment is detected when both primers and probes of the β-actin 136 bp assay and the β-actin 420 bp are able to amplify a long DNA fragment. This double positive signal is detected by the QuantStudio 3D Digital PCR System (van Dessel et al., 2017).*				Invitrogen
Plant	5'-GAT CTT CAA CCA GGA GAT CA-3'	5'-AGT GAC AGT GAG GAC AAT CC-3'	FAM-ACC CAT CTT CAC CGG A-BHQ1	70	Primers: IDT Probe: Sigma-Aldrich, Saint Louis, Missouri, USA

Table S2. Standard SNP genotyping assays

Assay ID	Assay name	Gene	Cosmic ID	Amino acid change	Nucleotide change
AHS1P6Q	NRAS_584	*NRAS*	584	p.Q61R	c.182A>G
AHD2BW0	KRAS_532	*KRAS*	532	p.G13D	c.38G>A
AHABHHX	PIK3CA_763	*PIK3CA*	763	p.E545K	c.1633G>A

Part 2
Liquid Biopsies

Chapter 4
High-throughput isolation of circulating tumor DNA:
a comparison of automated platforms

Table S3. Custom SNP genotyping assays

Assay ID	Assay name	Gene	Cosmic ID	Amino acid change	Nucleotide change	Forward primer	Reverse primer	Probe	Amplicon size (bp)	Manufacturer
ANNKR4W	BRAF_V600K_72bp	BRAF	473	p.V600K	c.1798_1799GT>AA	5'-TCA TGA AGA CCT CAC AGT AAA AAT AGG T-3'	5'-TGG GAC CCA CTC CAT CGA-3'	Variant: FAM-TGG TCT AGC TAC AAA GA-NFQ Wild-type : VIC- TTT TGG TCT AGC TAC AGT GA-NFQ	72	ThermoFisher Scientific
AN47WF2	BRAF_V600E_72bp-2	BRAF	475	p.V600E	c.1799_1800 TG>AA	5'-TCA TGA AGA CCT CAC AGT AAA AAT AGG T-3'	5'-TGG GAC CCA CTC CAT CGA-3'	Variant: FAM-TTG GTC TAG CTA CAG AAA-NFQ Wild-type : VIC- TTT TGG TCT AGC TAC AGT GA-NFQ	72	
AN9HJKW	KRAS_G12C_76bp	KRAS	516	p.G12C	c.34G>T	5'-TGC TGA AAA TGA CTG AAT ATA AAC TTG TG-3'	5'- AGC TGT ATC GTC AAG GCA CTC TT-3'	Variant: FAM-TTG GAG CTT GTG GCG TA-NFQ Wild-type : VIC- TTG GAG CTG GTG GCG T-NFQ	76	
ANU63FK	KRAS_G12D_76bp	KRAS	521	p.G12D	c.35G>A	5'-TGC TGA AAA TGA CTG AAT ATA AAC TTG TG-3'	5'-AGC TGT ATC GTC AAG GCA CTC TT-3'	Variant: FAM-TTG GAG CTG TTG GCG TA-NFQ Wild-type : VIC-TTG GAG CTG GTG GCG T-NFQ	76	
ANAAAYM	KRAS_G12V_76bp	KRAS	520	p.G12V	c.35G>T	5'-TGC TGA AAA TGA CTG AAT ATA AAC TTG TG-3'	5'-AGC TGT ATC GTC AAG GCA CTC TT-3'	Variant: FAM-TTG GAG CTG ATG GCG TA-NFQ Wild-type : VIC-TTG GAG CTG GTG GCG T-NFQ	76	

Chapter 5

Optimizing Nanopore sequencing-based detection of structural variants enables individualized circulating tumor DNA-based disease monitoring in cancer patients

Chapter 5

Jose Espejo Valle-Inclan*, Christina Stangl*, Anouk C. de Jong*, Lisanne F. van Dessel, Markus J. van Roosmalen, Jean C. Helmijr, Ivo Renkens, Roel Janssen, Sam de Blank, Chris J. de Witte, John W.M. Martens, Maurice P.H.M. Jansen, Martijn P. Lolkema, Wigard P. Kloosterman
* These authors contributed equally

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing–based detection of structural variants enables
individualized circulating tumor DNA–based disease monitoring in cancer patients

Abstract

Somatic genomic structural variations (SVs) are promising personalized biomarkers to quantify circulating tumor DNA (ctDNA) in liquid biopsies as they represent unique tumor derived molecules. However, in most solid malignancies these SVs are variable and can be located anywhere in the genome, thus the complexity of the identification of personalized SVs hinders routine use in the clinic. Here, we developed a novel approach for rapid discovery of a set of patient-specific somatic SVs. We combine low coverage cancer genome sketching using Oxford Nanopore sequencing with a machine learning approach to detect a set of somatic SVs. We analyzed tumor samples of high-grade ovarian and prostate cancer patients, successfully identified candidate SVs and validated on average ten somatic SVs per patient with breakpoint-spanning PCR mini-amplicons. These SVs could be quantified in ctDNA samples of patients with metastatic prostate cancer using a digital PCR assay. The SV quantification in these longitudinal samples suggest that indeed SV dynamics correlate with and may improve other response biomarkers such as PSA. Our work enables rapid and cost-effective identification of a set of patient-specific SVs that can be used to study ctDNA dynamics.

Introduction

The detection of cancer recurrence as well as accurate and fast monitoring of response to treatment currently lacks sensitivity for detection of changes over time[1,2]. Liquid biopsies, which can be used to detect circulating tumor DNA (ctDNA) from body fluids, such as blood, in a minimally invasive manner, are a promising approach to improve monitoring of tumor burden over time[3,4]. Circulating tumor DNA, which originates from apoptotic and necrotic tumor cells, has been shown to have a positive linear correlation with tumor burden[5]. In multiple cases, ctDNA analyses identified cancer recurrence months before clinical symptoms presented[6-8].

As ctDNA is only a fraction of the total circulating cell free DNA (cfDNA), it should be distinguished from cfDNA from normal cells by identification of ctDNA-specific genetic alterations. Genomic structural variations (SVs) represent tumor- and ctDNA-specific biomarkers to detect and quantify ctDNA with high sensitivity in liquid biopsies[7-10]. Most solid cancers contain dozens to hundreds of somatic SVs[11,12]. Besides some recurrent driver SV events that functionally impact tumorigenesis, the vast majority of these somatic SVs are patient- and tumor specific passenger events[13], which may nevertheless be good biomarkers for tumor load tracing. SVs form a unique breakpoint junction between two joined DNA strands and can be validated by straightforward junction-spanning (quantitative) PCR assays, which facilitates its applicability[8].

Somatic SVs are commonly detected with short-read, paired-end next generation sequencing (NGS). However, as SVs can be very large, short reads are less suited for SV detection[14-16]. Recently, long-read sequencing techniques from Oxford Nanopore Technologies (ONT) and Pacific Biosciences (PacBio) have emerged and their increased power for germline and somatic SV detection has been extensively demonstrated[15-19]. Moreover, ONT enables a short turnaround time and real-time data analysis[20].

To enable rapid and cost-efficient identification of a set of patient-specific somatic SVs for ctDNA monitoring, we developed a pipeline that leverages the long-read and fast sequencing capabilities of Nanopore sequencing in combination with a computational method that enables accurate selection of a subset of somatic SVs from low coverage Nanopore sequencing data. The method detects a subset of genomic SVs and can be applied to tumor tissue obtained from (needle) biopsy or resection. The computational approach combines SV calling with random forest classification and germline SV filtering against a blacklist to enrich for somatic SVs without the need of matching germline sequencing data, which reduces the cost and time of the assay. We were able to design SV-specific PCR-assays for ctDNA tracking within three days after obtaining a tumor biopsy. We validated the pipeline in multiple ovarian and prostate cancer samples. In addition, we demonstrate the clinical applicability of our pipeline by retrospectively tracking the identified somatic SVs in longitudinal cfDNA samples of patients with metastatic prostate cancer, by using digital PCR.

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing–based detection of structural variants enables
individualized circulating tumor DNA–based disease monitoring in cancer patients

Materials and Methods

DNA isolation and Nanopore sequencing

COLO829 (ATCC® CRL-1974™) cell line was obtained from the American Type Culture Collection (ATCC) and grown according to standard procedures as recommended by ATCC. DNA was isolated using a phenol chloroform protocol[21]. For some Nanopore sequencing runs, DNA was sheared using g-tubes (Covaris). DNA was size selected on the PippinHT (Sage Science). Library preparation was performed using the Lib SQK-LSK109 kit (Oxford Nanopore Technologies) and DNA was then sequenced in 49 separate runs using R9.4 flow cells (Oxford Nanopore Technologies) on the MinION (44), GridION (3) and PromethION (2) instruments (Table S1).

HGS-3 organoid line was cultured following the ovarian cancer organoid culture protocol[22]. DNA was isolated by using a phenol chloroform protocol[21]. DNA was size selected on the PippinHT (Sage Science). Library preparation was performed using the Lib SQK-LSK109 kit (Oxford Nanopore Technologies) and DNA was then sequenced in 40 separate runs using R9.4 (23) and R9.5 (17) flow cells (Oxford Nanopore Technologies) on the MinION (35) and GridION (5) instruments (Table S1).

Tumor DNA from patients with ovarian cancer was isolated with the Genomic-tip kit (Qiagen), following the manufacturer's protocol for tissue samples. DNA was prepared for Nanopore sequencing with the Lib SQK-LSK109 (Oxford Nanopore Technologies). The library from one tumor sample was loaded on one revD (Ova1) or R9.4 (Ova2-4) flow cell (Oxford Nanopore Technologies). Sequencing was performed on a MinION (Ova2, Ova4) or GridION (Ova1, Ova3) instrument (Oxford Nanopore Technologies) (Table S1). Lymphocyte DNA for PCR validation assays was isolated from blood with the DNeasy Blood & Tissue Kit (Qiagen).

Tumor and germline DNA from patients with prostate cancer were obtained from a fresh frozen core needle biopsy of a metastatic lesion and blood, respectively. DNA was isolated on an automated setup with the QIAsymphony according to the supplier's protocols (DSP DNA Midi kit for blood and DSP DNA Mini kit for tissue). In the context of the CPCT-02 study, whole-genome sequencing (WGS) was performed by the Hartwig Medical Foundation, Amsterdam, The Netherlands[23]. Residual tumor DNA (80-250 ng) was used for Nanopore sequencing. DNA was prepared for Nanopore sequencing with the Lib SQK-LSK109 (Oxford Nanopore Technologies). The library from one tumor sample was loaded on one R9.4 (Pros1), revD (Pros2,3) or high-sensitivity research prototype (Pros4-6) flow cell (Oxford Nanopore Technologies). Sequencing was performed on a GridION instrument (Oxford Nanopore Technologies) (Table S1).

Illumina sequencing and analysis (COLO829 and HGS-3)

Short read WGS was obtained for matched tumor and normal DNA from the COLO829 cell line[24] and the HGS-3 organoid line[22].

SV calling was performed by using GRIDSS (v. 2.0.1)[25] in joint calling mode (tumor + reference) for COLO829 and HGS-3 separately. Somatic SV calls were filtered as in[24] (https://github.com/hartwigmedical/pipeline/blob/master/scripts/gridss_somatic_filter.R)

Benchmarking somatic SV calling from low coverage Nanopore sequencing data

Nanopore data from COLO829 was randomly subsampled to 5x sequencing coverage three times independently with Sambamba[26]. SV calling was performed with NanoSV (v. 1.2.4)[17] with a 2-read support threshold; Sniffles (v. 1.0.12)[27] with parameters "--report_BND --genotype -s 2"; and NanoVar (v. 1.3.8)[28] with default parameters. In all cases 8 threads were used and computational resources were measured with GNU Time. True and false positives were calculated using the short-read somatic SV callset described above.

SV calling and filtering pipeline

The SHARC pipeline is available through https://github.com/UMCUGenetics/SHARC.

Mapping is performed in parallel for each FASTQ file by using minimap2 (v. 2.12)[29] with settings "-x map-ont -a --MD". The reference genome used is version GRCh37. Sorting and merging of BAM files was done by using sambamba (v. 0.6.5)[26]. SV calling was performed by using NanoSV (v. 1.1.2)[17]. Default NanoSV settings were used except a minimum read count of 2 (cluster_count=2) and minimum mapping quality of 20 (min_mapq=20).

VCFs are filtered by using the command `awk '$7 == "PASS" && $1 !~ /(Y|MT)/ && $5 !~ /(Y|MT):/ && $5 != "<INS>"'` to select PASS calls and remove insertions and SVs involving chromosomes Y or MT.

VCFs are then annotated with the distance to the closest single repeat element in the reference genome[30,31], the closest gap element in the reference genome[31,32], and the closest segmental duplication element in the reference genome[31]. These elements were taken from the UCSC genome browser (http://genome.ucsc.edu/)[31], using the GRCh37/hg19 genome version.

We trained a random forest (RF) model to filter out false-positive SV calls from Nanopore data, similarly as previously described[17]. We expanded the selection of input features for the RF, by including read length, SV calling features, and overlap with repeat features in the reference genome (Table S3). We trained the classifier on the well-characterized NA12878 Genome in a Bottle (GIAB) sample[33-35], for which high-quality germline SV call sets have been obtained by using Illumina[35], PacBio[34] and Nanopore[33] sequencing. The GIAB SV truth set was generated by intersecting these three GIAB SV sets resulting in a set of 1,515 germline SVs. We used ⅔ of the GIAB truth set as a training set and ⅓ as a test set. We established a precision-recall curve from 100 bootstrapping runs (Figure S4), where the training data were split into 90%-10% train-test subsets. Based on the precision-recall curve, we defined an operating point of 96% precision and 99.5% recall (Figure S4). The final model was then re-trained on the whole training set and tested on the ⅓ test set. The performance on the test set was 95.1% precision and 99.6% recall, representing an accuracy of 97.2% (Figure S4). SV candidates are classified as "true" or "false" based on this RF model.

We set up two databases of SV calls:

(i) SharcDB: containing raw NanoSV calls from Nanopore sequencing data of 14 samples, 11 of which belong to this study (COLO829, HGS-3, Ova1, Ova2, Ova3, Ova4, Pros1, Pros2, Pros4,

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing–based detection of structural variants enables
individualized circulating tumor DNA–based disease monitoring in cancer patients

Pros5 and Pros6; and three more for which we had SV calls from high coverage Nanopore data: COLO829BL (lymphoblastoid cell line, 50x sequencing depth), VCAP (prostate cancer cell line[36]) and the Genome in a Bottle (GIAB) SV calls[33]. For tests performed with the samples included in this study, the specific sample was excluded from blacklisting with SharcDB;

(ii) RefDB: containing germline calls obtained from WGS short-read data of 59 controls: 19 blood controls from patients with ovarian cancer[22] where germline SVs were called with Manta (v. 1.0.3)[37] with default parameters, and 40 healthy individuals (biological parents of individuals with congenital abnormalities)[38] where germline SVs were called with Manta (v. 0.29.5)[37] with default parameters.

SV calls from tumor samples are overlapped with those two databases by using VCF-explorer (https://github.com/UMCUGenetics/vcf-explorer).

Only samples classified as "true" by the RF model and that do not overlap with any sample in the databases qualify for primer design.

Primer design for filtered SV calls is automated by using Primer3 (v. 1.1.4)[39] with a product size range of 30-230 bp.

SVs with a successful primer design are ranked based on SV length and the 20 largest are selected for PCR validation. Insertions are filtered out early in the pipeline since the inserted sequence cannot be accurately inferred from the low coverage Nanopore sequencing data. Inter-chromosomal translocations are not present in the top 20 ranked SVs because the final ranking is based on SV size and this cannot be determined for inter-chromosomal SVs. However, they are available in the final VCF file and primers are designed by default, so they can be manually selected for PCR validation and assay development.

Breakpoint PCR

To validate SVs, breakpoint PCR with AmpliTaqGold (Applied Biosystems) was performed according to the manufacturer's protocol. 10 ng primary tumor DNA (somatic) and 10 ng lymphocyte DNA (germline) per primer-pair were used as input. PCR products were loaded and visualized on a 2% agarose gel.

cfDNA isolation

cfDNA was isolated from ascites fluid of Ova2 by using the QIAamp Circulating Nucleic Acid Kit (Qiagen) according to the manufacturer's protocol. Plasma samples from patients with prostate cancer were obtained longitudinally during treatment in 3x 10 ml CellSave preservative tubes (Menarini Silicon Biosystems, Huntingdon Valley, PA, USA) and processed within 96 hours as previously described[40]. Circulating DNA was isolated with the QIAsymphony® DSP Circulating DNA Kit (Qiagen) according to manufacturer's protocol with some minor modifications[41]. All cfDNA samples were quantified by Qubit[TM] fluorometric quantitation (Invitrogen).

Quantitative PCR

As primer specificity is essential for reliable interpretation of an end-point assay like digital PCR (dPCR), primers for the detection of structural variants were validated by quantitative PCR (qPCR) on whole genome amplified (WGA) tumor and germline DNA. In brief, qPCR was performed by using the CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad Laboratories) and the final reaction mix consisted of 10 µL SensiFAST[TM] SYBR ® Lo-Rox mix (Bioline), 0.5 µM forward and reverse primers, 10 ng of WGA DNA and Ultrapure DNas/RNAse free H_2O to bring up the reaction volume to 20 µL. The cycle conditions were as follows: 14 cycles of 10s at 95°C and 30s at 65-58°C (touchdown), followed by 20-40 cycles of 10s at 95°C and 30s at 60°C. In addition, a melt curve was generated from 56°C to 95°C to assess the generated PCR products. Based on qPCR results, two primer sets for the detection of SVs in each patient were selected for quantification by dPCR. Primer sets were excluded from use with dPCR when one of the following ocurred: > 1 PCR product, $Cq_{germline}$- Cq_{tumor} < 5 and/or Cq_{tumor} > 20.

DNA sonication and fragment size analysis

To mimic the length of cfDNA and improve DNA molecule partition, WGA DNA of both tumor and germline were sonicated to a peak size of ~ 150 bp with the S220 Focused-ultrasonicator (Covaris) according to the manufacturer's protocol. The sonication conditions were as follows; 200-250 ng WGA DNA (concentration determined by Qubit[TM] fluorometric quantitation) in 50 µL Ultrapure DNas/RNAse free H_2O, Peak Incident Power: 175 W, Duty Factor: 10%, Cycles per Burst: 200, Treatment Time: 280 s, Temperature: 7°C, and Water Level: 12. After sonication DNA fragment sizes were analyzed with the High Sensitivity DNA kit (Agilent Technologies) on the Bioanalyzer (Agilent Technologies) and the sample concentration was re-quantified by Qubit[TM] fluorometric quantitation (Invitrogen).

Design of digital PCR assays for absolute quantification of SVs in cfDNA

To quantify SVs in cfDNA, dPCR was performed. First, the exact position of the breakpoint as determined by Nanopore sequencing was validated. We used already available sequenced Illumina data from the CPCT-02 study (Pros1, Pros4, Pros5 and Pros6), but Sanger sequencing of the particular qPCR product could be used as well. To enable quantification of both mutant and wild-type alleles, additional primers for the detection of wild-type upstream (WT-U) allele and wild-type downstream (WT-D) allele of the breakpoint and fluorescent probes for both mutant and wild-type alleles were developed by using the Primer Express Software v3.0 (ThermoFisher) and the online tool Primer3Plus[39]. All primers and fluorescent probes (Table S4) were ordered from Eurogentec.

Preamplification of cfDNA

To enable sensitive detection of multiple SVs in limited amounts of cfDNA, two SVs per patient were preamplified with 0.2-1 ng of cfDNA. Pre-amplified tumor and germline DNA samples were

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing–based detection of structural variants enables
individualized circulating tumor DNA–based disease monitoring in cancer patients

used as respectively positive and negative control. Preamplification was performed by using 4 µL of TaqMan™ PreAmp Master Mix (cat.no: 4488593, Life Technologies), 2 µL primer pool (0.25 µM) consisting of SV forward (SV-F) and reverse (SV-R) primers and upstream (WT-U) and downstream (WT-D) wild-type primers, and 2 µL (cf)DNA for a total volume of 8 µL. Preamplification cycle conditions were: 10 min at 95°C followed by 14 cycles of 15s at 95°C and 4 min at 60°C, and finally pause at 4°C. After the preamplification reaction, 72 µL of Ultrapure DNase/RNAse free H_2O was added to each sample. Next, preamplified cfDNA was diluted 40x per 1 ng input, used for the preamplification, to prevent overloading of the dPCR chips.

Absolute quantification of SVs in cfDNA with digital PCR

For the quantification of SVs in (cf)DNA, dPCR was performed with the Naica Crystal PCR system (Stilla Technologies) by using the following optimized reaction mix: 1 µL of diluted preamplified (cf)DNA sample, 5.6 µL PerfeCTa Multiplex qPCR ToughMix (Cat.No: 733-2322PQ, Quantabio). 0.25 µM probes (SV^{FAM}, WT-U^{HEX}, WT-D^{CY5}), 0.75 µM of the SV forward (SV-F) and reverse primer (SV-R), 0.25 µM of the WT-U and WT-D primers, 0.1 µM Fluorescein (Cat.No: 0681-100G, VWR) and Ultrapure DNAse/RNAse free H_2O to bring up the total volume to 28 µL. Samples were loaded onto Stilla Sapphire chips (Cat.no. C13000, Stilla Technologies) and dPCR was performed with the same cycle conditions as for the primer validation with qPCR. Median number of analyzable droplets was 21,357, interquartile range 19,837-22,736. dPCR reactions were optimized with 10 ng sonicated tumor and germline WGA DNA. When a SV could be detected in preamplified cfDNA samples, a dPCR of all longitudinal cfDNA samples was performed on 5 ng of stock (no preamplification) cfDNA to enable absolute quantification of mutant molecules in plasma.

Statistical analysis

qPCR experiments were analyzed with Bio-Rad CFX Manager version 3.1. dPCR experiments were analyzed with Crystal Miner™ software, version 2.1.6 (Stilla Technologies). Thresholds for positive fluorescence were determined per primer pair based on positive and negative controls. Variant allele frequency (VAF) was calculated according to the following formula:

number of mutant molecules per µL in chip (as defined by Crystal Miner™ software) / (number of mutant molecules per µL in chip + number of wild-type molecules per µL in chip) * 100%.

Absolute number of mutant molecules per mL plasma was calculated as follows:

number of mutant molecules per µL in chip * 28 µL (input in chip) / (used eluate/total volume of eluate * volume of plasma used for isolation).

To correct for zero values on a log scale, +1 was counted to every value and axes were corrected with -1. Spearman's correlation coefficient was calculated for comparisons of VAF based on upstream wild-type allele vs. downstream wild-type allele, two replicates and preamplified vs. non-preamplified cfDNA samples. Corresponding slope was calculated by using linear regression analysis.

Ethics approval and consent to participate

Tumor samples of four patients with high-grade serous ovarian cancer (OC) and six patients with metastatic castration-resistant prostate cancer (PC) were used in this study. Patients with OC participated in the HUB-OVI study, in which tumor tissue and blood were obtained for organoid culture (tumor) and WGS (tumor and blood). Clinical data was extracted from the patient file in collaboration with the Dutch Cancer Registration. Patients with PC participated in both the CPCT-02 study (NCT01855477) and the CIRCUS study (NTR5732), in which tumor tissue from a metastatic lesion for WGS and longitudinal cfDNA samples were obtained. Longitudinal ctDNA quantification was performed for four patients with PC. Informed consent was obtained within all studies. Clinical data for patients with PC were collected in an electronic case report form (ALEA Clinical). All studies were performed according to the guidelines of the European Network of Research Ethics Committees (EUREC) following European, national and local law.

Results

Detection of somatic structural variations from low coverage Nanopore sequencing of tumor biopsies

The first step of our analysis involves low coverage Nanopore sequencing of genomic tumor-derived DNA (Figure 1A). A single Nanopore run on the MinION or GridION platforms typically generates between 5-15 Gbs of data[33], corresponding to 1.5-5x coverage of the human genome. Next, the low coverage sequencing data are mapped to the reference genome followed by the detection of SV breakpoint junctions from split read mappings (Figure 1B)[17]. Subsequently, a classification and filtering pipeline is applied to enrich for somatic SV breakpoints irrespective of corresponding germline data (Figure 1B). Finally, PCR assays with mini-amplicons are designed to validate the 20 most likely somatic SVs. SVs are confirmed as either somatic or germline by breakpoint PCR on tumor and corresponding lymphocyte DNA (Figure 1C). Successful breakpoint PCR assays for somatic SVs can then be utilized as biomarkers for ctDNA-based monitoring of treatment response and disease recurrence (Figure 1D).

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing–based detection of structural variants enables
individualized circulating tumor DNA–based disease monitoring in cancer patients
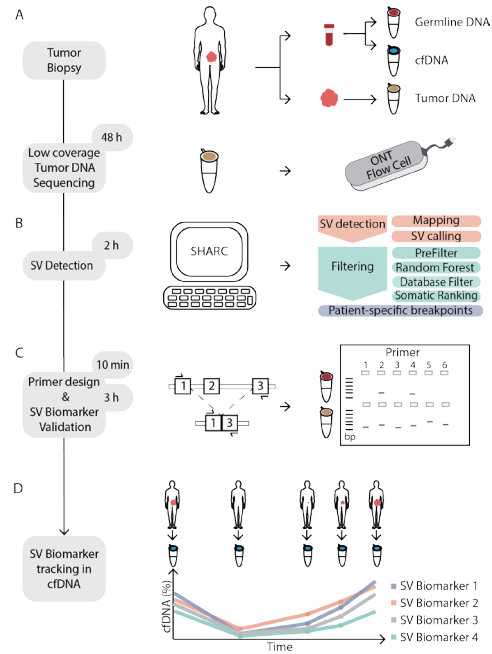
Figure 1. Schematic overview of SHARC

(A) (Needle) biopsy or resection from a tumor as well as blood are obtained from a patient at initial diagnosis. Germline DNA (red) and cfDNA (blue) isolated from blood and tumor DNA (brown) from tumor material. Tumor DNA is sequenced on one ONT flow cell.

(B) Tumor-specific SV detection and filtering is performed with the bioinformatic SHARC pipeline.

(C) SV-specific breakpoint spanning primers are designed. Breakpoint PCR with SV-specific primers is performed on germline and tumor DNA to confirm somatic SVs.

(D) Somatic SVs are used as biomarkers and traced within cfDNA from a patient to monitor disease dynamics in a longitudinal manner.

Establishment of a somatic SV reference set

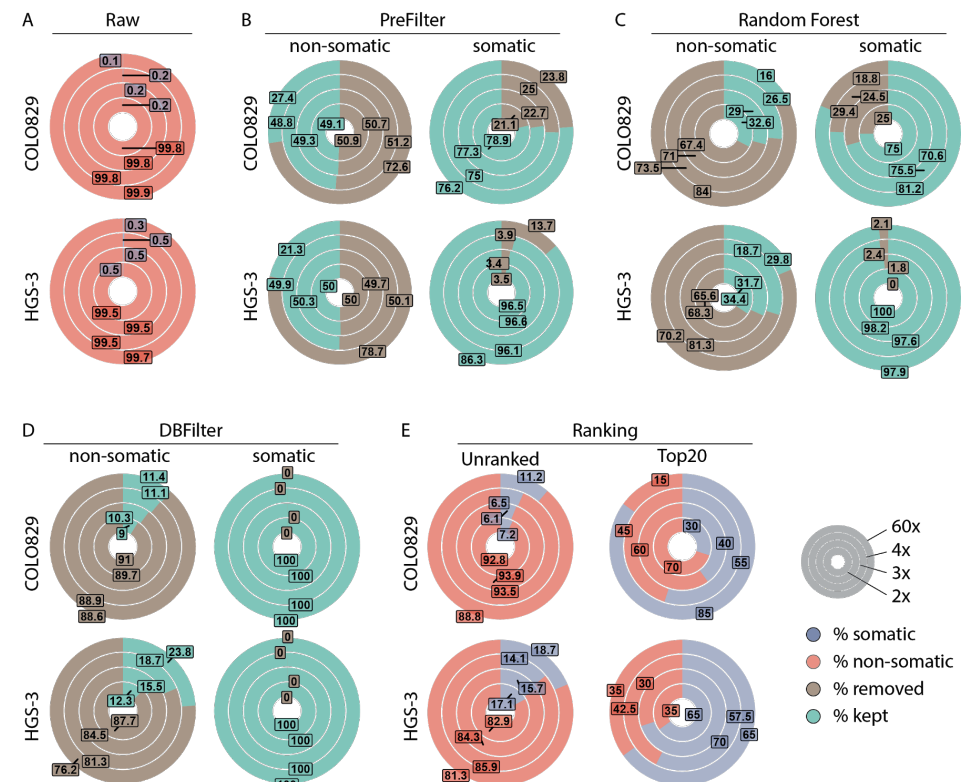To verify the ability of our pipeline to detect somatic SVs, we used genomic data from the melanoma cell line COLO829[42] and the ovarian cancer organoid line HGS-3[22]. We utilized short-read WGS data from both lines (90x and 30x coverage for COLO829 and HGS-3, respectively) and matching reference samples (30x coverage in both cases) to establish two reference sets of somatic SVs (Materials and Methods). By using a state-of-the-art somatic SV detection pipeline[24,43-45], we detected 92 and 295 somatic SVs in COLO829 and HGS-3, respectively. Additionally, we generated long-read Nanopore sequencing data for COLO829 and HGS-3, reaching high coverages of 59x (COLO829) and 56x (HGS-3) (Figure S1 and Table S1). To simulate low coverage long-read sequencing of tumor genomes, we randomly subsampled the Nanopore sequencing reads to coverages of 4x, 3x and 2x. The subsampling was performed 20 times independently for each case, to mitigate the effect of chance on the subsampling and subsequent analysis.

Next, we tested our ability to detect SVs from high and low coverage Nanopore sequencing data. We used NanoSV, a previously validated Nanopore SV caller[17,19], to call SVs from the Nanopore sequencing data. To maximize sensitivity, we performed SV calling using lenient settings on high and low coverage COLO829 and HGS-3 Nanopore datasets. (Table S2). Based on the overlap with the somatic short-read reference set, raw SV calls were classified as somatic (true-positives) or non-somatic (false-positives). As expected, the vast majority of the raw SV calls in all the different coverage datasets were non-somatic, on average 99.84% (range 99.81-99.9%, COLO829) and 99.55% (range 99.4-99.74%, HGS-3) (Figure 2A). In the high coverage Nanopore datasets, we validated 84 (91% of the short-read reference set) and 219 (74% of the short-read reference set) true-positive somatic SVs for COLO829 and HGS-3, respectively, representing a small fraction of the total number of raw SV calls (Figure 2A and Figure S2A). Similarly, we identified an average of 23 (25% of the short-read reference set) and 53 (18% of the short-read reference set) somatic SV breakpoints in each of the low coverage Nanopore sequencing datasets for COLO829 and HGS-3, respectively. (Figure 2A). Furthermore, we compared the performance of the SV callers NanoSV, Sniffles[27] and NanoVar[28]. Thus, we show that based on lenient SV calling of high- and low-coverage Nanopore sequencing data with NanoSV, somatic SVs can be identified.

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing–based detection of structural variants enables
individualized circulating tumor DNA–based disease monitoring in cancer patients

>>
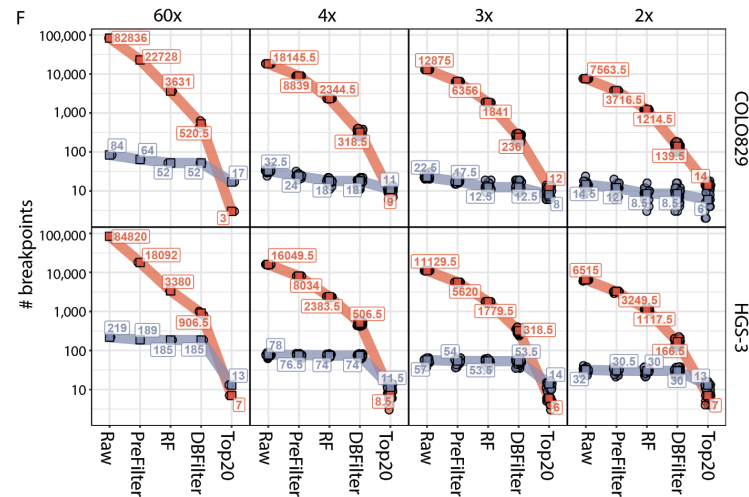
Figure 2. Detection of somatic SVs with the SHARC pipeline based on high and low coverage Nanopore data
High coverage Nanopore sequencing data from COLO829 (melanoma cell line) and HGS-3 (ovarian cancer organoid) were subsampled to low coverages. Outer circles represent the high coverage sets (59x for COLO829 and 56x for HGS-3) and inner circles represent low coverage subsets (4x 3x, 2x). The following filtering steps were applied in a cumulative manner in the order displayed.
(A) Median percentage of non-somatic (red) and somatic (blue) breakpoints in the raw NanoSV calls for COLO829 (top) and HGS-3 (bottom).
(B) Median percentage of non-somatic (left) and somatic (right) SV calls kept (green) or removed (brown) in the pre-filtering step for COLO829 and HGS-3.
(C) Median percentage of non-somatic (left) and somatic (right) SV calls kept (green) or removed (brown) by the Random Forest SV classifier for COLO829 and HGS-3.
(D) Median percentage of non-somatic (left) and somatic (right) SV calls kept (green) or removed (brown) by the database filtering for COLO829 and HGS-3.
(E) Median percentage of non-somatic (red) and somatic SV (blue) calls in the complete SHARC output (left) and top 20 largest SVs (right) for COLO829 and HGS-3.
(F) Total number of non-somatic (red) and somatic (blue) SV calls at each step of the pipeline for both COLO829 and HGS-3. In low coverage subsets, all data points are shown and the square box represents the median value.
RF: Random forest; DBFilter: Database filter.

Enrichment for somatic SV calls from Nanopore sequencing data

Since the somatic SVs identified among the SV call sets of the Nanopore data represent only a small fraction of the total raw SV calls, we implemented a panel of cumulative filtering steps to enrich for somatic SVs. First, we selected only "PASS" SV calls (based on default NanoSV filter flags[17] (Materials and Methods). Secondly, we excluded calls involving chromosome Y or the mitochondrial genome. Finally, we removed all insertions, since the exact inserted sequence cannot be accurately defined from low coverage Nanopore sequencing data, thus hampering the final PCR assay development at a later step. As a result of these filtering steps, 72.6% (COLO829) and 76.2% (HGS-3) false-positive calls were removed in the high coverage sets (Figure 2B and

Table S2). For the low coverage sets, the filtering removed on average 50.9% (COLO829) and 49.9% (HGS-3) of false-positive calls (Figure 2B and Table S2). In contrast, the vast majority of true-positive somatic SV calls were maintained following SV filtering (on average 76.9% in COLO829 and 93.9% in HGS-3, Figure 2B).

To further reduce the number of false-positive SV calls, we employed a random forest (RF) machine learning approach ((Materials and Methods)), similarly as previously described for SV calling of Nanopore data[17]. We applied the RF classifier to the filtered high and low coverage subsets of COLO829 and HGS-3. For the high coverage sets, the RF labelled 84% (COLO829) and 81.3% (HGS-3) of false-positive SV calls as false (Figure 2C). For the low coverage sets, on average 70.6% (COLO829) and 68% (HGS-3) of false-positive SV calls were labelled as false (Figure 2C). In addition, in the high coverage sets 81.25% (COLO829) and 97.88% (HGS-3) of true-positive somatic SV calls were labelled as true. Similar percentages of true-positive SV calls were labelled as true in the low coverage sets, on average 73.7% (COLO829) and 98.6% (HGS-3) (Figure 2C).

These results show that the RF classifier filters out the majority of non-somatic breakpoints, while maintaining true-positive somatic SV calls. However, germline SV calls are also maintained at this step, requiring further filtering to enrich for somatic SVs (Figure S2B).

To reduce the number of germline SVs, we implemented a blacklist filtering step. Therefore, the remaining SV calls were overlapped with two databases (DBFilter) as panel-of-normal (PON) filtering: (i) SharcDB, containing SV calls from Nanopore sequencing of 14 different samples, and (ii) RefDB, containing germline SV calls from 59 control samples previously sequenced using Illumina WGS in our group (Materials and Methods). Following this filtering step, 100% of true-positive somatic SV calls from both the COLO829 and HGS-3 high and low coverage sets were retained (Figure 2D). In contrast, 88.6% (COLO829, high coverage), 76.2% (HGS-3, high coverage) and on average 89.9% (COLO829, low coverage) and 84.5% (HGS-3, low coverage) of remaining false-positive SV calls were filtered out (Figure 2D). Due to this filtering, the fraction of true-positive somatic breakpoints among the remaining SV calls increased to 6.6%-18.7% for the low and high-coverage Nanopore datasets of COLO829 and HGS-3 (Figure 2E and Figure S2A).

To further enrich for somatic SVs, we implemented a ranking method, based on the observation that large SVs are more likely to be somatic than germline SVs (Figure S4). This increased the percentage of true-positive somatic SVs to 85% (COLO829) and 65% (HGS-3) in the high coverage sets, and to on average 43% (COLO829) and 64.1% (HGS-3) in the low coverage sets (Figure 2E).

Altogether, our SV filtering pipeline strongly enriches for true-positive somatic breakpoints and filters out the majority of false-positives and germline SVs. We demonstrate a total enrichment of true-positive somatic SV calls from 0.1% in the raw calls to 85% in the final top 20 ranked calls (17/20, COLO829, high coverage), 0.26% to 65% (13/20, HGS-3, high coverage), on average 0.18% to 41.7% (8.3/20, COLO829, low coverage) and on average 0.49% to 64.2% (12.8/20, HGS-3, low coverage) (Figure 2F). Of note, despite low coverage sequencing, each of the somatic SV calls

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing-based detection of structural variants enables
individualized circulating tumor DNA-based disease monitoring in cancer patients

identifies breakpoints at nucleotide resolution, providing immediate access to breakpoint PCR testing.

Validation in tumor tissue from patients with ovarian and prostate cancer

Next, we tested the pipeline on four high-grade serous ovarian cancer (Ova1-4) and six prostate cancer (Pros1-6) samples. We sequenced tumor DNA on one Nanopore flow cell per sample. The ovarian cancer samples and three prostate cancer samples (Pros1-3) were sequenced on commercial ONT flow cells. For the ovarian cancer samples, we started library preparation with minimally 1 µg of DNA. For the prostate cancer samples limited material was available, and we started library preparation with 250 ng of DNA. For one sample (Pros3), not enough sequencing data was produced to confidently detect somatic SVs and this sample was therefore excluded from all subsequent analyses (Table S1). Three additional prostate cancer samples (Pros4-6) were sequenced on ONT research prototype flow cells with higher sequencing sensitivity, thus requiring less DNA input material. In these cases, library preparation was started with an average of 108 ng (80-128 ng) of DNA and an average of 10 ng of library was loaded for sequencing (Table S1). We obtained an average sequence coverage of 2.3x (range: 1.8-4.0) (Figure 3A and Table S1) and average read lengths of 7.8 Kbp (range: 4.2-16.3 Kbp) (Figure 3B and Table S1). The sequencing throughput was not affected by the lower DNA input when using the high-sensitivity prototype flow cells. (Table S1).

Following the lenient SV calling, pre-filtering, RF classification, the database filtering and ranking steps, an average 2.8% (range of 1.0-4.4%) of SVs per sample were retained (Figure 3C). We performed breakpoint PCR assays on lymphocyte and tumor DNA for the top 20 ranked SVs and validated an average of 10 (50%, range 25-80%) somatic SVs per sample (Figure 3D). Therefore, despite not having enough sequencing depth to provide a complete genome construction, we were able to identify several somatic SV biomarkers in each of the tumor samples.

We investigated the recall of validated somatic SVs at different timepoints during the sequencing run. We found that, on average, 81.6% (range 50-100%) of validated somatic SVs were already detected within the first 24 hours of sequencing (Figure S6). This offers the opportunity to reduce the sequencing time, accelerating tumor biomarker discovery with one day.
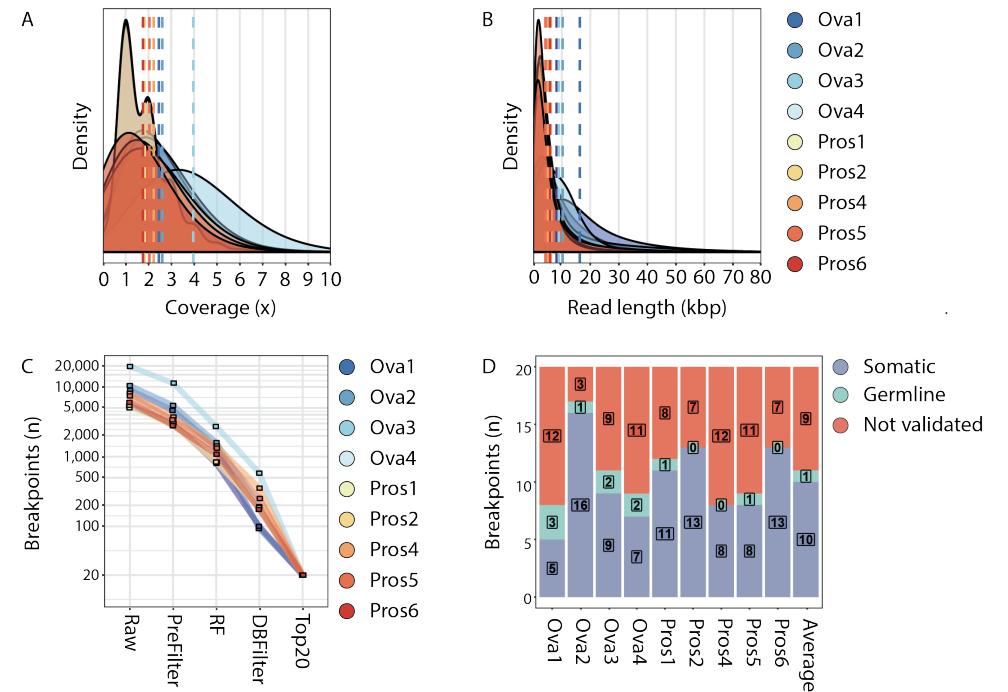
Figure 3. SHARC identifies and validates tumor-specific SV biomarkers from low-pass Nanopore tumor sequencing data
(A-B) Plots showing the distribution of coverage (A) and read length (B) read length for the nine tumor samples sequenced on one flow cell each. Dashed lines represent averages for each sample.
(C) Total number of somatic SVs present at each of the steps throughout the SV calling and filtering pipeline.
(D) The top 2020 ranked breakpoints for each sample were tested by breakpoint PCR using tumor and germline DNA. Graph depicts the number of breakpoints validated as somatic (blue), germline (green) or breakpoints that could not be validated (red).
RF: Random forest; DBFilter: Database filter.

Detection of somatic SVs in cfDNA from patients with ovarian and prostate cancer

To show the applicability of the pipeline to detect clinically relevant biomarkers, we next tested if we could detect the validated somatic SVs in cfDNA of patients. Ascites fluid, which is known to contain cfDNA and ctDNA[46] was available for Ova2 at time of disease recurrence. We extracted cfDNA from the ascites and tested the 16 validated somatic SVs out of the Top20 by PCR. 100% of somatic SVs could be detected within the cfDNA from ascites (Figure S7), and not in the germline or water controls. Next, we tested whether validated SVs could be detected in cfDNA from blood. Therefore, we selected two patient-specific SVs for four prostate cancer patients (Pros1, 4, 5 and 6) based on a high signal to noise ratio observed in qPCR assays for SV breakpoints (Figure 4A and Materials and Methods).

To enable sensitive and quantitative detection, we designed digital PCR (dPCR) assays for the eight selected SVs (Figure 4B). For each SV, we aimed to design a probe for both wild-type alleles

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing–based detection of structural variants enables
individualized circulating tumor DNA–based disease monitoring in cancer patients

(up- and downstream) and for the mutant allele (across the breakpoint junction). For five SVs we could design an assay that quantified both the upstream and downstream wild-type allele. For the three other SVs, primers/probes for only one of the wild-type alleles were designed, as appropriate primer design for the other allele was hindered by repetitive sequences at the target site. As the amount of cfDNA within one liquid biopsy is limited, we used a conditional breakpoint detection approach: (i) if dPCR on preamplified cfDNA (input preamplification: 0.2-1 ng cfDNA) confirmed the presence of the SV within cfDNA, (ii) then subsequent dPCR on non-preamplified cfDNA (stock cfDNA) (input dPCR: 5 ng cfDNA) was performed. The latter enabled calculation of both the variant allele frequency (VAF) and the number of mutant molecules per milliliter plasma (MM/mL plasma). First, we selected two timepoints per patient, one at baseline and one at progression of disease and confirmed the presence of all eight SVs with dPCR on preamplified cfDNA (Figure S8). Thereafter, dPCR on the stock cfDNA successfully detected all SVs in the four patients, both in baseline and progression samples (Figure 4C and 4D). Despite the fact that the VAF in preamplified cfDNA correlates to the VAF in stock cfDNA ($r_s = 0.928$), they should be considered two separate outcome measurements (regression coefficient = $0.72 \neq 1$) (Figure S9A). Moreover, VAF based on the wild-type upstream allele was highly similar to VAF based on the wild-type downstream allele in stock cfDNA ($r_s = 0.996$, regression coefficient = 1.05) (Figure S9B), suggesting no significant imbalances between the two sides of the breakpoint.
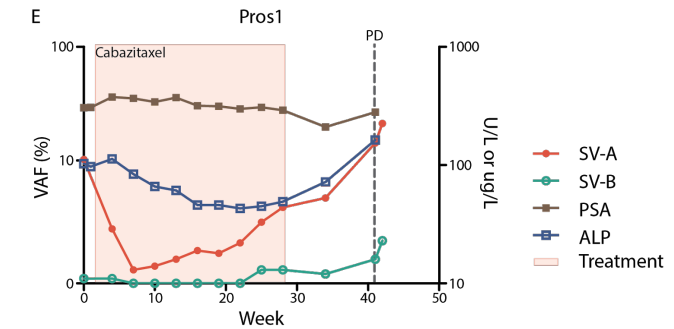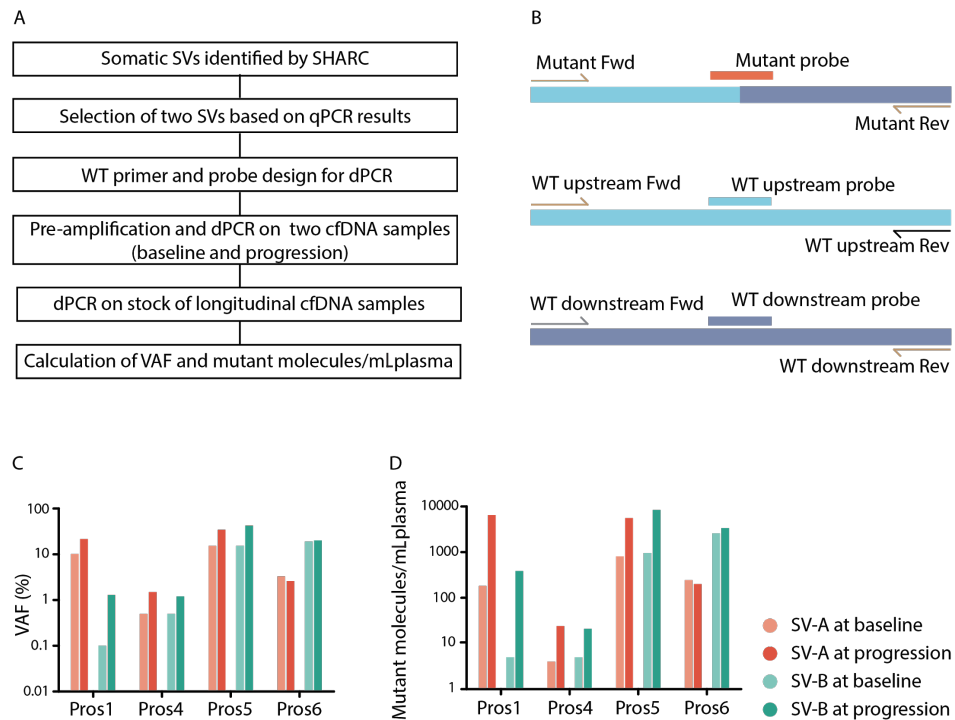
Figure 4. Digital PCR-based quantification of SVs in blood
(A) Schematic overview of quantification of tumor-specific SVs, identified by SHARC, in cfDNA from blood by using qPCR and dPCR.
(B) Primer and probe design for dPCR. The wild-type upstream and wild-type downstream allele share each one primer with the mutant allele. Three probes with different fluorescents were designed to specifically detect the mutant allele or one of the wild-type alleles.
(C-D) Detection of two tumor-specific SVs in cfDNA from blood from four patients with prostate cancer at baseline and at progression of disease with dPCR. Shown are VAF (C) and mutant molecules per mL plasma (D).
(E) Quantification of SVs in longitudinal cfDNA samples from blood of patient Pros1. Graph depicts VAFs of SVs, treatment, laboratory parameters (prostate specific antigen (PSA), alkaline phosphatase (ALP)) and clinical progression of disease (PD).

Monitoring treatment response in patients with prostate cancer

In addition to the detection of SVs in cfDNA at baseline and progression of disease, we explored the capacity to use SVs to monitor treatment response over time. To enable reliable response monitoring, measurements should be accurate and repeatable. As VAFs are ratios and in principle not influenced by technical variations between timepoints, we chose to report VAFs only. To verify the accuracy of dPCR, we performed two technical replicates for all preamplified samples of Pros5 and Pros6 and confirmed a high correlation of VAFs between the replicates ($r_s = 0.987$, regression coefficient = 0.918) (Figure S9C). Finally, we quantified the eight SVs of the four prostate cancer patients in the longitudinally collected samples from before, during and after treatment. For Pros1, SV-A shows the potential to improve response evaluation as its dynamics correspond to the expected response to treatment with cabazitaxel and increases towards the end of treatment, resulting in the highest levels at clinical progression of disease (Figure 4E). These changes also seem to correlate with other blood biomarkers, including PSA and ALP. In addition, SV-B in Pros1 similarly correlates with response to treatment (Figure 4E). Also, for Pros5 both SV-A and SV-B show clear changes over time correlating with clinical parameters, and Pros4 and Pros6 have less compelling dynamics of the detected SVs (Figure S10A-C).

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing–based detection of structural variants enables
individualized circulating tumor DNA–based disease monitoring in cancer patients

Discussion

Recent studies have utilized somatic SVs for tracking tumor burden from liquid biopsies[7-10]. Although these studies showed the potential of this methodology, they lacked sufficient turn-around time to provide personalized biomarkers before the initiation of patient treatment. This is due to lengthy short-read WGS approaches for SV detection and an associated substantial number of false-positive somatic SVs, requiring laborious testing Nanopore sequencing combined with a machine learning approach is capable to efficiently identify a set of somatic SVs from tumor tissue within three days. The rapid and simple workflow offers great potential for routine monitoring of cancer dynamics. We illustrate the applicability of our method to measure tumor burden by using a series of longitudinally gathered blood samples from metastatic prostate cancer patients.

Obtaining enough tumor material for DNA isolation is often a limiting factor for next-generation sequencing assays. We show that Nanopore sequencing and somatic SV detection is possible from limited amounts of DNA that can be extracted from a metastatic tumor needle biopsy, which is an important requisite for clinical viability. DNA input can be decreased even further to as little as 80 ng when using flow cells with increased sensitivity for DNA (research prototype flow cells provided by ONT).

Long-read sequencing is an excellent method for the detection of SVs at nucleotide resolution, even at low sequencing depth, because each long-read that bridges a breakpoint-junction provides direct information on the breakpoint position and sequence[17]. Sequencing of a tumor sample on a single GridION/MinION Nanopore flow cell generates insufficient sequencing data to accurately establish a complete genomic profile. However, using the pipeline developed here, we efficiently enriched for patient-specific somatic SV events - irrespective of their functional impact on tumor biology. Despite the very low coverage, the computational method functions independently of corresponding germline sequencing data. These assets make our pipeline a cost-efficient assay for detection of personalized somatic SV biomarkers. Furthermore, on average 50% of the detected SVs are somatic, which minimizes the hands-on effort needed for validation purposes. For all analyzed tumors, we identified at least five somatic SV biomarkers per patient, an amount within the range of biomarkers used to trace ctDNA in previous work[7,9,47]. With expected increases in sequencing throughput from ONT sequencing, the performance of the pipeline will improve significantly. Furthermore, the use of cheap disposable flow cells (Flongle) could reduce assay costs to ⅛ of the current sequencing price of 800€[48]. The minimal costs of this assay would enable the broader application of such individualized SV monitoring in cancer patients.

We retrospectively traced levels of ctDNA with two SVs per patient for four prostate cancer patients and compared tumor dynamics to clinical biomarkers such as PSA and ALP. The quantitative measurement of SVs in ctDNA suggests that VAFs of SVs correlate with tumor load (Pros1 and Pros5). Moreover, the SVs would have indicated progression of disease earlier than PSA did in some patients (Pros1 and Pros 4). Even though we only tested two SVs per patient, this clearly illustrates the potential clinical utility of quantifying ctDNA with SVs to monitor response to treatment. The assay could be optimized by not only identifying the tumor-specific SVs, but also SVs that represent the dominant disease clone and upcoming, targetable subclones. In addition, larger prospective studies should confirm that indeed measuring SVs improves clinical decision making in patients with metastatic prostate, and other cancer types.

Conclusions

Clinicians are well aware of the dynamic response of cancer to treatment but lack the tools to monitor these changes in real-time and thus generally respond to alterations too late for true treatment success. We present a method to overcome these limitations and provide a solution to immediate individualized disease monitoring. This approach could increase sensitivity of disease monitoring to such levels that more intelligent treatment approaches could be envisioned.

Acknowledgements

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing–based detection of structural variants enables
individualized circulating tumor DNA–based disease monitoring in cancer patients
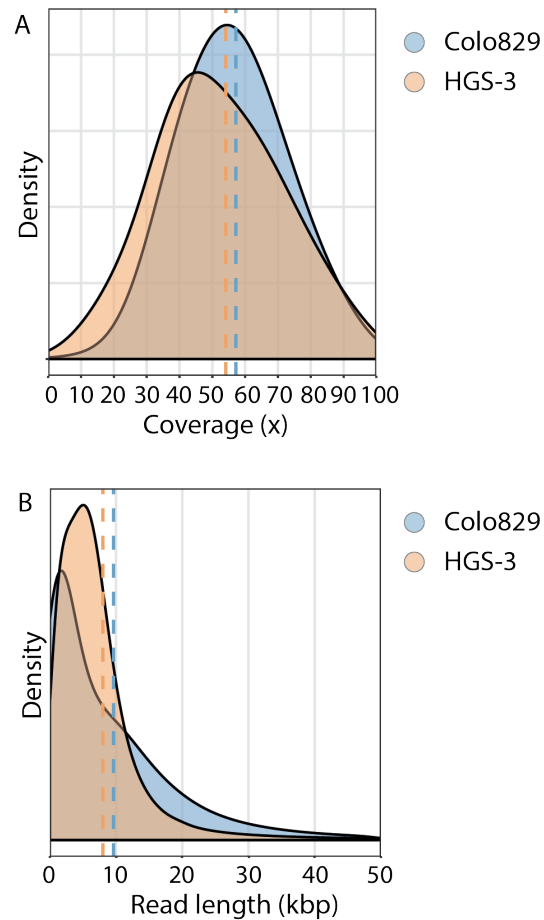
References

1. Turkbey B, Pinto PA, Choyke PL. Imaging techniques for prostate cancer: implications for focal therapy. Nat Rev Urol 2009;6(4):191-203. DOI: 10.1038/nrurol.2009.27.
2. Gerwing M, Herrmann K, Helfen A, et al. The beginning of the end for conventional RECIST - novel therapies require novel imaging approaches. Nat Rev Clin Oncol 2019;16(7):442-458. DOI: 10.1038/s41571-019-0169-5.
3. Wan JCM, Massie C, Garcia-Corbacho J, et al. Liquid biopsies come of age: towards implementation of circulating tumor DNA. Nat Rev Cancer 2017;17(4):223-238. DOI: 10.1038/nrc.2017.7.
4. Heitzer E, Haque IS, Roberts CES, Speicher MR. Current and future perspectives of liquid biopsies in genomics-driven oncology. Nat Rev Genet 2019;20(2):71-88. DOI: 10.1038/s41576-018-0071-5.
5. Schwarzenbach H, Hoon DS, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. Nat Rev Cancer 2011;11(6):426-37. DOI: 10.1038/nrc3066.
6. Diehl F, Schmidt K, Choti MA, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med 2008;14(9):985-90. (In eng). DOI: nm.1789 [pii] 10.1038/nm.1789.
7. Olsson E, Winter C, George A, et al. Serial monitoring of circulating tumor DNA in patients with primary breast cancer for detection of occult metastatic disease. EMBO Mol Med 2015;7(8):1034-47. DOI: 10.15252/emmm.201404913.
8. McBride DJ, Orpana AK, Sotiriou C, et al. Use of cancer-specific genomic rearrangements to quantify disease burden in plasma from patients with solid tumors. Genes Chromosomes Cancer 2010;49(11):1062-9. DOI: 10.1002/gcc.20815.
9. Leary RJ, Kinde I, Diehl F, et al. Development of personalized tumor biomarkers using massively parallel sequencing. Sci Transl Med 2010;2(20):20ra14. DOI: 10.1126/scitranslmed.3000702.
10. Klega K, Imamovic-Tuco A, Ha G, et al. Detection of Somatic Structural Variants Enables Quantification and Characterization of Circulating Tumor DNA in Children With Solid Tumors. JCO Precis Oncol 2018;2018. DOI: 10.1200/PO.17.00285.
11. Priestley P, Baber J, Lolkema MP, et al. Pan-cancer whole-genome analyses of metastatic solid tumors. Nature 2019;575(7781):210-216. (In eng). DOI: 10.1038/s41586-019-1689-y 10.1038/s41586-019-1689-y [pii].
12. Li Y, Roberts ND, Wala JA, et al. Patterns of somatic structural variation in human cancer genomes. Nature 2020;578(7793):112-121. DOI: 10.1038/s41586-019-1913-9.
13. Macintyre G, Ylstra B, Brenton JD. Sequencing Structural Variants in Cancer for Precision Therapeutics. Trends Genet 2016;32(9):530-542. DOI: 10.1016/j.tig.2016.07.002.
14. Dixon JR, Xu J, Dileep V, et al. Integrative detection and analysis of structural variation in cancer genomes. Nat Genet 2018;50(10):1388-1398. DOI: 10.1038/s41588-018-0195-8.
15. Chaisson MJP, Sanders AD, Zhao X, et al. Multi-platform discovery of haplotype-resolved structural variation in human genomes. Nat Commun 2019;10(1):1784. DOI: 10.1038/s41467-018-08148-z.
16. Huddleston J, Chaisson MJP, Steinberg KM, et al. Discovery and genotyping of structural variation from long-read haploid genome sequence data. Genome Res 2017;27(5):677-685. DOI: 10.1101/gr.214007.116.
17. Cretu Stancu M, van Roosmalen MJ, Renkens I, et al. Mapping and phasing of structural variation in patient genomes using Nanopore sequencing. Nat Commun 2017;8(1):1326. DOI: 10.1038/s41467-017-01343-4.
18. Nattestad M, Goodwin S, Ng K, et al. Complex rearrangements and oncogene amplifications revealed by long-read DNA and RNA sequencing of a breast cancer cell line. Genome Res 2018;28(8):1126-1135. DOI: 10.1101/gr.231100.117.
19. De Coster W, De Rijk P, De Roeck A, et al. Structural variants identified by Oxford Nanopore PromethION sequencing of the human genome. Genome Res 2019;29(7):1178-1187. DOI: 10.1101/gr.244939.118.
20. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of Nanopore sequencing to the genomics community. Genome Biol 2016;17(1):239. DOI: 10.1186/s13059-016-1103-0.
21. Quick J. Ultra-long read sequencing protocol for RAD004 V.3. 2018. DOI: https://doi.org/10.17504/protocols.io.mrxc57n.
22. Kopper O, de Witte CJ, Lohmussaar K, et al. An organoid platform for ovarian cancer captures intra- and interpatient heterogeneity. Nat Med 2019;25(5):838-849. DOI: 10.1038/s41591-019-0422-6.
23. van Dessel LF, van Riet J, Smits M, et al. The genomic landscape of metastatic castration-resistant prostate cancers reveals multiple distinct genotypes with potential clinical impact. Nature Communications 2019;10(1):5251. DOI: 10.1038/s41467-019-13084-7.
24. Cameron DL BJ, Shale C, Espejo Valle-Inclan J, Besselink N, Cuppen E, Priestley P, Papenfuss AT. GRIDSS2: harnessing the power of phasing and single breakends in somatic structural variant detection. bioRxiv 2020.
25. Cameron DL, Schroder J, Penington JS, et al. GRIDSS: sensitive and specific genomic rearrangement detection using positional de Bruijn graph assembly. Genome Res 2017;27(12):2050-2060. DOI: 10.1101/gr.222109.117.
26. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. Bioinformatics 2015;31(12):2032-4. (In eng). DOI: btv098 [pii] 10.1093/bioinformatics/btv098.
27. Sedlazeck FJ, Rescheneder P, Smolka M, et al. Accurate detection of complex structural variations using single-molecule sequencing. Nat Methods 2018;15(6):461-468. DOI: 10.1038/s41592-018-0001-7.
28. Tham CY, Tirado-Magallanes R, Goh Y, et al. NanoVar: accurate characterization of patients' genomic structural variants using low-depth Nanopore sequencing. Genome Biol 2020;21(1):56. DOI: 10.1186/s13059-020-01968-7.
29. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 2018;34(18):3094-3100. DOI: 10.1093/bioinformatics/bty191.
30. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 1999;27(2):573-80. DOI: 10.1093/nar/27.2.573.
31. Haeussler M, Zweig AS, Tyner C, et al. The UCSC Genome Browser database: 2019 update. Nucleic Acids Res 2019;47(D1):D853-D858. DOI: 10.1093/nar/gky1095.
32. Bailey JA, Gu Z, Clark RA, et al. Recent segmental duplications in the human genome. Science 2002;297(5583):1003-7. DOI: 10.1126/science.1072047.
33. Jain M, Koren S, Miga KH, et al. Nanopore sequencing and assembly of a human genome with ultra-long reads. Nat Biotechnol 2018;36(4):338-345. DOI: 10.1038/nbt.4060.
34. Pendleton M, Sebra R, Pang AW, et al. Assembly and diploid architecture of an individual human genome via single-molecule technologies. Nat Methods 2015;12(8):780-6. DOI: 10.1038/nmeth.3454.
35. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. Nature 2015;526(7571):68-74. DOI: 10.1038/nature15393.
36. Korenchuk S, Lehr JE, L MC, et al. VCaP, a cell-based model system of human prostate cancer. In Vivo 2001;15(2):163-8. (https://www.ncbi.nlm.nih.gov/pubmed/11317522).
37. Chen X, Schulz-Trieglaff O, Shaw R, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics 2016;32(8):1220-2. DOI: 10.1093/bioinformatics/btv710.
38. Middelkamp S, Vlaar JM, Giltay J, et al. Prioritization of genes driving congenital phenotypes of patients with de novo genomic structural variants. Genome Med 2019;11(1):79. DOI: 10.1186/s13073-019-0692-0.
39. Untergasser A, Cutcutache I, Koressaar T, et al. Primer3--new capabilities and interfaces. Nucleic Acids Res 2012;40(15):e115. DOI: 10.1093/nar/gks596.
40. van Dessel LF, Beije N, Helmijr JC, et al. Application of circulating tumor DNA in prospective clinical oncology trials - standardization of preanalytical conditions. Mol Oncol 2017;11(3):295-304. (In eng). DOI: 10.1002/1878-0261.12037.
41. van Dessel LF, Vitale SR, Helmijr JCA, et al. High-throughput isolation of circulating tumor DNA: a comparison of automated platforms. Mol Oncol 2019;13(2):392-402. DOI: 10.1002/1878-0261.12415.
42. Pleasance ED, Cheetham RK, Stephens PJ, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. Nature 2010;463(7278):191-6. DOI: 10.1038/nature08658.
43. Gong T, Hayes VM, Chan EKF. Detection of somatic structural variants from short-read next-generation sequencing data. Brief Bioinform 2020. DOI: 10.1093/bib/bbaa056.
44. Cameron DL, Di Stefano L, Papenfuss AT. Comprehensive evaluation and characterisation of short read general-purpose structural variant calling software. Nat Commun 2019;10(1):3240. DOI: 10.1038/s41467-019-11146-4.
45. Kosugi S, Momozawa Y, Liu X, Terao C, Kubo M, Kamatani Y. Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. Genome Biol 2019;20(1):117. DOI: 10.1186/s13059-019-1720-5.
46. Husain H, Nykin D, Bui N, et al. Cell-Free DNA from Ascites and Pleural Effusions: Molecular Insights into Genomic Aberrations and Disease Biology. Mol Cancer Ther 2017;16(5):948-955. DOI: 10.1158/1535-7163.MCT-16-0436.
47. Harris FR, Kovtun IV, Smadbeck J, et al. Quantification of Somatic Chromosomal Rearrangements in Circulating Cell-Free DNA from Ovarian Cancers. Sci Rep 2016;6:29831. DOI: 10.1038/srep29831.
48. Gilpatrick T, Lee I, Graham JE, et al. Targeted Nanopore sequencing with Cas9-guided adapter ligation. Nat Biotechnol 2020;38(4):433-438. DOI: 10.1038/s41587-020-0407-5.

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing–based detection of structural variants enables
individualized circulating tumor DNA–based disease monitoring in cancer patients

Supplementary information

Figure S1. Coverage and read length of COLO829 and HGS-3
Coverage (A) and read length (B) distribution for COLO829 and HGS-3 Nanopore sequencing data. Dashed lines represent average.
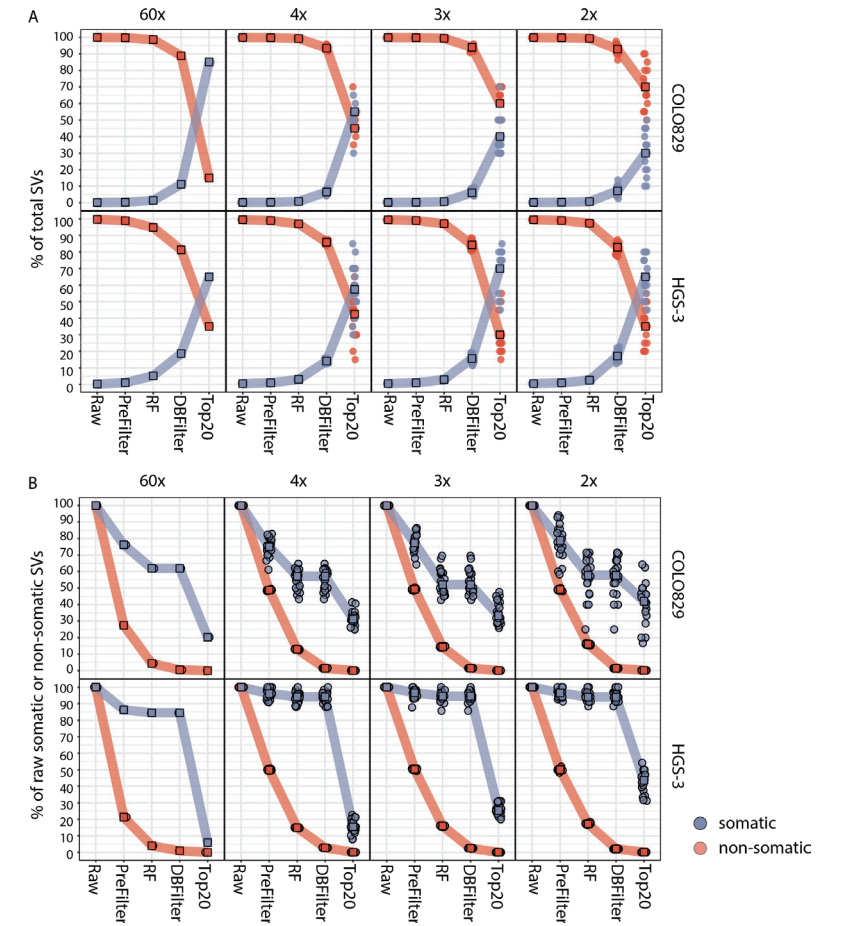
Figure S2. Enrichment of somatic SV calls of COLO829 and HGS-3 after subsequent steps of the SHARC pipeline

The filtering steps were applied in a cumulative manner in the order displayed for subsampled Nanopore sequencing datasets from COLO829 and HGS-3. For each level of coverage, 20 independent subsampled datasets were generated and subjected to each filtering step in a cumulative manner.

(A) Enrichment for somatic SVs after subsequent steps of the SHARC filtering pipeline. The blue and red lines/dots indicate the percentage of somatic and non-somatic SV calls after each filtering step of the pipeline for both COLO829 and HGS-3. The percentage of somatic and non-somatic SV calls is calculated relative to the sum of remaining somatic and non-somatic SV calls after each filtering step. Thus, 100% represents the total number of SV calls (somatic plus non-somatic) present at each step.

(B) This figure panel is based on the same underlying data as for panel A, but here the percentage of somatic (blue) and non-somatic (red) SV calls is plotted relative to the total number of somatic and non-somatic SV calls detected at the first step, respectively. Thus, 100% represents the total number of non-somatic or somatic SV calls found initially in the raw data prior to filtering. While the percentage of non-somatic SV calls (red line/dots) decreases rapidly to very low percentages, the percentage of true positive somatic SV calls (blue line/dots) remains substantial (around 20%, depending on the sequence coverage). In low coverage subsets, all data points are shown and the square box represents the median value.

RF: Random forest; DBFilter: Database filter.

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing–based detection of structural variants enables
individualized circulating tumor DNA–based disease monitoring in cancer patients
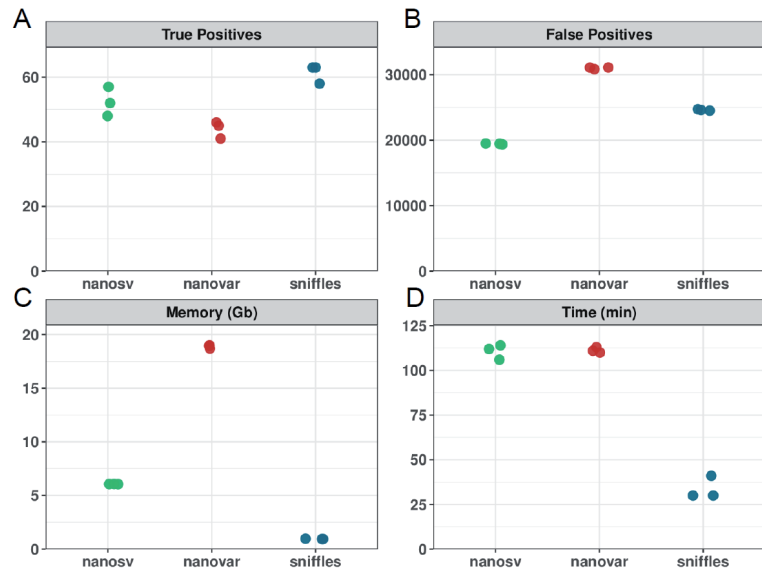
Figure S3. Benchmarking of Nanopore SV callers on low coverage Nanopore sequencing data
The SV callers NanoSV, Sniffles and NanoVar are compared in terms of true positives (A), false positives (B) and
required computation memory (C) and time (D). Triplicates of 5x randomly subsampled COLO829 data were
used, and comparisons were performed against a short-read somatic SV reference set.
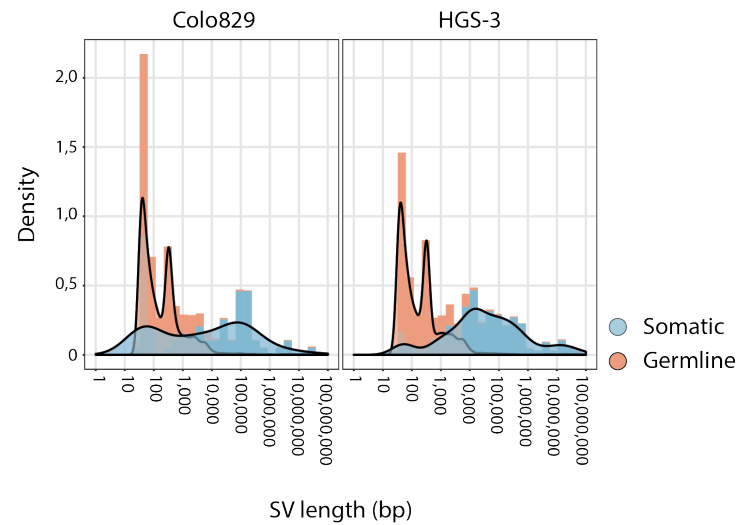
Figure S4. Somatic vs. germline SV length
Histogram and density plot of SV lengths of somatic and germline SVs from short-read data of COLO829 and
HGS-3.
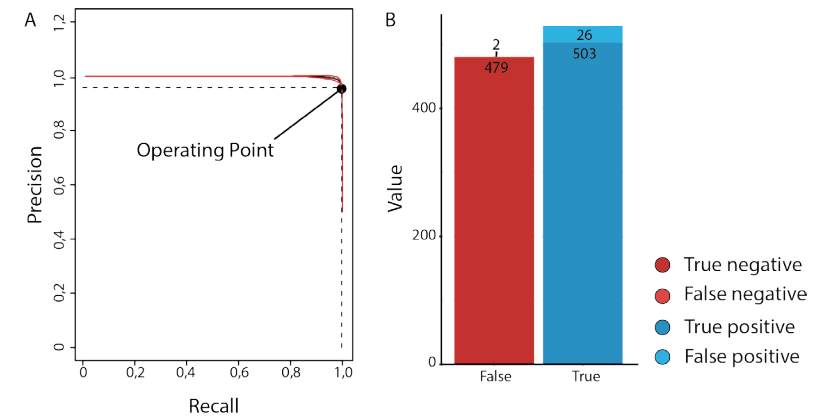
Figure S5. Random forest performance on the Genome in a Bottle sample (GIAB)
(A) Precision vs. recall curve on the training set. Depicted is the operating point selected of 96% precision and
99.5% recall.
(B) Random forest performance on the hold-out set.
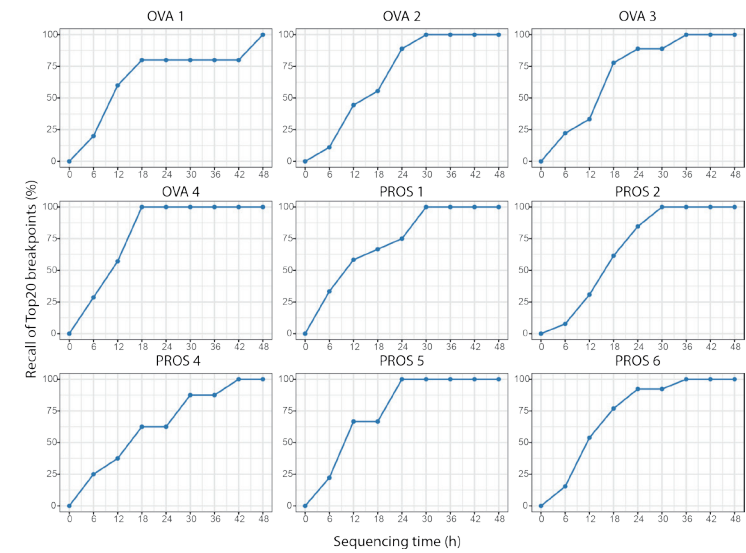
Figure S6. Nanopore sequencing time vs. somatic SV detection
Plots showing the sequencing time and the recall of validated somatic SVs in 6 hour cumulative bins.

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing-based detection of structural variants enables
individualized circulating tumor DNA-based disease monitoring in cancer patients
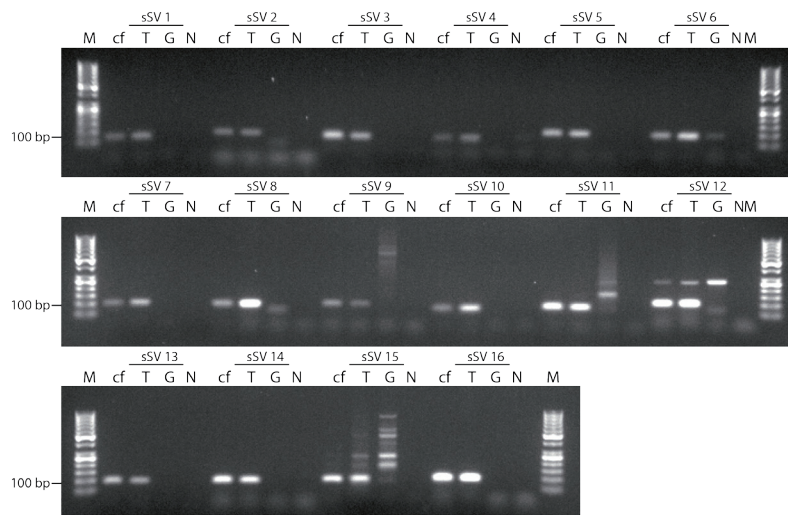
Figure S7. Validation of somatic SV of Ova2 biomarkers in cfDNA
Somatic SVs (sSV) of patient Ova2 were tested on cfDNA from ascites (cf), tumor DNA (T), germline DNA (G) and water control (N). M = DNA ladder
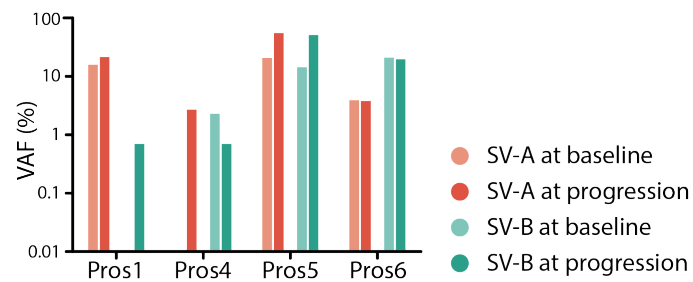
Figure S8. Confirmation of presence of SVs in preamplified cfDNA
Detection of two patient-specific SVs in cfDNA from blood from four prostate cancer patients at baseline and at progression of disease with dPCR. Shown are VAFs.
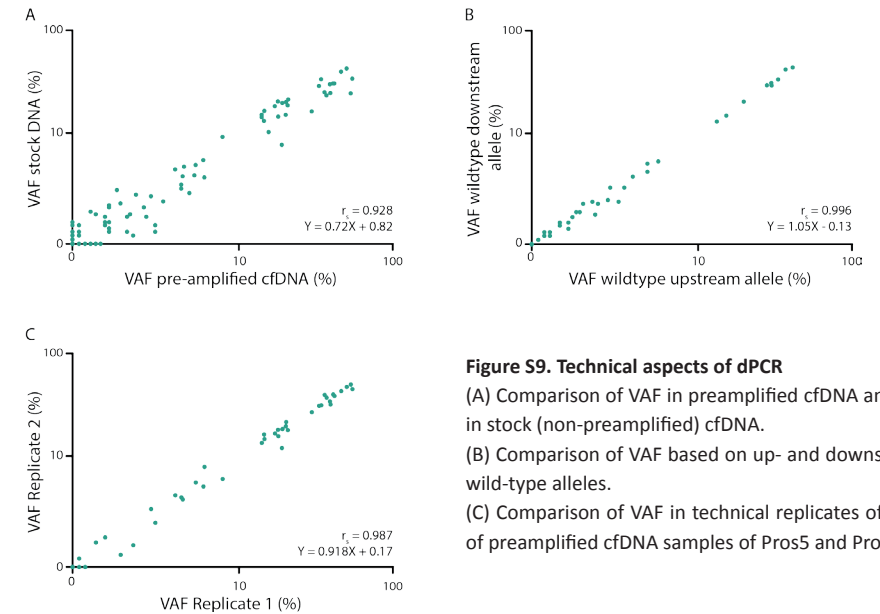
Figure S9. Technical aspects of dPCR
(A) Comparison of VAF in preamplified cfDNA and VAF in stock (non-preamplified) cfDNA.
(B) Comparison of VAF based on up- and downstream wild-type alleles.
(C) Comparison of VAF in technical replicates of dPCR of preamplified cfDNA samples of Pros5 and Pros6.
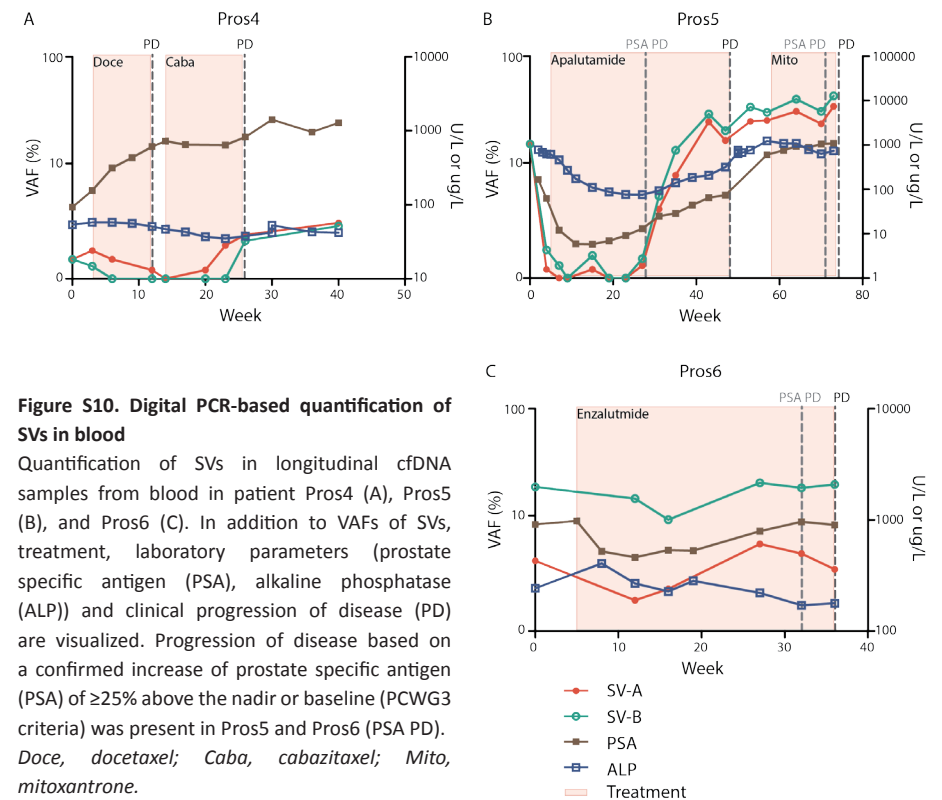
Figure S10. Digital PCR-based quantification of SVs in blood
Quantification of SVs in longitudinal cfDNA samples from blood in patient Pros4 (A), Pros5 (B), and Pros6 (C). In addition to VAFs of SVs, treatment, laboratory parameters (prostate specific antigen (PSA), alkaline phosphatase (ALP)) and clinical progression of disease (PD) are visualized. Progression of disease based on a confirmed increase of prostate specific antigen (PSA) of ≥25% above the nadir or baseline (PCWG3 criteria) was present in Pros5 and Pros6 (PSA PD).
Doce, docetaxel; Caba, cabazitaxel; Mito, mitoxantrone.

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing-based detection of structural variants enables
individualized circulating tumor DNA-based disease monitoring in cancer patients

Table S1. Sequencing statistics

Sample	Number of runs	Flow cell version	Sequencing device	DNA input library prep (ng)	DNA input flow cell (ng)	Number of reads	Total throughput (Gbp)	Mean read length (bp)	N50 read length (bp)	Average mapped coverage
Ova1	1	revD	GridION	1,620	704	590,395	4.83	8,181	14,840	1.83
Ova2	1	R9.4	MinION	1,800	977	444,225	7.25	16,317	24,178	2.46
Ova3	1	R9.4	GridION	4,800	750	756,013	7.82	10,337	25,564	2.61
Ova4	1	R9.4	MiniON	3,800	1,661	1,540,517	13.72	8,904	12,853	3.96
Pros1	1	R9.4	GridION	250	49	723,128	4.47	6,183	12,999	1.80
Pros2	1	revD	GridION	250	36	923,602	4.41	4,773	7,886	1.83
Pros3*	1	revD	GridION	250	45	209,846	1.15	5,462	10,691	0.36
Pros4	1	RP-HS**	GridION	128	6	1,232,000	6.15	4,992	7,367	2.11
Pros5	1	RP-HS**	GridION	80	10	1,472,000	6.29	4,270	6,690	2.11
Pros6	1	RP-HS**	GridION	117	15	927,313	5.56	5,993	14,347	1.76
COLO829	49	R9.4	Promethion (2), GridION (3), MinION (44)	variable	variable	18,688,674	196.09	10,492	18,687	57.19
HGS-3	40	R9.4 (23), R9.5 (17)	GridION (5), MinION (35)	variable	variable	23,464,214	193.08	8,228	10,725	54.04

Discarded (too low sequencing yield)
**Research protoype-high sensitivity

Table S2. Cumulative filtering steps to enrich for somatic SVs

This table is available from the author on request as it is to extensive to publish here.

Table S3. Random forest model

Name	Description	Gini importance
svlength	Length of the SV event	304,976939
sr_distance1	Distance to closest single repeat element in the reference genome from the left side of the breakpoint	206,414102
sr_distance2	Distance to closest single repeat element in the reference genome from the right side of the breakpoint	178,889844
vaf	Variant allele frequency	107,06722
pid2	% identity on right side of the breakpoint	26,652871
ciend2	Right confidence interval of breakpoint on the right side of the breakpoint	23,179306
pid1	% identity on left side of the breakpoint	19,069325
gap	Segment of the read that is unmapped	17,791556
plength2	Segment of the read mapped on the right side of the breakpoint	16,248287
gap_distance1	Distance to closest gap in the reference genome from the left side of the breakpoint	15,558552
plength1	Segment of the read mapped on the left side of the breakpoint	13,394601
gap_distance2	Distance to closest gap in the reference genome from the right side of the breakpoint	12,030576
sd_distance1	Distance to closest segmental duplication in the reference genome from the left side of the breakpoint	10,694879
sd_distance2	Distance to closest segmental duplication in the reference genome from the right side of the breakpoint	10,223092
total_cov_norm	Coverage around the breakpoint normalized for total coverage in the sample	10,206684
rlength	Read length	9,133996
cipos2	Right confidence interval of breakpoint on the left side of the breakpoint	8,485721
cipos1	Left confidence interval of breakpoint on the left side of the breakpoint	6,394172
ciend1	Left confidence interval of breakpoint on the right side of the breakpoint	5,883401
mapq2	Mapping quality on the rigth side of the breakpoint	3,772137
mapq1	Mapping quality on the left side of the breakpoint	3,448196

Part 2
Liquid Biopsies

Chapter 5
Optimizing Nanopore sequencing–based detection of structural variants enables
individualized circulating tumor DNA–based disease monitoring in cancer patients

Table S4. Design of digital PCR assays for absolute quantification of SVs

Assay	Type		Sequence
Pros1 SV-A (G8) Translocation Breakpoint positions: chr17: 20919555, chr8: 6557627	Structural variant	Probe	6-FAM_AAA-GAA-CAT-GGA-CAC-CA_MGB-EDQ
		SV-F	CTG-GGC-AAC-ATA-GCT-AGA-CC
		SV-R	AGT-GAT-GTG-CTT-CCC-TTT-G
	Wild-type Upstream	Probe	.
		SV-F	.
		WT-U	.
	Wild-type Downstream	Probe	Cy5_CCA-GGA-CAC-CAG-CAC_MGB-EDQ
		SV-R	AGT-GAT-GTG-CTT-CCC-TTT-G
		WT-D	GCA-CTA-GCC-AAA-CCA-CTC-A
Pros1 SV-B (C7) Duplication Breakpoint positions: chr8: 92013487, chr8: 114058951	Structural variant	Probe	6-FAM_ACA-AGG-ACA-GAG-TAT-GGA-T_MGB-EDQ
		SV-F	CTC-TCC-ACA-GCC-TCT-GC
		SV-R	CAA-TAG-ATA-CAT-GAA-CCT-AAA-CGT-G
	Wild-type Upstream	Probe	HEX_GTA-TAT-GAG-TAT-GGA-TTT_MGB-EDQ
		SV-F	CTC-TCC-ACA-GCC-TCT-GC
		WT-U	TCC-ACC-AAT-ATT-CAG-ATT-TCT-CAG
	Wild-type Downstream	Probe	Cy5_TGA-ATC-TTG-TCC-TTG-TAT-C_MGB-EDQ
		SV-R	CAA-TAG-ATA-CAT-GAA-CCT-AAA-CGT-G
		WT-D	CAA-GTC-CCT-GAC-TCC-CTT
Pros4 SV-A (C2) Duplication Breakpoint positions: chr8: 92013487, chr8: 114058951	Structural variant	Probe	6-FAM_CTT-GTA-AAC-CAT-AAA-CAA-G _MGB-EDQ
		SV-F	TCT-CAG-GTA-TGA-GCA-CCA-TC
		SV-R	CTT-CAT-TCA-CAC-AAG-CTT-CC
	Wild-type Upstream	Probe	HEX_ AGC-CTC-TTA-CCT-GGA-CAC_MGB-EDQ
		SV-F	TCT-CAG-GTA-TGA-GCA-CCA-TC
		WT-U	CAT-GGC-CCT-TAC-CAC-ATC
	Wild-type Downstream	Probe	Cy5_ AGG-CCT-GTT-TCA-CAG-AGA _MGB-EDQ
		SV-R	CTT-CAT-TCA-CAC-AAG-CTT-CC
		WT-D	GTT-ATG-TTC-TCC-ACC-CTA-CTC
Pros1 SV-B (E2) Deletion Breakpoint positions: chr4: 88945002, chr4: 89615623	Structural variant	Probe	6-FAM_AGC-CTA-GAC-GTG-TTA-TT_MGB-EDQ
		SV-F	TCT-TGC-CAA-AGA-TTG-TTG-AG
		SV-R	CAG-GGA-ATG-GCT-AGG-ATT-AG
	Wild-type Upstream	Probe	HEX_AGA-TTC-ATT-TGG-AGC-AGC-T_MGB-EDQ
		SV-F	TCT-TGC-CAA-AGA-TTG-TTG-AG
		WT-U	TAA-GCT-CTT-GGT-GAG-CAG
	Wild-type Downstream	Probe	Cy5_TAG-AGT-GTG-TCT-TTA-ACT-CT_MGB-EDQ
		SV-R	CAG-GGA-ATG-GCT-AGG-ATT-AG
		WT-D	TCC-TCT-GTT-TTT-GGT-TTT-GAA
Pros5 SV-A (A4) Duplication Breakpoint positions: chr4: 68016545, chr4: 105752220	Structural variant	Probe	6-FAM_TTC-ACA-ACT-GCT-GAG-CCA_MGB-EDQ
		SV-F	CAT-CTT-GCC-AAT-GAA-CAG-AG
		SV-R	TTT-CAC-CAT-TAC-AGT-ACA-GTG-C
	Wild-type Upstream	Probe	HEX_TGG-CAT-GTT-ATG-GCT-TG_MGB-EDQ
		SV-F	CAT-CTT-GCC-AAT-GAA-CAG-AG
		WT-U	CTT-GAC-CCT-GCC-TTC-AG
	Wild-type Downstream	Probe	.
		SV-R	.
		WT-D	.
Pros5 SV-B (C4) Inversion Breakpoint positions: chr6: 115647533, chr6: 134275493	Structural variant	Probe	6-FAM_CAG-TCA-AAC-TGG-GAA-AA_MGB-EDQ
		SV-F	CAT-CTT-GCC-AAT-GAA-CAG-AG
		SV-R	TTT-CAC-CAT-TAC-AGT-ACA-GTG-C
	Wild-type Upstream	Probe	HEX_AGC-CAC-CTT-TGA-CTG-AT_MGB-EDQ
		SV-F	CAT-CTT-GCC-AAT-GAA-CAG-AG
		WT-U	GAT-CAT-GCT-CTG-AAA-TTC-AAG-GTT-T
	Wild-type Downstream	Probe	Cy5_AGA-ATC-CAC-ATT-GCT-TAC-AC_MGB-EDQ
		SV-R	TTT-CAC-CAT-TAC-AGT-ACA-GTG-C
		WT-D	GGG-TTT-AGA-ACA-GTG-CCT-TA
Pros6 SV-A (A6) Inversion Breakpoint positions: chr14: 19788077, chr14: 55164423 Remark: Wild-type Upstream has an identical sequence on chromosome 22.	Structural variant	Probe	6-FAM_AGC-GTC-GGT-GCT-GGA_MGB-EDQ
		SV-F	CTC-CCT-CAA-ACT-GGG-AAT-G
		SV-R	TTT-ACC-TAG-AGG-GAC-AGT-TGG
	Wild-type Upstream	Probe	HEX_ACA-ACC-AGC-CTC-ATA-C_MGB-EDQ
		SV-F	CTC-CCT-CAA-ACT-GGG-AAT-G
		WT-U	CCT-TTG-CTT-CCC-AAG-ACA
	Wild-type Downstream	Probe	Cy5_ACA-TGT-TAC-ACA-GTT-TCT-GT_MGB-EDQ
		SV-R	TTT-ACC-TAG-AGG-GAC-AGT-TGG
		WT-D	CCT-AAG-CCA-GAA-TCT-CTC-ATA-TTT
Pros6 SV-B (E6) Translocation Breakpoint positions: chr14: 121469145, chr8: 127512920	Structural variant	Probe	6-FAM_TGA-ACT-GAA-TTT-TTA-TAT-TTA-TTC-AC_MGB-EDQ
		SV-F	AGC-CTA-CAA-ACC-ACT-TAT-TGC
		SV-R	TGC-AGT-GGA-CAT-TGG-ATG
	Wild-type Upstream	Probe	.
		SV-F	.
		WT-U	.
	Wild-type Downstream	Probe	Cy5_AGC-ACA-CAA-CAA-ATC_MGB-EDQ
		SV-R	TGC-AGT-GGA-CAT-TGG-ATG
		WT-D	GAA-CAC-AGA-GGC-AGG-AAA-T

SV-F Structural Variant Forward primer
SV-R Structural Variant Reverse primer
PB probe
WT-U Wild-type Upsteam primer
WT-D Wild-type Downstream primer
All sequences depicted in the 5'-3' read direction

Chapter 6

Generating human prostate cancer
organoids from leukapheresis enriched
circulating tumor cells

Chapter 6

Lisanne Mout*, Lisanne F. van Dessel*, Jaco Kraan, Anouk C. de Jong, Rui P.L. Neves, Sigrun Erkens-Schulze,
Corine M. Beaufort, Anieta M. Sieuwerts, Job van Riet, Thomas L.C. Woo, Ronald de Wit, Stefan Sleijfer, Paul
Hamberg, Yorick Sandberg, Peter A.W. te Boekhorst, Harmen J.G. van de Werken, John W.M. Martens, Nikolas
H. Stoecklein, Wytske M. van Weerden, Martijn P. Lolkema
* These authors contributed equally

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
circulating tumor cells

Abstract

Background
Circulating tumor cell (CTC) derived organoids provide a powerful tool for personalized cancer therapy, but are restrained by low CTC numbers provided by blood samples. Here, we used diagnostic leukapheresis (DLA) to enrich CTCs from metastatic prostate cancer (mPCa) patients and explored whether organoids provide a platform for *ex vivo* treatment modelling.

Materials and Methods
We prospectively screened 102 mPCa patients and performed DLA in 40 patients with ≥5 CTCs/7.5 mL blood. We enriched CTCs from DLA using white blood cell (WBC) depletion alone or combined with EpCAM selection. The enriched CTC samples were cultured in 3D to obtain organoids and used for downstream analyses.

Results
The DLA procedure resulted in a median yield of 5312 CTCs as compared to only 22 CTCs in 7.5 mL of blood. Using WBC depletion, we recovered 46% of the CTCs, which reduced to 12% with subsequent EpCAM selection. From the isolated and enriched CTC samples, organoid expansion succeeded in 35% of the patient samples. Successful organoid cultures contained significantly higher CTC numbers at initiation. Moreover, we performed treatment modelling in one organoid cell line and investigated tumor heterogeneity in CTCs using single cell DNA sequencing.

Conclusions
DLA is an efficient method to enrich CTCs. Although the modest success rate of culturing these cells precludes large scale clinical application, these data do suggest that DLA and subsequent processing provides a rich source of viable tumor cells. Therefore, DLA provides a promising alternative to biopsy procedures to obtain sufficient number of tumor cells to study disease kinetics in mPCa patients.

Introduction

Translational studies on advanced prostate cancer (PCa) have often been limited to static sources, such as resection of the primary tumor or biopsies[1-4]. However, cancer is an ever moving target, as dynamic evolution drives spatial and temporal heterogeneity allowing tumors to adapt and escape therapeutic interventions. We therefore require new methods that provide real-time insights into evolving cancer biology for treatment tailoring. Circulating tumor cells (CTCs) could serve as a dynamic tumor source which captures the genetic and phenotypic heterogeneity of cancer and can be obtained at multiple time-points during the disease course to assess clinical progression[5]. Additionally, CTCs can be obtained from peripheral blood in a relatively non-invasive manner, thus providing an easily accessible source of metastatic cells as an alternative to tumor biopsies. This is particularly beneficial for patients with primarily bone metastasis such as in metastatic PCa (mPCa). In mPCa, CTC numbers have already been shown to harbor significant prognostic and predictive value[6-10]. The possibility to obtain viable CTCs also provides the opportunity to propagate CTCs *ex vivo.* The development of cancer organoids has allowed us to better capture the tumor-specific characteristics than standard 2D culture methods[11]. Together, this provides the opportunity to use CTC derived organoids as a representative model of the current disease status and use for drug discovery and sensitivity-screening[12]. PCa CTCs can be cultured as organoids, although the reports thus far suggest a very low efficiency rate for success[13,14]. Previous reports suggest that one important obstacle is the high number of CTCs needed to initiate organoid propagation[13,15]. Since the median CTC count in mPCa patients is 2-20 CTCs per 7.5 mL of blood[10], *ex vivo* expansion of CTCs is unlikely to be successful in most patients. Diagnostic leukapheresis (DLA) is a standardized procedure to enrich for mononuclear cells by continuous centrifugation of blood. Since CTCs have a similar density to mononuclear cells, they are enriched as well[16]. Importantly, DLA is a minimally-invasive and safe procedure, that is well tolerated by patients[17]. Within this prospective study we set out to isolate CTCs from mPCa patients by DLA. We hypothesized that the increased number of CTCs obtained by DLA, will allow us to culture CTC derived organoids and potentially provide a platform for individualized disease modelling. Using optimized methods we validated DLA as a feasible and safe method to enrich for CTCs in mPCa patients. CTCs could be propagated as short-term organoid cultures in 35% of the samples, from which we could obtain one stable organoid cell line. These short-term organoid cultures expressed the classical markers of PCa and maintained genomic variants previously identified in metastatic samples. Overall, our study provides an important step forward in implementing CTCs in individualized disease modelling, nevertheless identifies several challenges that require further optimization.

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
circulating tumor cells

Materials and Methods

Study design

This prospective study was conducted in accordance with the Declaration of Helsinki and was approved by the medical ethical committee of the Erasmus Medical Center Rotterdam (EMC-16-449). The study was performed using a Simon's two-stage design[18] with an interim analysis after 18 DLAs were performed. Continuation to the second stage could only continue if seven out of 18 samples yielded successful organoid cultures meaning that organoids could be propagated for six weeks or longer and PCa origin could be validated. The study protocol, including sample size calculations, has been added as supplementary information.

Patients

From November 2016 till July 2020, patients with mPCa were prospectively included (NL6019; https://www.trialregister.nl/trial/6019). All patients provided written informed consent before any study procedure took place. Patients ≥ 18 years in age were eligible if they had histologically or cytologically confirmed PCa, with measurable metastatic lesion(s) (according to PCWG2 and/or RECIST 1.1 criteria)[19,20] and intended to start a new line of systemic treatment. Patients should have at least two adequate peripheral veins as access point for DLA. Additionally, patients should be in good physical condition (WHO performance status ≤ 2), have an adequate hematology, coagulation status and liver/renal function as assessed by routine laboratory tests (see study protocol for details). Exclusion criteria were CTC count of < 5 per 7.5 mL of blood at screening, current androgen deprivation therapy treatment for hormone-sensitive mPCa patients, known hypersensitivity to anticoagulant used during DLA, hemorrhage disease or coagulation disorders, and chronic viral infections. Patients were allowed to enter the study multiple times during their treatment course. All patients provided written informed consent before any study procedure. The study procedures included a blood draw at baseline to screen for eligibility criteria, a blood draw before start of the DLA procedure and the DLA procedure itself. Clinical data on disease characteristics, previous therapies and response on the subsequent systemic therapy was collected in an electronic case report form (ALEA Clinical).

Primary and secondary endpoints

The primary endpoint of this study was the rate of successfully cultured organoids from CTCs obtained by DLA with successfully being defined as i) propagation for at least 6 weeks in culture and ii) proven PCa characteristics. PCa characteristics include epithelial origin CK8/18, lack of p63 (not basal cells), and (combined) expression of *TMPRSS2-ERG* fusion, AR (splice variants) expression, PTEN loss, MYC amplification (together indicative of PCa) or the detection of patient-specific somatic variants (known from previously characterized tumor material).

Study procedures

The study procedures included a blood draw at baseline to screen for eligibility criteria, including CTC count, a blood draw before start of the DLA procedure and the DLA procedure itself. All study procedures were performed before start of a new line of treatment.

Circulating tumor cell enumeration using CellSearch

For CTC enumeration, 7.5 mL of blood, 1 mL of DLA product diluted in 6.5 mL of PBS (Thermo Fisher Scientific, Waltham, MA), or a proportion of the enriched fraction was collected in a CellSave Preservative tube (Menarini Sillicon Biosystems, Castel Maggiore, BO, Italy). All samples were processed within 96 hours and CTC enumeration was performed using the Circulating Epithelial Cell Kit on the CellSearch® system (Menarini Sillicon Biosystems) by certified personnel. Briefly, samples were subjected to immunomagnetic capture using ferrofluids coupled to anti-EpCAM antibodies and enriched cells were stained with antibodies specific for cytokeratin (CK) 8, 18, and 19, CD45 and nucleic acid dye (DAPI). Images were captured using the CellTracks Analyzer II (Menarini Silicon Biosystems) and manually examined to determine the presence of CTCs. Cells were defined as CTCs when positive for CK and DAPI and negative for CD45 according to the manufacturer's instructions. CellSearch cartridges were stored in the dark at 4°C before further analyses.

Diagnostic leukapheresis

DLA was performed at the department of Hematology at the Erasmus Medical Center using the Spectra Optia Cell Separator machine (Terumo BCT, Lakewood, CO). Only peripheral venous access was used to process a maximum volume of 10 L of circulating blood. Citrate dextrose solution A was used as anticoagulant. For CTC isolation the standard settings for white blood cell (WBC) isolation were applied, only the plasma pump rate was increased to collect a slightly higher cell density with an approximate hematocrit of 5% instead of 2%. In addition, extra blood was drawn for the collection of red blood cells (40 ml in EDTA tubes), CTC count of peripheral blood (20 ml in CellSave tubes) and routine hematological tests. Fresh DLA product was immediately processed for CTC enrichment. Excess DLA product was stored in liquid nitrogen at 2 mL aliquots containing 50% autologous plasma and 10% DMSO (Sigma-Aldrich, Saint Louis, MO). All grade ≥ 3 (serious) adverse events during and within 48 hours after DLA were registered.

Circulating tumor cell enrichment

CTCs from DLA were enriched by negative depletion of WBC using RosetteSep™ CTC Human CD45 Depletion Cocktail (STEMCELL Technologies, Vancouver, BC, Canada) with the following modifications. WBC concentration of the DLA product was measured using a hematology analyzer (Beckman Coulter, Brea, CA) and diluted with phosphate buffered saline (PBS) to a final concentration of $50x10^6$ WBC/mL. Red blood cells (RBCs) were collected in K2EDTA vacutainer® tubes (BD, Franklin Lakes, NJ) before start of the DLA procedure and concentrated by centrifugation

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
circulating tumor cells

at 800g for 8 minutes. These autologous RBCs were added to the DLA product to achieve a 1:50 WBC:RBC ratio before incubation with the RosetteSep™ Cocktail to induce crosslinking of RBCs with WBCs. RBC/WBC rosettes were depleted using Sepmate™-50 tubes and Lymphoprep™ as density gradient medium (both STEMCELL Technologies), by centrifugation at 1200g for 12 minutes at room temperature with the brake off. Enriched CTCs were washed using PBS. For 15 DLA samples additional manual positive immunomagnetic enrichment was performed. Ten mL of CD45 depleted fraction was incubated with 150 µL CellSearch EpCAM ferrofluids and 150 µL capture enhancement reagent for 10 minutes, at room temperature on a roller mixer. To separate the magnetic labeled CTCs, the tube was placed next to a MagCellect magnet (R&D systems, Minneapolis, MN) and the supernatant aspirated. A second magnetic separation was performed for optimized enrichment. Finally, the enriched CTCs were collected in culture medium and used for organoid culture.

Organoid culture from circulating tumor cells

The enriched CTC suspension was collected by centrifugation at 1200g for 8 minutes. The cell pellet was cooled on ice and mixed with Matrigel (MG; Corning, cat. no. 356231, Corning, New York). The cell suspension-MG mix was dispensed in a pre-warmed 24-wells plate (Corning, cat. no. 3527), in 30-40 µL droplets. Subsequently, the plate was placed upside-down at 37˚C and incubated for 15 minutes to let the MG solidify. Next, 500 µL of medium was added; CTCs were cultured in prostate growth medium (PGM) and adjusted prostate cancer organoid medium (APCOM) in parallel (overview in Table S1)[21]. APCOM was based on the previously published organoid culture media described by Gao et al. and Beshiri et al.[13,22]. Weekly images were obtained to monitor organoid growth using the Zeiss Axiovert 25 equipped with 10x and 20x Plan-Neofluar objectives, a AxioCam ICc1 camera and AxioVision imaging software (version 4.8.2.0, Oberkochen, Germany) and Nikon Eclipse TS2 equipped with 4x, 10x CFI Achro brightfield objectives and a 20x Fluor ELWD objective, a DS-Fi3 camera and NIS-Elements imaging software (Minato, Tokyo, Japan). Cell culture media was either replaced or supplemented every 3-4 days.

DNA/RNA isolation and PCR analyses of CTC derived organoids

For DNA and RNA isolation, organoids were collected, lysed with RLT buffer (Qiagen, cat. no. 79216, Hilden, NRW, Germany) and stored at -80°C until further processing. DNA and RNA isolation using the QIAcube system (Qiagen, cat. no. 9001293) in combination with the AllPrep DNA/RNA Micro Kit (Qiagen, cat. no. 80284,) was performed as described previously[23]. cDNA was generated from RNA using the RevertAid H Minus First Strand cDNA Synthesis Kit according to the manufacturer's protocol (Thermo Fisher Scientific, cat. no. K1632). Subsequently, cDNA (0.1 to 1 ng/µL) was pre-amplified for patient-specific targets and/or a multiplex prostate gene expression panel[23] with a Taqman assay covering wild-type and mutant molecules during 15 cycles using TaqMan PreAmp Master Mix (Thermo Fisher Scientific) as recommended by the manufacturer. Prior to downstream processing, the pre-amplified product was diluted 15-fold in LoTE buffer

(3 mM Tris-HCl/0.2 mM EDTA, pH 8.0). Gene expression levels were measured in real-time in the preamplified samples by qRT-PCR and subsequently analyzed as described previously[23]. For validation of single nucleotide variants (SNVs) in organoids (Table S2), digital PCR (dPCR) reactions were performed with either the QuantStudio 3D Digital PCR System (Thermo Fisher Scientific) or the Naica™ Crystal Digital PCR System (Stilla Technologies, Villejuif, France) according to the manufacturer's protocol. For the former, each pre-amplified cDNA sample or DNA sample was partitioned into 20,000 wells of a QuantStudio 3D Digital PCR v2 Chip and run on a ProFlex 2x Flat PCR System (Thermo Fisher Scientific). The target-specific optimized PCR program was: 10 min at 96°C, followed by 40 cycles of 30 sec incubation at 98°C and 2 min at 55°C and a final pause for up to 16 hours at 10°C. Chips were read in a QuantStudio 3D Digital PCR Instrument and analyzed with the web-based Quantstudio 3D dPCR Analysis Software version 3.01 (Thermo Fisher Scientific). For the latter, each pre-amplified cDNA sample or DNA sample was partitioned into 30,000 crystal droplets and amplified by the following program: 45 cycles of 30 sec at 95°C and 15 sec at 58°C. Chips were read in a Naica Prism3 instrument and analyzed with the Crystal Miner software (Stilla Technologies). For both, at least one positive and one negative control sample was included in every run.

Results

Patient characteristics and CTC screening

We screened 102 mPCa patients for eligibility and selected 45 patients who had ≥ 5 CTC in 7.5 mL of peripheral blood (PB) for CTC collection by DLA (Figure 1A). Metastatic castration-resistant PCa (mCRPC) patients screened for study participation tended to have a higher CTC burden compared to the metastatic hormone-sensitive PCa (mHSPC) patients (median of 5 CTC/7.5 mL vs. 1 resp. P = 0.07, Figure S1A). Moreover, 50% of the mCRPC patients had 5 or more CTCs while 39% of the mHSPC patients were included. The overall median CTC count of the included patient population was 22 CTCs per 7.5 mL PB prior to DLA, with 10 CTCs/7.5 mL and 26 CTCs/7.5 mL for mHSPC and mCRPC patients resp. (Figure 1B). The CTC burden was monitored in 24 patients and was found to remain stable in between screening, prior to the DLA and after completion of the DLA procedure (P = 0.37, Figure S1B).

Diagnostic leukapheresis

After initial screening, a total of 40 patients successfully underwent DLA, as four patients withdrew and the DLA procedure was terminated in one patient because of an adverse event (grade 3 vasovagal reaction directly after start of the procedure). No other grade ≥ 3 adverse events were observed. The baseline characteristics of the patient population and DLA procedure are described in Table 1. Four patients were included twice at separate time points during their treatment course, resulting in serial samples for subjects 9/19, 11/15, 24/34 and 22/74. We

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
circulating tumor cells

examined the impact of DLA density settings on CTC enrichment in four patients, by performing the DLA procedure in two fractions of both 5 L blood at 2% and 5% hematocrit respectively. As there was a trend towards higher CTC yield when DLA was performed at 5% hematocrit (Figure S1C, P = 0.125), we continued the DLA procedure using high DLA density settings and processed 5 L of total blood volume, which limited the procedure time to approximately 2 hours. The median CTC concentration of the DLA product was 64/mL compared to 2.5/mL in PB samples pre-DLA (P < 0.001), with an estimated median yield of 5312 CTCs in 96 mL of DLA product (Figure 1C). Using these DLA settings, we were able to retrieve a median of 36% of the estimated CTCs available given the processed blood volume and the CTC count in PB (Figure S1D).
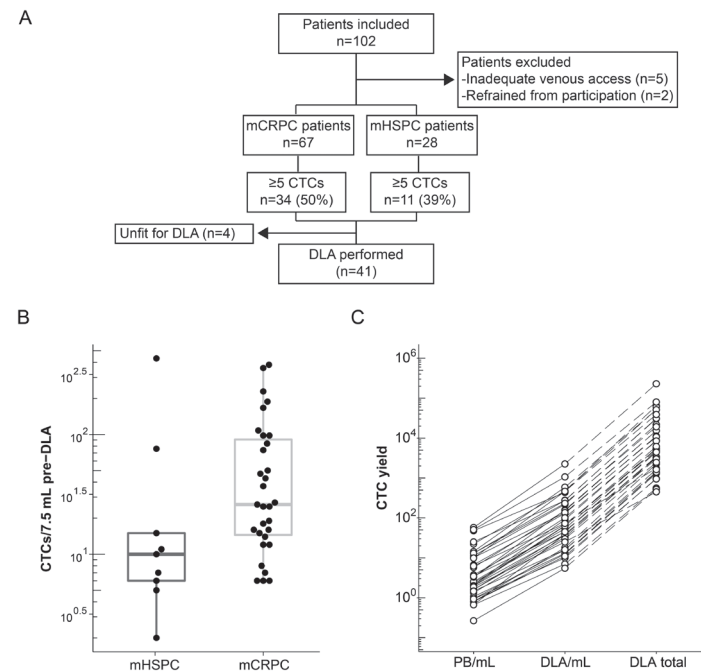
Figure 1. Overview of patient inclusion, diagnostic leukapheresis procedure and subsequent circulating tumor cell enrichment methods

(A) Patients were eligible for diagnostic leukapheresis (DLA) if they had adequate venous access and ≥ 5 circulating tumor cells (CTCs). Two patients refrained from participation because of urge-incontinence and because of pain. Five patients were not screened for the presence of CTCs in peripheral blood due to inadequate venous access. After screening and inclusion four patients refrained from DLA because of progressive disease.
(B) CTC count in 7.5 mL peripheral blood (PB) obtained pre-DLA in patients where DLA was successfully performed. CTC count in the metastatic castration-resistant (mCRPC, n = 31) and hormone-sensitive prostate cancer (mHSPC, n = 9) patients is shown.
(C) Absolute CTC count in PB and DLA product. The X-axis shows CTC count per mL PB, per mL DLA product and extrapolated to total DLA volume. Each dot represents an individual subject (n = 40) and for all subjects we show the results of the 5% RBC density setting in 5 L processed blood volume. Samples from the same subject are connected by a (dashed) line. Y-axis is a logarithmic scale. Statistical comparison of CTC yield per mL DLA versus PB was performed by a paired two-sided Wilcoxon rank test P < 0.0001.

Table 1. Baseline characteristics of subjects who underwent DLA
Baseline characteristics are shown for patients who underwent the diagnostic leukapheresis (DLA) procedure. In addition, we show the duration, processed blood volume, and collected volume of the DLA procedure.

|	N = 37*
Patient and tumor characteristics	
Age, years (median, range)	70 (49-83)
WHO status at registration	
0	12
1	23
2	2
Hormone status at time of inclusion	
HSPC	12
CRPC	25
Gleason score at diagnosis	
≤ 6	2
7	8
8	9
9-10	13
Missing	5
M-stage at diagnosis	
M0	5
M1	21
Mx	9
Missing	2
Type of prior therapy	
Local therapy (i.e. radical prostatectomy or RT on prostate)	16
ADT (i.e. chemical or surgical)	25
Chemotherapy	17
Hormonal therapy (other than ADT)	16
Radionucleotide therapy	8
Other	5
Diagnostic leukapheresis characteristics	
Total duration of DLA (minutes; median, range)	104 (25-925)
Total processed blood volume (mL; median, range)	5112 (1153-10001)
Volume of collected DLA product (mL; median, range)	96 (18-178)
Missing	3

*4 subjects underwent DLA twice
HSPC, hormone-sensitive prostate cancer; CRPC, castration-resistant prostate cancer; ADT, androgen deprivation therapy

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
circulating tumor cells

### CTC enrichment from DLA material

We compared two methods to enrich and isolate CTCs from the DLA product: depletion of WBCs using the RosetteSep™ method, with or without subsequent positive selection using EpCAM directed antibodies. For the RosetteSep™ method to function, WBCs are cross-linked with RBCs leading to erythrocyte rosetting of WBC allowing for gradient separation. Since the DLA product contains a high WBC and relative low RBC concentration, we needed to add RBCs to achieve an optimal WBC:RBC ratio for depletion of WBCs. Using the autologous RBCs from 40 mL of PB, we were able to effectively process a median of 25.5 mL out of the 96 mL DLA obtained. We investigated the impact of WBC depletion alone and subsequent EpCAM enrichment in twelve paired samples (Figure S2). WBC depletion alone was found to reduce the WBC concentration from $93.6*10^6$/mL to $0.12*10^6$/mL, resulting in a WBC depletion factor of 3.1 $\log_{10}$-fold. With subsequent EpCAM enrichment of CTCs, we further reduced the WBC concentration to 4.1 $\log_{10}$-fold. However, this was at expense of a substantial CTCs loss, as the median CTC recovery reduced from 54% with WBC depletion alone to 11.5% with additional EpCAM selection ($P < 0.001$). We therefore chose to use WBC depletion alone for the majority of the remaining samples and only applied additional EpCAM selection if WBC depletion insufficiently enriched the sample. In the overall population, we reduced the WBC concentration by 3.21 $\log_{10}$-fold and recovered 46% from the CTCs with WBC depletion alone (Figure 2).
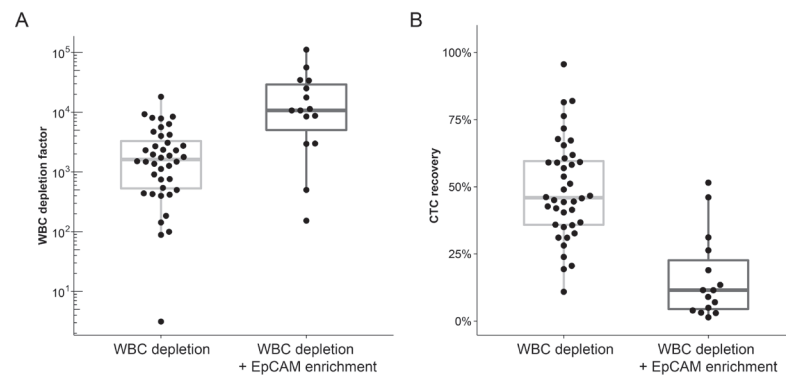
Figure 2. Efficiency of CTC enrichment and isolation techniques for diagnostic leukapheresis
(A-B) Two CTC enrichment methods were compared (depicted on the X-axis): (1) white blood cell (WBC) depletion (n = 40) and (2) WBC depletion followed by EpCAM enrichment (n = 15). The boxplot depicts the median, upper and lower quartiles, whiskers indicate 1.5 times the interquartile range (IQR). Individual data points are shown.
(A) WBC depletion factor after CTC enrichment. To calculate the WBC depletion factor, the number of WBCs before enrichment was divided by the number of WBCs after enrichment. WBC concentration was measured by a hematology analyzer. Y-axis is a logarithmic scale.
(B) Relative CTC recovery (%) after CTC enrichment. To calculate CTC recovery, the absolute CTC count after the enrichment was divided by the absolute CTC count before the enrichment. Absolute CTC counts were extrapolated from 1 mL samples.

CTC derived organoids

After 18 DLA samples we performed an interim analysis to assess the success rate of organoid cultures from isolated and enriched CTC samples. In nine out of eighteen samples we obtained organoid samples of which seven could be maintained for at least six weeks and thus we continued with the second stage of our prospective study. In total, we established CTC derived organoids in 14 out of 40 DLA samples (35% success rate, Figure S3). Both the absolute CTC yield after enrichment and the tumor cell percentage in culture was found to be higher in samples that resulted in organoid propagation *ex vivo* ($P < 0.001$ and $P < 0.01$ resp, Figure 3A and B). Moreover, only one out of nine mHSPC samples, with the highest CTC yield within the population (subject 38; 64,155), could be propagated as organoid. The majority of the organoid cultures could be maintained for six to eight weeks until proliferation stalled, thus providing limited number of organoids for downstream applications (Figure 3C). Two organoid cultures (EMC-PCa-25 and EMC-PCa-41) could be expanded and maintained for over six months, the latter yielding a stable cell line. Validation of the PCa origin of the organoids was shown by quantitative real-time PCR (qRT-PCR) analysis of prostate (cancer) specific transcripts (Table 2, Figure S4). The vast majority of the isolated samples were positive for *AR* and/or *KLK3* (PSA) while expression of *AR-V7* was identified in only one sample. Three out of 14 organoid cultures (Subject 16, 41 and 93) expressed the *TMPRSS2-ERG* fusion transcript. We performed additional validation of PCa origin using patient-specific somatic single nucleotide variants (SNVs) previously identified in metastatic biopsies (Table 2, Table S2 and 3)[4]. We validated PCa origin in the matched organoid cultures from subject 9/19, 79 and 24/34 by detection of the *TP53* and *PTEN* SNVs resp. Subject 79 was a patient with neuro-endocrine PCa (NEPC), of whom the CTC derived organoids maintained their NEPC features as they lacked *AR* and *KLK3* expression.
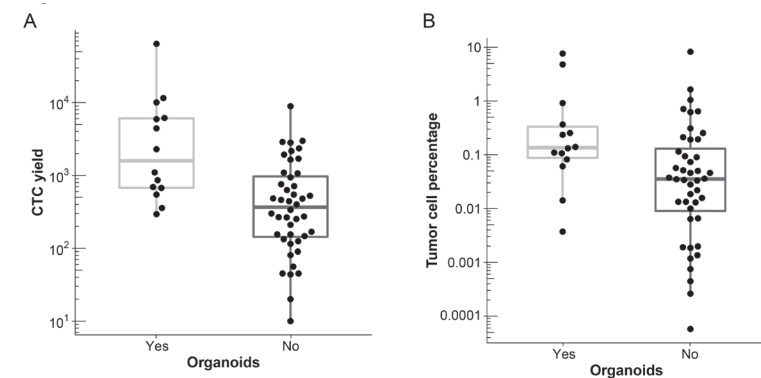
Figure 3. Overview of successful CTC derived organoid cultures
(A-B) Results of CTC enrichment and isolation of samples that generated organoids (n = 14) and those that did not (n = 44). From 11 DLA products we cultured two samples, as white blood cell (WBC) depletion alone and subsequent EpCAM enrichment was performed.
(A) Estimated CTC yield after CTC enrichment and isolation as determined by CTC count from 1 mL sample and extrapolated to the entire product after processing. Y-axis is a logarithmic scale.
(B) Tumor cell percentage in sample after CTC enrichment and isolation. Percentage of tumor cells was calculated by dividing CTC count by WBC count in sample after enrichment and isolation. Y-axis is a logarithmic scale.

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
circulating tumor cells

>>
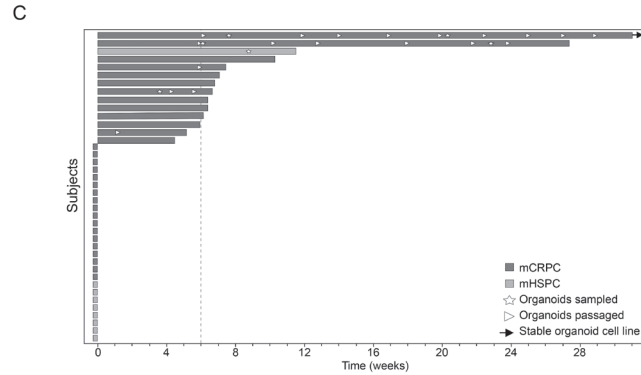
(C) Swimmersplot of all patient samples used to generated CTC derived organoids (marked by ticks), for confirmed organoid samples the time in culture is shown. Most organoid cultures could be maintained as short-term cultures and were subsequently isolated for genetic and/or transcriptional analysis. Organoid sampling is marked by a star and passaging of organoids is marked by arrow heads. Light grey bars represent CTC samples from metastatic hormone-sensitive prostate cancer (mHSPC) patients while dark grey is used for metastatic castration-resistant prostate cancer (mCRPC). Subject 41 was deemed a stable organoid cell line (EMC-PCa-41) after 10 passages. X-axis depicts the time in weeks since initiating organoid culture.

Table 2. Validation of prostate (cancer) transcripts and patient-specific somatic variants in CTC derived organoids
Overview of genomic and transcriptomic characteristics of the isolated organoids. Expression of prostate (cancer) transcripts in the organoid samples was acquired by qRT-PCR. Positive expression was determined by a ΔΔCq above -8.5 (normalized to EpCAM/KRT19 and VCaP RNA used as calibrator). Patient-specific somatic single nucleotide variants (SNVs) were selected using whole-genome sequencin data from metastatic biopsies (CPCT-02 study) and validated using dPCR. Shown are the SNVs and variant allele frequency as defined by dPCR in the organoid samples.

Subject ID	Expression prostate (cancer) transcripts	Somatic variant (variant allele frequency)
9	AR	TP53 c.407A>T (99.7%)
16	AR, KLK3, and TMPRSS2-ERG	NT
19	NT	TP53 c.407A>T (83.9-99.9%)
21	AR and KLK3	NT
24	KLK3	PIK3CA c.3140A>G (1.9%)
25	AR and KLK3	PLCG2 c.655G>A (27.5%)
34	AR and KLK3	PIK3CA c.3140A>G (45%)
38	AR and KLK3	NT
41	AR, KLK3, and TMPRSS2-ERG	NT
79	–	TP53 c.733G>A (97%)
91	AR and KLK3	NT
93	AR, KLK3, and TMPRSS2-ERG	NT
94	AR and KLK3	NT
97	AR, KLK3, and AR-V7	NT

NT, not tested

Characterization of CTC derived organoid cell line

From the organoid culture samples we were able to generate one stable organoid cell line, which enabled us to perform in-depth genomic and phenotypical characterization. Whole-genome sequencing (WGS) of EMC-PCa-41 revealed a triploid genome with an estimated tumor cell purity of 99%, an overall tumor mutational burden of 2.13 somatic mutations per mega base pairs and no predominant mutational signature (Figure 4A). We identified multiple copy-number alterations (CNAs) including a focal amplification on chromosome Xq, encompassing the *AR* locus and a focal deep deletion on 10q causing loss of *PTEN*. Furthermore, EMC-PCa-41 was characterized by multiple inter- and intra-chromosomal rearrangements, including the interstitial deletion leading to the *TMPRSS2-ERG* fusion as was identified by qRT-PCR (Table 2, Figure S3B). Moreover, ERG expression in the organoids was validated by immunohistochemistry (Figure 4B). Overall, EMC-PCa-41 harbors genomic features which are frequently identified in mCRPC tumors[4]. Next, we determined the sensitivity of EMC-PCa-41 to commonly used treatments for mCRPC; enzalutamide and taxane chemotherapeutics (Figure 4C and 4D). Both androgen depleted culture conditions (minus R1881) and enzalutamide treatment could only partially inhibit cell proliferation of EMC-PCa-41, suggesting resistance. Interestingly, subject 41 started with enzalutamide after the DLA procedure and switched treatment after only two months due to rising PSA levels (Figure S5).

>>

>>
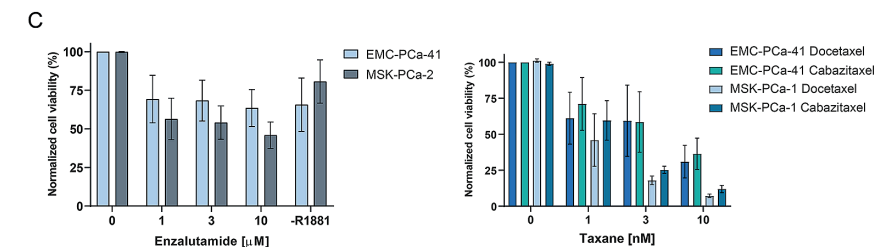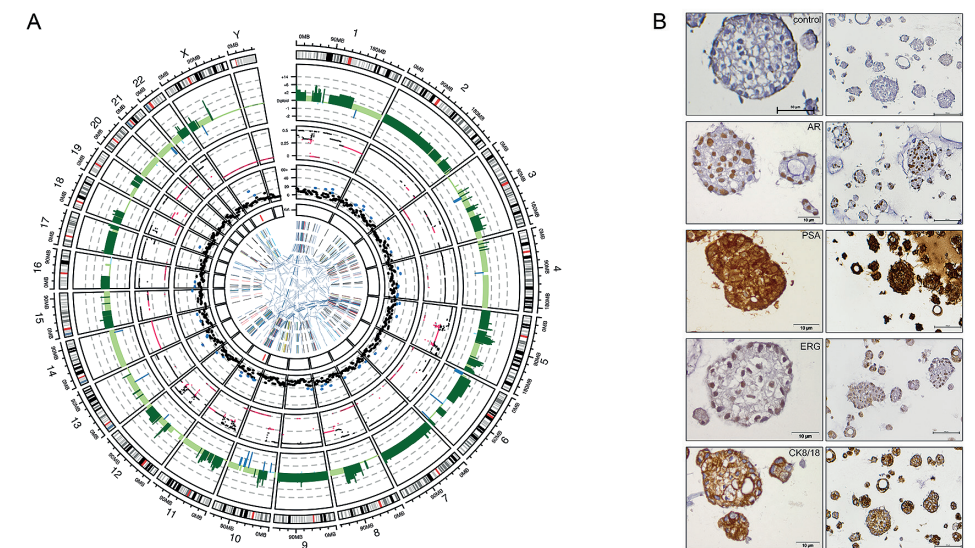
Figure 4. Genetic and phenotypic characterization of the CTC derived stable organoid cell line, EMC-PCa-41

(A) CIRCOS plot representing the whole genome characterization of EMC-PCa-41 as obtained by whole genome sequencing. The first outer track depicts the genomic ideogram. The second track displays the copy-number profile with amplifications marked in light-green, deep amplification in dark-green, deletions in blue and deep deletions in dark-blue. The third track depicts the minor-allele frequency (MAF) values of individual copy-number segments (MAF values ≤ 0.33 in pink and MAF values > 0.33 in black). The fourth track displays the number of mutations per 5 megabase pairs (Mb), with regions of mutational frequency above 20 Mb marked in blue. The fifth track highlight regions marked by regional hypermutation (kataegis). Inner circle displays structural variants, with deletions in black, translocations in dark blue, insertions in yellow, inversions in light-blue and tandem duplication in red.

(B) Representative overview and detailed images of immunohistochemical staining on EMC-PCa-41 organoids of AR, PSA, cytokeratin 8/18 (CK8/18) and ERG, and counterstained with hematoxylin. Top row depicts negative control in which the primary antibody was omitted. Scale bar represents 10 μm in size in the detailed images and 100 μm in the overviews.

(C) Drug sensitivity of the organoid cell line EMC-PCa-41 towards the anti-androgen treatment enzalutamide and taxane chemotherapeutics, as compared to the established PCa organoid cell lines. Data shown is the average of 3 individual cell viability experiments with three technical replicates, scale bars represent SEM. Hormone sensitivity of EMC-PCa-41 was determined by cell viability in androgen depleted conditions (minus R1881) and compared to the AR positive cell organoid cell line MSK-PCa-2. Chemosensitivity was compared to the AR negative cell line MSK-PCa-1.

CTC and organoid heterogeneity by single cell copy-number alterations

To investigate the heterogeneity within CTCs and early-stage organoid cultures, we performed low-pass WGS on multiple single cells obtained from the two long-term mCRPC organoid samples with matched CTCs and WBCs that served as normal controls (subject 25 and 41). The genomic profiles revealed several single cells without any distinct CNA that clustered together with WBC controls. Additional validation excluded these single cells, as they lacked tumor-specific SNVs and were likely non-malignant (Figure S6). We then performed a t-SNE analysis to identify clusters of tumor cells bearing similar copy-number profiles to assess heterogeneity and extract consensus copy-number profiles. The CTCs and organoid cells isolated from subject 25 separated into five distinct clusters (Figure 5A). The organoid cells isolated from early cultures clustered separately (cluster 1) from the CTCs and displayed unique focal amplifications on chromosome 11p, 14 and 15 (Figure S7). Tumor cells obtained from subject 41 clustered into three distinct groups (Figure 5B) and showed distinct heterogeneity and ploidy between clusters. Moreover, the consensus plot from cluster 1 revealed a baseline copy-number of 6 with focal amplifications on chromosome 4, 13 and 20. Cluster 3, encompassing 8 of 18 CTCs, harbored a triploid genome and closely resembled the focal amplifications previously identified in the matched organoid cell line EMC-PCa-41.
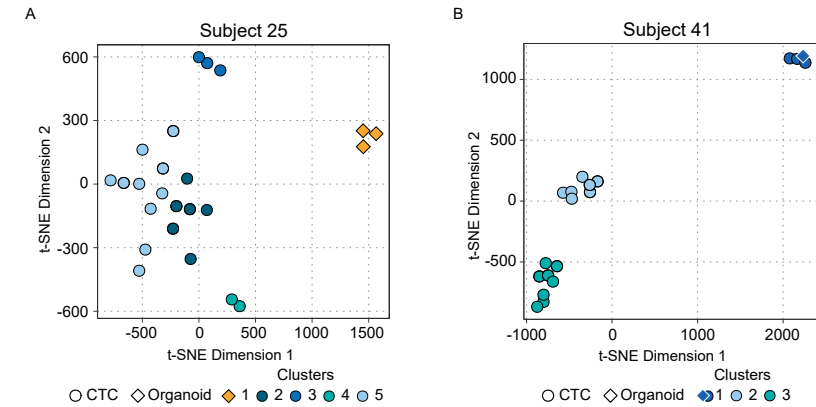
Figure 5. Clustering of individual circulating tumor cells and organoid cells based on low-pass whole genome sequencing data

t-SNE plot (k-nearest neighbor algorithm; Louvain method) of tumor cells, clustering by absolute copy-number values (0.01 Mb). Copy-number values were obtained by low-pass whole genome sequencing of single tumor cells after whole genome amplification, white blood cells were taken along as negative control (not displayed in t-SNE).

Discussion

This study confirms that DLA is a safe and efficient method to harvest large amounts of CTCs from mPCa patients. Furthermore, we have optimized WBC depletion methods to efficiently recover CTCs and remove WBCs from DLA samples. From the isolated and enriched CTC fractions we were able to establish organoid cultures in 35% of the samples which were mostly of short-term nature, although one sample lead to a stable organoid line. Our study shows that DLA is a promising method to obtain viable tumor cells from mPCa patients for subsequent downstream analyses such as single cell sequencing. Unfortunately, the modest success rate to expand organoid cultures precluded us from using CTC derived organoid cultures as a platform to select personalized treatment options for now.

The use of living cells, directly obtained from patients as 'real-life' drug screening models, is an appealing prospect in our quest to improve personalized cancer treatment. Indeed, phenotyping living tumor cells has the advantage to directly measure the response to treatment compared to 'phenotyping after fixation' based stratification[24]. Previous reports on CTC cultures indicated that one of the main factor for success is the number of cells to initiate expansion[13,15]. Indeed, we found that samples that were successfully propagated as organoids contained significantly higher number and percentage of CTCs after enrichment and isolation. While DLA allows us to obtain vast numbers of CTCs, processing of the DLA material provides new challenges due to the excess of WBCs present in the sample. The currently available methods for WBC depletion or CTC enrichment require an excess of magnetic beads or RBCs to achieve the appropriate

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
circulating tumor cells

ratio to capture all cells by antibodies. This limited the DLA volume that can be (cost)effectively processed. In our study, we were limited to one-fourth of the DLA sample from 5 L blood and thus the number of CTCs obtained for culture. Further optimization of DLA sample processing could therefore tremendously impact the amount of viable tumor cells obtained and benefit downstream applications.

We were able to generate organoids in 35% of the samples, which is a substantial improvement compared to previous reports that used CTCs for *ex vivo* organoid cultures (5-7% success-rate)[13,14]. Unfortunately, most organoids could be maintained as short-term cultures with limited proliferative capacity, implicating that the current culture techniques do not provide an optimal environment for sustained viability. We were able to maintain two samples for several passages (subject 25 and 41), eventually leading to one stable organoid line (EMC-PCa-41). Obtaining preclinical models of PCa has been notoriously difficult, presumably due to the low proliferative capacity of PCa as well as overgrowth of benign epithelial and stromal cells. Therefore, establishing PCa cell lines from liquid biopsy samples provides several advantages, including lack of normal epithelial cells and the possibility to obtain metastatic samples from patients with bone only disease in a minimally-invasive manner. The CTC derived organoid cell line EMC-PCa-41 provides a unique novel model for mCRPC with enzalutamide resistance similarly to the patient. Moreover, EMC-PCa-41 harbors genomic alterations similar to a large subset of mCRPC patients as well as the *TMPRSS2-ERG* fusion, which is underrepresented in the currently available models of PCa[4,25].

Treatment modalities for mPCa have profoundly changed and expanded during the last decade. Understanding how PCa cells adapt to the selective pressure of treatment is becoming increasingly important to further improve treatment outcome. Therefore, we need patient derived materials that reflect the current status of the patient's cancer, including the spatial and temporal tumor heterogeneity. Here we show that DLA enables in-depth studies into intra-tumor heterogeneity in mPCa by performing single cell whole genome DNA sequencing on CTCs. The single cell analysis distinguished clusters of cells with unique copy-number alteration which is in line with a previous study[14]. Within our study we were also able to obtain matched samples in four patients, which shows that longitudinal CTC sampling by DLA is feasible, providing a platform to study clonal evolution and adaptation to treatment. Overall, our study provides an important step forward in implementing CTCs in individualized disease modelling, nevertheless identifies several challenges that require further optimization to enable the development of a personalized drug screening platform.

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
circulating tumor cells

References

1. Armenia J, Wankowicz SAM, Liu D, et al. The long tail of oncogenic drivers in prostate cancer. Nat Genet 2018;50(5):645-651. (In eng). DOI: 10.1038/s41588-018-0078-z 10.1038/s41588-018-0078-z [pii].

2. Wedge DC, Gundem G, Mitchell T, et al. Sequencing of prostate cancers identifies new cancer genes, routes of progression and drug targets. Nat Genet 2018;50(5):682-692. (In eng). DOI: 10.1038/s41588-018-0086-z 10.1038/s41588-018-0086-z [pii].

3. Carreira S, Romanel A, Goodall J, et al. Tumor clone dynamics in lethal prostate cancer. Sci Transl Med 2014;6(254):254ra125. (In eng). DOI: 6/254/254ra125 [pii] 10.1126/scitranslmed.3009448.

4. van Dessel LF, van Riet J, Smits M, et al. The genomic landscape of metastatic castration-resistant prostate cancers reveals multiple distinct genotypes with potential clinical impact. Nature Communications 2019;10(1):5251. DOI: 10.1038/s41467-019-13084-7.

5. Miyamoto DT, Zheng Y, Wittner BS, et al. RNA-Seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance. Science 2015;349(6254):1351-6. (In eng). DOI: 349/6254/1351 [pii] 10.1126/science.aab0917.

6. de Bono JS, Scher HI, Montgomery RB, et al. Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. Clin Cancer Res 2008;14(19):6302-9. (In eng). DOI: 14/19/6302 [pii] 10.1158/1078-0432.CCR-08-0872.

7. Scher HI, Lu D, Schreiber NA, et al. Association of AR-V7 on Circulating Tumor Cells as a Treatment-Specific Biomarker With Outcomes and Survival in Castration-Resistant Prostate Cancer. JAMA Oncol 2016;2(11):1441-1449. (In eng). DOI: 2526995 [pii] 10.1001/jamaoncol.2016.1828.

8. Antonarakis ES, Lu C, Wang H, et al. AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. N Engl J Med 2014;371(11):1028-38. DOI: 10.1056/NEJMoa1315815.

9. Onstenk W, Sieuwerts AM, Kraan J, et al. Efficacy of Cabazitaxel in Castration-resistant Prostate Cancer Is Independent of the Presence of AR-V7 in Circulating Tumor Cells. Eur Urol 2015;68(6):939-45. DOI: 10.1016/j.eururo.2015.07.007.

10. Heller G, McCormack R, Kheoh T, et al. Circulating tumor cell number as a response measure of prolonged survival for metastatic castration-resistant prostate cancer: A comparison with prostate-specific antigen across five randomized phase III clinical trials. J Clin Oncol 2018;36(6):572-580. (Conference Paper) (In English). DOI: 10.1200/jco.2017.75.2998.

11. Drost J, Clevers H. Organoids in cancer research. Nat Rev Cancer 2018;18(7):407-418. (In eng). DOI: 10.1038/s41568-018-0007-6 10.1038/s41568-018-0007-6 [pii].

12. Yu M, Bardia A, Aceto N, et al. Cancer therapy. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. Science 2014;345(6193):216-20. (In eng). DOI: 345/6193/216 [pii] 10.1126/science.1253533.

13. Gao D, Vela I, Sboner A, et al. Organoid cultures derived from patients with advanced prostate cancer. Cell 2014;159(1):176-187. (In eng). DOI: S0092-8674(14)01047-2 [pii] 10.1016/j.cell.2014.08.016.

14. Lambros MB, Seed G, Sumanasuriya S, et al. Single-Cell Analyses of Prostate Cancer Liquid Biopsies Acquired by Apheresis. Clin Cancer Res 2018;24(22):5635-5644. (In eng). DOI: 1078-0432.CCR-18-0862 [pii] 10.1158/1078-0432.CCR-18-0862.

15. Cayrefourcq L, Mazard T, Joosse S, et al. Establishment and characterization of a cell line from human circulating colon cancer cells. Cancer Res 2015;75(5):892-901. (In eng). DOI: 0008-5472.CAN-14-2613 [pii] 10.1158/0008-5472.CAN-14-2613.

16. Stoecklein NH, Fischer JC, Niederacher D, Terstappen LW. Challenges for CTC-based liquid biopsies: low CTC frequency and diagnostic leukapheresis as a potential solution. Expert Rev Mol Diagn 2016;16(2):147-64. (In eng). DOI: 10.1586/14737159.2016.1123095.

17. Stenzinger M, Bonig H. Risks of leukapheresis and how to manage them-A non-systematic review. Transfus Apher Sci 2018;57(5):628-634. (In eng). DOI: S1473-0502(18)30359-8 [pii] 10.1016/j.transci.2018.09.008.

18. Simon R. Optimal two-stage designs for phase II clinical trials. Controlled Clinical Trials 1989;10(1):1-10.

19. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45(2):228-47. (In eng). DOI: S0959-8049(08)00873-3 [pii] 10.1016/j.ejca.2008.10.026.

20. Scher HI, Morris MJ, Basch E, Heller G. End points and outcomes in castration-resistant prostate cancer: from clinical trials to clinical practice. J Clin Oncol 2011;29(27):3695-704. (In eng). DOI: JCO.2011.35.8648 [pii] 10.1200/JCO.2011.35.8648.

21. Marques RB, Erkens-Schulze S, de Ridder CM, et al. Androgen receptor modifications in prostate cancer cells upon long-termandrogen ablation and antiandrogen treatment. Int J Cancer 2005;117(2):221-9. (In eng). DOI: 10.1002/ijc.21201.

22. Beshiri ML, Tice CM, Tran C, et al. A PDX/organoid biobank of advanced prostate cancers captures genomic and phenotypic heterogeneity for disease modeling and therapeutic screening. Clin Cancer Res 2018 (In eng). DOI: 1078-0432.CCR-18-0409 [pii] 10.1158/1078-0432.CCR-18-0409.

23. Sieuwerts AM, Mostert B, van der Vlugt-Daane M, et al. An In-Depth Evaluation of the Validity and Logistics Surrounding the Testing of AR-V7 mRNA Expression in Circulating Tumor Cells. Journal of Molecular Diagnostics 2018;20(3):316-325. (In English). DOI: 10.1016/j.jmoldx.2018.01.008.

24. Ooft SN, Weeber F, Dijkstra KK, et al. Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. Sci Transl Med 2019;11(513) (In eng). DOI: 11/513/eaay2574 [pii] 10.1126/scitranslmed.aay2574.

25. Navone NM, van Weerden WM, Vessella RL, et al. Movember GAP1 PDX project: An international collection of serially transplantable prostate cancer patient-derived xenograft (PDX) models. Prostate 2018;78(16):1262-1282.

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
circulating tumor cells

Supplementary information

Supplementary materials and methods

Passaging and isolating organoids for phenotyping

For passaging and isolation of early stage organoids, MG droplets were mechanically disrupted and dispensed in a pre-cooled 96-wells plate (Corning, cat. no. 3595). TryplE (Thermo Fisher Scientific, cat. no. 12605010) was added to the wells and incubated at 37˚C and organoids were resuspended regularly until the organoids were fractioned. Organoids were collected in cold medium (AdMEM/F12+++), centrifuged at 1200*g* for 5 minutes at 4°C and plated as described previously. For cell viability assays, enzymatic disruption was prolonged to acquire mostly single cells and plated at a density of 2500 cells per well in 8 µL of MG in a 96-wells plate. To allow for organoid formation, cells were incubated for seven days with 100 µL media. We used APCOM without Y-27632 and R1881, as this would interfere with treatment induced cell-death and anti-androgen response. Subsequently medium was replaced to contain the appropriate drug/androgen concentration and incubated for another seven days (Enzalutamide Axon Medchem, Groningen, the Netherlands, cat. no 1613; R1881 details in Table S1 and Taxanes were provided by Sanofi, Paris, France). Cell viability was measured using the CellTiter-Glo® 3D (Promega, cat. no. G9681, Madison, WI), and normalized to untreated controls. MSK-PCa1 and MSK-PCa2 were maintained in the organoid culture media described by Gao et al.[1] and cell viability assays were performed as described above. For cryopreservation, isolated organoids were mixed with 1 mL cooled Recovery ™ Cell Culture Freezing Medium (Thermo Fisher Scientific, cat. no. 12648010), stored in a Styrofoam container (Westburg CoolCell LX, cat. no. BCS-405, Leusden, The Netherlands) overnight at -80°C and subsequently transferred to liquid nitrogen storage. For immunohistochemistry, organoids in MG were formalin fixed for four hours, and subsequently embedded in 4% agarose and paraffin. Four µm section were stained for the expression of the androgen receptor (AR; 1:200, SP107, Cell Marque, Rocklin, CA), prostate specific antigen (PSA ;1:500, N1517, Dako, Santa Clara, CA), CK8/18 (1:150, Ma5-14088, Thermo Fisher Scientific) and ERG (1:100, EPR3864, Abcam, Cambridge, UK) and visualized with DAB/H_2O_2 (EnVision kit, Dako). Images were obtained using the Olympus BX41 microscope equipped with 2x, 10x, 20x and 40x UPlanFL N objectives, a ColorView III camera and Cell[B] imaging software (version 3.4, Olympus Shinjuku, Tokyo, Japan).

Whole genome sequencing and variant calling of EMC-PCa-41

From six patients, whole genome sequencing (WGS) data from metastatic tissue was present as part of the CPCT-02 study (NCT01855477) and used to identify somatic SNVs for validation studies (Table S3)[2]. EMC-PCa-41 organoids were collected for WGS (passage no 5) in a similar fashion as passaging, washed with cold PBS and dry cell pellets were stored at -80°C. WBCs isolated from

DLA were used as matched normal cells for sequencing. DNA isolation and WGS was performed by the Hartwig Medical foundation as previously described[2]. In short, Illumina technology was used for WGS of DNA libraries on the HiSeq X Ten system using paired-end (2x 150 bp) sequencing (Illumina, San Diego, CA) to a minimum sequencing depth of 30x and 60x for the matched normal and tumor sample respectively. The human reference genome (GRCh37) was used for alignment and post-processing and subsequent somatic analysis for variants (single nucleotide variants (SNVs), small insertion/deletions and multi-nucleotide variants), copy-number alterations (CNA), structural variants was performed as previously described[2,3].

Single cell isolation, whole genome amplification and sequencing of CTCs and single cells from organoids

For single cell isolation we used fresh DLA product diluted in CellSearch dilution buffer, collected in a CellSave tube and processed within 96 hours, or stored DLA product which was thawed before processing. Samples were processed on the CellSearch system as described above. CellSearch cartridges were stored in the dark at 4°C before further analyses. Isolation of single CTCs, defined as DAPI^pos/CK^pos/CD45^neg cells, and single white blood cells WBCs, defined as DAPI^pos/CK^neg/CD45^pos, was performed by flow cytometry using MoFlo XDP sorter (Beckman Coulter, Germany) as previously described[4]. Single cells were sorted into individual empty PCR tubes and stored at -20 °C until analysis. Single cell isolation of organoids was performed using the VyCAP Puncher System (VyCAP, Enschede, The Netherlands), combining a silicon chip with microwells, fluorescence imaging, and a punching method to isolate and transfer the single cells to standard reaction tubes as previously described[5]. Subsequent whole genome amplification (WGA) for all picked/sorted single cells was performed using adapter-linker PCR as previously described[6,7] , and commercialized as *Ampli*1™ WGA Kit by Menarini Sillicon Biosystems. *Ampli*1™ LowPass kit for Illumina (Menarini Silicon Biosystems, Bologna, Italy) was used for preparing low-pass WGS libraries at Menarini Silicon Biosystem facilities. For high-throughput processing, the manufacturer procedure was implemented in a fully automated workflow on a STARlet Liquid Handling Robot (Hamilton, Reno, NV, USA). Resulting libraries were sequenced on HiSeq instrument (Illumina, Hayward, CA, USA) and the obtained FASTQ files were aligned to the human reference genome (GRCH37) sequence using Burrows-Wheeler Aligner version 0.7.12 (BWA)[8]. Quality control (QC) included read count distribution and derivative log ratio spread as described previously[9], two CTCs from subject 41 failed to pass QC. CNA in the data were identified using Control-FREEC software (version 11.0) and ploidy level was analyzed using the MSBiosuite pipeline based on best fitting of profiles to underlying copy-number levels[10].

Validation of single cell sequencing

Validation of single cell sequencing was performed by Sanger sequencing of known SNVs from subject 25 and 41, identified in WGS of the patient's tumor or patient derived organoid cell line resp. In short, a nested PCR strategy was used to amplify genomic regions with known SNVs from

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
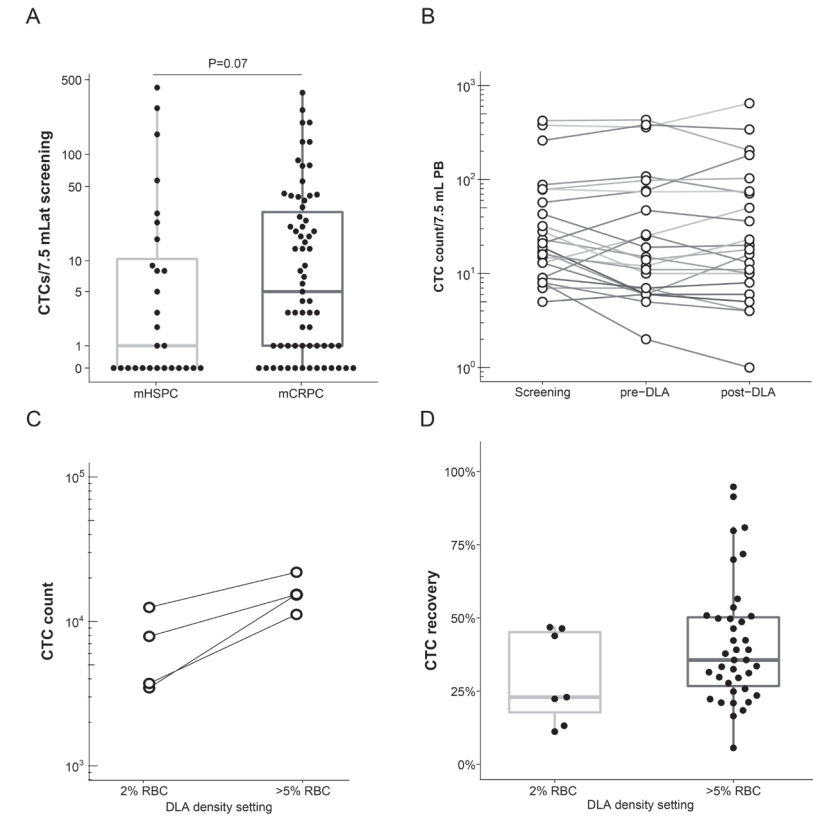circulating tumor cells

the single cell WGA product (Table S4). PCRs were performed using Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific, cat. no. F530) and the final product was purified (QIAquick PCR purification kit, Qiagen, cat. no. 28104). The amplicons were Sanger sequenced (Macrogen Europe, Amsterdam, The Netherlands) and results were analyzed using pairwise alignment with the reference gene (BLAST, NCBI). CTCs and organoids samples that did not harbor any tumor specific SNVs were excluded, WBC served as negative controls.

Bioinformatics analysis of single cell copy-number alterations

For the clustering of the complete dataset (Figure S6) we used the median ratio obtained by the Control-FREEC software, normalized each profile on fixed bins length (weighted mean on 0.25 Megabase (Mb) bins) and calculate the $\log_2$ values. Subsequently, after setting the minimum value to -2 and maximum value to 2, hierarchical clustering of profiles was performed, using "Euclidean" distance metric and "ward" clustering method. For clustering the validated tumor cell profiles, absolute copy-numbers were binned into 0.01 Mb bins which were subsequently annotated by the mean copy-number of overlapping copy-number segments and rounded to the nearest integer. These bins (0.01 Mb) were used as input for *t*-SNE ($\theta = 0.5$ with a perplexity of 2) using the Rtsne package (v0.15)[11]. Subsequent clustering of the *t*-SNE results, outputted as two dimensions, was performed using the k-nearest neighbor algorithm (k ≤ 15) and the Louvain method for community detection, as implemented by the igraph package (v1.2.5) from the statistical platform R (v3.6.1)[12,13]. Per cluster, a consensus copy-number profile was generated by adopting the median copy-number value per bin (0.01 Mb) over all samples captured within the respective cluster. Within the consensus copy-number profiles, the maximum absolute copy-number was capped to 8.

Statistical analysis

Sample size calculation for the patient inclusion is described in the attached clinical protocol. For comparison of CTC count in PB samples obtained at screening versus before the DLA procedure, CTC yield in low versus high density fraction enrichment by DLA, CTC yield and WBC depletion after WBC depletion with or without EpCAM enrichment we used a paired, two-sided Wilcoxon-rank test due to the non-normality of the data. Data visualization and statistical testing were performed using the statistical platform R (version 3.6.1)[13] or Graphpad Prism (version 5.01, GraphPad Software, San Diego, CA), we considered statistical significance with P < 0.05.

Figure S1. Circulating tumor cell burden in metastatic prostate cancer patients and optimization of diagnostic leukapheresis

(A) Circulating tumor cell (CTC) count at screening of all metastatic prostate cancer patients (mPCa) included and screened within this study (n = 98). The boxplots depict the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range (IQR). Y-axis is a pseudo log scale. CTC count per 7.5 mL peripheral blood is shown for mPCa patients with hormone-sensitive (n = 28) and castration-resistant (n = 67) disease and compared using an unpaired two-sided Wilcoxon sign rank test (P = 0.07).

(B) CTC count per 7.5 mL peripheral blood (PB) as obtained during different stages of study participation. For 24 patients samples were collected at screening, just prior to diagnostic leukapheresis (pre-DLA) and post-DLA. Y-axis is a log-scale, matched samples are connected with a line. CTC burden in the different samples was compared by a Friedman test, no statistical significant impact was observed (P = 0.37).

(C) Absolute CTC count in DLA product obtained at different densities. X-axis displays the two tested density conditions: 2% red blood cells (RBC) and > 5% RBC. Each dot represents an individual subject (n = 4), and samples from the same subject are connected. Y-axis is a logarithmic scale. Statistical comparison was performed by a paired two-sided Wilcoxon rank test (P = 0.125).

(D) Relative CTC recovery (%) in DLA product processed at different densities. The X-axis displays the two tested density conditions: 2% and > 5% RBC (n = 7 and n = 46 fractions resp.). In six patients two DLA fractions were collected. The boxplots depict the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the IQR.

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
circulating tumor cells
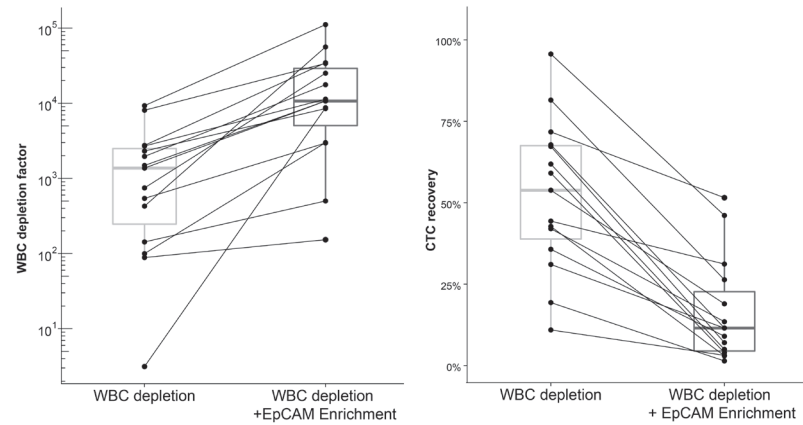
Figure S2. Comparison of white blood cell depletion with or without EpCAM enrichment for the efficacy to enrich for circulating tumor cells

(A-B) Two CTC enrichment methods were compared in twelve paired samples (depicted on the X-axis): (1) white blood cell (WBC) depletion and (2) WBC depletion followed by EpCAM enrichment. The boxplots depict the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the IQR. Paired samples are connected with a line. Statistical comparison was performed by a paired two-sided Wilcoxon rank test, P < 0.001 for both datasets.

(A) Displays the WBC depletion factor after CTC enrichment. To calculate the WBC depletion factor, the number of WBCs before enrichment was divided by the number of WBCs after enrichment. Y-axis is a logarithmic scale.

(B) Displays the relative CTC recovery (%) after CTC enrichment. To calculate relative CTC recovery, the absolute CTC count after the enrichment was divided by the absolute CTC count before the enrichment. Absolute CTC counts were extrapolated from 1 mL samples.
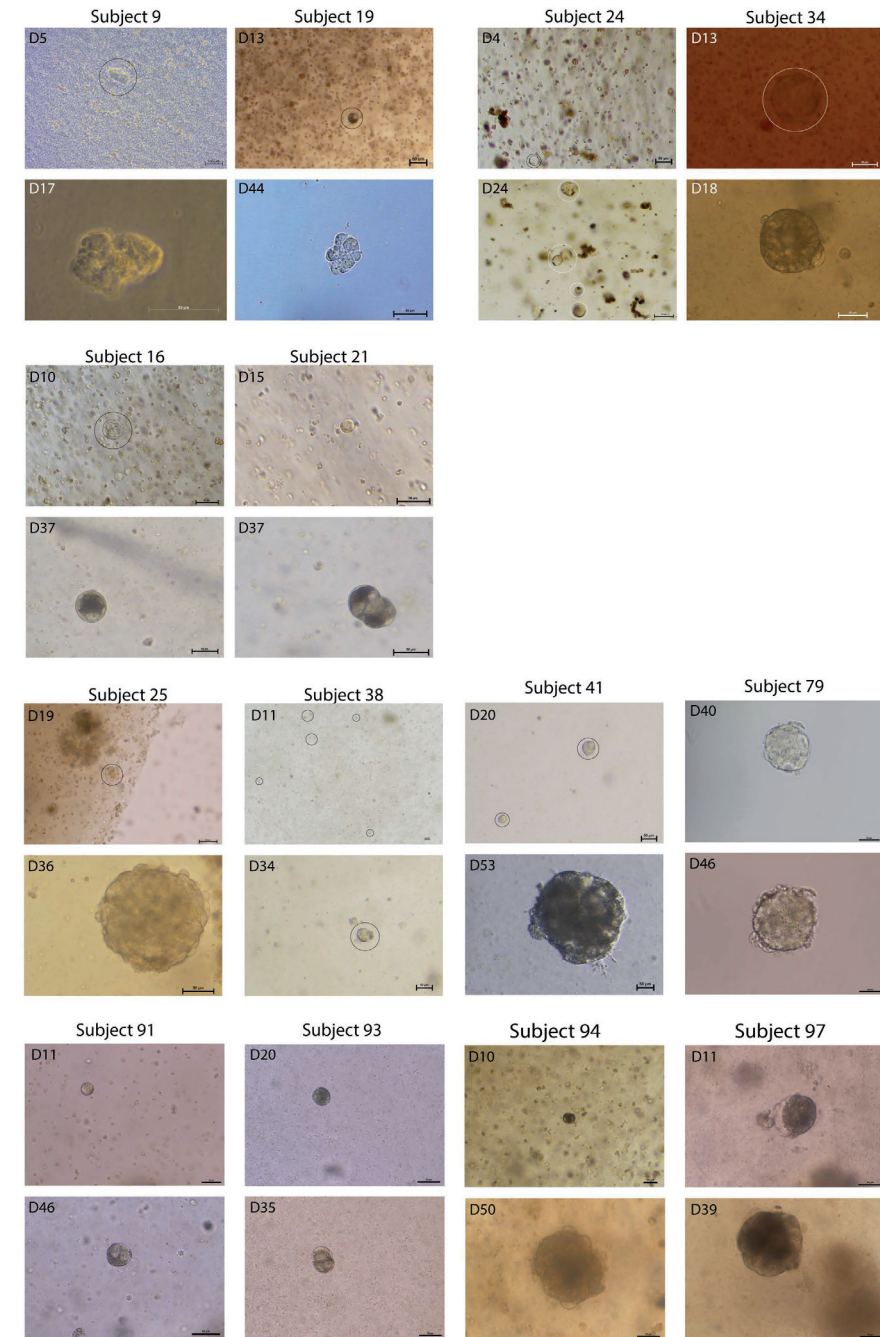
Figure S3. Overview of successful CTC derived organoid cultures

Overview of all successful circulating tumor cells (CTC) derived organoid cultures obtained (n = 14), and organoid expansion observed over time. Timespan (in days) of organoid cultures are indicated in the top left corners, scale bars depicted are 50 μm in size with the exception of the day 5 image of subject 9 (100 μm).

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
circulating tumor cells

Figure S4. Results of qPCR validation of prostate (cancer) associated genes of circulating tumor cell derived organoids

Heatmap of androgen receptor full length (AR wt), AR splice variant 7 (AR-V7), KLK3 (PSA) and *TMPRSS2-ERG* expression as determined by qPCR in circulating tumor cell (CTC) derived organoid cultures. Expression is depicted by ΔΔCq which is defined by target Ct values subtracted with the average EpCAM/KRT19 Ct value. If several organoid samples were obtained, median expression per culture media (PGM or APCOM) was calculated and depicted in the heatmap. Organoids were deemed positive for expression of prostate (cancer) associated genes if ΔΔCq were above -8.5.
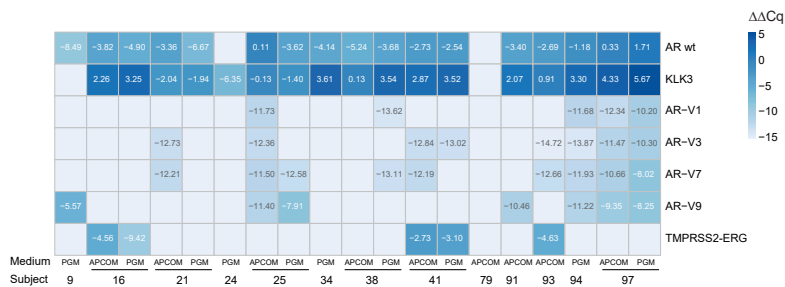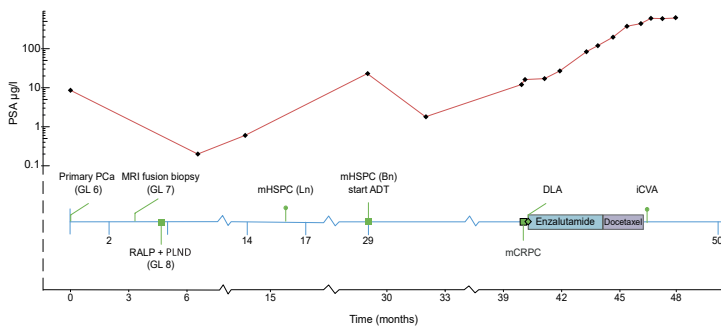

Figure S5. Clinical overview of subject 41

Clinical overview of subject 41, including PSA levels during the disease course (top part). Initially, a Gleason (GL) 6 adenocarcinoma of the prostate was diagnosed by systematic biopsies at an initial PSA of 8.6 μg/L (while the patient was using dutasteride). MRI showed a PIRADS 5 lesion which was target biopsied and revealed a Gleason 7 adenocarcinoma of the prostate. After robot-assisted laparoscopic prostatectomy (RALP) with pelvic lymph node dissection (PLND), PSA remained detectable at 0.2 μg/L. Upon PSA progression metastases in lymph node (Ln) and bone (Bn) were detected and androgen deprivation therapy (ADT) was started. Before start of enzalutamide, the patient underwent diagnostic leukapheresis (DLA). The patient received four cycles of docetaxel which was discontinued because of an ischemic cerebrovascular accident (iCVA).
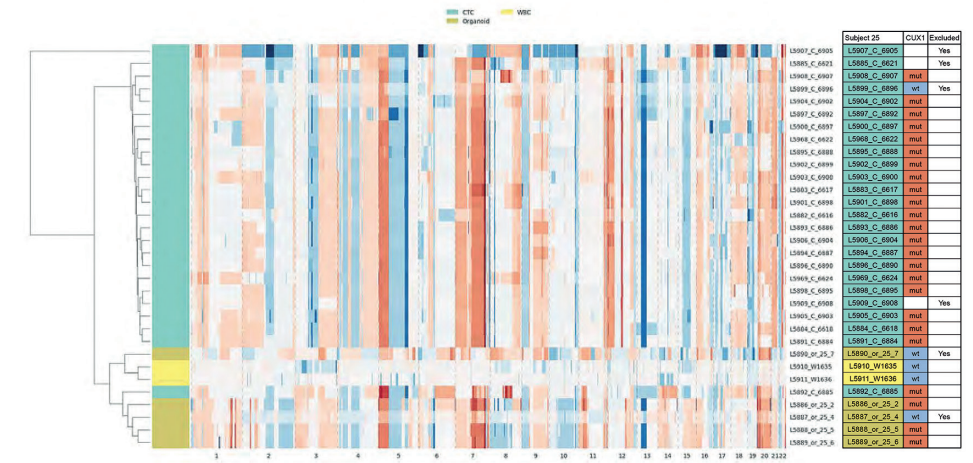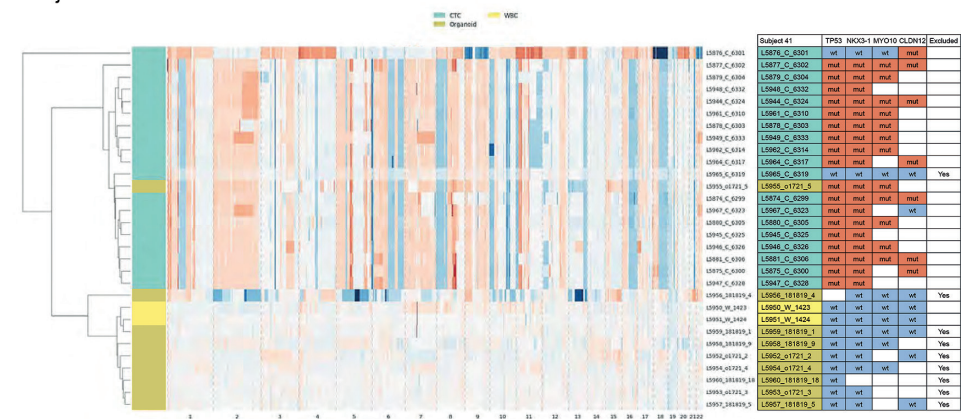
Figure S6. Initial clustering of copy-number alterations in circulating tumor cells and early organoids from two metastatic castration-resistant prostate cancer samples and subsequent validation

Unsupervised hierarchical clustering with dendogram (Euclidean distance; ward method) using $\log_2$ ratio of normalized counts per 0.25 megabase, depicting copy-number alterations (CNA) of individual circulating tumor cells (CTCs, aqua), organoid cells (beige) and white blood cells (WBCs, yellow) for subject 25 (top panel) and 41 (bottom panel). CNA segments for each individual cell (column) are shown from left to right and ordered on chromosome (indicated below). The color gradient represents copy-number deletions in dark blue (-2 to 0) and amplification in dark red (0 to 2). The tables alongside the hierarchical clustering represent the results from the validation experiments. Sanger sequencing was performed to identify somatic variants in whole genome amplified DNA from the CTCs, organoids and WBCs from subject 25 and 41. Empty wells represent missing data, shown are the organoid and CTC samples excluded based on validation.

Part 2
Liquid Biopsies

Chapter 6
Generating human prostate cancer organoids from leukapheresis enriched
circulating tumor cells

A
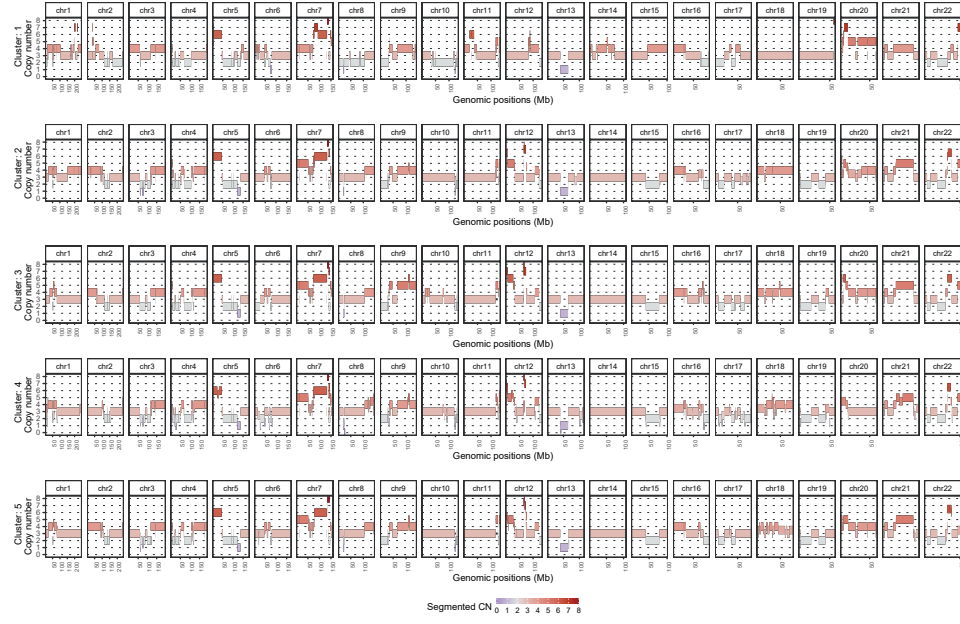
Subject 25

B

Subject 41

Figure S7. Copy-number heterogeneity in circulating tumor cells and organoids
(A-B) Consensus profiles per t-SNE cluster, depicting median absolute copy-numbers in (A) subject 25 (n = 24) and (B) subject 41 (n = 19). The Y-axis displays the absolute copy-number ranging from 0 to ≥ 8 copies. X-axis displays the genomic position of the copy-number segments, ticks indicate 50 megabase (Mb). The color gradient represents absolute copy-number, with deletions in blue (0 to < 2) and amplification in red (> 2 to ≥ 8).

Table S1. Description of adjusted prostate cancer organoid media (APCOM) and prostate growth media (PGM) used to culture circulating tumor cell derived organoids

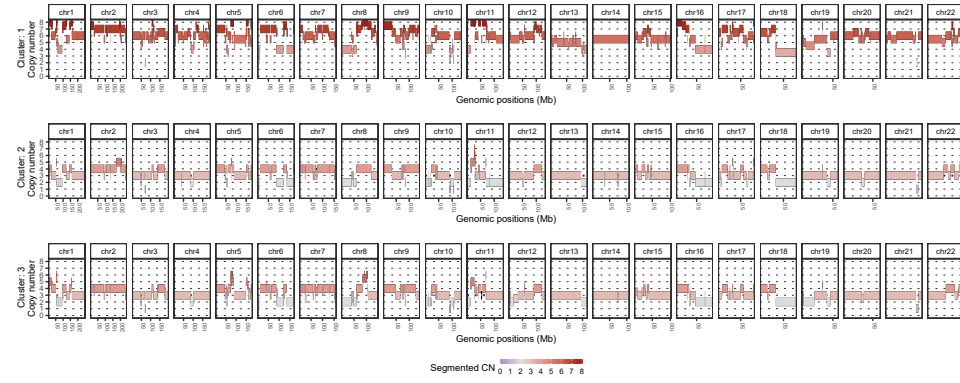| Media	Reagent	Concentration in media	Stock concentration	Supplier (info supplier, catalogue number)
APCOM	Advanced DMEM/F12[1] (AdMEM/F12)	NA	NA	ThermoFisher Scientific (Waltham, Massachusetts, USA, 12634028)
APCOM	Hepes[1]	10 mM	1M	ThermoFisher Scientific (15630056)
APCOM	L-Glutamine[1]	2 mM	200 mM	Lonza (Basel, Swizerland, 17-605E)
APCOM	Noggin[2]	-	-	Conditioned media from Hek293T-Noggin-Fc (Heijmans et al.)
APCOM	R-spondin[2]	-	-	Conditioned media from Hek293T-hRSpo1 (Kim et al.)
APCOM	A-83-01	500 nM	25 mM	Tocris Bioscience (Bristol, UK, 2939)
APCOM	Prostaglandin E2 (PGE2)	1 µM	10 mM	Tocris Bioscience (2296)
PGM	DMEM/F12	NA	NA	ThermoFisher Scientific (11330032)
PGM	Insulin-Transferrin-Selenium (ITS)	10 µg/mL, 5.5 µg/mL and 6.7 ng/mL	1 mg/mL, 0.55 mg/mL and 0.67 µg/mL	ThermoFisher Scientific (41400045)
PGM	Bovine serum albumine Fraction V (BSA)	0.01% (w/v)	10% (w/v)	Sigma-Aldrich (Saint Louis, Missouri, USA, 10735094001)
PGM	Fetuin	20 µg/mL	8 mg/mL	Sigma-Aldrich (F3385)
PGM	Fibronectin	100 ng/mL	100 µg/mL	Alfa Aesar (Haverhill, Massachusetts, USA, J65460)
PGM	Triiodo-L-thyronine (T3)	0.6 ng/mL	600 ng/mL	Sigma-Aldrich (T6397)
PGM	Phosphoethanolamine (PEA)	100 µM	100 mM	Sigma-Aldrich (P0503)
PGM	Choleratoxin	50 ng/mL	50 µg/mL	Sigma-Aldrich (C8052)
PGM	Fetal Bovine serum (FBS)	2% (v/v)		ThermoFischer Scientific (10270106)
PGM	Hydrocortisone	0.5 µg/mL	0.5 mg/mL	Sigma-Aldrich (H0888)
Both	Epithelial growth factor (EGF)	20 ng/mL or 10 ng/mL	10 µg/mL	Sigma-Aldrich (F29644)
Both	Fibroblast growth factor 2 (FGF-2)	5 ng/mL	50 µg/mL	R&D Systems (Minneapolis, Minnesota, USA, 233-FB-025)
Both	Fibroblast growth factor 10 (FGF-10)	10 ng/mL	100 µg/mL	PeproTech (Rocky Hill, New Jersey, USA, 100-26)
Both	Y-27632 dihydrochloride	10 µM	10 mM	Abmole Bioscience (Houston, Texas, USA,M1817)
Both	R1881	0.1 nM	0.1 µM	Sigma-Aldrich (D5027)
Both	Penicilline/streptomycine[1]	100 U and 100 µg/mL	10.000 U and 10.000 µg/mL	Lonza (17-602E)

[1] Components for AdMEM/F12[+++]. APCOM additves are diluted in AdMEM/F12[+++], and top up till desired volume. [2] Desired supplement concentration is determined by an ELISA assay (Kim et al.)

Table S2. Primer and probe design for digital PCR
Custom and commercial digital polymerase chain reaction assays used for validation of CTC derived organoids from subject 9/19 (TP53), subject 24/34 (PIK3CA) and subject 25 (PLCG2).

Gene	Cosmic ID	Amino acid variant	Somatic variant	Forward primer	Reverse primer	Probe*
TP53	COSV53122800	p.Q136L	c.407A>T	5'-ACT CCC CTG CCC TCA ACA A-3'	5'- AAG ACC TGC CCT GTG CAG C-3'	5' –TGT TTT GCC [A/T]AC TGG C-3'
TP53	COSV52661877	p.G245S	c.733G>A	5'-TGT TAT CTC CTA GGT TGG CTC TGA-3'	5'-AGG ATG GGC CTC CGG TT-3'	5'-CTG CAT GGG C[G/A]A GCA-3'
PIK3CA	COSV55873195	p.H1047R	c.3140A>G	Assay ID**: AHPAVCD		
PLCG2	COSV63869132	p.D219N	c.655G>A	5'-TGC ATT AAG TGA CTT GTC TAA GGT TCT TT-3'	5'-CCA GGA TGA ACA CGG ACG AAT C-3'	5'-TTT CAG ATT CTC [G/A]AT GAA TT-3'

Reporter dye/quencher: wild-type, VIC; mutant, FAM / NFQ
***Validated TaqMan® SNP genotyping assay (ThermoFisher Scientific)*

Table S3. Genomic characteristics of metastatic prostate cancer patients who underwent diagnostic leukapheresis
Actionable and cancer associated genomic alteration detected in metastatic biopsies from included patients whose whole genome sequencing (WGS) data was used to validate organoid cultures. WGS data was generated as part of the CPCT-02 study. Shown are the somatic variants, gain/losses, gene-fusions and gene-disruptions in cancer associated genes (as described in Priestley et al.).

Subject ID	Somatic variants	Somatic gains and losses	Somatic gene fusions	Somatic gene disruption	Microsatalite status	Tumor mutational load
9/19	TP53 c.407A>T; CDK12 c.2350C>T; CDK12 c.3002T>A	Copy-gain: MYC, CCND1, MDM2, PTPN11 and CDK4	-	?	?	Low
16	TP53 c.243_249delACCGGCG; AR c.2623C>T; CTNNB1 c.133T>C	Copy-loss: PTEN	CDC27-ETV4; TMPRSS2-ERG	PMS2:INV intron 12	Stable	Low
22	-	Copy-gain: AR Copy-los: PTEN	-	PTEN: INV intron 2/3 SMAD4: DUP promotor region	stable	Low
24/34	MYCN c.1226C>T; MSH2 c.1699delA; MSH6 c.2554_2556delAAG; PIK3CB c.2056G>A ; PIK3CA c.3140A>G; APC c.1660C>T; APC c.4385_4386delAG; EZH2 c.1730C>T; PTEN c.734A>C; ATM c.7927+5_7927+6insT; MDM2 c.1181C>A; HNF1A c.521C>T; FLT3 c.1419-4dupT; TSC2 c.348delG; AR c.2105T>A; AR c.2623C>T	-	-	MSH2: INV intron 7	Unstable	High
25	-	-	-	BRAF: INV intron 1 (2x) FANCG: INV promotor region	Stable	Low
79	TP53 c.733G>A; ERBB2 c.1796G>A; ESR1 c.1514C>T; RASA1 c.1211C>T; ASXL c.1264A>T	Copy-loss: PTEN, RB1	FASN-ETV4	CUX1: DEL intron 1-> intron 22; BRCA: INV intron 20; CBFB: BND intron 3	Stable	Low

Threshold for high tumor mutational load was 140 and MSI samples were defined by a MSiSeq score above 4; - indicated that this aberration was not detected and ? if not reported.

Table S4. Single nucleotide variants and primer design for validation of single cell sequencing
Single nucleotide variants of subject 25 and 41 previously identified in whole genome sequencing of metastatic biopsy or matched organoid cell line resp. A nested polymerase chain reaction was used for amplification and subsequent Sanger sequencing.

| Subject ID	Gene	Somatic variant	Cosmic ID	Outer primers		Inner primers	
				Forward	Reverse	Forward	Reverse
25	CUX1	c.1381G>C	-	AGCTCCTTCTCCTACCCAGG	GCTCCAGGTCCTTTAGCTGG	ATAGCCAGCACTCTCACAGC	TGTCACCAGCTGCCTGATAC
41	TP53	c.927dup	COSM6951762	AGGACCTGATTTCCTTACTGCC	TTAGTTAGCTACAACCAGGAGCC	AGTGGTAATCTACTGGGACGG	CCCCAATTGCAGGTAAAACAGTC
41	NKX3-1	c.406G>A	COSM5733098	CGATGACAGTGGGCTGTTTG	TAGAGACACCCTGGGGAAGG	CCTTCCCACTCTCCTTTCC	CCCACGCAGTACAGGTATGG
41	MYO10	c.4798C>T	COSM4874050	GCATTGCATCCGTCCTCCTC	TAAGCTTGCTCAGGCAACCC	GCCCTATTCTGAACCCCTCAG	CAGTGCCCCACCTGTGTAAG
41	CLDN12	c.109C>T	COSM1248386	TCTGAGGTCACACACACCTG	GCCCAGGAAAAATAGCAGCCC	GTGTGTCACCCCCTAGTCTG	GTTTGATGTTGGGCACCGAG

References supplementary information

1. Gao D, Vela I, Sboner A, et al. Organoid cultures derived from patients with advanced prostate cancer. Cell 2014;159(1):176-187. (In eng). DOI: S0092-8674(14)01047-2 [pii] 10.1016/j.cell.2014.08.016.

2. Priestley P, Baber J, Lolkema MP, et al. Pan-cancer whole-genome analyses of metastatic solid tumors. Nature 2019;575(7781):210-216. (In eng). DOI: 10.1038/s41586-019-1689-y 10.1038/s41586-019-1689-y [pii].

3. van Dessel LF, van Riet J, Smits M, et al. The genomic landscape of metastatic castration-resistant prostate cancers reveals multiple distinct genotypes with potential clinical impact. Nature Communications 2019;10(1):5251. DOI: 10.1038/s41467-019-13084-7.

4. Neves RP, Raba K, Schmidt O, et al. Genomic high-resolution profiling of single CKpos/CD45neg flow-sorting purified circulating tumor cells from patients with metastatic breast cancer. Clin Chem 2014;60(10):1290-7. (In eng). DOI: clinchem.2014.222331 [pii] 10.1373/clinchem.2014.222331.

5. Andree KC, Abali F, Oomens L, et al. Self-Seeding Microwells to Isolate and Assess the Viability of Single Circulating Tumor Cells. Int J Mol Sci 2019;20(3) (In eng). DOI: ijms20030477 [pii] 10.3390/ijms20030477.

6. Klein CA, Schmidt-Kittler O, Schardt JA, Pantel K, Speicher MR, Riethmuller G. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. Proc Natl Acad Sci U S A 1999;96(8):4494-9. (http://www.ncbi.nlm.nih.gov/pubmed/10200290).

7. Stoecklein NH, Erbersdobler A, Schmidt-Kittler O, et al. SCOMP is superior to degenerated oligonucleotide primed-polymerase chain reaction for global amplification of minute amounts of DNA from microdissected archival tissue samples. Am J Pathol 2002;161(1):43-51. (In eng). DOI: S0002-9440(10)64155-7 [pii] 10.1016/S0002-9440(10)64155-7.

8. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25(14):1754-60. DOI: 10.1093/bioinformatics/btp324.

9. Ferrarini A, Forcato C, Buson G, 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(3):e0193689. (http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=medc1&AN=29494651).

10. Boeva V, Popova T, Bleakley K, et al. Control-FREEC: a tool for assessing copy-number and allelic content using next-generation sequencing data. Bioinformatics 2012;28(3):423-5. DOI: 10.1093/bioinformatics/btr670.

11. van der Maaten L. Accelerating t-SNE using Tree-Based Algorithms. J Mach Learn Res 2014;15:3221-3245. (In English) (<Go to ISI>://WOS:000344638800013).

12. Csardi G, Nepusz T. The Igraph Software Package for Complex Network Research. InterJournal 2005;Complex Systems:1695.

13. R Core Team. R: A Language and Environment for Statistical Computing. Version 1.1.463 ed. Vienna, Austria: R Foundation for Statistical Computing; 2019.

Part 3

Steroidomics
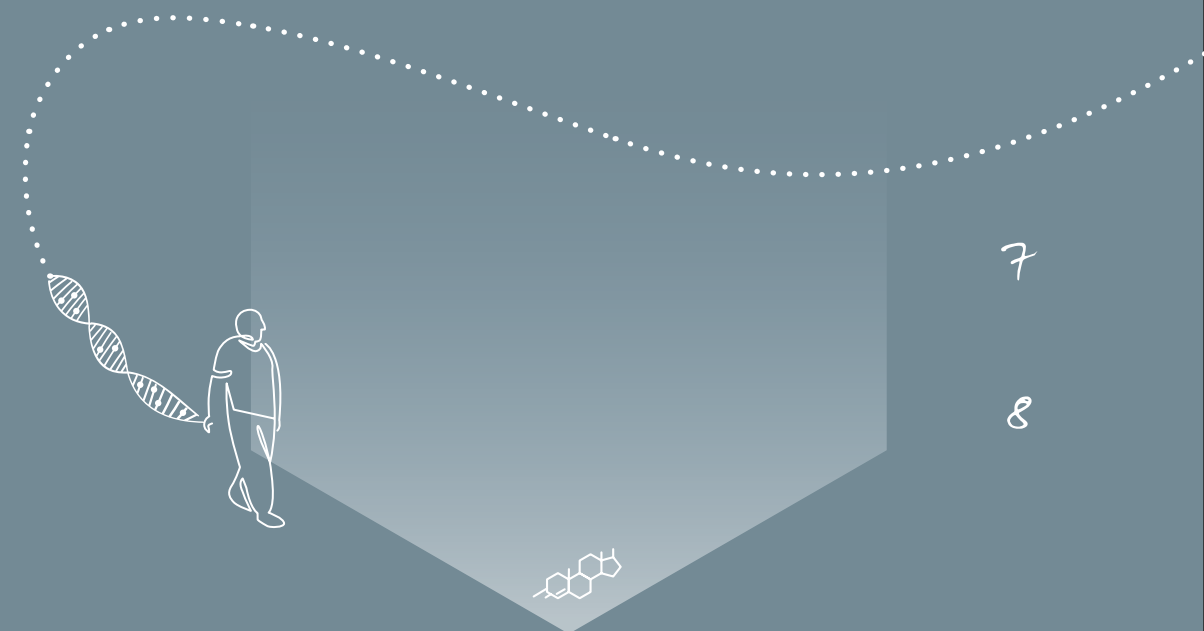
Chapter 7

Validation of circulating steroid hormone measurements across different matrices by liquid chromatography-tandem mass spectrometry

Chapter 7

Gido Snaterse*, Lisanne F. van Dessel*, Angela E. Taylor, Jenny A. Visser, Wiebke Arlt, Martijn P. Lolkema, Johannes Hofland
* These authors contributed equally

Part 3
Steroidomics

Chapter 7
Validation of circulating steroid hormone measurements across different matrices
by liquid chromatography–tandem mass spectrometry

Abstract

Background

Steroid hormones are essential signalling molecules in prostate cancer (PC). However, many studies focusing on liquid biomarkers fail to take the hormonal status of these patients into account. Steroid measurements are sensitive to bias caused by matrix effects, thus assessing potential matrix effects is an important step in combining circulating tumor DNA (ctDNA) analysis with hormone status.

Materials and Methods

We investigated the accuracy of multi-steroid hormone profiling in mechanically-separated plasma (MSP) samples and in plasma from CellSave Preservative (CS) tubes, that are typically used to obtain ctDNA, compared to measurements in serum. We performed multiplex steroid profiling by liquid chromatography-tandem mass spectrometry (LC-MS/MS) in samples obtained from ten healthy controls and ten castration-resistant prostate cancer (CRPC) patients.

Results

Steroid measurements were comparable between MSP and serum. A small but consistent decrease of 8-21% compared to serum was observed when using CS plasma, which was considered to be within the acceptable margin. The minimal residual testosterone levels of CRPC patients could be sensitively quantified in both MSP and CS samples.

Conclusions

We validated the use of MSP and CS samples for multi-steroid profiling by LC-MS/MS. The optimised use of these samples in clinical trials will allow us to gain further insight into the steroid metabolism in PC patients.

Abbreviations

ADT	Androgen deprivation therapy
ANOVA	Analysis of variance
AR	Androgen receptor
CRPC	Castration-resistant prostate cancer
CS	CellSave Preservative
CTC	Circulating tumor cell
ctDNA	circulating tumor DNA
DHEA	Dehydroepiandrosterone
DHEA-S	Dehydroepiandrosterone-sulphate
DHT	5α-dihydrotestosterone
HC	Healthy control
LC-MS/MS	Liquid-chromatography tandem mass spectrometry
LLOQ	Lower limits of quantification
mCRPC	Metastatic castration-resistant prostate cancer
MS	Mass spectrometry
MSP	Mechanically-separated plasma
MTBE	Methyl-tert butyl ether
PC	Prostate cancer
PBS-BSA	Phosphate-buffered saline with bovine serum albumin

Part 3
Steroidomics

Chapter 7
Validation of circulating steroid hormone measurements across different matrices
by liquid chromatography–tandem mass spectrometry

Introduction

Prostate cancer (PC) is a steroid hormone dependent disease where androgens play a pivotal role in the evolution of the disease. Targeting the androgen receptor (AR) signalling pathway through androgen deprivation therapy (ADT) in locally advanced and metastatic PC is a highly effective way to inhibit tumor growth[1]. However, tumor cells will eventually become resistant to these low androgen concentrations and show disease progression. Resistance mechanisms include AR modifications, like mutations and overexpression[2,3], and changes in androgen biosynthesis and metabolism, thereby increasing intratumoral androgen availability[4-6]. The continued importance of the AR signalling pathway in castration-resistant prostate cancer (CRPC) is underlined by the survival benefits observed with second-line therapies such as the anti-androgen enzalutamide and adrenal steroidogenesis inhibitor abiraterone[7-10]. Circulating steroid levels are measured to verify efficacy of hormonal treatment and have a prognostic value in patients with PC[11-13].

The assessment of circulating steroid hormones relies heavily on sensitive, specific and accurate measurement techniques, especially at castrate levels. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) combines multi-steroid profiling capabilities with superior sensitivity and specificity[14,15] over older techniques[16,17], while maintaining high sample throughput[18,19]. Multi-steroid assays for LC-MS/MS have been successfully developed in recent years to improve the detection and diagnosis of disorders associated with abnormal steroid hormones concentrations[20-23].

Steroid measurements are predominantly performed in serum samples. Previous LC-MS/MS studies have shown that steroids can be quantified reliably across different blood matrices[24-26], but there are differences observed between plasma and serum, use of glass or plastic tubes in the analytic process, or when using tubes with different stabilizing agents or with gel-separators for blood collection[24-29]. Consequently, alternative collection tubes and extraction methods must be validated before they can reliably be used for steroid profiling.

The use of CellSave Preservative (CS) tubes in population- and patient-based cohorts has grown as this specialized 'cell-stabilizing' blood collection tube preserves both circulating tumor cells (CTCs) and cell-free circulating tumor DNA (ctDNA)[30,31]. These biomarkers allow for the assessment of tumor cell genomic characteristics such as genomic instability in these patients[32]. These samples are now extensively collected in (cancer) biobanks and could potentially also be used to measure patient steroid profiles.

Mechanical blood separation methods are similarly gaining popularity in clinical chemistry due to their easy applicability compared to the use of a separation gel. The BD Vacutainer® Barricor™ is a mechanically-separated plasma (MSP) tube, and it has shown no obvious bias in steroid hormone measurements versus a gel-based plasma tube in a single study[33], but this was confined to a select number of five steroids, warranting further investigation.

In this study, we aimed to determine if plasma obtained with MSP and CS tubes is suitable for multiplex steroid profiling, which, if confirmed, would streamline biomaterial collection for ctDNA and steroid profiling in cancer patients to the use of a single tube. To this end, we performed LC-MS/MS analysis on plasma samples obtained with MSP and CS tubes in comparison to serum obtained with standard SST™ II Advance Vacutainer® tubes, collecting blood from healthy control (HC) subjects and patients with metastatic CRPC (mCRPC).

Materials and Methods

Subjects

At the Erasmus MC Cancer Institute in Rotterdam, The Netherlands, HCs and patients were included within study EMC-2016-761, which was approved by the medical ethical committee of our institute. HCs were all adult male subjects. Patients were adult subjects with mCRPC treated with ADT. Patients were eligible to start treatment with or were currently treated with second-line hormonal therapy (abiraterone with prednisone, enzalutamide or apalutamide). For all subjects the following exclusion criteria were applied: 1) an endocrine disease with altered activity of the hypothalamic-pituitary-adrenal or hypothalamic-pituitary-gonadal axis; and 2) the use of medications, excluding those used to treat PC, that interfered with circulating steroid levels or dysregulated the hypothalamic-pituitary-adrenal or hypothalamic-pituitary-gonadal axis. All subjects provided written informed consent before any study procedure.

Samples

Blood was collected from HCs and mCRPC patients in SST™ II Advance Vacutainer® (serum; BD, Franklin Lakes, NJ, USA), Vacutainer® Barricor™ (BD) and CellSave Preservative (Menarini Sillicon Biosystems Inc, Huntington Valley, PA, USA) blood collection tubes. All HC samples were processed within 6 hours after blood collection, and all mCRPC patient samples within two days. All tubes were centrifuged at 1,711g for 10 minutes at room temperature. Plasma from CS tubes was subsequently centrifuged at 12,000g for 10 minutes at 4°C. Samples were stored at -80°C until extraction.

Steroid extraction

Calibration series (0.25 ng/mL – 500 ng/mL for HC, and 0.01 ng/mL – 500 ng/mL for mCRPC) were prepared in phosphate buffered saline (PBS) with 0.1% bovine serum albumin (BSA) or in charcoal-stripped pooled human serum (Goldenwest Diagnostics, Temecula, CA, USA). Steroids investigated were 17-hydroxyprogesterone, androstenedione, cortisol, cortisone, corticosterone, dehydroepiandrosterone (DHEA), dihydrotestosterone (DHT) and testosterone. The stripped-serum calibration series was used to quantify all steroids with the exception of androstenedione, due to a high background signal in stripped serum but not in PBS-BSA. An internal standard solution was prepared in methanol/water 50/50 with equal concentrations (1 μg/mL) of the

Part 3
Steroidomics

Chapter 7
Validation of circulating steroid hormone measurements across different matrices
by liquid chromatography–tandem mass spectrometry

following deuterated steroids: 17-hydroxyprogesterone-d8, cortisol-d4, corticosterone-d8, DHEA-d6, DHT-d3, testosterone-d3. All steroids were obtained from Sigma Aldrich, UK.

400 μL of sample was transferred to hexamethyldisilazane-treated (Thermo Fisher) glass tubes (VWR, Amsterdam, The Netherlands). 20 μL of the internal standard solution was added and all samples were thoroughly vortexed. Liquid-liquid extraction was performed as previously described[34] by adding 2 mL methyl-tert butyl ether (MTBE, Sigma Aldrich, Zwijndrecht, The Netherlands) to each tube and vortexing. The samples were left at room temperature for 30 minutes to allow phase separation. The upper organic layer was transferred and the MTBE was evaporated under nitrogen at 50°C. The samples underwent a second liquid-liquid extraction with 2 mL MTBE. Samples were reconstituted in 125 μL LC-MS grade 50% methanol (CHROMASOLV, Sigma Aldrich, Zwijndrecht, The Netherlands) before measurement.

Steroid analysis by tandem mass spectrometry

Steroid concentrations were measured by mass spectrometry (Xevo TQ-XS, Waters, Milford, MA, USA) after injection of 20 μL sample volume and separation on an ACQUITY uPLC (Waters) with a Waters HSS T3 column (2.1 mm x 50 mm, 1.8 μm, Waters). The mobile phases consisted of water (A) and methanol (B) both with 0.1% formic acid and a 5-minute linear gradient was used (45-75% B) with a flow rate of 0.6 mL/min. The quantification of androgens[35] and glucocorticoids[36] in serum was previously reported, and multiple reaction monitoring settings were reported by Quanson et al.[37] and Jühlen et al.[38]. The current method represents an optimisation of these methods as steroid transitions were re-tuned for optimum response. The updated reaction settings, retention times and lower limits of quantification (LLOQ) can be found in Table S1. A chromatographic separation of the eight steroids in methanol/water can be found in Figure S1, chromatograms of the LLOQ can be found in Figure S2 and chromatograms of the internal standards and associated analyte channels, showing no interference, can be found in Figure S3.

Steroids were quantified against the linear calibration series relative to an internal standard and were only included in the final analysis if the calibration series R^2 was > 0.99 and appropriate lower limits of quantification were reached. The LLOQ was set to the lowest calibration concentration that had a clearly defined peak and a signal-to-noise ratio > 10. Samples with concentrations below the LLOQ were detectable, but quantification was less accurate. Accuracy of the method was verified in a separate measurement by quantification of Cerilliant analytical reference standards (Sigma) containing 17-hydroxyprogesterone, cortisol corticosterone and testosterone. Additionally, samples consisting of PBS-BSA or stripped serum spiked with corticosterone, cortisol, cortisone, testosterone and DHT (0.03, 0.3, 3 and 30 ng/mL) and ran on three separate days were used to calculate accuracy and precision.

Statistical analysis

LC-MS/MS raw data was processed using MassLynx (v4.1, Waters). Statistical analysis was performed using GraphPad Prism (Version 6). Normality of the data was analysed with a D'Agostino & Pearson's test. Comparisons of steroid hormone concentrations between the blood collection tubes were performed with Bland-Altman difference analysis and repeat measurements 1-way ANOVA with post hoc Dunnett's test. Correlations in hormone levels were determined by Deming regression. Group concentrations and differences are shown as mean ± SD, unless specified otherwise. P values were considered significant if < 0.05.

Results

Comparison of blood collection tubes

Analysis of the Cerilliant analytical reference samples showed quantification bias of less than 5% of the nominal concentration for the four steroids included, and accuracy in the spiked samples was > 90% (Figure S4). Baseline characteristics of the study participants are shown in Table 1. Circulating steroid levels of 8 steroids were determined by LC-MS/MS in plasma collected with CS and MSP tubes, respectively, and serum, all collected from 10 HCs and 10 mCRPC patients. Concentration ranges and deviation in CS and MSP tubes are reported in Table 2. Representative chromatograms of the different sample tubes are shown in Figure S5. Serum values below the LLOQ were detected for DHEA (HC: n = 1, mCRPC: n = 4) 17-hydroxyprogesterone (mCRPC: n = 2) and testosterone (mCRPC: n = 2) and excluded from further analysis. Low signal-to-noise ratios limited the reliability of DHT quantification, which could not be accurately quantified in two healthy control subjects as well as the mCRPC subjects.

Table 1. Characteristics of HCs and mCRPC patients

	HCs	mCRPC patients
N	10	10
Age (median (range))	32 (25 - 56)	64 (59 - 76)
Androgen deprivation therapy (n)		**10**
Leuproreline		4
Gosereline		4
Bilateral orchiectomy		2
Current second-line treatment (n)		**6**
Enzalutamide		3
Apalutamide		1
Abiraterone + Prednisone		1
Prednisone		1

HCs, healthy controls
mCRPC, metastatic castration-resistant prostate cancer

Part 3
Steroidomics

Chapter 7
Validation of circulating steroid hormone measurements across different matrices
by liquid chromatography–tandem mass spectrometry

Table 2. Relative differences in CS and MSP samples compared to serum samples
Relative differences are shown as mean ± SD. Statistical comparison was performed by repeated measurement 1-way ANOVA with post hoc Dunnett's Test. *P < 0.05, ** P < 0.01, ***P < 0.001.
Serum values below the LLOQ were detected for 17-hydroxyprogesterone (mCRPC: n = 2), cortisone (mCRPC: n = 1), DHEA (HC: n = 1, mCRPC: n = 4), DHT (HC: n = 2, mCRPC: n = 10) and testosterone (mCRPC: n = 2). Additional measurements below the LLOQ were detected in CS samples for DHEA (CRPC: n = 1) and DHT (HC: n = 1) and in MSP samples for DHT (n = 1).

	Healthy controls (n = 10)			mCRPC (n = 10)		
	Serum Range nmol/L	CS Rel. Difference mean (SD) %	MSP Rel. Difference mean (SD) %	Serum Range nmol/L	CS Rel. Difference mean (SD) %	MSP Rel. Difference mean (SD) %
Corticosterone	2.32 - 32.71	-20.08 (16.6) **	2.6 (11)	0.55 - 25.34	-12.3 (13.3) *	-2.8 (9.8)
17-hydroxyprogesterone	0.37 - 3.75	-14.4 (15.3) *	3.16 (16.26)	0.1 - 1.75	-3.2(24.1)	-2.2 (9.4)
Cortisol	149.1 - 475.7	-13.7 (6.1) ***	-0.14 (3.5)	7.43 - 504.6	-12.7 (17.9) **	-1.3 (16.1)
Cortisone	36.35 - 79.4	-10.8 (6.3) ***	-2 (9.2)	0.1 - 80.81	-17.9 (17.2) **	-1.5 (10.5)
DHEA	<LOQ – 36.14	-0.17 (18.8)	28.7 (46.6)*	<LOQ – 1.7	-17.7 (16.5)	6.0 (12.2)
Androstenedione	2.33 - 6.05	-15.85 (13.5) **	6.9 (9.5)	0.48 - 3.19	- 21.2 (14.01) **	6.34 (10.4)
Testosterone	7.47 - 17.74	-11.48 (2.8) ***	0.82 (4.7)	0.13 - 0.78	-16.9 (24.3) *	-2.8 (30.6)
DHT	<LOQ - 1.64	-11.26 (29.2)	-3.5 (30.4)	<LOQ	<LOQ	<LOQ

CS, CellSave Preservative
DHEA, dehydroepiandrosterone
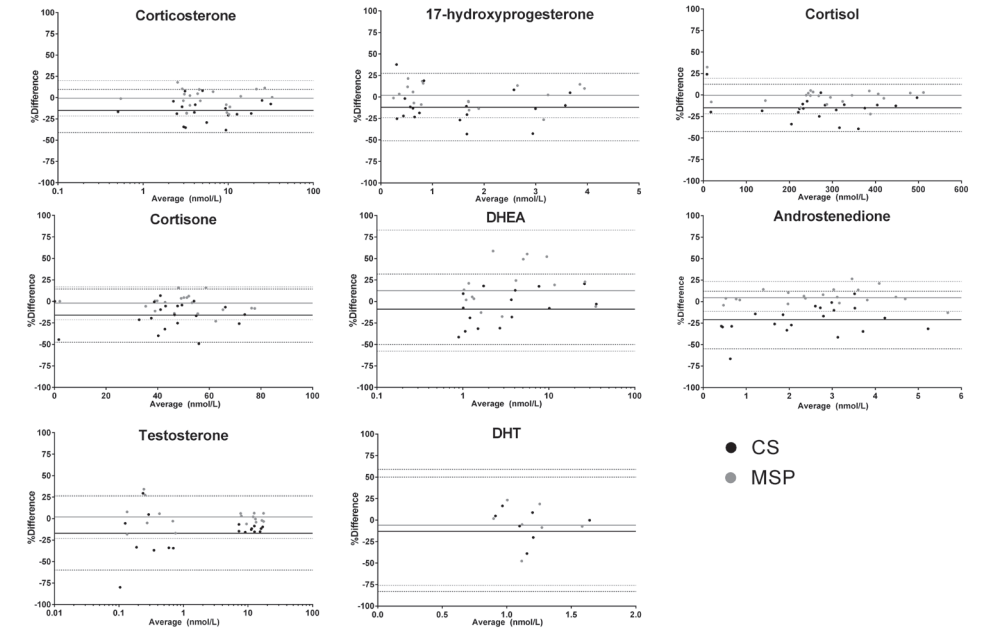DHT, 5α-dihydrotestosterone
HC, healthy control
mCRPC, metastatic castration-resistant prostate cancer
MSP, mechanically separated plasma

The values observed in MSP samples were comparable to those found in serum samples for most steroids. The only exception was DHEA, which was higher (21.3% ± 32.9%, P < 0.05) in MSP samples than in serum (Figure 1). In CS samples, significantly lower concentrations compared to serum samples were observed for corticosterone (-16.2% ± 15.2%, P < 0.001), 17-hydroxyprogesterone (-9.4% ± 19.9%, P < 0.05), cortisol (-13.2% ± 13.0%, P < 0.001), cortisone (-14.4% ± 13.1%, P < 0.001), and androstenedione (-18.4% ± 13.6%, P < 0.001). No significant differences were found for DHEA compared to serum.
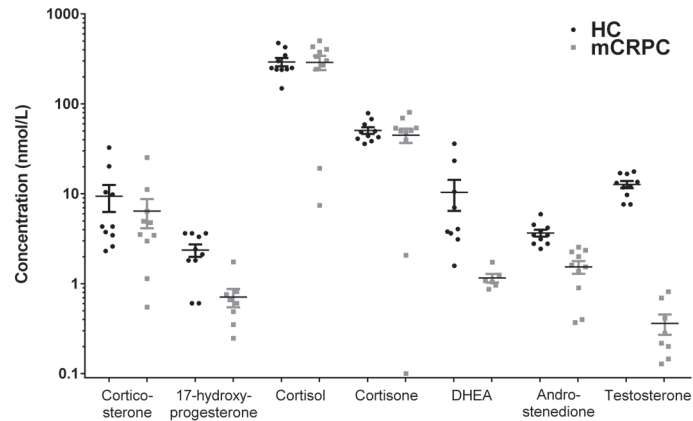

Figure 1. CS and MSP samples compared to serum samples in healthy controls and mCRPC patients
Steroid concentrations from serum, CS and MSP samples obtained from ten HCs and ten mCRPC patients measured by LC-MS/MS. Bland-Altman plots show the relative difference of CS (black) and MSP (gray) measurements compared to serum. Continuous lines show the mean difference and dotted lines show the upper- and lower limits of the 95% confidence interval.
Abbreviations: CS – CellSave Preservative, DHEA – dehydroepiandrosterone, HC – healthy control, mCRPC – metastatic castration-resistant prostate cancer, MSP – mechanically seperated plasma

Similar steroid concentrations between HC and mCRPC subjects were observed for corticosterone, cortisol and cortisone (Figure 2). Lower concentrations were observed for 17-hydroxyprogesterone, androstenedione, DHEA and testosterone in mCRPC subjects. This was likely due to a combination of castration (testosterone) and age-related effects, as the mCRPC subjects were older than the HC subjects.

Circulating androgen concentration in mCRPC patients are > 10-fold lower than in healthy men due to ADT, requiring highly sensitive techniques to accurately measure residual androgens. Therefore, the calibration series was expanded to include lower concentrations (0.01 – 0.25 ng/mL) to allow quantification of castrate testosterone levels. Accurate quantification at low concentration was achieved, with an analytic LLOQ for testosterone of 0.1 nmol/L. Similar to the other steroids, lower testosterone concentrations compared to serum were detected in CS samples, but not in MSP samples, at normal HC concentrations (-11.5% ± 2.8%, P < 0.001) and at castrate concentrations (-16.9% ± 24.3%, P < 0.05) (Figure 1). Low signal-to-noise ratios limited the reliability of DHT quantification which could not be accurately quantified in the mCRPC subjects with our assay.

Part 3
Steroidomics

Chapter 7
Validation of circulating steroid hormone measurements across different matrices
by liquid chromatography–tandem mass spectrometry

Figure 2. Circulating steroid concentrations in serum of HCs and mCRPC subjects
Steroid hormone concentrations in healthy controls (n = 10) and mCRPC patients (n = 10). Four patients received no additional treatment, four received antiandrogens (enzalutamide (n = 3) or apalutamide (n = 1)), one received abiraterone and prednisone and one received prednisone. Line and error represent mean ± SEMs.
Abbreviations: DHEA – dehydroepiandrosterone, mCRPC – metastatic castration-resistant prostate cancer, HC – Healthy control

Correlation of steroid measurements between matrices

The correlations between results obtained in MSP and CS samples compared to those results obtained in serum were independently determined by Deming regression (Figure 3). For DHT analysis only HC samples were included. Corticosterone, DHEA and testosterone were normally distributed after log-transformation. Significant correlations (all P < 0.001) between both matrices and serum was observed for 17-hydroxyprogesterone, androstenedione, corticosterone, cortisol, cortisone, DHEA and testosterone. The analysis also revealed a poor correlation for DHT between CS and serum (R^2 = 0.60) and between MSP and serum (R^2 = 0.45), whereas steroid concentrations measured in MSP and CS samples, respectively, correlated more closely (R^2 = 0.89, y = 0.89x + 0.15) (data not shown).
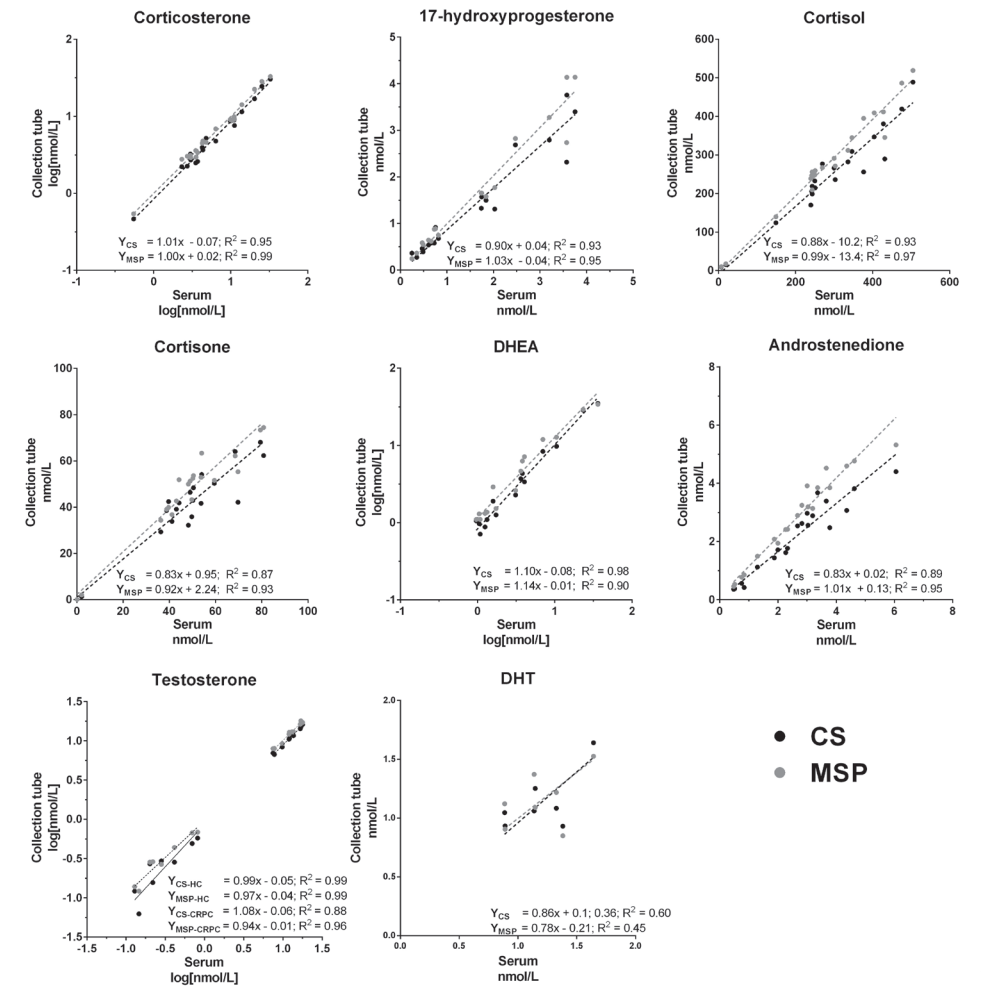
Figure 3. Correlations between CS samples and MSP samples with serum samples
Deming regression analysis between measurements from serum and CS samples and between serum and MSP samples in HC and mCRPC patients combined. Corticosterone, DHEA and testosterone values did not pass normality testing (D'Agostino and Pearson), but did so after logarithmic transformation. Regression equations for testosterone are presented separately for HCs and mCRPC patients due to the bimodal distribution resulting from ADT. For DHT, only HC values were used as values in CRPC patients were below the LLOQ.
Abbreviations: ADT – androgen deprivation therapy, CS – CellSave Preservative, DHEA – dehydroepiandrosterone, DHT – dihydrotestosterone, HC – healthy control, LLOQ – lower limit of quantification, mCRPC – metastatic castration-resistant prostate cancer, MSP – mechanically separated plasma

Part 3
Steroidomics

Chapter 7
Validation of circulating steroid hormone measurements across different matrices
by liquid chromatography–tandem mass spectrometry

Discussion

In this study, we investigated whether MSP- or CS-derived plasma, samples that are abundantly present in biobanks obtained from PC patients, are suitable for multiplex steroid profiling by LC-MS/MS. We compared them to the current standard collection method using serum samples collected with SST™ II Advance Vacutainers®. We showed that measurements in MSP are equal to serum. When collecting plasma in CS tubes, steroid concentrations were 8-21% lower than measurements using serum or MSP. The Guideline on Bioanalytical Method Validation (2011) of the European Medicines Agency advises that accuracy must be within 15% of the nominal concentration and within 20% at the LLOQ[39]. The decreases observed for most steroids in CS plasma, using serum as a reference, were within this acceptable range (Table 2). Therefore, we conclude that CS plasma samples are suitable for steroid profiling, which can be combined with analysis of CTCs or ctDNA. However, caution is advised when interpreting results obtained in CS plasma samples against reference values that were obtained in serum, and direct comparison to samples collected in other tubes should be avoided. Nevertheless, these findings may reduce patient burden and open up the possibility to undertake detailed steroid profiling of large collections of biomaterial already collected for ctDNA analysis.

In this study, most steroids could be quantified accurately within the range of the calibration series. Most of the Δ4-steroids, such as cortisol or testosterone, ionise more easily and so can be accurately quantified at low levels (0.03 – 0.15 nM). This allows for the quantification of testosterone in mCRPC patients (typically < 0.5 nM)[11]. Δ5-steroids such as DHEA and saturated steroids such as DHT are poor ionisers and quantification at lower concentrations is beyond the sensitivity of the mass spectrometer[22]. Like testosterone, DHT levels are suppressed in castrated patients and the concentrations in these patients could not be accurately assessed. DHEA levels in men decline with age[40], and the mCRPC subjects in this study were older than the HC subjects. Consequently, values below the LLOQ were detected in four mCRPC patients and were excluded from the analysis. DHEA has been quantified in smaller serum volume (≤ 100 µL) to reduce ion suppression, but that would decrease sensitivity for other steroids, such as testosterone or DHT, which carried greater significance in our analysis. Alternatively, derivatization, for example to form an oxime, increases ionisability, and therefore sensitivity allowing low level quantification of Δ5- and 5α-reduced steroids[22]. However, the derivatization method is not suited for routine clinical diagnostic measurements due to increased sample preparation time and cost. In addition to this, fragmentation produces multiple derivatives for some steroids adding to the complexity of the analysis.

Conjugated steroids, such as DHEA-sulphate (DHEA-S) or androsterone-glucuronide were not included in the method. While DHEA-S is the most abundant steroid in the circulation, it can only be converted into active androgens after deconjugation into DHEA. Conjugated steroids downstream of testosterone and DHT are rapidly excreted with the urine, and quantification in serum may be less reliable compared to quantification in urine. Additionally, measurement of

these steroids would further add to the complexity of the analysis and decrease throughput, as additional time-consuming enzymatic digestion and purification steps would be necessary to obtain the deconjugated steroids.

Matrix effects and cross-reactivity are established sources of interference in steroid hormone profile studies with immunoassay and LC-MS/MS[24-29]. Previous studies have identified the type of blood sample tube as a potential source of interference. MSP tubes like BD Vacutainer® Barricor™ utilise mechanical separation of plasma, which makes them easily applicable, but the accuracy of steroid hormone measurements in these tubes has not been fully validated yet. Our study shows that multiplex steroid quantification in MSP samples is comparable to the serum collected with the reference tube, in line with a previous study that detected no bias versus gel-based plasma tubes using an immunoassay platform[33]. The only significant difference between steroid measurements in MSP and serum related to DHEA. While Fournier et al.[33] reported comparable DHEA concentrations between MSP and plasma tubes, it is possible than minor differences in deconjugation compared to serum could affect DHEA concentrations. Quantification of DHEA at low concentrations with LC-MS/MS remains a challenge as its structure contributes to poor MS ionisation[22,41]. This challenge could be overcome by using a more sensitive mass spectrometer or with the use of derivatization. Prior to using MSP tubes for clinical studies, however, full validation considering accuracy, precision and recovery with larger sample size is recommended.

CS tubes are optimized for the measurement of circulating nucleic acids or tumor cells[30,31]. The use of this matrix for liquid biopsies has increased exponentially over the last years due to the successful genomic characterization of CTCs or free circulating nucleic acids, but no studies have investigated if plasma from CS tubes are suitable for quantification of circulating steroid hormone levels. Our experiments indicate that steroid measurements in CS samples are affected by a mild bias, which resulted in an approximate 8-21% decrease compared to serum. We observed similar effects in both HCs and mCRPC subjects. CS tubes contain 300 µL of Na_2EDTA anticoagulant as well as an undisclosed preservative to stabilise cells in the sample. Due to the presence of the Na_2EDTA there may be an inherent dilution of the sample, which amounts to approximately 3-4% on a 7.5-10 mL volume. This dilution factor is insufficient to account for the difference in circulating steroid hormone levels however, and it is possible that other factors also contribute towards the observed difference. Sample processing time is another potential factor that may affect steroid quantification. While the CS user manual indicates that samples may be stored for up to 96 hours for genomic analysis, the concentration of certain steroids, such as androstenedione, may not remain stable over longer storage periods[25].

This decrease was observed across a variety of different polarity steroids with different molecular weights. It is therefore unlikely that the preservative co-elutes with one of the steroids and suppresses the MS signal. Either there is an unidentified contaminant in the tubes which affects all steroids or the steroids themselves are being retained/bound to the tube itself. Steroids have been long recognised to bind to plastics[42], which may contribute to the lower values in the CS samples.

Part 3
Steroidomics

Chapter 7
Validation of circulating steroid hormone measurements across different matrices
by liquid chromatography–tandem mass spectrometry

Currently, CS tubes are most commonly utilized in oncological studies to obtain CTCs and ctDNA. Hormonal treatment options in breast- and prostate cancer involve potent suppression of oestrogens or androgens. For example, inhibition of testicular steroidogenesis by ADT will typically lower testosterone levels by > 90%[11]. Interpretation of such changes is unlikely to be affected by the difference observed in CS samples. Especially within the context of a single study the relative difference should affect all samples identically as long as a single collection tube is used. As such, the observed difference is acceptable for most clinical purposes, including the use of CS samples for steroid profiling in PC patients. Combined investigation of steroid profiles and analysis of CTCs or ctDNA will decrease costs and reduce patient burden.

In conclusion, MSP samples are suitable for steroid quantification, including castrate range of androgens. Similarly, CS samples are suitable for steroid measurements, although there is a consistent bias of 8-21% lower steroid hormone levels. Therefore, all samples in a research study should be collected in the same sample tubes to avoid potential variation due to effects from the tubes themselves.

Acknowledgements

References

1. Sharifi N, Gulley JL, Dahut WL. Androgen deprivation therapy for prostate cancer. JAMA 2005;294(2):238-44. (In eng). DOI: 294/2/238 [pii]
10.1001/jama.294.2.238.
2. Linja MJ, Savinainen KJ, Saramaki OR, Tammela TL, Vessella RL, Visakorpi T. Amplification and overexpression of androgen receptor gene in hormone-refractory prostate cancer. Cancer Res 2001;61(9):3550-5. (In eng) (https://www.ncbi.nlm.nih.gov/pubmed/11325816).
3. Taylor BS, Schultz N, Hieronymus H, et al. Integrative genomic profiling of human prostate cancer. Cancer Cell 2010;18(1):11-22. (In eng). DOI: S1535-6108(10)00238-2 [pii]
10.1016/j.ccr.2010.05.026.
4. Stanbrough M, Bubley GJ, Ross K, et al. Increased expression of genes converting adrenal androgens to testosterone in androgen-independent prostate cancer. Cancer Res 2006;66(5):2815-25. (In eng). DOI: 66/5/2815 [pii]
10.1158/0008-5472.CAN-05-4000.
5. Montgomery RB, Mostaghel EA, Vessella R, et al. Maintenance of intratumoral androgens in metastatic prostate cancer: a mechanism for castration-resistant tumor growth. Cancer Res 2008;68(11):4447-54. (In eng). DOI: 68/11/4447 [pii]
10.1158/0008-5472.CAN-08-0249.
6. Mohler JL, Gregory CW, Ford OH, 3rd, et al. The androgen axis in recurrent prostate cancer. Clin Cancer Res 2004;10(2):440-8. (In eng). DOI: 10.1158/1078-0432.ccr-1146-03.
7. Scher HI, Fizazi K, Saad F, et al. Increased survival with enzalutamide in prostate cancer after chemotherapy. N Engl J Med 2012;367(13):1187-97. (In eng). DOI: 10.1056/NEJMoa1207506.
8. de Bono JS, Logothetis CJ, Molina A, et al. Abiraterone and increased survival in metastatic prostate cancer. N Engl J Med 2011;364(21):1995-2005. (In eng). DOI: 10.1056/NEJMoa1014618.
9. Fizazi K, Tran N, Fein L, et al. Abiraterone plus Prednisone in Metastatic, Castration-Sensitive Prostate Cancer. N Engl J Med 2017;377(4):352-360. (In eng). DOI: 10.1056/NEJMoa1704174.
10. Davis ID, Martin AJ, Stockler MR, et al. Enzalutamide with Standard First-Line Therapy in Metastatic Prostate Cancer. N Engl J Med 2019;381(2):121-131. (In eng). DOI: 10.1056/NEJMoa1903835.
11. Snaterse G, Visser JA, Arlt W, Hofland J. Circulating steroid hormone variations throughout different stages of prostate cancer. Endocr Relat Cancer 2017;24(11):R403-R420. (In eng). DOI: ERC-17-0155 [pii]
10.1530/ERC-17-0155.
12. Sakamoto S, Maimaiti M, Xu M, et al. Higher serum testosterone levels associated with favorable prognosis in enzalutamide- and abiraterone-treated castration-resistant prostate cancer. J Clin Med 2019;8(4) (Article) (In English). DOI: 10.3390/jcm8040489.
13. Attard G, Reid AH, A'Hern R, et al. Selective inhibition of CYP17 with abiraterone acetate is highly active in the treatment of castration-resistant prostate cancer. J Clin Oncol 2009;27(23):3742-8. (In eng). DOI: JCO.2008.20.0642 [pii]
10.1200/JCO.2008.20.0642.
14. Wang C, Catlin DH, Demers LM, Starcevic B, Swerdloff RS. Measurement of total serum testosterone in adult men: comparison of current laboratory methods versus liquid chromatography-tandem mass spectrometry. J Clin Endocrinol Metab 2004;89(2):534-43. (In eng). DOI: 10.1210/jc.2003-031287.
15. Taieb J, Mathian B, Millot F, et al. Testosterone measured by 10 immunoassays and by isotope-dilution gas chromatography-mass spectrometry in sera from 116 men, women, and children. Clin Chem 2003;49(8):1381-95. (In eng) (https://www.ncbi.nlm.nih.gov/pubmed/12881456).
16. Handelsman DJ, Newman JD, Jimenez M, McLachlan R, Sartorius G, Jones GR. Performance of direct estradiol immunoassays with human male serum samples. Clin Chem 2014;60(3):510-7. (In eng). DOI: clinchem.2013.213363 [pii]
10.1373/clinchem.2013.213363.
17. Krasowski MD, Drees D, Morris CS, Maakestad J, Blau JL, Ekins S. Cross-reactivity of steroid hormone immunoassays: clinical significance and two-dimensional molecular similarity prediction. BMC Clin Pathol 2014;14:33. (In eng). DOI: 10.1186/1472-6890-14-33
1472-6890-14-33 [pii].
18. Taylor AE, Keevil B, Huhtaniemi IT. Mass spectrometry and immunoassay: how to measure steroid hormones today and tomorrow. Eur J Endocrinol 2015;173(2):D1-12. (In eng). DOI: EJE-15-0338 [pii]
10.1530/EJE-15-0338.

Part 3
Steroidomics

Chapter 7
Validation of circulating steroid hormone measurements across different matrices
by liquid chromatography–tandem mass spectrometry

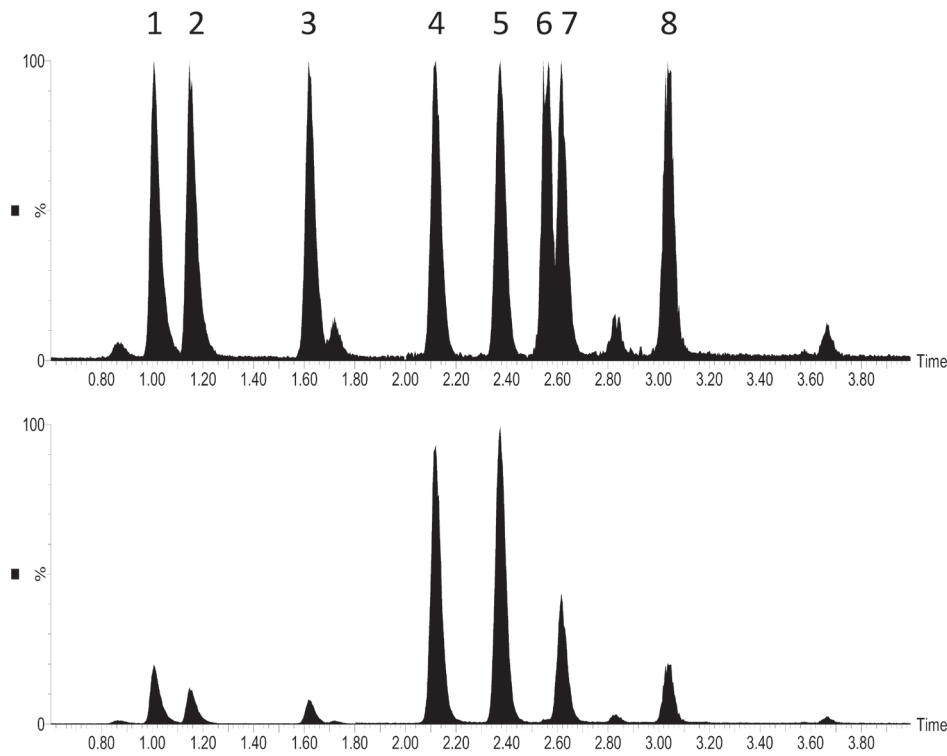19.   Monaghan PJ, Keevil BG, Trainer PJ. The use of mass spectrometry to improve the diagnosis and the management of the HPA axis. Rev Endocr Metab Disord 2013;14(2):143-57. (In eng). DOI: 10.1007/s11154-013-9240-1.

20.   Arlt W, Lang K, Sitch AJ, et al. Steroid metabolome analysis reveals prevalent glucocorticoid excess in primary aldosteronism. JCI Insight 2017;2(8) (In eng). DOI: 93136 [pii]
10.1172/jci.insight.93136.

21.   Arlt W, Biehl M, Taylor AE, et al. Urine steroid metabolomics as a biomarker tool for detecting malignancy in adrenal tumors. J Clin Endocrinol Metab 2011;96(12):3775-84. (In eng). DOI: jc.2011-1565 [pii]
10.1210/jc.2011-1565.

22.   Hakkinen MR, Murtola T, Voutilainen R, et al. Simultaneous analysis by LC-MS/MS of 22 ketosteroids with hydroxylamine derivatization and underivatized estradiol from human plasma, serum and prostate tissue. J Pharm Biomed Anal 2019;164:642-652. (In eng). DOI: S0731-7085(18)32517-2 [pii]
10.1016/j.jpba.2018.11.035.

23.   Storbeck KH, Schiffer L, Baranowski ES, et al. Steroid metabolome analysis in disorders of adrenal steroid biosynthesis and metabolism. Endocr Rev 2019 (In eng). DOI: pii: er.2018-00262. doi: 10.1210/er.2018-00262.

24.   Coburn SB, Stanczyk FZ, Falk RT, et al. Comparability of serum, plasma, and urinary estrogen and estrogen metabolite measurements by sex and menopausal status. Cancer Causes Control 2019;30(1):75-86. (In eng). DOI: 10.1007/s10552-018-1105-1
10.1007/s10552-018-1105-1 [pii].

25.   Hepburn S, Wright MJ, Boyder C, et al. Sex steroid hormone stability in serum tubes with and without separator gels. Clin Chem Lab Med 2016;54(9):1451-9. (In eng). DOI: 10.1515/cclm-2015-1133
/j/cclm.ahead-of-print/cclm-2015-1133/cclm-2015-1133.xml [pii].

26.   Raff H, Sluss PM. Preanalytical issues for testosterone and estradiol assays. Steroids 2008;73(13):1297-304. (In eng). DOI: S0039-128X(08)00143-8 [pii]
10.1016/j.steroids.2008.05.005.

27.   Morovat A, James TS, Cox SD, et al. Comparison of Bayer Advia Centaur immunoassay results obtained on samples collected in four different Becton Dickinson Vacutainer tubes. Ann Clin Biochem 2006;43(Pt 6):481-7. (In eng). DOI: 10.1258/000456306778904713.

28.   Smets EM, Dijkstra-Lagemaat JE, Blankenstein MA. Influence of blood collection in plastic vs. glass evacuated serum-separator tubes on hormone and tumor marker levels. Clin Chem Lab Med 2004;42(4):435-9. (In eng). DOI: 10.1515/CCLM.2004.076.

29.   Schouwers S, Brandt I, Willemse J, et al. Influence of separator gel in Sarstedt S-Monovette(R) serum tubes on various therapeutic drugs, hormones, and proteins. Clin Chim Acta 2012;413(1-2):100-4. (In eng). DOI: S0009-8981(11)00486-4 [pii]
10.1016/j.cca.2011.08.037.

30.   Rothwell DG, Smith N, Morris D, et al. Genetic profiling of tumors using both circulating free DNA and circulating tumor cells isolated from the same preserved whole blood sample. Mol Oncol 2016;10(4):566-74. (In eng). DOI: S1574-7891(15)00205-7 [pii]
10.1016/j.molonc.2015.11.006.

31.   van Dessel LF, Beije N, Helmijr JC, et al. Application of circulating tumor DNA in prospective clinical oncology trials - standardization of preanalytical conditions. Mol Oncol 2017;11(3):295-304. (In eng). DOI: 10.1002/1878-0261.12037.

32.   van Dessel LF, van Riet J, Smits M, et al. The genomic landscape of metastatic castration-resistant prostate cancers reveals multiple distinct genotypes with potential clinical impact. Nat Commun 2019;10(1):5251. DOI: 10.1038/s41467-019-13084-7.

33.   Fournier JE, Northrup V, Clark C, et al. Evaluation of BD Vacutainer(R) Barricor blood collection tubes for routine chemistry testing on a Roche Cobas(R) 8000 Platform. Clin Biochem 2018;58:94-99. (In eng). DOI: S0009-9120(17)31202-X [pii]
10.1016/j.clinbiochem.2018.06.002.

34.   O'Reilly MW, Taylor AE, Crabtree NJ, et al. Hyperandrogenemia predicts metabolic phenotype in polycystic ovary syndrome: the utility of serum androstenedione. J Clin Endocrinol Metab 2014;99(3):1027-36. (In eng). DOI: 10.1210/jc.2013-3399.

35.   O'Reilly MW, Kempegowda P, Jenkinson C, et al. 11-Oxygenated C19 Steroids Are the Predominant Androgens in Polycystic Ovary Syndrome. J Clin Endocrinol Metab 2017;102(3):840-848. (In eng). DOI: 10.1210/jc.2016-3285.

36.   Prete A, Taylor AE, Bancos I, et al. Prevention of Adrenal Crisis: Cortisol Responses to Major Stress Compared to Stress Dose Hydrocortisone Delivery. J Clin Endocrinol Metab 2020;105(7). DOI: 10.1210/clinem/dgaa133.

37.   Quanson JL, Stander MA, Pretorius E, Jenkinson C, Taylor AE, Storbeck KH. High-throughput analysis of 19 endogenous androgenic steroids by ultra-performance convergence chromatography tandem mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 2016;1031:131-138. (In eng). DOI: S1570-0232(16)30486-X [pii]
10.1016/j.jchromb.2016.07.024.

38.   Juhlen R, Idkowiak J, Taylor AE, et al. Role of ALADIN in human adrenocortical cells for oxidative stress response and steroidogenesis. PLoS One 2015;10(4):e0124582. DOI: 10.1371/journal.pone.0124582.

39.   Agency EM. Guideline on Bioanalytical Method Validation. 2011.

40.   Kaufman JM, Vermeulen A. The decline of androgen levels in elderly men and its clinical and therapeutic implications. Endocr Rev 2005;26(6):833-76. (In eng). DOI: er.2004-0013 [pii]
10.1210/er.2004-0013.

41.   Keski-Rahkonen P, Huhtinen K, Poutanen M, Auriola S. Fast and sensitive liquid chromatography-mass spectrometry assay for seven androgenic and progestagenic steroids in human serum. J Steroid Biochem Mol Biol 2011;127(3-5):396-404. (In eng). DOI: S0960-0760(11)00129-4 [pii]
10.1016/j.jsbmb.2011.06.006.

42.   Bruning PF, Jonker KM, Boerema-Baan AW. Adsorption of steroid hormones by plastic tubing. J Steroid Biochem 1981;14(6):553-5. (In eng). DOI: 10.1016/0022-4731(81)90029-7.

Part 3
Steroidomics

Chapter 7
Validation of circulating steroid hormone measurements across different matrices
by liquid chromatography–tandem mass spectrometry

Supplementary information

Figure S1. Chromatograph of steroids

Chromatographic separation of 8 steroids in methanol/water. (1) cortisone, (2) cortisol, (3) corticosterone, (4) androstenedione, (5) testosterone, (6) DHEA, (7) 17-hydroxyprogesterone and (8) DHT. Top chromatogram shows separation with normalized intensity, bottom graph with actual response, note in lower chromatogram DHEA signal is indistinguishable from 17-hydroxyprogesterone due to poor ionization.
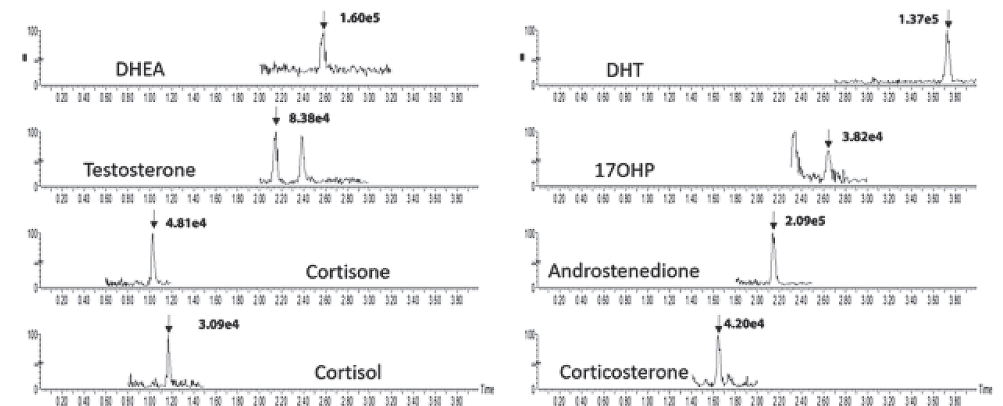
Figure S2. Chromatograms generated from a single mass transition for each steroid at the limit of quantification

Steroid chromatograms obtained at the lower limits of quantification (LLOQ) in spiked stripped serum. Arrows indicate the corresponding peak, and signal intensity is noted.
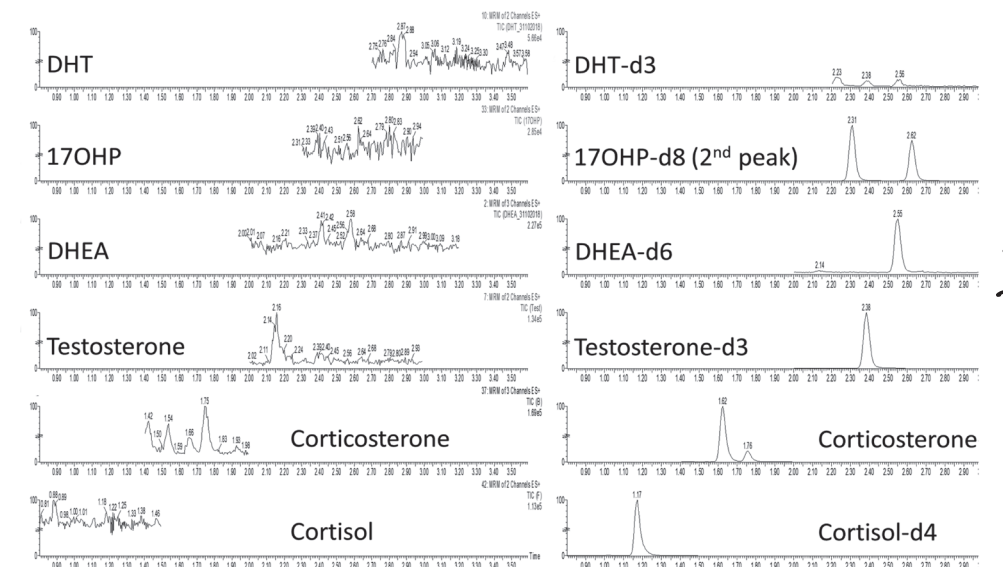
Figure S3. Chromatograms for internal standards (right) and corresponding chromatograms for their associated steroid (left) showing no significant cross-reactivity

Part 3
Steroidomics

Chapter 7
Validation of circulating steroid hormone measurements across different matrices
by liquid chromatography–tandem mass spectrometry

Figure S4. Accuracy and precision
Quantification of eight analytical reference samples revealed accuracy of > 95% for the steroids present in the samples, shown below. The dashed lines indicate the mean deviation from the nominal concentra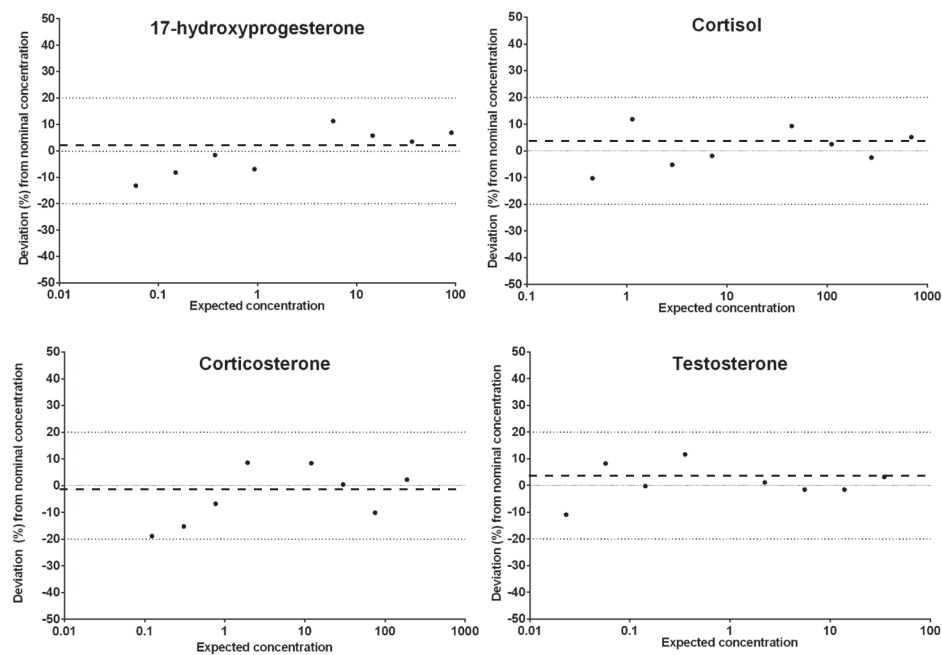tion. Analysis of spiked control samples revealed that an accuracy of ≥ 90% was achieved for T and cortisone at all concentrations, and ≥ 90% for cortisol, corticosterone and DHT at 0.3, 3 and 30 ng/mL. Inter-assay precision was calculated based on 4-6 biological replicates ran on three separate days, and the residual standard deviation was < 10% for all five steroids.
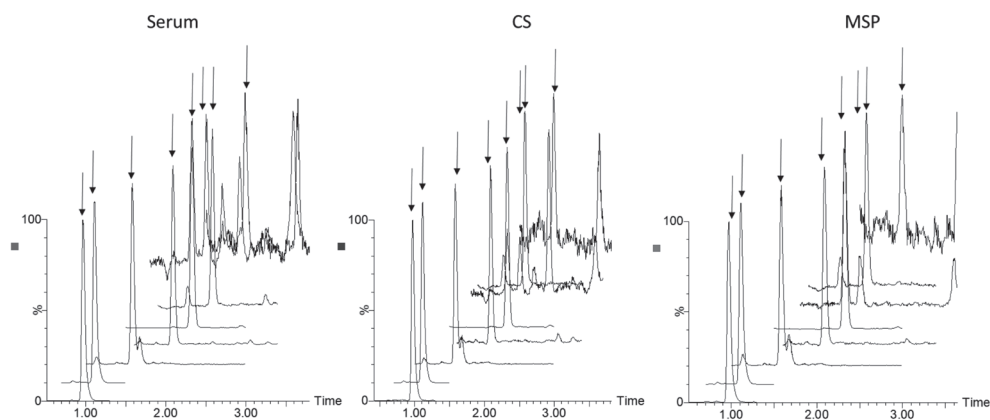
Figure S5. Chromatograms in serum, CS and MSP sample
Chromatograms of steroid quantification in serum, CellSave Preservative (CS) and mechanically-separated plasma (MSP) samples obtained from a healthy control subject. Mass transitions are overlaid with an offset of 10%. Largest peak in each chromatogram is normalised to 100%. Steroid elution order (left to right) cortisone, cortisol, corticosterone, androstenedione, testosterone, DHEA, 17OHP and DHT.

Table S1. Overview of mass transitions and retention times of measured steroids

Steroid	Mass transition	Retention time	LOQ
17-hydroxyprogesterone	331.1 > 96.9	2.62	0.30
17-hydroxyprogesterone-d8	339.1 > 100.0	2.57	-
Androstenedione	287.1 > 96.9	2.13	0.35
Corticosterone	347.1 > 121.0	1.65	0.29
Corticosterone-d8	355.1 > 124.9	1.62	-
Cortisol	363.1 > 90.9	1.18	0.14
Cortisol-d4	367.2 > 121.1	1.18	-
Cortisone	361.0 > 121.0	1.03	0.14
DHEA	289.2 > 213.1	2.57	0.86
DHEA-d6	277.2 > 219.1	2.53	-
DHT	291.1 > 159.1	3.04	0.86
DHT-d3	294.1 > 259.1	3.00	-
Testosterone	289.1 > 96.9	2.38	0.10
Testosterone-d3	292.1 > 96.9	2.35	-

Chapter 8

11-Ketotestosterone is the predominant androgen in castration-resistant prostate cancer patients

Chapter 8

Gido Snaterse*, Lisanne F. van Dessel*, Job van Riet, Angela E. Taylor, Michelle van der Vlugt-Daane, Paul Hamberg, Ronald de Wit, Jenny A. Visser, Wiebke Arlt, Martijn P. Lolkema, Johannes Hofland
* These authors contributed equally
Submitted

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients

Abstract

Background

Androgen receptor (AR) signaling plays an important role in metastatic castration-resistant prostate cancer (CRPC). Studies have identified adrenal-derived 11-ketotestosterone (11KT) as a potent AR agonist, but it is unknown if 11KT is present at physiologically relevant concentrations in CRPC patients.

Objective

To investigate the steroid hormone profiles of CRPC patients at baseline, during treatment and after clinical progression.

Design, setting, and participants

Twenty-nine metastatic CRPC patients who intended to start a new line of systemic therapy were included. Plasma samples were obtained at baseline, during treatment and after progression. Metastatic tumor biopsy samples were obtained at baseline.

Interventions

Treatment with enzalutamide, apalutamide, docetaxel with prednisone or cabazitaxel with prednisone.

Measurements

Circulating steroid concentrations were determined by multi-steroid profiling liquid chromatograph tandem mass spectrometry (LC-MS/MS). Next-generation sequencing and RNA-sequencing were performed on metastatic tumor biopsy samples.

Results and limitations

11KT was the most abundant circulating androgen in CRPC patients (median 0.39 nmol/L, range: 0.03–2.39 nmol/L), constituting 66% (IQR 54-80%) of the total androgen (TA) pool. Treatment with glucocorticoids reduced 11KT by 84% (49-89%), testosterone by 68% (38-79%) and 11KT precursors steroids by 66-92%. Circulating TA concentrations at baseline were associated with progression-free survival and with differential expression of genes within the tumors, which included AR-regulated genes. The small sample size of this study is a limitation.

Conclusions

This study has identified 11KT, a potent AR agonist, as the major circulating androgen in CRPC patients, and therefore as one of the potential drivers of AR activation in CRPC. Assessment of androgen status should be extended to include 11KT, as current clinical approaches likely underestimate androgen abundance in CRPC patients.

Patient summary

11KT is the most abundant circulating androgen in CRPC patients, and androgen measurements should be extended to include 11KT.

Abbreviations

General

ADT	Androgen deprivation therapy
ANOVA	Analysis of variance
AR	Androgen receptor
CRPC	Castration-resistant prostate cancer
eGC	Exogenous glucocorticoids
HPA	Hypothalamus-pituitary-adrenal
IQR	Interquartile range
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LOQ	Limit of quantification
OT	On treatment
PC	Prostate cancer
PD	Progressive disease
PFS	Progression-free survival
TA	Total androgens
WGS	Whole-genome sequencing

Steroids

11KA4	11-ketoandrostenedione
11KT	11-ketotestosterone
11OHA4	11β-hydroxyandrostenedione
11OHT	11β-hydroxytestosterone
DHT	Dihydrotestosterone
T	Testosterone

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients

Introduction

Targeting the androgen receptor (AR) pathway through androgen deprivation therapy (ADT) is the mainstay of treatment in metastatic prostate cancer (PC)[1]. Eventually, most tumors will evolve from hormone-sensitive to castration-resistant prostate cancer (CRPC) and show progression despite suppressed testosterone (T) levels. The continued importance of the AR pathway in tumor growth and progression has been underlined by the efficacy of novel drugs targeting the AR pathway[2-5]. AR upregulation[6], increased intratumoral conversion of adrenal androgen precursors[7-9], non-canonical dihydrotestosterone (DHT) synthesis[10], and downregulation of androgen inactivating enzymes[11,12] may all contribute to AR pathway reactivation.

In recent years, novel androgenic steroids have been identified with significant AR activation potential[13]. The adrenal-derived steroid 11-ketotestosterone (11KT) is of particular interest, as it activates the AR at concentrations comparable to T and DHT[13,14,15]. In healthy adult men, circulating T concentrations exceed those of 11KT[16]. Of note, while T decreases with ageing in men, 11KT increases[17]. During ADT, gonadal steroidogenesis is inhibited and T concentrations are typically below 0.5 nmol/L[18], which is lower than 11KT concentrations in healthy men (0.77±0.16 nmol/L)[16].

We hypothesized that due to their adrenal origin, 11KT may persist after castration and may therefore exceed the residual concentrations of T and DHT in CRPC patients. Androgen abundance after castration is associated with outcome and predicts response to AR pathway inhibition[19,20]. Thus, persistence of a previously overlooked, potent androgen class would be of major clinical significance in CRPC patients.

In this study, we assessed the plasma steroid profile of 29 patients before, during and after treatment with second-line therapies for CRPC. We report the abundance of circulating active androgens in these patients as well as the effects imposed on the steroid metabolome by treatment with exogenous glucocorticoids (eGC). Finally, we show that 11-oxygenated androgen levels may have a potential prognostic value in CRPC, linked to differential gene expression in tumor biopsies.

Materials and Methods

Patients and samples

From April 2016 onwards, metastatic CRPC patients who continued ADT and intended to start a new line of systemic therapy were included in the CIRCUS study (Netherlands Trial Registry ID: NL5625), which was approved by the medical ethics board of the Erasmus Medical Center (MEC-2016-081). Metastatic disease and progression were defined according to the PCWG2 and/or RECIST 1.1 criteria[21,22]. The treatments included in this study were antiandrogens, docetaxel with prednisone or cabazitaxel with prednisone (Figure 1). Concurrent participation in the CPCT-02 study (NCT01855477) was required to obtain a tumor biopsy before start of therapy. All patients provided written informed consent before any study procedure; this involved blood collection at baseline, on treatment (OT), at progressive disease (PD), and collection of clinical data. Blood was collected in three CellSave preservative tubes (Menarini Sillicon Biosystems Inc, Huntington Valley, PA, USA) every three to four weeks. Plasma was isolated and stored as previously described[23,24].

Measurement of circulating steroids

The extraction and quantification of steroids was previously described[24-28]. Briefly, steroids were extracted from 400 μL plasma by liquid-liquid extraction. Multi-steroid profiling was performed by liquid-chromatography tandem mass spectrometry (LC-MS/MS; Xevo TQ-XS, Waters, Milford, MA, USA) after separation on an ACQUITY UPLC (Waters) with UPLC HSS T3 column (21 mm x 50 mm, 1.8 μm, Waters). A representative chromatogram is shown in Figure S1 and an overview of included steroids, their mass transitions, internal standards and limits of quantification (LOQ) can be found in Table S1. The Supplementary Materials and Methods section contains additional information regarding extraction, accuracy, precision and values below LOQ.

Whole-genome sequencing and RNA sequencing

As part of the CPCT-02 study (NCT01855477), whole-genome sequencing (WGS) was performed by the Hartwig Medical foundation on 180 metastatic CRPC biopsy samples obtained at baseline and matched blood samples (WGS only) as previously described[29]. Briefly, paired-end (2x 150 bp) sequencing of DNA-libraries on the HiSeq X Ten Illumina system (Illumina, San Diego, California, USA) was performed for WGS. Further processing of the data is detailed in the Supplementary Materials and Methods.

RNA-sequencing was performed according to the manufacturer protocols using a minimum of 100 ng total RNA input. Total RNA was extracted using the Qiagen QIAsymphony kit (Qiagen, FRITSCH GmbH, Idar-Oberstein, Germany). Paired-end sequencing of (m)RNA was performed on the Illumina NextSeq 550 platform (2x 75 bp; Illumina) and NovaSeq 6000 platform (2x 150 bp; Illumina). Downstream data processing and analysis is detailed in the Supplementary Materials and Methods. Briefly, CRPC tumor biopsies from the entire CPCT-02 study (n = 180) were used to identify biopsy-site specific genes. Subsequently, an untargeted approach was used to analyze gene expression across TA concentrations in biopsy samples of patients included in this study only (n = 15), excluding biopsy-site related genes. Genomic alterations and expression of steroid hormone receptors and genes involved in steroid metabolism was assessed using a targeted approach[30].

Statistical analysis

Masslynx (v4.1, Waters) was used to process LC-MS/MS data. Statistical analyses were performed with Graphpad Prism (version 6.01, La Jolla, California, USA), SPSS (version 26, IBM Corp., Armonk, New York, USA) and R (version 3.6.1, Vienna, Austria). Logarithmic transformation was applied

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients

if obtained steroid concentrations did not pass D'Agostino and Pearson's test for normality. We performed one-way ANOVA with post hoc Dunnett test to compare circulating androgen concentrations. Wilcoxon's signed-rank test was performed to assess the effects of treatment. Mann-Whitney tests were used to compare the difference between eGC treated and untreated patients after 12 weeks of treatment. Linear models were used to assess how the individual androgens were associated with TA. The associations between progression-free survival (PFS) and TA were investigated at baseline and during treatment. Group steroid concentrations and changes were reported as median with interquartile ranges (IQR), unless stated otherwise.
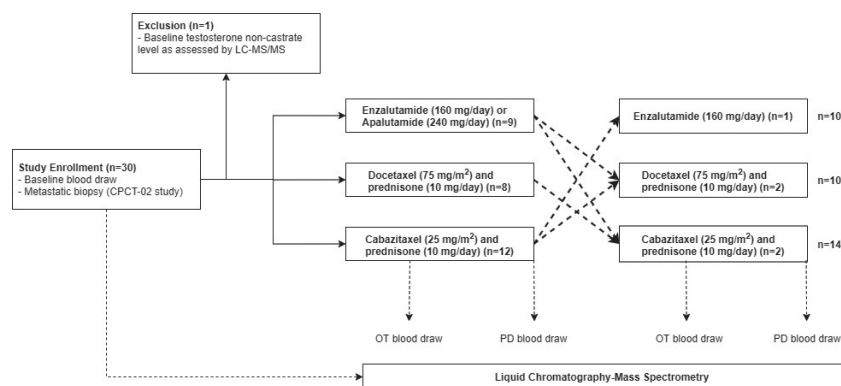
Results

Patients

Samples used in this study were obtained between May 2016 and July 2018 from patients with metastatic CRPC who were scheduled to start a new line of treatment. In total, 29 patients who completed 34 treatments were included in our analysis (Figure 1); five patients completed two treatments during their enrollment. Patient characteristics, disease and treatment history are shown in Table 1. For five patients with early progression (PFS: 22–82 days) no separate OT sample was available. These subjects were included in the comparison between baseline and OT samples, but not between OT and PD samples.

Figure 1. Patient and sample selection
Selection and exclusion of CIRCUS study samples for multi-steroid profiling, glucocorticoid quantification, survival analysis and tumor biopsy analysis.
Abbreviations: LC-MS/MS – liquid-chromatography tandem mass spectrometry, OT – on treatment, PD – progressive disease

Table 1. Patient characteristics, disease history and treatment history

	N	%	Median	Range (min-max)	IQR
Age at registration CIRCUS	29		67	48-81	62.0-75.0
WHO at registration CIRCUS	29		1	0-1	0-1
Testosterone level (nmol/L)	29		0.12	0.02 – 0.64	0.05 – 0.2
Prior ADT					
Drug-based	26	89.7			
Surgery-based (orchiectomy)	3	10.3			
With upfront docetaxel	5	17.2			
Prior systemic therapy (other than ADT)					
0	3	10.3			
1	5	17.2			
2	16	55.2			
3	4	13.8			
4	1	3.4			
Type of prior systemic therapy (other than ADT)					
Hormonal therapy only	3	10.3			
Chemotherapy only	7	24.1			
Immunotherapy only	1	3.4			
Hormonal and chemotherapy	12	41.4			
Chemotherapy and radionuclide therapy	1	3.4			
Hormonal, radionuclide and chemotherapy	1	3.4			
Hormonal, radionuclide and other therapy	1	3.4			
PSA at baseline (µg/L)*	34		75	1.5-913	19.3-173.3
PSA at progression (µg/L)*	34		123.5	5-1286	30.8-280.8

* For some patients with multiple treatments progression and subsequent baseline sample was identical
ADT, androgen deprivation therapy
IQR, interquartile range
PSA, prostate specific antigen

Androgen abundance

The circulating androgen concentrations in CRPC patients were determined in samples obtained before the start of the first treatment after enrollment (n = 29). The median concentration of 11KT (0.33 nmol/L, range 0.03-2.39 nmol/L) was significantly higher than T (0.12 nmol/L; range 0.03-0.64 nmol/L, P < 0.001) in CRPC patients at baseline (Figure 2A). 11KT constituted 66% (43-79%) of the TA pool whereas T was 24% (15-32%) (Figure 2B). Although T was below the castrate cutoff (1.74 nmol/L, 50 ng/mL) in all baseline samples (n = 34), the TA concentration (0.59 nmol/L,

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients

0.23-1.27 nmol/L) exceeded 1.74 nmol/L in six patients (Figure 2C). DHT was below the LOQ in most patients (range: < LOQ-0.27 nmol/L). Using linear models, 11KT was a major predictor of the TA concentration (R^2 = 0.945), whereas T predicted TA to a lesser extent (R^2 = 0.511).
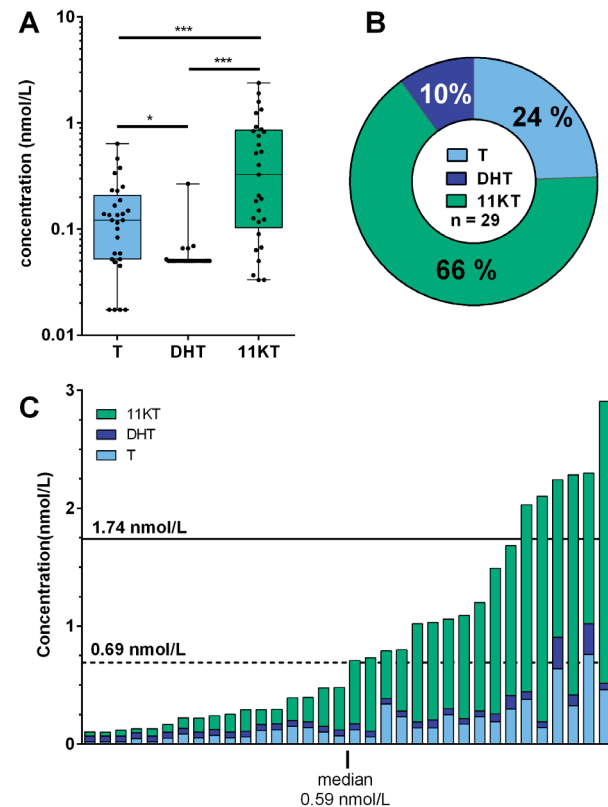
Figure 2. 11-Ketotestosterone is the most abundant circulating androgen in castration-resistant prostate cancer patients at baseline

(A) Active androgen concentrations of all castration-resistant prostate cancer patients before the start of the first treatment after enrollment (n = 29). Boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate the range. Dots indicate individual data points. Statistical analysis was performed by one-way ANOVA (P < 0.0001) with Tukey's Multiple Comparison test. * P < 0.05, *** P < 0.001.

(B) The relative abundance of the median androgen concentrations is shown as a percentage of the total androgen pool.

(C) Androgen concentrations are shown for all baseline samples (n = 34). Values below the analytical limit of quantification are shown if relevant calibrator and spiked quality control samples were accurate and reproducible with signal to noise >10:1. Samples with undetectable concentrations were set to 0.5 times the lowest accurate calibration sample for statistical purposes. Conventional clinical cut-off values for castrate testosterone levels (0.69 and 1.74 nmol/L (or 20 and 50 ng/dL) testosterone) are indicated on the y-axis for reference.

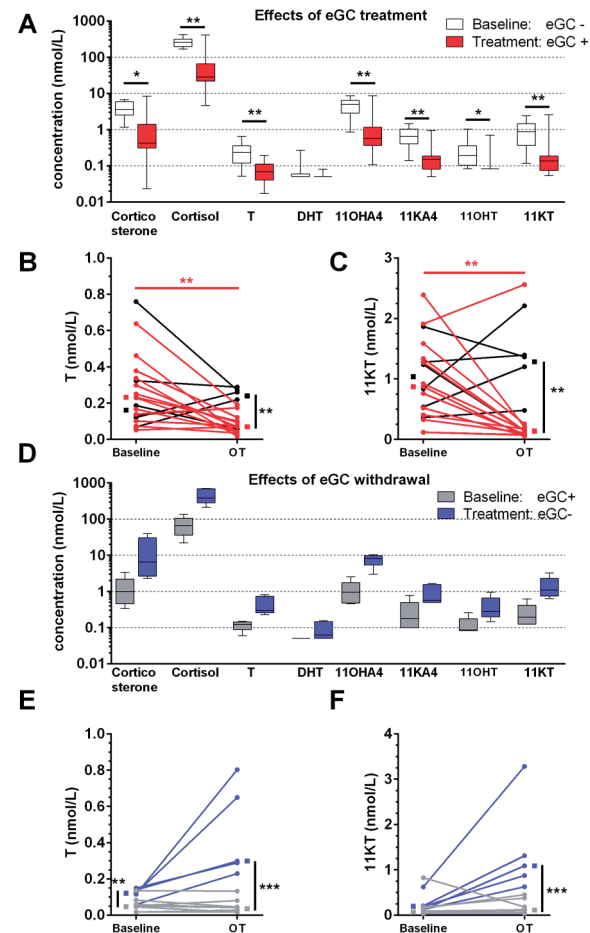*Abbreviations: 11KT – 11-ketotestosterone, DHT – dihydrotestosterone, T - testosterone*

Effects of treatment

Subjects included in this study started treatment with antiandrogens (n = 10), docetaxel with prednisone (n = 10) or cabazitaxel with prednisone (n = 14). Steroid hormone concentrations at baseline stratified for the different treatments are shown in Figure S2A-D. Significant suppression of adrenal-derived steroids was observed after 12 weeks of cabazitaxel with prednisone treatment. In the docetaxel group, a similar suppression was observed in a subset of patients, but this did not reach significance. In the antiandrogen group, increased steroid concentrations were observed after 12 weeks of treatment.

Low baseline cortisol concentrations were detected in a subset of patients, suggestive of hypothalamus-pituitary-adrenal (HPA) axis suppression by eGC. Post hoc exogenous glucocorticoids quantification by LC-MS/MS was performed to detect prednisone, prednisolone and dexamethasone in all samples (Figure S3A–C). Samples were classified as eGC+ if prednisolone (≥ 20.7 ng/mL) and/or dexamethasone (≥ 16.1 ng/mL) were detected. Cortisol was suppressed (< 140 nmol/L) in all eGC+ baseline samples. (Figure S3D).

A significant reduction in circulating glucocorticoid, T and 11KT concentrations was observed in patients starting treatment with eGC (Figure 3A). Circulating T (Figure 3B) and 11KT (Figure 3C) concentrations were lowered by 68% (38-79%) and 84% (49-89%), respectively, in patients starting eGC. Decreases of similar magnitude were observed for 11β-hydroxyandrostenedione (11OHA4), 11β-hydroxytestosterone (11OHT) and 11-ketoandrostenedione (11KA4, medians 66-92%; Figure 3A). In a subset of eGC+ patients, glucocorticoid treatment was withdrawn. The group size was insufficient to detect a statistical difference between baseline and treatment (n = 5, Figure 3D). Compared to patients who continued eGC (n = 10), withdrawn patients had eight-fold higher T (0.30 nmol/L (0.26–0.73 nmol/L) vs. 0.04 nmol/L (0.02–0.05 nmol/L)) and ten-fold higher 11KT (1.09 nmol/L (0.75-2.30 nmol/L) vs. 0.11 nmol/L (0.04–0.23 nmol/L)) (Figures 3E and 3F). An overview of steroid concentrations at baseline can be found in Table S2. Additionally, glucocorticoid treatment was withdrawn in six patients before progression. Again, higher median circulating concentrations of T (0.20 nmol/L (0.09–0.38 nmol/L) vs. 0.05 nmol/L (0.02–0.08 nmol/L), P < 0.01) and higher median 11KT (0.90 nmol/L (0.52–1.46 nmol/L) vs. 0.10 nmol/L (0.06–0.29 nmol/L), P = 0.001) were observed after withdrawal, compared to patients that continued GC treatment (n = 14, Figure S4).

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients

Figure 3. Effects of exogenous glucocorticoid treatment on circulating steroid concentrations
(A-C) Differences between steroid concentrations at baseline (white boxes) and on treatment (OT) (red boxes)
were assessed in patients who were exogenous glucocorticoid (eGC) untreated at baseline and who started
treatment with eGC (n = 13) by Wilcoxon signed-rank test (A). The individual data points are shown for
testosterone (T, B) and 11-ketotestosterone (11KT, C) at baseline and OT in patients who started therapy with
glucocorticoids (red lines, n = 13) or without glucocorticoids (black lines, n = 6).
(D-F) Differences between concentrations at baseline (gray boxes) and OT (blue boxes) were assessed in
patients who were eGC+ at baseline and discontinued glucocorticoid treatment (D, n = 5). The individual
data points are shown for T (E) and 11KT (F) at baseline and OT in patients who continued treatment with
glucocorticoids (gray lines, n = 10) or were withdrawn from glucocorticoids (blue lines, n = 5).
Boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate
the range. Effects of treatment were assessed by Wilcoxon ranked-sum test, while group differences were
assessed by Mann-Whitney test. Lines connect individual patients and group medians (squares) are shown
beside the individual data points. * P < 0.05, ** P < 0.01, *** P < 0.001.
Abbreviations: 11KA4 – 11-ketoandrostenedione, 11KT – 11-ketotestosterone, 11OHA4 -
11β-hydroxyandrostenedione, 11OHT - 11β-hydroxytestosterone, DHT – dihydrotestosterone, eGC –
exogenous glucocorticoids, OT – on treatment, T – testosterone

Clinical outcomes

A post hoc survival analysis was performed on this limited patient group. PFS was significantly
longer in patients with TA concentrations above the median (209 vs. 133.5 days, P < 0.05, Figure
4A). Stratification based on 11KT alone similarly showed this association (Figure 4B), whereas T
alone did not (Figure 4C). Overall survival was not significantly different between the patients
with high or low TA pool (14.7 months vs. 12.3 months, P > 0.05).
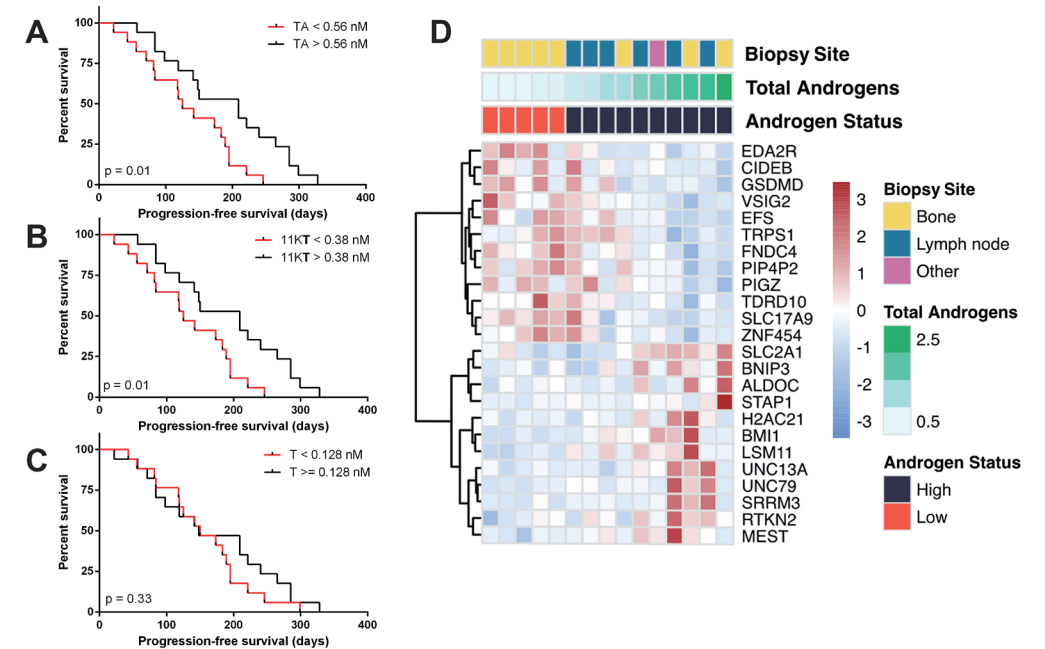
**Figure 4. Effects of total androgen concentration on progression-free survival and intratumoral gene
expression**
(A-C) Progression-free survival curves are shown for patients stratified into two groups with concentrations
above or below median total androgen (TA, A), 11-ketotestosterone (11KT, B) or testosterone (T, C). Log-rank
test for survival was used to determine difference between the low and high TA groups.
(D) Heatmap of differentially expressed genes (n = 24) across TA concentration in the tumor samples.
Differential gene expression was determined using TA concentration as a continuous variable. Heatmap
displays mean-centered and normalized (variance-stabilizing transformation) read counts. Unsupervised
hierarchical clustering (Euclidean distance and Ward.D2 method) was performed on genes and samples.
Upper tracks display biopsy site, TA concentration and androgen status of the samples.
Abbreviations: 11KT – 11-ketotestoterone, T – testosterone, TA – total androgens

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients

RNA sequencing analysis

Gene expression profile analysis of the complete CPCT-02 CRPC cohort (n = 180) revealed significant biopsy-site related bias, which we attempted to limit through exclusion of 5232 biopsy-site specific genes (Figure S5). Next, patients included in this study were assessed by RNA sequencing (n = 15) for further analysis, excluding genes which were biopsy-site related (Figure S6).

We observed androgen-mediated differential expression of 24 genes (Figure 4D), including several known androgen-regulated genes, using TA concentration as a continuous variable. Of these genes, 12 were upregulated in the high TA environment and 12 genes in the low TA environment. Known AR target genes *BMI1*[31] and *SLC2A1*[32] were upregulated in the high TA environment, whereas the androgen-repressed gene *TRPS1*[33] was upregulated in the low TA environment. Furthermore, a trend towards increased *AR* expression in the low TA environment (5.4-fold higher) was observed, although this did not reach statistical significance.

Genomic alteration and expression of steroid pathway genes

Whole-genome sequencing of the tumor tissue (n = 18) showed that somatic alterations within the *AR* or *AR* enhancer locus were highly prevalent (16/18 tumors) (Figure S7A). Interestingly, in 10/18 patients we observed high-level amplification of the *CYP11B1* gene, encoding 11β-hydroxylase, which is essential in the synthesis of cortisol and 11-oxygenated androgens. Amplification of other genes involved in androgen metabolism (*HSD3B1/2*, *SRD5A1/2*, *AKR1C3*, *HSD17B6* and *HSD17B10*) was rare, although copy-number loss of *HSD17B2*, which inactivates androgens, was observed in two subjects.

Several enzymes related to androgen biosynthesis and (re-)activation were highly expressed in all tumor biopsy samples, including *HSD17B10*, *STS*, *SRD5A1*, *AKR1C3* and *HSD11B2* (Figure S7B). This suggests that the enzymes required for activation of both androgen and 11-oxygenated androgen precursors are expressed in CRPC tumors. Expression of *CYP11B1* was absent in all but two samples, one of which did not contain a genomic abnormality.

Discussion

Using plasma multi-steroid profiling by LC-MS/MS, we show that 11KT, a potent AR agonist[13], is the predominant circulating androgen in CRPC patients and therefore needs to be considered when assessing the hormonal status of these patients. AR pathway reactivation after ADT is an important process that leads to tumor progression and our data suggest that 11KT may be an important contributor to AR reactivation in CRPC. 11KT constituted a median 66% of the TA pool. Androgen abundance in CRPC patients has probably been underestimated until now, as the quantification of T alone detects only approximately 25% of all androgens. Moreover, circulating T concentrations do not accurately predict TA abundance and thus does not allow for optimal

patient stratification. Androgen abundance after castration could be assessed more accurately by including 11KT. Additionally, this study shows the suppressive effects of eGC on circulating androgen concentrations, highlighting the potential therapeutic role of glucocorticoids in CRPC, which crucially includes suppression of adrenal-derived 11-oxygenated androgens. Finally, this study suggests that circulating TA abundance may correlate with intratumoral gene expression and PFS.

In addition to 11KT, precursor steroids such as 11OHA4 and 11KA4 may also contribute to AR activation through intratumoral conversion to 11KT. In line with previous findings, this study finds that 11OHA4 is abundant in the circulation with a median concentration of 2.7 nmol/L (0.48 − 5.1 nmol/L). Storbeck et al.[34] previously showed that the PC cell line LNCaP converts 11OHA4 and 11OHT into 11KT, requiring the enzymes HSD11B2 and AKR1C3. We confirm substantial expression of *HSD11B2* and *AKR1C3* in nearly all CRPC tumor samples (93% and 100%, respectively). Expression of enzymes that inactivate (11-oxygenated) androgens to their upstream precursors, such as *HSD17B2* and *HSD11B1,* was lower. A higher ARK1C3:HSD17B2 ratio favors production of 11KT especially, as AKR1C3 has a significantly higher substrate preference for 11KA4 than androstenedione[35]. Despite frequent genomic amplification of *CYP11B1*, which can convert androstenedione into 11OHA4, we detected *CYP11B1* expression in only two samples. Amplification of the *CYP11B1* locus therefore does not appear to result in detectable *CYP11B1* expression, and may be a passenger mutation. Consequently, intratumoral synthesis of 11OHA4 from androstenedione due to CYP11B1 amplification is unlikely to occur; thus, the majority of 11-oxygenated androgen precursors are likely derived from adrenal steroidogenesis.

Glucocorticoid treatment significantly decreased circulating T and 11-oxygenated androgens through suppression of the HPA-axis. Potent suppression of 11-oxygenated androgens has similarly been observed in patients treated with abiraterone acetate and prednisone[36]. The potent suppression of adrenal androgens, and 11KT in particular, may partially explain the clinical benefits of glucocorticoid treatment and should be considered when designing trials with glucocorticoid treatment in the control arm[37,38]. Withdrawal of glucocorticoid treatment may inadvertently lead to an increase in circulating androgens, although the effect on CRPC tumor growth is still unknown.

PFS was longer in our patients with high TA, in line with earlier findings that higher T concentrations are associated with a more favorable outcome[19,20]. Interestingly, in our study this association was attributed to 11KT, as T alone was not associated with PFS. This indicates that 11KT may be a stronger prognostic marker than T. In the presence of adrenal androgens, AR activation may remain (partially) ligand-dependent. AR pathway inhibition may exert greater effects in androgen-AR interaction dependent tumor cells, providing a possible explanation for the association between TA, 11KT and PFS. It is, however, important to consider that our study was not designed to detect differences in survival, and the limited sample size and patient heterogeneity do not permit a definitive conclusion in this regard. Most patients in the CIRCUS study had previously received treatment for CRPC. No association between the number of

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients

treatment lines and PFS was found, nor was the number of treatment lines different between the low and high TA groups. Further investigation into the actions and consequences of circulating 11KT is warranted, especially as a potential biomarker to select patients more likely to respond to AR-targeting therapies.

Gene expression analysis of the tumor biopsy samples identified 24 genes differentially expressed between patients across the TA spectrum, including three androgen-regulated genes[31-33]. Compensatory AR upregulation may explain the absence of classic AR-regulated genes, such as *KLK3* and *TMPRSS2*. *TRPS1* and *SLC2A1* were previously implicated in PC and AR action[39,40], and decreased *EFS* expression has been associated with more advanced PC and tumor recurrence[41,42]. Other genes differentially expressed in our cohort (*EDA2R, SLC17A9, TDRD10, ALDOC, SRRM3, MEST* and *RTKN2*) have been implicated in other malignancies, but not in PC specifically[43-47]. Androgen-mediated regulation of these may contribute to the observed association between TA and PFS.

Conclusion

This study demonstrates that 11KT is the predominant circulating androgen in CRPC patients. Paired with evidence from previously published findings, our results position 11KT as one of the potential drivers of AR activation in CRPC. Both T and 11KT can be suppressed by glucocorticoid treatment, providing a possible explanation why glucocorticoids are beneficial in CRPC patients. Future studies should consider TA as a potential biomarker in patients that have undergone ADT.

Acknowledgements

References

1. Sartor O, de Bono JS. Metastatic Prostate Cancer. N Engl J Med 2018;378(7):645-657. (In eng). DOI: 10.1056/NEJMra1701695.
2. Scher HI, Fizazi K, Saad F, et al. Increased survival with enzalutamide in prostate cancer after chemotherapy. N Engl J Med 2012;367(13):1187-97. (In eng). DOI: 10.1056/NEJMoa1207506.
3. Attard G, Reid AH, Yap TA, et al. Phase I clinical trial of a selective inhibitor of CYP17, abiraterone acetate, confirms that castration-resistant prostate cancer commonly remains hormone driven. J Clin Oncol 2008;26(28):4563-71. (In eng). DOI: JCO.2007.15.9749 [pii] 10.1200/JCO.2007.15.9749.
4. Smith MR, Saad F, Chowdhury S, et al. Apalutamide Treatment and Metastasis-free Survival in Prostate Cancer. N Engl J Med 2018;378(15):1408-1418. (In eng). DOI: 10.1056/NEJMoa1715546.
5. de Bono JS, Logothetis CJ, Molina A, et al. Abiraterone and increased survival in metastatic prostate cancer. N Engl J Med 2011;364(21):1995-2005. (In eng). DOI: 10.1056/NEJMoa1014618.
6. Taylor BS, Schultz N, Hieronymus H, et al. Integrative genomic profiling of human prostate cancer. Cancer Cell 2010;18(1):11-22. (In eng). DOI: S1535-6108(10)00238-2 [pii] 10.1016/j.ccr.2010.05.026.
7. Stanbrough M, Bubley GJ, Ross K, et al. Increased expression of genes converting adrenal androgens to testosterone in androgen-independent prostate cancer. Cancer Res 2006;66(5):2815-25. (In eng). DOI: 66/5/2815 [pii] 10.1158/0008-5472.CAN-05-4000.
8. Montgomery RB, Mostaghel EA, Vessella R, et al. Maintenance of intratumoral androgens in metastatic prostate cancer: a mechanism for castration-resistant tumor growth. Cancer Res 2008;68(11):4447-54. (In eng). DOI: 68/11/4447 [pii] 10.1158/0008-5472.CAN-08-0249.
9. Mohler JL, Gregory CW, Ford OH, 3rd, et al. The androgen axis in recurrent prostate cancer. Clin Cancer Res 2004;10(2):440-8. (In eng). DOI: 10.1158/1078-0432.ccr-1146-03.
10. Chang KH, Li R, Papari-Zareei M, et al. Dihydrotestosterone synthesis bypasses testosterone to drive castration-resistant prostate cancer. Proc Natl Acad Sci U S A 2011;108(33):13728-33. (In eng). DOI: 1107898108 [pii] 10.1073/pnas.1107898108.
11. Friedlander TW, Roy R, Tomlins SA, et al. Common structural and epigenetic changes in the genome of castration-resistant prostate cancer. Cancer Res 2012;72(3):616-25. (In eng). DOI: 0008-5472.CAN-11-2079 [pii] 10.1158/0008-5472.CAN-11-2079.
12. Ji Q, Chang L, VanDenBerg D, Stanczyk FZ, Stolz A. Selective reduction of AKR1C2 in prostate cancer and its role in DHT metabolism. Prostate 2003;54(4):275-89. (In eng). DOI: 10.1002/pros.10192.
13. Pretorius E, Africander DJ, Vlok M, Perkins MS, Quanson J, Storbeck KH. 11-Ketotestosterone and 11-Ketodihydrotestosterone in Castration-resistant Prostate Cancer: Potent Androgens Which Can No Longer Be Ignored. PLoS One 2016;11(7):e0159867. (In eng). DOI: 10.1371/journal.pone.0159867 PONE-D-16-19863 [pii].
14. Turcu AF, Nanba AT, Chomic R, et al. Adrenal-derived 11-oxygenated 19-carbon steroids are the dominant androgens in classic 21 hydroxylase deficiency. Eur J Endocrinol 2016;174(5):601-9. (In eng). DOI: EJE-15-1181 [pii] 10.1530/EJE-15-1181.
15. Rege J, Turcu AF, Kasa-Vubu JZ, et al. 11-Ketotestosterone Is the Dominant Circulating Bioactive Androgen During Normal and Premature Adrenarche. J Clin Endocrinol Metab 2018;103(12):4589-4598. (In eng). DOI: 5075164 [pii] 10.1210/jc.2018-00736.
16. Imamichi Y, Yuhki KI, Orisaka M, et al. 11-Ketotestosterone Is a Major Androgen Produced in Human Gonads. J Clin Endocrinol Metab 2016;101(10):3582-3591. (In eng). DOI: 10.1210/jc.2016-2311.
17. Davio A, Woolcock H, Nanba AT, et al. Sex Differences in 11-Oxygenated Androgen Patterns Across Adulthood. J Clin Endocrinol Metab 2020;105(8). DOI: 10.1210/clinem/dgaa343.
18. Snaterse G, Visser JA, Arlt W, Hofland J. Circulating steroid hormone variations throughout different stages of prostate cancer. Endocr Relat Cancer 2017;24(11):R403-R420. (In eng). DOI: ERC-17-0155 [pii] 10.1530/ERC-17-0155.
19. Attard G, Reid AH, A'Hern R, et al. Selective inhibition of CYP17 with abiraterone acetate is highly active in the treatment of castration-resistant prostate cancer. J Clin Oncol 2009;27(23):3742-8. (In eng). DOI: JCO.2008.20.0642 [pii] 10.1200/JCO.2008.20.0642.

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients

20. Sakamoto S, Maimaiti M, Xu M, et al. Higher serum testosterone levels associated with favorable prognosis in enzalutamide- and abiraterone-treated castration-resistant prostate cancer. J Clin Med 2019;8(4) (Article) (In English). DOI: 10.3390/jcm8040489.

21. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45(2):228-47. (In eng). DOI: S0959-8049(08)00873-3 [pii] 10.1016/j.ejca.2008.10.026.

22. Scher HI, Morris MJ, Basch E, Heller G. End points and outcomes in castration-resistant prostate cancer: from clinical trials to clinical practice. J Clin Oncol 2011;29(27):3695-704. (In eng). DOI: JCO.2011.35.8648 [pii] 10.1200/JCO.2011.35.8648.

23. van Dessel LF, Beije N, Helmijr JC, et al. Application of circulating tumor DNA in prospective clinical oncology trials - standardization of preanalytical conditions. Mol Oncol 2017;11(3):295-304. (In eng). DOI: 10.1002/1878-0261.12037.

24. Snaterse G, van Dessel LF, Taylor AE, et al. Validation of multiplex steroid hormone measurements in prostate cancer using plasma for multimodality biomarker studies. medRxiv 2020:2020.08.05.20164202. DOI: 10.1101/2020.08.05.20164202.

25. O'Reilly MW, Taylor AE, Crabtree NJ, et al. Hyperandrogenemia predicts metabolic phenotype in polycystic ovary syndrome: the utility of serum androstenedione. J Clin Endocrinol Metab 2014;99(3):1027-36. (In eng). DOI: 10.1210/jc.2013-3399.

26. O'Reilly MW, Kempegowda P, Jenkinson C, et al. 11-Oxygenated C19 Steroids Are the Predominant Androgens in Polycystic Ovary Syndrome. J Clin Endocrinol Metab 2017;102(3):840-848. (In eng). DOI: 10.1210/jc.2016-3285.

27. van der Pas R, Hofland LJ, Hofland J, et al. Fluconazole inhibits human adrenocortical steroidogenesis in vitro. J Endocrinol 2012;215(3):403-12. (In eng). DOI: JOE-12-0310 [pii] 10.1530/JOE-12-0310.

28. Quanson JL, Stander MA, Pretorius E, Jenkinson C, Taylor AE, Storbeck KH. High-throughput analysis of 19 endogenous androgenic steroids by ultra-performance convergence chromatography tandem mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 2016;1031:131-138. (In eng). DOI: S1570-0232(16)30486-X [pii] 10.1016/j.jchromb.2016.07.024.

29. Priestley P, Baber J, Lolkema MP, et al. Pan-cancer whole-genome analyses of metastatic solid tumors. Nature 2019;575(7781):210-216. (In eng). DOI: 10.1038/s41586-019-1689-y 10.1038/s41586-019-1689-y [pii].

30. Schiffer L, Barnard L, Baranowski ES, et al. Human steroid biosynthesis, metabolism and excretion are differentially reflected by serum and urine steroid metabolomes: A comprehensive review. J Steroid Biochem Mol Biol 2019;194:105439. (In eng). DOI: S0960-0760(19)30279-1 [pii] 10.1016/j.jsbmb.2019.105439.

31. Zhu S, Zhao D, Li C, et al. BMI1 is directly regulated by androgen receptor to promote castration-resistance in prostate cancer. Oncogene 2020;39(1):17-29. (In eng). DOI: 10.1038/s41388-019-0966-4 10.1038/s41388-019-0966-4 [pii].

32. Ragnum HB, Roe K, Holm R, et al. Hypoxia-independent downregulation of hypoxia-inducible factor 1 targets by androgen deprivation therapy in prostate cancer. Int J Radiat Oncol Biol Phys 2013;87(4):753-60. (In eng). DOI: S0360-3016(13)02861-7 [pii] 10.1016/j.ijrobp.2013.07.023.

33. Chang GT, van den Bemd GJ, Jhamai M, Brinkmann AO. Structure and function of GC79/TRPS1, a novel androgen-repressible apoptosis gene. Apoptosis 2002;7(1):13-21. (In eng). DOI: 10.1023/a:1013504710343.

34. Storbeck KH, Bloem LM, Africander D, Schloms L, Swart P, Swart AC. 11beta-Hydroxydihydrotestosterone and 11-ketodihydrotestosterone, novel C19 steroids with androgenic activity: a putative role in castration-resistant prostate cancer? Mol Cell Endocrinol 2013;377(1-2):135-46. (In eng). DOI: S0303-7207(13)00289-X [pii] 10.1016/j.mce.2013.07.006.

35. Barnard M, Quanson JL, Mostaghel E, Pretorius E, Snoep JL, Storbeck KH. 11-Oxygenated androgen precursors are the preferred substrates for aldo-keto reductase 1C3 (AKR1C3): Implications for castration-resistant prostate cancer. J Steroid Biochem Mol Biol 2018;183:192-201. (Article) (In English). DOI: 10.1016/j.jsbmb.2018.06.013.

36. Wright C, O'Day P, Alyamani M, Sharifi N, Auchus R. Abiraterone acetate treatment lowers 11-oxygenated androgens. Eur J Endocrinol 2020 (In eng). DOI: 10.1530/EJE-19-0905 EJE-19-0905.R2 [pii].

37. Venkitaraman R, Thomas K, Huddart RA, Horwich A, Dearnaley DP, Parker CC. Efficacy of low-dose dexamethasone in castration-refractory prostate cancer. BJU Int 2008;101(4):440-3. (In eng). DOI: BJU7261 [pii] 10.1111/j.1464-410X.2007.07261.x.

38. Tannock I, Gospodarowicz M, Meakin W, Panzarella T, Stewart L, Rider W. Treatment of metastatic prostatic cancer with low-dose prednisone: evaluation of pain and quality of life as pragmatic indices of response. J Clin Oncol 1989;7(5):590-7. (In eng). DOI: 10.1200/JCO.1989.7.5.590.

39. Vaz CV, Marques R, Alves MG, et al. Androgens enhance the glycolytic metabolism and lactate export in prostate cancer cells by modulating the expression of GLUT1, GLUT3, PFK, LDH and MCT4 genes. J Cancer Res Clin Oncol 2016;142(1):5-16. DOI: 10.1007/s00432-015-1992-4.

40. Chang GT, Jhamai M, van Weerden WM, Jenster G, Brinkmann AO. The TRPS1 transcription factor: androgenic regulation in prostate cancer and high expression in breast cancer. Endocr Relat Cancer 2004;11(4):815-22. DOI: 10.1677/erc.1.00853.

41. Sertkaya S, Hamid SM, Dilsiz N, Varisli L. Decreased expression of EFS is correlated with the advanced prostate cancer. Tumor Biol 2015;36(2):799-805. (In eng). DOI: 10.1007/s13277-014-2703-5.

42. Vanaja DK, Ehrich M, Van den Boom D, et al. Hypermethylation of genes for diagnosis and risk stratification of prostate cancer. Cancer Invest 2009;27(5):549-60. (In eng). DOI: 908866522 [pii] 10.1080/07357900802620794.

43. Li J, Su T, Yang L, Deng L, Zhang C, He Y. High SLC17A9 expression correlates with poor survival in gastric carcinoma. Future Oncol 2019;15(36):4155-4166. (In eng). DOI: 10.2217/fon-2019-0283.

44. Yang L, Chen Z, Xiong W, et al. High expression of SLC17A9 correlates with poor prognosis in colorectal cancer. Hum Pathol 2019;84:62-70. (In eng). DOI: S0046-8177(18)30359-9 [pii] 10.1016/j.humpath.2018.09.002.

45. de Almeida BP, Apolonio JD, Binnie A, Castelo-Branco P. Roadmap of DNA methylation in breast cancer identifies novel prognostic biomarkers. BMC Cancer 2019;19(1):219. (In eng). DOI: 10.1186/s12885-019-5403-0 10.1186/s12885-019-5403-0 [pii].

46. Pedersen IS, Dervan PA, Broderick D, et al. Frequent loss of imprinting of PEG1/MEST in invasive breast cancer. Cancer Res 1999;59(21):5449-51. (In eng) (https://www.ncbi.nlm.nih.gov/pubmed/10554015).

47. Kim MS, Lee HS, Kim YJ, Lee DY, Kang SG, Jin W. MEST induces Twist-1-mediated EMT through STAT3 activation in breast cancers. Cell Death Differ 2019;26(12):2594-2606. (In eng). DOI: 10.1038/s41418-019-0322-9 10.1038/s41418-019-0322-9 [pii].

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients

Supplementary information

Supplementary materials and methods

Extraction and measurement of endogenous steroids

Calibration series (0.01–100 ng/mL) were prepared in (1) phosphate-buffered saline + bovine serum albumin (0.1%) and in (2) charcoal-stripped human serum (Goldenwest Diagnostics, Temecula, CA, USA). Steroids were extracted from 400 µL plasma by liquid-liquid extraction using methyl-tert-butyl ether (MTBE, Sigma Aldrich, Zwijndrecht, the Netherlands) and evaporated under a nitrogen manifold at 50°C. Samples were reconstituted in 125 µL LC-MS grade 50% methanol (CHROMASOLV, Sigma Aldrich). Multi-steroid profiling was performed by tandem mass spectrometry (Xevo TQ-XS, Waters, Milford, MA, USA) after separation on an ACQUITY UPLC (Waters) with UPLC high-strength silica T3 column (21 mm x 50 mm, 1.8 µm, Waters) as previously described[1-5]. Spiked analytical controls showed accuracy of ≥ 90% at very low (0.03 ng/mL), low (0.3 ng/mL), medium (3 ng/mL) and high (30 ng/mL) concentrations for testosterone, 11KT, 11KA4 and cortisone. For 11OHA4, controls showed accuracy ≥ 80% at very low (0.03 ng/mL) and ≥ 90% at higher concentrations. Cortisol, corticosterone, DHT and 11OHT showed accuracy of ≥ 90% at low to high concentrations. Inter-assay precision was calculated based on 4-6 biological replicates ran on 3 days, with RSD < 10% for testosterone, DHT, 11KT, cortisol, cortisone and corticosterone, < 20% for 11OHT, and 22% for 11OHA4 and 32% for 11KA4. Calibration of androstenedione was not successful due to a matrix contaminant and was excluded from further analysis.

Data has been included where steroid concentrations were below the analytical LOQ, but where calibration and spiked QC samples were still accurate with signal to noise > 10:1. Samples with undetectable concentrations were set to 0.5 times the lowest accurate calibration point for statistical purposes. The total androgen (TA) pool was defined as the sum of testosterone, DHT and 11KT as these steroids have been confirmed to directly activate the AR[6].

Quantification of exogenous glucocorticoids

eGC were measured in the original sample plates by LC-MS/MS using a protocol optimized for separation of the exogenous and endogenous glucocorticoids[7,8]. To quantify the steroids a calibration series (1–100 ng/mL) containing prednisone, prednisolone and dexamethasone was prepared. Cutoff values for prednisolone (20.7 ng/mL) and dexamethasone (16.1 ng/mL) were determined based on suppression of cortisol (< 140 nmol/L) and were used to distinguish eGC+ from eGC- samples.

Whole-genome sequencing and RNA sequencing

WGS of metastatic tumor biopsies and matched normal blood was performed using the HiSeq X Ten system (Illumina) with a sequencing depth of 30x for blood and 60x for tumor tissue. The genome reference build 37 (GRCh37) was used for alignment. Subsequently, mutational analysis was performed to call single and multi-nucleotide variants (SNV and MNV) and small insertion/deletions (InDels) as previously described[9]. Somatic SNV, InDels and MNV were further annotated with Variant Effect Predictor[10] (VEP; v99) using GENCODE v33 annotations.

Discovery of somatic structural variants (SV) and copy-number alterations was performed using the GRIDDS, PURPLE and LINX suite[11]. During the downstream analyses, we only retained somatic structural variants passing all default QC filters (PASS-only) and with an upstream and/or downstream Tumor Allele Frequency (TAF) ≥ 0.1.

Putative protein-altering (coding) or high-impact (e.g. splicing) mutations were aggregated per sample and gene by selecting the most deleterious annotated effect (from VEP) on any known overlapping gene-wise transcript (except those transcripts flagged as retained intron and nonsense mediated decay). In addition, structural variants with a TAF ≥ 0.1 that overlapped partly or completely with the respective coding sequences, were annotated as 'structural variant' mutations. Multiple coding mutations and/or SV per gene were annotated as 'multiple mutations'.

RNA sequencing

Sequencing libraries were prepared using the KAPA RNA HyperPrep Kit with RiboErase (HMR) KR1351 (Roche, Indianapolis, IN, USA) according to the manufacturer protocols. Paired-end sequencing of (m)RNA was performed on the Illumina NextSeq 550 platform (2x 75 bp; Illumina, San Diego, California, USA) and Illumina NovaSeq 6000 platform (2x 150 bp; Illumina. Raw sequencing reads were trimmed (paired-end) with fastp[12] (v0.20.0) using default settings to perform adapter, low-quality and low-complexity trimming using the following command:

fastp --detect_adapter_for_pe -L --html --thread 5 --in1 <fq.R1> --in2 <fq.R2> --out1 <fq.R1.out> --out2 <fq.R1.out>

Trimmed paired-end reads were subsequently aligned against the human genome reference build 37 (GRCh37) with STAR[13] (v2.7.3a) on GENCODE annotations[14] (v33). Samples sequenced on multiple sequencing lanes were aligned simultaneously with respective read-group information using the following command:

STAR --genomeDir <GRCh37> --readFilesIn <fq.R1> <fq.R2> --readFilesCommand zcat --outFileNamePrefix <prefix> --outSAMtype BAM SortedByCoordinate --outSAMunmapped Within --twopassMode Basic --twopass1readsN -1 --runThreadN 10 --limitBAMsortRAM 10000000000 --quantMode TranscriptomeSAM --outSAMattrRGline <readgroup>

Marking of duplicate reads, sorting, indexing and retrieving flagstat information were performed using Sambamba[15] (v0.7.1; Figure S5A). Overlapping primary-aligned reads per exon were summarized per gene by Subread featurecounts[15] (v1.6.3) using reversely-stranded modus on GENCODE annotations (v33):

featureCounts -T 50 -t exon -g gene_id --primary -p -s 2 -a <gencodev33.gtf> -o <output> <BAM files>

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients

Read counts from protein coding genes present in the GENCODE annotations (v33; n = 20084) were inputted into DESeq2[16] (v1.24.0) for all 180 RNA-sequenced samples and normalized for library-size using default settings.

To identify possible batch-effects in expression due to biopsy localization (Figure S5B) or sequencing platform (Illumina NextSeq / Illumina NovaSeq), we first performed t-SNE[17] analysis using Rtsne (v0.15; with Θ set to 0.5, a perplexity of 30, 1000 iterations and two dimensions) on normalized reads counts (variance stabilizing transformation (VST)) to reveal signs of possible batch-effects (Figure S5D). This analysis revealed several clusters attributed to biopsy location and only limited correlation to sequencing technique. To reduce the batch-effect seen by biopsy localization, we performed differential analysis using DESeq2 (v1.24.0; Wald method) with additional LFC shrinkage (ashr[18]; v2.2.47) per major biopsy site and comparing against all others (Figure S5B); bone (n = 89 vs. 91), lymph node (n = 42 vs. 138) and liver (n = 31 vs. 149). Genes with the following criteria were designated as differentially-expressed for each major biopsy site (Figure S5C): adjusted P ≤ 0.05 & average read count over all 180 samples ≥ 50 and a $\log_2$ fold change < 0.1 (up-regulated in biopsy site). In downstream analyses, we marked these genes as 'putative biopsy site-associated genes' (n = 7527) and discarded these as candidates. Masking these genes and performing an identical t-SNE analysis on all 180 RNA-sequenced samples revealed reduced contribution of biopsy site (Figure S5E).

Subsequently, we performed differential expression analysis using DESeq2 (v.1.24.0; Wald method) for CPCT-02 RNA-sequenced samples which were also included into the CIRCUS study and of which the date of biopsy and androgen measurement was less than 30 days apart, using total androgen concentration as a continuous variable (n = 15; Figure S6A). Genes with the following criteria (and not present in the aforementioned 'putative batch-effect genes'; Figure S6B and 6C) were considered to be differentially expressed (n = 24): |$\log_2$ fold change| ≥ 0.5 (with SE ≤ 1.25) & adjusted P ≤ 0.05 and an average read count ≥ 50 over all 15 samples.
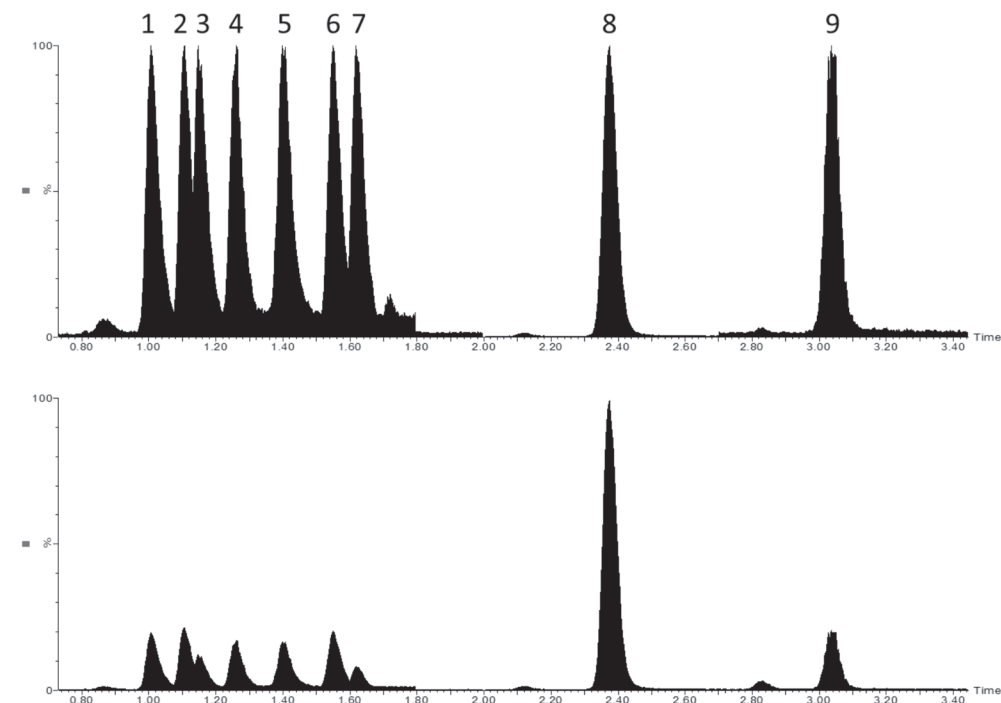
Figure S1. Chromatographic separation of endogenous steroids
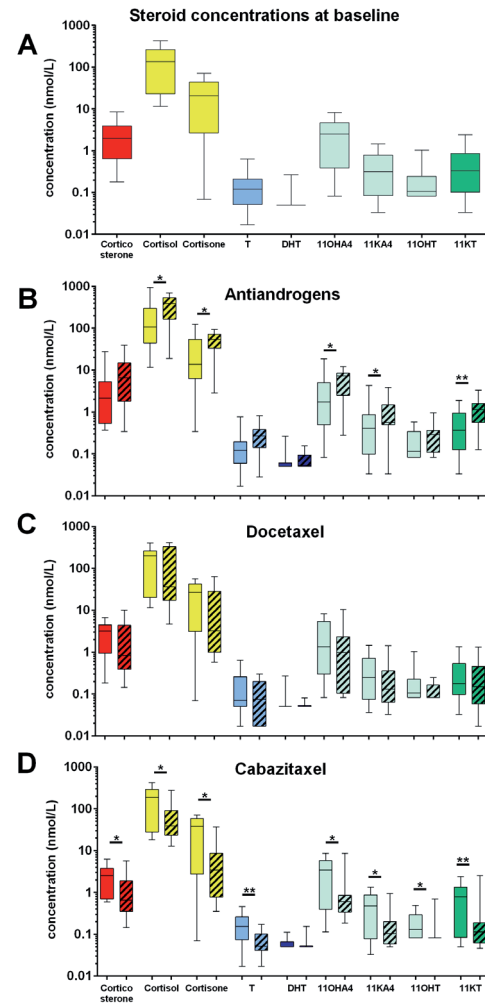Representative chromatographic separation of 9 steroids: (1) cortisone, (2) 11-ketoandrostenedione, (3) cortisol, (4) 11-ketotestosterone, (5) 11β-hydroxyandrostenedione, (6) 11β-hydroxytestosterone, (7) corticosterone, (8) testosterone, (9) dihydrotestosterone. The top chromatogram shows separation with normalized intensity, whereas bottom graph shows actual response.

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients

Figure S2. Baseline circulating steroid concentrations and effects of 12-week treatment
Steroid profiles were obtained for all castration-resistant prostate cancer patients before the start of first treatment (A, n = 29). Steroid profiles at baseline and on treatment are shown for CRPC patients treated with antiandrogens (B, n = 10), docetaxel with prednisone (C, n = 10) and cabazitaxel with prednisone (D, n = 14). Samples with undetectable concentrations were set to 0.5 times the lowest accurate calibration sample for statistical purposes. Differences between baseline (clear boxes) and on treatment (striped boxes) were assessed by Wilcoxon signed-rank test. Boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate the range. * P < 0.05, ** P < 0.01.
Boxes are colored to reflect steroid synthesis pathways: mineralocorticoid (red), glucocorticoid (yellow), androgen (blue) and 11-oxygenated androgen (green). Lighter colors are used to indicate precursor steroids.
Abbreviations: 11KA4 – 11-ketoandrostenedione, 11KT – 11-ketotestosterone, 11OHA4 -
11β-hydroxyandrostenedione, 11OHT – 11β-hydroxytestosterone, CRPC – castration-resistant prostate
cancer, DHT – dihydrotestosterone, T – testosterone
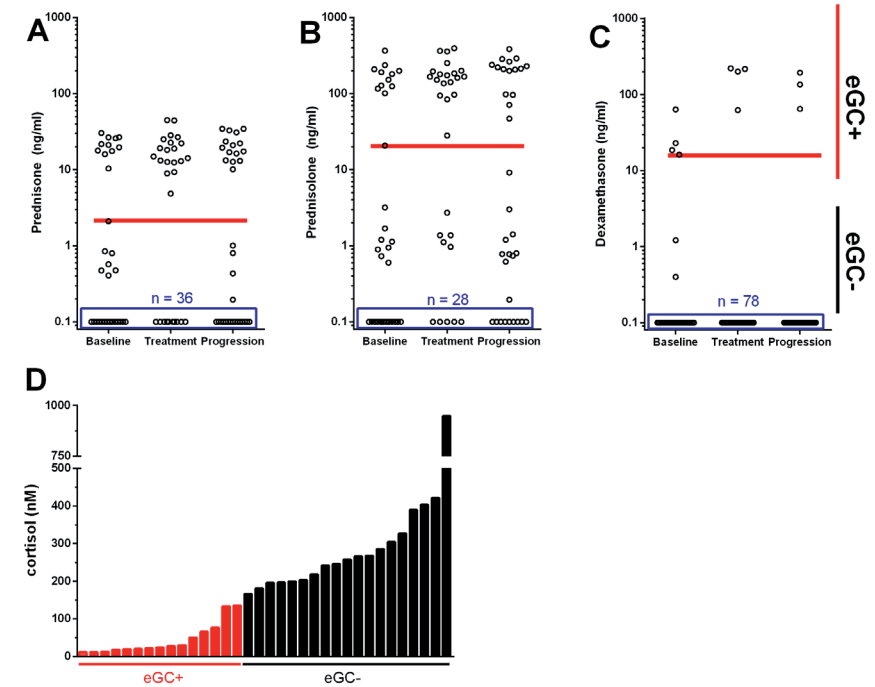
Figure S3. Quantification of prednisone, prednisolone and dexamethasone identifies glucocorticoid suppressed patient samples
Quantification of prednisone (A), prednisolone (B) and dexamethasone (C) was performed using LC-MS/MS to identify the use of exogenous glucocorticoids (eGC) in patient samples for classification. Values that were obtained during this measurement should be considered in the context of this study only, as appropriate internal standards had not been included in the initial extraction. No suppression of adrenal steroids was observed in patients with eGC values below the red lines. Classification of the exogenous glucocorticoids successfully identified samples with suppressed cortisol values at baseline (D).
Abbreviations: eGC – exogenous glucocorticoids

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients
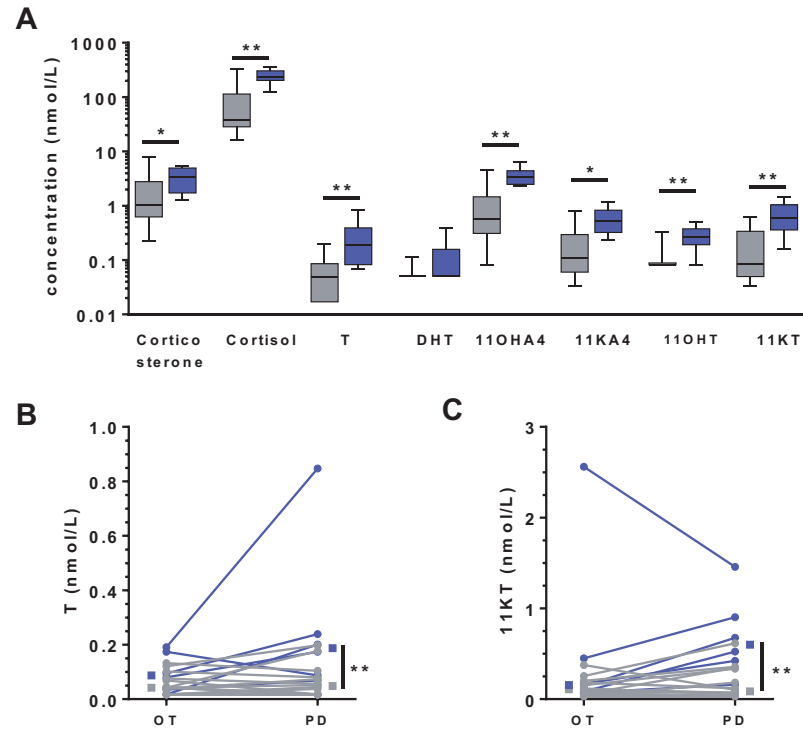
Figure S4. Circulating androgens are higher in patients withdrawn before progression

(A) The effects of glucocorticoid withdrawal were studied in patients who were exogenous glucocorticoid (eGC)+ on treatment (n = 20). Differences between steroid concentrations upon progression were assessed in patients in whom glucocorticoid treatment was continued (gray boxes, n = 14) or discontinued (blue boxes, n = 6) by Wilcoxon signed-rank test. Boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate the range. * P < 0.05, ** P < 0.01.

(B-C) The individual data points are shown for testosterone (T, B) and 11-ketotestosterone (11KT, C) for patients that were eGC+ on treatment and continued (gray lines, n = 14) or discontinued glucocorticoid treatment (blue lines, n = 6). Samples with undetectable concentrations were set to 0.5 times the lowest accurate calibration sample for statistical purposes. Effects of treatment were assessed by Wilcoxon ranked-sum test, while group differences were assessed by Mann-Whitney test. Lines connect individual patients and group medians (squares) are shown beside the individual data points. ** P < 0.01.

Abbreviations: 11KA4 – 11-ketoandrostenedione, 11KT – 11-ketotestosterone, 11OHA4 –
11β-hydroxyandrostenedione, 11OHT - 11β-hydroxytestosterone, DHT – dihydrotestosterone, eGC –
exogenous glucocorticoids, OT – on treatment, PD – progressive disease, T – testosterone
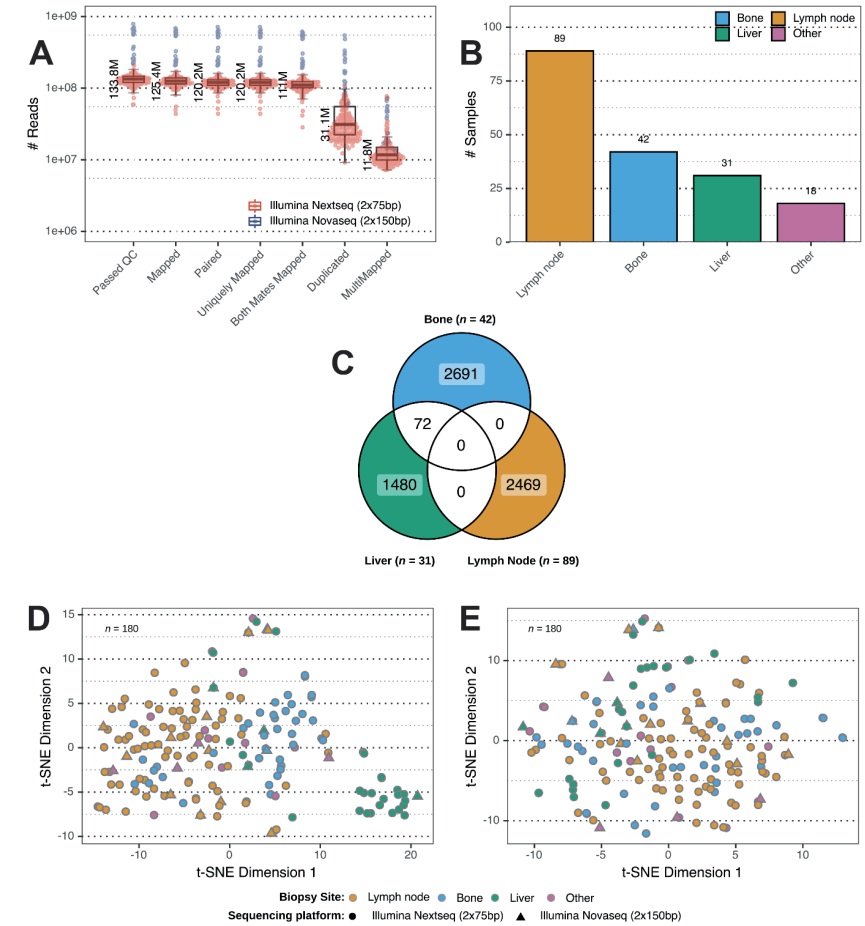
Figure S5. Overview of the CPCT-02 (m)RNA-sequenced metastatic prostate cancers (n = 180) including selection of biopsy site-specific expression

(A) Sequencing and alignment metrics (flagstat) highlighting samples (n = 180) performed on the Illumina NextSeq (orange) and Illumina NovaSeq (blue) platform. Y-axis depicts number of reads per sample and X-axis depicts various alignment metrics. Boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range. Individual data points are shown. Median per alignment metric (per million) is shown on the left-side of each boxplot.

(B) Number of RNA-sequenced samples per biopsy site. Biopsy sites with fewer than 3 samples were categorized into the 'Other' category. Y-axis depicts number of samples and X-axis depicts the major biopsy sites (bone, lymph node and liver). Number of samples is shown above each bar per biopsy site.

(C) Venn-diagram of differentially expressed protein-coding genes (adjusted P ≤ 0.05 & average read count over all 180 samples ≥ 50 and a $\log_2$ fold change < 0.1) per biopsy site.

(D) t-SNE analysis of all 180 samples over the read counts of all protein-coding genes (n = 20084). Samples are colored per biopsy site and shaped per sequencing platform (circle for Illumina NextSeq and triangle for Illumina NovaSeq).

(E) t-SNE analysis of all 180 samples over the read counts of all protein-coding genes (n = 20084), except those marked as 'putative biopsy site associated genes' (n = 7527). Samples are colored per biopsy site and shaped per sequencing platform (circle for Illumina NextSeq and triangle for Illumina NovaSeq).

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients
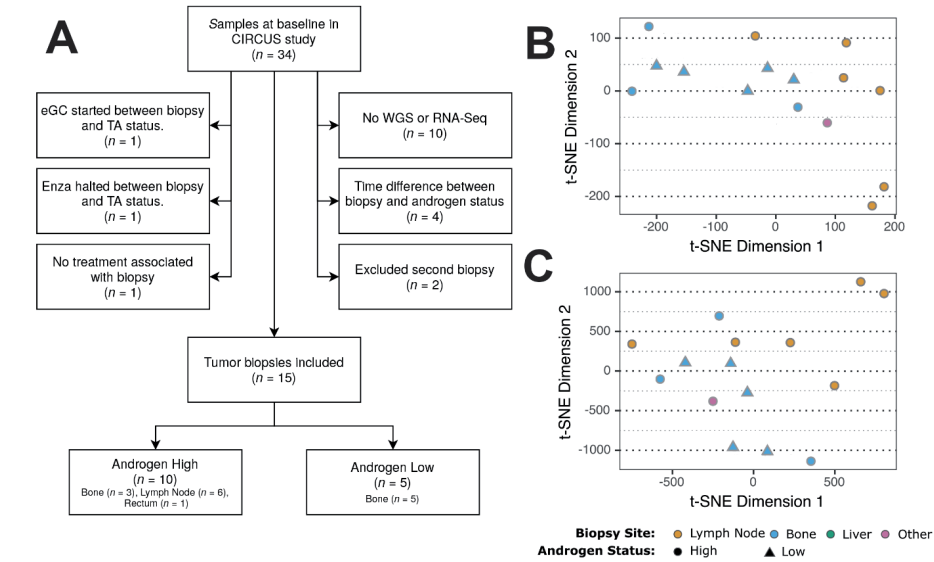
Figure S6. Overview of the CPCT-02 (m)RNA-sequenced metastatic prostate cancers (n = 15) present in the CIRCUS study

(A) Overview of the inclusion criteria of CIRCUS samples present within the CPCT-02 RNA-sequenced metastatic prostate cancer cohort (n = 180) and categorization of these samples (n = 15) into androgen status high and low samples.

(B) t-SNE analysis of the 15 included samples over the read counts of all protein-coding genes (n = 20084). Samples are colored per biopsy site and shaped per androgen status (circle for high androgen status and triangle for low androgen status).

(C) t-SNE analysis of all 15 samples over the read counts of all protein-coding genes (n = 20084), except those marked as 'putative biopsy site-associated genes' for lymph node or bone (n = 5232). Genes associated with the liver biopsy site were not excluded (n = 1480) as none of the biopsies originated from liver metastasis. Samples are colored per biopsy site and shaped per androgen status (circle for high androgen status and triangle for low androgen status).

Abbreviations: eGC - exogenous glucocorticoids, Enza – enzalutamide, RNA-seq - RNA sequencing, TA – total androgens, WGS - whole-genome sequencing
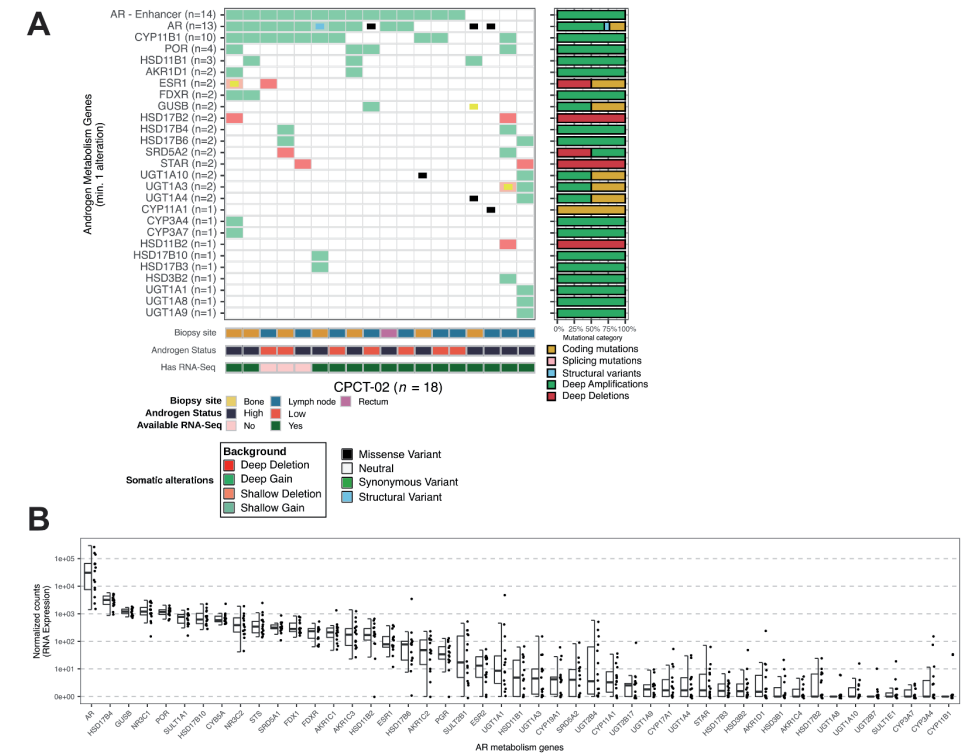
Figure S7. Overview of genome-wide characteristics and intratumoral expression of genes involved in steroid metabolism and steroid hormone receptors in metastatic castration-resistant prostate cancer

(A) Overview of genome-wide characteristics of the samples present within the CPCT-02 cohort with both whole-genome sequencing (WGS, n = 18). Samples are ordered by mutually-exclusive mutations in steroid metabolism-associated genes. In decreasing order, the following tracks show in following order: somatic mutations for genes of interest, biopsy site, androgen status and availability of RNA sequencing data

(B) Normalized (VST-transformed) RNA-seq read counts of steroid hormone receptors and enzymes involved in the conversion and metabolism of steroid hormones. Boxplot depicts the upper and lower quartiles, with the median shown as a solid line; whiskers indicate 1.5 times the interquartile range. Individual data points are shown.

Abbreviations: AR – androgen receptor

Part 3
Steroidomics

Chapter 8
11-Ketotestosterone is the predominant androgen in castration-resistant
prostate cancer patients

Table S1. Overview of steroids and their mass transitions, retention times and analytical limits of quantification

Analyte	Abbreviation	Ion (m/z)	Retention time	Internal standard	LOQ (nmol/L)
11-ketoandrostenedione	11KA4	301.0 > 121.0	1.11	11-ketotestosterone-d3	0.5
11-ketotestosterone	11KT	303.1 > 121.0	1.25	11-ketotestosterone-d3	0.2
11β-hydroxyandrostenedione	11OHA4	303.1 > 267.1	1.40	11β-hydroxyandrostenedione-d7	0.5
11β-hydroxytestosterone	11OHT	305.1 > 121.0	1.56	11β-hydroxyandrostenedione-d7	0.5
Corticosterone	-	347.1 > 121.0	1.65	Corticosterone-d8	0.2
Cortisol	-	363.1 > 90.9	1.18	Cortisol-d4	0.1
Cortisone	-	361.0 > 121.0	1.03	Cortisone-d7	0.1
Dihydrotestosterone	DHT	291.3 > 159.1	3.03	DHT-d3	0.2
Testosterone	T	289.1 > 96.9	2.38	Testosterone-d3	0.1

LOQ, limits of quantification
m/z, mass per charge

Table S2. Baseline steroid concentrations of patients with or without exogenous glucocorticoids
Undetectable concentrations were set to 0.5 times the lowest accurate calibration sample for statistical purposes and are italicized below. All concentrations are in nmol/L.

Steroids	Without glucocorticoids (n = 19) Median (IQR)	Range	With glucocorticoids (n = 15) Median (IQR)	Range	Median difference (%)
11-ketoandrostenedione	0.64 (0.43 - 1.07)	0.14 - 4.29	0.09 (0.04 - 0.22)	*0.03* - 0.96	86
11-ketotestosterone	0.87 (0.40 - 1.34)	0.12 - 2.39	0.12 (0.05 - 0.19)	*0.03* - 0.83	86
11β-hydroxyandrostenedione	4.96 (3.05 - 6.14)	0.86 - 18.77	0.41 (0.12 - 0.95)	*0.08* - 3.26	92
11β-hydroxytestosterone	0.23 (0.11 - 0.40)	*0.08* - 1.03	*0.08* (*0.08* - 0.09)	*0.08* - 0.26	66
Corticosterone	3.60 (2.50 – 6.00)	1.15 - 27.57	0.42 (0.43 - 1.90)	0.18 - 3.69	83
Cortisol	257.50 (199.30 - 326.90)	166.50 - 949.1	23.80 (18.20 - 65.80)	11.46 - 135.50	91
Cortisone	43.68 (34.26 - 58.44)	24.88 - 124.10	3.19 (0.34 - 11.56)	*0.07* - 20.71	93
Dihydrotestosterone	*0.05* (*0.05* - 0.07)	*0.05* - 0.27	*0.05* (*0.05* - 0.05)	0.05 - 0.07	0
Testosterone	0.23 (0.12 - 0.34)	0.05 - 0.76	0.06 (*0.02* - 0.12)	*0.02* - 0.15	74

IQR, interquartile range

References supplementary information

1. van der Pas R, Hofland LJ, Hofland J, et al. Fluconazole inhibits human adrenocortical steroidogenesis in vitro. J Endocrinol 2012;215(3):403-12. (In eng). DOI: JOE-12-0310 [pii] 10.1530/JOE-12-0310.
2. Snaterse G, van Dessel LF, Taylor AE, et al. Validation of multiplex steroid hormone measurements in prostate cancer using plasma for multimodality biomarker studies. medRxiv 2020:2020.08.05.20164202. DOI: 10.1101/2020.08.05.20164202.
3. O'Reilly MW, Kempegowda P, Jenkinson C, et al. 11-Oxygenated C19 Steroids Are the Predominant Androgens in Polycystic Ovary Syndrome. J Clin Endocrinol Metab 2017;102(3):840-848. (In eng). DOI: 10.1210/jc.2016-3285.
4. O'Reilly MW, Taylor AE, Crabtree NJ, et al. Hyperandrogenemia predicts metabolic phenotype in polycystic ovary syndrome: the utility of serum androstenedione. J Clin Endocrinol Metab 2014;99(3):1027-36. (In eng). DOI: 10.1210/jc.2013-3399.
5. Quanson JL, Stander MA, Pretorius E, Jenkinson C, Taylor AE, Storbeck KH. High-throughput analysis of 19 endogenous androgenic steroids by ultra-performance convergence chromatography tandem mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 2016;1031:131-138. (In eng). DOI: S1570-0232(16)30486-X [pii] 10.1016/j.jchromb.2016.07.024.
6. Pretorius E, Africander DJ, Vlok M, Perkins MS, Quanson J, Storbeck KH. 11-Ketotestosterone and 11-Ketodihydrotestosterone in Castration-resistant Prostate Cancer: Potent Androgens Which Can No Longer Be Ignored. PLoS One 2016;11(7):e0159867. (In eng). DOI: 10.1371/journal.pone.0159867 PONE-D-16-19863 [pii].
7. Hassan-Smith ZK, Morgan SA, Sherlock M, et al. Gender-Specific Differences in Skeletal Muscle 11beta-HSD1 Expression Across Healthy Aging. J Clin Endocrinol Metab 2015;100(7):2673-81. (In eng). DOI: 10.1210/jc.2015-1516.
8. Richards J, Lim AC, Hay CW, et al. Interactions of abiraterone, eplerenone, and prednisolone with wild-type and mutant androgen receptor: a rationale for increasing abiraterone exposure or combining with MDV3100. Cancer Res 2012;72(9):2176-82. (In eng). DOI: 0008-5472.CAN-11-3980 [pii] 10.1158/0008-5472.CAN-11-3980.
9. Priestley P, Baber J, Lolkema MP, et al. Pan-cancer whole-genome analyses of metastatic solid tumors. Nature 2019;575(7781):210-216. (In eng). DOI: 10.1038/s41586-019-1689-y 10.1038/s41586-019-1689-y [pii].
10. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. Genome Biol 2016;17(1):122. (In eng). DOI: 10.1186/s13059-016-0974-4 10.1186/s13059-016-0974-4 [pii].
11. Cameron DL, Baber J, Shale C, et al. GRIDSS, PURPLE, LINX: Unscrambling the tumor genome via integrated analysis of structural variation and copy-number. bioRxiv 2019:781013. DOI: 10.1101/781013.
12. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 2018;34(17):i884-i890. (In eng). DOI: 5093234 [pii] 10.1093/bioinformatics/bty560.
13. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29(1):15-21. (In eng). DOI: bts635 [pii] 10.1093/bioinformatics/bts635.
14. Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 2012;22(9):1760-74. (In eng). DOI: 22/9/1760 [pii] 10.1101/gr.135350.111.
15. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. Bioinformatics 2015;31(12):2032-4. (In eng). DOI: btv098 [pii] 10.1093/bioinformatics/btv098.
16. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15(12):550. (In eng). DOI: s13059-014-0550-8 [pii] 10.1186/s13059-014-0550-8.
17. Laurens Van Der M. Accelerating t-SNE using tree-based algorithms. J Mach Learn Res 2014;15(1):3221–3245.
18. Stephens M. False discovery rates: a new deal. Biostatistics 2016;18(2):275-294. DOI: 10.1093/biostatistics/kxw041.

Part 4

Chapter 9

Discussion and general conclusion
Published in part as:

Fundamentals of liquid biopsies
in metastatic prostate cancer:
from characterization to stratification

Chapter 9

Lisanne F. van Dessel, John W.M. Martens, Martijn P. Lolkema

Introduction

From recent genomic analyses we know that metastatic prostate cancer has a highly complex genomic makeup and that only few distinctive characteristics recur, like androgen receptor (AR) alterations and *TMPRSS2-ERG* gene fusion rearrangements[1-3]. This challenges our understanding of the biology of metastatic prostate cancer and of what alterations might be actionable targets or predictive markers to guide treatment decisions. In this thesis I unraveled the genomic landscape of metastatic prostate cancer using whole-genome sequencing on a large patient cohort. In the future, this might guide researchers and clinicians to set up clinical trials to personalize cancer treatment. Next to better understanding the biology of the disease, this thesis focused on new blood-based biomarkers with prognostic and predictive properties to further optimize treatment selection and evaluation. This research is the first step to transit from metastatic biopsies, that only reflect a spatial and temporal snapshot of tumor heterogeneity, to liquid biopsies, that enables the study of cancer biology over time. For this I first optimized the preanalytical conditions and next, I demonstrated the potential applications of liquid biopsies in metastatic prostate cancer. Furthermore, I investigated the role of 11-ketotestosterone in metastatic prostate cancer, a circulating steroid that is a potent AR agonist.

Part I: Genomic landscape of metastatic prostate cancer

Metastatic prostate cancer is characterized by marked heterogeneity at clinical and genomic level. As the available treatment options are rapidly evolving for metastatic prostate cancer, it is imperative to identify predictive markers to select subgroups of patients that are likely to have an (enduring) response. This will improve patient outcome and reduce cost and excessive side-effects. Comprehensive genomic analyses in primary prostate cancer was able to classify 74% of analyzed patients into seven predefined subtypes based on ETS fusions and mutations in *SPOP*, *FOXA1* and *IDH1*[4]. Since radical treatment options like surgery and radiotherapy is the current standard for localized disease, this classification is mostly of prognostic value and does not guide treatment decision. In **chapter 2** we describe the classification of metastatic castration-resistant prostate cancer (mCRPC) patients using whole-genome sequencing. Using unsupervised clustering based on genomic events, we defined eight distinct genomic clusters:

A) Microsatellite Instability (MSI) signature positive phenotype with high tumor mutational burden (TMB) and association with mismatch repair deficiency; 6.6% of our cohort;

B) Tandem duplication (> 100 kbp) phenotype associated with biallelic *CDK12* inactivation; 6.6% of our cohort;

D) Homologous Recombination Deficiency (HRD) features with many (> 100 kbp) deletions and association with (somatic) mutations in BRCAness-associated genes; 11.2% of our cohort;

F) Cases enriched for chromothripsis; 10.2% of our cohort;

C, E, G, H) Non-significant genomic signature without any currently known biological association; 65.4% of our cohort.

Our clustering is in line with recent reports demonstrating that structural variations arise from specific alterations such as *CDK12*[-/-] and *BRCA2*[-/-] genotypes[5-7]. Several clusters include patients that may be eligible for targeted therapy. The first cluster (A) contains patients with MSI and high TMB who might be eligible for therapy with immune check-point inhibitors, like pembrolizumab. For other tumor types with similar characteristics a high sensitivity to immune check-point inhibitors is reported[8-10]. Two clinical trials, the KEYNOTE-028 and the KEYNOTE-199, have demonstrated durable antitumor activity in a subset of metastatic prostate cancer patients[11,12]. Patient selection was predominantly based on PD-L1 expression, the target of pembrolizumab. Unfortunately, the genomic landscape of the tumors of these patients is not reported up to now. Cluster D includes patients with HRD features who may be eligible for treatment with poly(ADP-ribose) polymerase (PARP) inhibitors that block DNA damage response (DDR) pathways. The TOPARP-B and the PROfound trial showed that the PARP inhibitor olaparib had antitumor activity and improved overall survival in mCRPC patients with DDR gene aberrations, including *BRCA2* deficiency[13,14]. These promising results should be further validated in large prospective trials. Nevertheless, a large population of mCRPC patients (cluster C, E, G, H; approx. 66% of patients) do not fall into an as-of-yet clinically-relevant or biologically-clear genotype and further research should focus on identifying oncogenic driver events and investigate new actionable targets. Importantly, we established a new CTC-derived cell line that is similar to Cluster E at a genomic level as is shown in **chapter 6**. This cell line can help further research to elucidate the biology behind this cluster of mCRPC patients.

Part II: Liquid biopsies

Advances in standardization of preanalytical conditions

To implement liquid biopsies in routine cancer care, it is vital to standardize the preanalytical conditions and workflows to guarantee accurate and consistent results. Moreover, there is an urgent need for prospective clinical studies for analytical and clinical validation of liquid biopsy-based assays.

Sample collection
In **chapter 3** we investigated several preanalytical conditions to optimally preserve cell-free DNA/circulating tumor DNA (cfDNA/ctDNA) for downstream analyses. We found that using different blood collection tubes did not affect ctDNA detection, but did stabilize cfDNA concentrations over time. Since a lot of studies report variant allele frequencies (VAF) this could potentially result in an underestimation of the VAF when a sample is not processed in due time. Several studies have tested different available blood collection tubes to optimally preserve liquid biopsies as well, and

in line with our results, it appears that when processing samples within a few days, the selected tube is trivial[15-23]. Table 1 provides an overview of the most commonly used blood collection tubes. For RNA-based circulating tumor cell (CTC) analysis an EDTA tube processed within 48 hours is the optimal matrix[24,25]. For cfDNA analysis, EDTA tubes processed within six hours, or a cell-stabilization tube processed within 2-7 days perform equally well. Further essential improvements are to use a two-step high-speed protocol and limit to one freeze-thaw cycle before proceeding with cfDNA extraction[26-30]. The timing of the second high-speed centrifugation step (immediately or after storage at -80°C) does not affect cfDNA levels[16].

Table 1. Most commonly used blood collection tubes for liquid biopsy

Tube type	Specimen	Analyte	Cell stabilization
K$_2$EDTA	Plasma	cfDNA CTC	Non-stabilized
Cell-Free DNA BCT®	Plasma	cfDNA CTC	Stabilized
PAXgene Blood ccfDNA	Plasma	cfDNA	Stabilized
Cell-Free DNA BCT®	Plasma	cfDNA	Stabilized
CellSave Preservation Tubes	Plasma	cfDNA CTC	Stabilized
TransFix/EDTA Vacuum BCT	Plasma	CTC	Stabilized

CTC, circulating tumor cell
cfDNA, cell-free DNA

Optimization of cell-free DNA extraction and quantification

The next step in cfDNA analysis is the extraction of DNA from plasma, for this the QIAamp Circulating Nucleic Acid kit (Qiagen, Venlo, The Netherlands) is considered to be the gold standard method with the highest cfDNA recovery[31,32]. However, automated approaches are preferred, as clinical routine use of liquid biopsies requires high throughput and limited operator-based variability. In **chapter 4** we describe a head-to-head comparison of the QIAamp and two automated isolation methods, the QIAsymphony (Qiagen) and Maxwell platform (Promega, Madison, WI, USA). In our hands, the QIAsymphony performed comparable to QIAamp. Although the cfDNA yield was lowest using the Maxwell platform, the VAF was comparable across all platforms. Data from a multicenter evaluation study showed that the QIAsymphony and Maxwell platform are both acceptable automated isolation methods, which is in line with our results[21].

By increasing plasma input volume for cfDNA extraction and minimizing the extraction volume, an improvement in cfDNA concentration can be achieved. In **chapter 4**, we further optimized the QIAamp protocol by re-eluting three times and thereby improving cfDNA quantity. Accurate cfDNA quantification is key for downstream analysis and the use of a quantitative polymerase chain reaction (qPCR)-based assay over spectrophotometry or fluorometric assay allows for optimal accuracy and consistent measurements and is therefore recommended[21,28].

Despite ongoing efforts to optimize cfDNA/ctDNA quantity and quality, recovery efficiency and associated sensitivity to detect and analyze cfDNA/ctDNA remains the biggest hurdle to overcome. Considerable progress has been made in standardization of preanalytical conditions and future research should concentrate on technical advancements like molecular barcoding to improve cfDNA/ctDNA detection sensitivity. Furthermore, our research should focus on unravelling the biology behind cfDNA/ctDNA to gain more insight into the seemingly volatile levels of cfDNA/ctDNA. We need to better understand what drives the release and clearance of cfDNA/ctDNA, which cells and cellular processes are involved, and learn more about the behavior of cfDNA/ctDNA in blood and other body fluids. This knowledge can be used to direct future research and improve our interpretation of results leading to more accurate decisions affecting patient outcome.

Methods for CTC enrichment

More than a dozen CTC enrichment and detection platforms for CTC detection and quantification are available today, but the CellSearch system is the only FDA-approved platform (Table 2)[33]. The CellSearch system magnetically separates CTCs from other blood cells using EpCAM-coated ferrofluids, which is an epithelial marker. The limited sensitivity of the available CTC enrichment and detection platforms promotes the development of alternative methods for CTC enrichment as in 80% of metastatic prostate cancer patients at least one CTC can be detected, with a median count of 9 CTCs in 7.5 mL of blood[34].

To isolate CTCs for purposes such as *ex vivo* expansion and to assess heterogeneity, filtration-based microwells and diagnostic leukapheresis are interesting unbiased methods based on biophysical characteristics, including cell size and density. Filtration-based microwells select for single and viable CTCs based on their size and rigidity, however recovery efficiency needs to be optimized as only ~ 2% of CTCs was retrieved in a recent study[35]. As described in **chapter 6**, the product of diagnostic leukapheresis yields thousands of viable CTCs and the use of large blood volumes increases the yield compared to a single tube of blood, which is in line with previous reports[36,37]. However, sufficient white blood cell depletion still poses a challenge when using the RosetteSep depletion cocktail with or without subsequent EpCAM-based selection of CTCs as demonstrated in **chapter 6**. Another promising approach to isolate CTCs in an unbiased manner is the use of oncofetal chondroitin sulfate (ofCS), a universally expressed cancer cell surface marker on cells of both epithelial and mesenchymal origin. A feasibility study efficiently isolated CTCs based on magnetic Dynabeads coated with recombinant VAR2CSA protein, specifically targeting ofCS, in patients with various types and stages of cancer, including prostate cancer[38]. This and other methods should be investigated further to improve CTC enrichment.
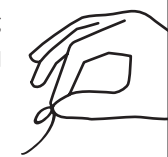
Table 2. FDA-approved liquid biopsy tests

Analyte	Platform/assay	Test characteristic	Marker type	Tumor type
CTC	CellSearch system	Isolation/enumeration	Prognostic	Metastatic breast, prostate and colorectal cancer
cfDNA	Cobas EGFR Mutation Test v2	Qualitative detection of specific mutations of the *EGFR* gene	Predictive	Non-small cell lung cancer
cfDNA	Therascreen PIK3CA RGQ PCR Kit	Qualitative detection of specific mutations of the *PIK3CA* gene	Predictive	Breast cancer
cfDNA	Epi proColon	Qualitative detection of the methylation status of the *SEPT9* promotor	Diagnostic (screening)	Colorectal cancer

CTC, circulating tumor cell
cfDNA, cell-free DNA

Liquid biopsy as proxy for tissue biopsy

A tissue biopsy might fail to capture all genomic aberrations as heterogeneity exists within a particular tumor site but also between multiple tumor loci. A liquid biopsy has the potential to better represent spatial tumor heterogeneity as it is assumed that it arises from multiple if not all metastatic sites in the patient. For a liquid biopsy to serve as a proxy for tissue biopsy, it is a prerequisite that they are an accurate reflection of the tumor's molecular aberrations. Interestingly, success rates for obtaining tumor tissue from bone and detection of ctDNA in prostate cancer is fairly similar and can vary from 25% to 75%[39-43].

Concordance between liquid vs. tissue biopsies

Genomic analyses of ctDNA in mCRPC have shown to recapitulate the genomic landscape of metastatic prostate cancer as established by comparison to (matched) tissue biopsies[44-47]. The most commonly affected genes in ctDNA are *TP53*, *AR*, *FOXA1*, *APC*, *PTEN*, *RB1*, *MYC*, *BRCA2* and *SPOP*. Interestingly, 33% of somatic mutations in ctDNA were not detected in matched tissue biopsy and some of those were clear cancer drivers[45]. In part, this was traced back to technical reasons, like the use of different sequencing techniques for tissue and liquid biopsies and a difference in sequencing depth. In a cohort of newly diagnosed metastatic hormone-sensitive prostate cancer patients *TP53* mutations and DNA damage repair defects (*ATM*, *ATR*, *BRCA2*, *CDK12*, *MSH2* and *RAD51C*) could be identified in cfDNA in respectively 47% and 21% of patients using targeted next-generation sequencing (NGS)[48]. No *AR* alterations were observed. This was for 80% concordant with matched prostate tumor tissue.

Molecular analyses of CTCs showed that CTCs can provide a reliable overview of the tumor's genomic landscape[37,49]. CTCs with unique mutations might be derived from minor subclones that are undetectable in tissue biopsies[49]. Only a few CTCs appear to be sufficient to represent up to half of the mutations detected in matched metastatic biopsies[49].

Overall, evidence so far indicates that both ctDNA and CTCs are an equivalent alternative for tissue biopsies. With the superior detection rate of ctDNA currently, it might be more sensible to use ctDNA over CTCs for genomic analyses, especially in the setting of minimal residual disease detection where CTCs are rare.

Concordance among liquid biopsies

Assessing the concordance between CTCs and cfDNA in mCRPC identified shared alterations in *TP53*, *FOXA1*, *AR*, *MYC*, *BRCA1*, *PTEN*, and *RB1*[50,51]. Furthermore, cfDNA concentration, ctDNA fraction and CTC number appear to be associated[50,52]. However, some genomic alterations, like *MYCN* gain, showed considerable discordance between CTCs and cfDNA and may suggest biological differences between these entities[50]. Some differences may be traced back to their origin, as CTCs are shed by tumor sites and ctDNA is shed from apoptotic or necrotic tumor cells, or to differences in sensitivity between methods, but further research is needed to elucidate this. In **chapter 6**, analysis of concordance between individual CTCs within a patient demonstrates both shared and unique copy-number alterations which could indicate different subclones with a shared genetic background and reflects the intra-patient and inter-tumor site heterogeneity. These results are in line with other reports[37,51,53].

The prognostic value of liquid biopsy burden

CTCs are the most studied type of liquid biopsy in metastatic prostate cancer and CTC counts using the CellSearch system have proven its prognostic value in mCRPC, whereby patients with ≥ 5 CTCs/7.5 mL of blood have an unfavorable prognosis[54]. Increasing numbers of CTCs over time also have prognostic significance[55].

Evidence for the prognostic value of cfDNA and ctDNA levels is also accumulating, but clear cut-off levels have not been established yet. CtDNA level and fraction (% of total cfDNA) both associate with clinical features of tumor burden, including metastasis location and levels of alkaline phosphatase, lactate dehydrogenase, hemoglobin and prostate specific antigen (PSA). Furthermore, ctDNA fraction gradually increases with number of lines of therapy[52]. Importantly, increased ctDNA fraction, defined as at least ≥ 10% ctDNA fraction or VAF ≥ 7%, was associated with worse PFS in mCRPC patients treated with AR-targeted therapy (ARTT, e.g. enzalutamide, abiraterone, apalutamide)[44,47,58]. This has to be further validated in prospective clinical trials.

Circulating tumor cell-based stratification using organoids

The use of living cells as 'real life' drug screening models, is a captivating outlook to personalize cancer treatment. Several clinical trials (e.g. TUMOROID NCT03821870) are currently including patients to investigate if treatment response of tumor biopsy derived organoids can predict treatment response in patients. Patient derived organoids of luminal-breast and metastatic gastro-intestinal cancers have been shown to recapitulate patient responses in the clinic[59,60]. Phenotyping living tumor cells has the advantage to directly measure the response to treatment compared to 'phenotyping after fixation' based stratification[61]. In **chapter 6**, we aimed to generate CTC derived organoids from metastatic prostate cancer patients. We obtained organoids in 9/18 samples after short-term culture, in which prostate cancer origin could be confirmed. Furthermore, we were able to establish one long-term organoid cell line. Up to now, efforts to establish prostate cancer cell lines have proven that prostate cancer is an inherently difficult tumor type to culture *in*

vitro and *in vivo*, so our long-term cell line is a valuable contribution to the scientific community. Interestingly, we could confirm the patient's clinical response to enzalutamide in the organoid *ex vivo*, thereby providing proof-of-concept that individualized CTC derived organoids may serve as 'real life' drug sensitivity screening model in metastatic prostate cancer. However, the current take rate and growth speed of CTC derived organoids is too limited to allow guided treatment in the clinic. Our work showed that using diagnostic leukapheresis can overcome the main obstacle, which is the recovery of a sufficient number of CTCs to start organoid culture. Future efforts should investigate how to further improve the enrichment process to obtain a purer CTC sample using an unbiased method. In addition, culturing conditions need to be further improved and the use of short-term cultures need to be maximized by investing in the development of methods for visualization and cell viability assays.

Circulating tumor DNA-based stratification using molecular markers

Up to now, there are four FDA approved tests for the analysis of liquid biopsies that can be used for specific tumor types and applications (Table 2). Various assays are being developed that mostly focus on the detection of specific alterations using PCR or (targeted) NGS[62-65].

For the analysis of ctDNA in metastatic prostate cancer the development of commercially available targeted panels lags behind, as only few driver genes are known, compared to lung cancer or colorectal cancer for example, for which actionable driver genes are known. CtDNA analysis is technically challenging, therefore Schweizer et al. investigated clinical determinants to improve ctDNA detection rate. They found that disease burden and CRPC status was significantly associated with successful detection of ctDNA alterations by NGS approaches[66]. These results have to be interpreted with caution since false positive results can occur due to low VAFs and clonal hematopoiesis[52].

AR alterations (amplifications and/or mutations) in ctDNA were predictive for poor response to ARTT and had a negative prognostic value for overall survival (OS) in mCRPC patients treated with ARTT[44,67,68], and could be detected in 39% of pretreated mCRPC patients using a targeted NGS approach[69]. In a retrospective analysis in mCRPC patients, *AR* gain in ctDNA was associated with a longer progression-free survival (PFS) to docetaxel[70]. Furthermore, data suggest that *AR* copy-number-neutral mCRPC patients might benefit more from ARTT[70-72] and missense mutations in the ligand-binding domain of *AR* was associated with a shorter PFS[58].

Recent reports showed that alterations in *BRCA2*, *TP53*, *ATM*, and *PI3K* pathway detected in ctDNA have an unfavorable prognosis in CRPC patients, independently of clinical prognostic factors and ctDNA level[44,58,73]. These alterations were found in patients with poor response to ARTT, suggesting primary resistance[44]. Interestingly, *TP53* inactivation outperformed any AR-derived biomarker tested in CTCs and ctDNA[73].

Detection of DNA damage repair alterations in ctDNA might be predictive for response and stratify patients for targeted therapies. In a prospective cohort of mCRPC patients before start of ARTT, > 25% of patients had deleterious mutations in *BRCA1/2* or *ATM*[58]. Interestingly, another

trial targeting HRD with the PARP inhibitor olaparib detected at disease progression multiple aberrations in cfDNA that restored the function of targeted genes *(BRCA2, PALB2)* as mechanisms of resistance[74,75]. MSI signature with high TMB and association with mismatch repair deficiency can be identified in ctDNA as well. This genomic phenotype has been identified in ~ 4% of ctDNA from patients with metastatic prostate cancer and might stratify patients for immunotherapy[52,76]. Prospective randomized trials are needed to confirm the utility of ctDNA analysis as a prognostic and/or predictive biomarker. Ideally, an unbiased NGS approach should be used to characterize ctDNA upfront and the overall molecular makeup of the tumor, rather than just a single driver gene, should be used to stratify patients for the appropriate treatment.

Response evaluation using liquid biopsies

In CRPC it is challenging to accurately monitor treatment response with the current available response markers. PSA and radiological imaging are poor indicators of tumor load and reflect late response.

Both CTC counts and CTC-expression markers may inform on response. Heller et al. investigated CTC count and PSA response at treatment week 13 in mCRPC patients from five prospective trials (COU-AA-301, AFFIRM, ELM-PC-5, ELM-PC-4, and COMET-1)[77]. The endpoints 'CTC0' (≥ 1 CTCs at baseline and 0 at 13 weeks) and 'CTC conversion' (≥ 5 CTCs at baseline and ≤ 4 at 13 weeks) provided greater discriminatory power for OS than the endpoints percent change in CTC count or PSA response. Using the 'CTC0' endpoint 75% of eligible patients could be evaluated against 51% for the 'CTC conversion' endpoint, making this a clinically meaningful response measure. Lorente et al. demonstrated that 'CTC progression' (< 5 CTCs at baseline and increasing) during the first 12 weeks of treatment was independently associated with reduced survival in mCRPC patients from the COU-AA-301 and IMMC-38 trials[78]. Paller et al. observed dynamic changes in prostate specific membrane antigen (PSMA) expression on CTCs during therapy and a decrease in PSMA expression was associated with concurrent decreases in serum PSA[79].

Quantification of ctDNA may be a biomarker for response as well. Promising results from the TOPARP-A trial (NCT01682772), a study of the PARP inhibitor olaparib in mCRPC, showed that VAF of somatic DNA repair mutations in cfDNA declined in responders[74]. Furthermore, responders showed > 50% decrease in cfDNA level after 4 weeks of treatment and a decrease in cfDNA levels was independently associated with OS. In line with this, patients responding to ARTT appear to have lower ctDNA fractions[68,80]. Mehra et al. found that changes in cfDNA levels correlate with both radiologic PFS and OS in mCRPC patients on chemotherapy[57].

In **chapter 5**, we aimed to detect structural variants in ctDNA of mCRPC patients with the ultimate goal to evaluate treatment response. Structural variants are genomic rearrangements of ≥ 50 base pairs and are widespread in cancer. In prostate cancer predominant alterations in involved signaling pathways are often caused by structural variants like amplification of *AR* and *MYC*, and ETS-fusion genes. Large structural variants, such as inter-chromosomal translocations

or long-range intra-chromosomal rearrangements, are of special interest since they represent unique DNA-molecules that are only present in tumor cells. These properties make them a specific cancer biomarker. To detect structural variants in ctDNA, we used a novel methodology called Oxford Nanopore portable sequencing to first identify structural variants in tumor tissue obtained from metastases. After a bioinformatic filtering pipeline and using Breakpoint PCR, we were able to detect structural variants in ctDNA. Our data suggest that the quantitative measurement of structural variants in ctDNA correlates with tumor load and that it may indicate disease progression earlier than PSA. Eventually, future studies should inform whether early intervention based on liquid biopsy dynamics improves patient outcome compared to awaiting changes in PSA level and radiological evaluation. Nevertheless, our selection of structural variants needs to be further optimized, as in some patients we could not demonstrate the correlation with tumor load and disease progression, which may indicate that we did not select structural variants that represent the dominant tumor driving disease clone or that we missed the emergence of new resistant clones as we used a targeted approach for the longitudinal samples. When these new resistant clones would provide new targets for treatment, this would be of great interest to try to improve patient outcome.

Part III: Steroidomics

In CRPC, tumor cells evolve to escape androgen suppression and disease progression occurs at castrate levels of testosterone. Nevertheless, several reports from Japan showed that the level of testosterone may be a prognostic and predictive marker in CRPC. Patients with high testosterone levels (> 0.05 ng/mL) before start of ARTT tended to have a better PSA response and PFS[81-83]. An inverse association was observed for taxane-based chemotherapy[82,84]. Interestingly, another study showed that a further decline of testosterone levels during docetaxel was associated with a longer PSA response and PFS[85]. This supports the concept of the continued importance of the AR pathway in CRPC and renewed interest has grown for steroidogenesis as it modulates circulating steroid levels.

In vitro studies investigated the function of isozymes 17βHSD2 and 17βHSD4, whose activities leads to androgen inactivation, and 11βHSD2, which converts adrenal hydroxy-steroids (C11-oxy C_{19}) to more potent keto-steroids (C11-keto C_{19}, including the AR agonists 11-ketotestosterone (11KT) and 11-ketodihydrotestosterone (11KDHT)). 17βHSD2 and 17βHSD4 variant 2 was functionally silenced in prostate cancer cell lines and a xenograft model, indicating its tumor suppressor effect on androgen conversion[86,87]. By contrast, 11βHSD2 activity in prostate cancer cell lines probably contributes to prostate cancer progression by directly increasing 11KT concentration and through an interaction with cytochrome P450 17A1, thereby converting C11-oxy C_{21} steroids to C11-oxy C_{19} steroids via the C11-oxy backdoor pathway[88].

In **chapter 7** we show that CellSave preservation tubes, validated for CTC and cfDNA analysis, can also be used for steroid measurements, thereby optimizing the samples used in clinical trials and limiting the number of tubes that need to be draw from a patient. In **chapter 8**, we demonstrate that 11KT is the most abundant androgen in mCRPC patients and that its circulating levels exceed those of testosterone. In line with above mentioned results, we found that PFS was longer in our patients with high total androgen concentrations. Our and other reported data support the concept of continued importance of the AR pathway in CRPC. We believe that the 11-oxygenated androgens are a major contributor to AR activation after castration. Our findings indicate that we have structurally underestimated androgen levels in prostate cancer patients and this implies that we should reconsider our definition for 'castration'.

General conclusion

This thesis describes the genomic landscape of metastatic prostate cancer in more detail. We have identified several distinct genomic clusters looking at specific genomic events or so-called 'genomic scars'. This knowledge was incorporated to give a prospect on patient stratification using liquid biopsies.

I believe that the use of liquid biopsies in metastatic prostate cancer has great potential as prognostic, predictive and therapy response biomarker. The main advantage of liquid biopsy analysis is that it is acquired by a simple blood draw, which is safe, minimally invasive and repeatable over time. Moreover, it provides a real-time snapshot of the biologic variables that affect tumor growth and progression in metastatic prostate cancer. Our work on standardization of preanalytical conditions that affect liquid biopsy analysis already has and continues to drive guidelines and consensus statements to implement liquid biopsies in clinical practice.

Comprehensive analyses of ctDNA and CTCs demonstrate that they adequately reflect the genomic makeup of the tumor and may thus complement or even replace tumor biopsies. The assessment of genomic aberrations like structural variants in ctDNA can potentially predict therapy response and detect mechanisms of resistance. CTCs cannot only be enumerated as prognosticator in metastatic prostate cancer, but can also be used to culture *ex vivo* thereby providing novel disease models for metastatic prostate cancer and to predict therapy response.

Furthermore, steroidomics has proven its continued importance in metastatic prostate cancer with the establishment of clinical relevance of significant levels of circulating 11-oxygenated androgens.

The clinical utility of blood-based markers like liquid biopsies and circulating steroids as predictive and response markers has to be further validated in prospective clinical trials and should therefore be routinely sampled as PSA and radiological imaging is. Essential to the work on identifying novel biomarkers, is to further elucidate the complex biology of metastatic prostate cancer and to decode the oncogenic driver events and potential actionable targets of the large population of mCRPC without characteristic genomic scars.

References

1. Robinson D, Van Allen EM, Wu YM, et al. Integrative Clinical Genomics of Advanced Prostate Cancer. Cell 2015;161(5):1215-1228. (In English). DOI: 10.1016/j.cell.2015.05.001.

2. Fraser M, Sabelnykova VY, Yamaguchi TN, et al. Genomic hallmarks of localized, non-indolent prostate cancer. Nature 2017;541(7637):359-364. (In eng). DOI: nature20788 [pii]
    10.1038/nature20788.

3. van Dessel LF, van Riet J, Smits M, et al. The genomic landscape of metastatic castration-resistant prostate cancers reveals multiple distinct genotypes with potential clinical impact. Nat Commun 2019;10(1):5251. DOI: 10.1038/s41467-019-13084-7.

4. Cancer Genome Atlas Research N. The Molecular Taxonomy of Primary Prostate Cancer. Cell 2015;163(4):1011-25. (In eng). DOI: S0092-8674(15)01339-2 [pii]
    10.1016/j.cell.2015.10.025.

5. Viswanathan SR, Ha G, Hoff AM, et al. Structural Alterations Driving Castration-Resistant Prostate Cancer Revealed by Linked-Read Genome Sequencing. Cell 2018;174(2):433-447.e19. (Article) (In English). DOI: 10.1016/j.cell.2018.05.036.

6. Quigley DA, Dang HX, Zhao SG, et al. Genomic Hallmarks and Structural Variation in Metastatic Prostate Cancer. Cell 2018;175(3):889. (In eng). DOI: S0092-8674(18)31328-X [pii]
    10.1016/j.cell.2018.10.019.

7. Wu YM, Cieslik M, Lonigro RJ, et al. Inactivation of CDK12 Delineates a Distinct Immunogenic Class of Advanced Prostate Cancer. Cell 2018;173(7):1770-1782 e14. (In eng). DOI: S0092-8674(18)30565-8 [pii]
    10.1016/j.cell.2018.04.034.

8. Yarchoan M, Hopkins A, Jaffee EM. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. N Engl J Med 2017;377(25):2500-2501. (In eng). DOI: 10.1056/NEJMc1713444.

9. Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science 2015;348(6230):124-8. (In eng). DOI: science.aaa1348 [pii]
    10.1126/science.aaa1348.

10. Samstein RM, Lee CH, Shoushtari AN, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet 2019;51(2):202-206. (In eng). DOI: 10.1038/s41588-018-0312-8
    10.1038/s41588-018-0312-8 [pii].

11. Hansen AR, Massard C, Ott PA, et al. Pembrolizumab for advanced prostate adenocarcinoma: findings of the KEYNOTE-028 study. Ann Oncol 2018;29(8):1807-1813. (In eng). DOI: S0923-7534(19)34146-8 [pii]
    10.1093/annonc/mdy232.

12. Antonarakis ES, Piulats JM, Gross-Goupil M, et al. Pembrolizumab for Treatment-Refractory Metastatic Castration-Resistant Prostate Cancer: Multicohort, Open-Label Phase II KEYNOTE-199 Study. J Clin Oncol 2020;38(5):395-405. (In eng). DOI: 10.1200/JCO.19.01638.

13. Mateo J, Porta N, Bianchini D, et al. Olaparib in patients with metastatic castration-resistant prostate cancer with DNA repair gene aberrations (TOPARP-B): a multicentre, open-label, randomised, phase 2 trial. Lancet Oncol 2020;21(1):162-174. (Article) (In English). DOI: 10.1016/s1470-2045(19)30684-9.

14. Hussain M MJ, Fizazi K, Saad F, Shore ND, Sandhu S, Chi KN, Sartor O, Agarwal N, Olmos D, Thiery-Vuillemin A, Twardowski P, Mehra P, Goessl C, Kang J, Burgents J, Wu W, Kohlmann A, Adelman CA, De Bono J. PROFOUND: phase 3 study of olaparib versus enzalutamide or abiraterone for metastatic castration-resistant prostate cancer with homologous recombination repair gene alterations. Annals of Oncology 2019;30(5):5:v851-v934.

15. Barták BK, Kalmár A, Galamb O, et al. Blood Collection and Cell-Free DNA Isolation Methods Influence the Sensitivity of Liquid Biopsy Analysis for Colorectal Cancer Detection. Pathol Oncol Res 2019;25(3):915-923. (Article) (In English). DOI: 10.1007/s12253-018-0382-z.

16. Cavallone L, Aldamry M, Lafleur J, et al. A study of preanalytical variables and optimization of extraction method for circulating tumor DNA measurements by digital droplet PCR. Cancer Epidemiol Biomarkers Prev 2019;28(5):909-916. (Article) (In English). DOI: 10.1158/1055-9965.Epi-18-0586.

17. Gahlawat AW, Lenhardt J, Witte T, et al. Evaluation of storage tubes for combined analysis of circulating nucleic acids in liquid biopsies. Int J Mol Sci 2019;20(3) (Article) (In English). DOI: 10.3390/ijms20030704.

18. Parackal S, Zou D, Day R, Black M, Guilford P. Comparison of Roche Cell-Free DNA collection Tubes® to Streck Cell-Free DNA BCT®s for sample stability using healthy volunteers. Prac Lab Med 2019;16 (Article) (In English). DOI: 10.1016/j.plabm.2019.e00125.

19. Sorber L, Zwaenepoel K, Jacobs J, et al. Specialized Blood Collection Tubes for Liquid Biopsy: Improving the Preanalytical Conditions. Mol Diagn Ther 2019 (Article in Press) (In English). DOI: 10.1007/s40291-019-00442-w.

20. Zhao Y, Li Y, Chen P, Li S, Luo J, Xia H. Performance comparison of blood collection tubes as liquid biopsy storage system for minimizing cfDNA contamination from genomic DNA. J Clin Lab Anal 2019;33(2) (Article) (In English). DOI: 10.1002/jcla.22670.

21. Lampignano R, Neumann MHD, Weber S, et al. Multicenter evaluation of circulating cell-free DNA extraction and downstream analyses for the development of standardized (Pre)analytical work flows. Clin Chem 2020;66(1):149-160. (Article) (In English). DOI: 10.1373/clinchem.2019.306837.

22. Ilie M, Hofman V, Leroy S, et al. Use of circulating tumor cells in prospective clinical trials for NSCLC patients-standardization of the preanalytical conditions. Clin Chem Lab Med 2018;56(6):980-989. (Article) (In English). DOI: 10.1515/cclm-2017-0764.

23. Rodríguez-Lee M, Kolatkar A, McCormick M, et al. Effect of Blood Collection Tube Type and Time to Processing on the Enumeration and High-Content Characterization of Circulating Tumor Cells Using the High-Definition Single-Cell Assay. Arch Pathol Lab Med 2018;142(2):198-207. (Article) (In English). DOI: 10.5858/arpa.2016-0483-OA.

24. Zavridou M, Mastoraki S, Strati A, Tzanikou E, Chimonidou M, Lianidou E. Evaluation of preanalytical conditions and implementation of quality control steps for reliable gene expression and DNA methylation analyses in liquid biopsies. Clin Chem 2018;64(10):1522-1533. (Article) (In English). DOI: 10.1373/clinchem.2018.292318.

25. Luk AWS, Ma Y, Ding PN, et al. CTC-mRNA (AR-V7) Analysis from Blood Samples-Impact of Blood Collection Tube and Storage Time. Int J Mol Sci 2017;18(5). DOI: 10.3390/ijms18051047.

26. Merker JD, Oxnard GR, Compton C, et al. Circulating Tumor DNA Analysis in Patients With Cancer: American Society of Clinical Oncology and College of American Pathologists Joint Review. J Clin Oncol 2018;36(16):1631-1641. (http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=medl&AN=29504847).

27. Compton CC, Robb JA, Anderson MW, et al. Preanalytics and Precision Pathology: Pathology Practices to Ensure Molecular Integrity of Cancer Patient Biospecimens for Precision Medicine. Arch Pathol Lab Med 2019;143(11):1346-1363. (In eng). DOI: 10.5858/arpa.2019-0009-SA.

28. Meddeb R, Pisareva E, Thierry AR. Guidelines for the preanalytical conditions for analyzing circulating cell-free DNA. Clin Chem 2019;65(5):623-633. (Article) (In English). DOI: 10.1373/clinchem.2018.298323.

29. Sorber L, Zwaenepoel K, Jacobs J, et al. Circulating cell-free DNA and RNA analysis as liquid biopsy: Optimal centrifugation protocol. Cancers 2019;11(4) (Article) (In English). DOI: 10.3390/cancers11040458.

30. Shishido SN, Welter L, Rodriguez-Lee M, et al. Preanalytical variables for the genomic assessment of the cellular and acellular fractions of the liquid biopsy in a cohort of breast cancer patients. J Mol Diagn 2020 (Article in Press) (In English). DOI: 10.1016/j.jmoldx.2019.11.006.

31. Diefenbach RJ, Lee JH, Kefford RF, Rizos H. Evaluation of commercial kits for purification of circulating free DNA. Cancer Genet 2018;228-229:21-27. (In eng). DOI: S2210-7762(18)30250-3 [pii]
    10.1016/j.cancergen.2018.08.005.

32. Warton K, Graham LJ, Yuwono N, Samimi G. Comparison of 4 commercial kits for the extraction of circulating DNA from plasma. Cancer Genet 2018;228-229:143-150. (In eng). DOI: S2210-7762(17)30267-3 [pii]
    10.1016/j.cancergen.2018.02.004.

33. Morrison GJ, Goldkorn A. Development and Application of Liquid Biopsies in Metastatic Prostate Cancer. Curr Oncol Rep 2018;20(4) (Review) (In English). DOI: 10.1007/s11912-018-0683-0.

34. Thalgott M, Rack B, Maurer T, et al. Detection of circulating tumor cells in different stages of prostate cancer. J Cancer Res Clin Oncol 2013;139(5):755-63. (In eng). DOI: 10.1007/s00432-013-1377-5.

35. Andree KC, Abali F, Oomens L, et al. Self-seeding microwells to isolate and assess the viability of single circulating tumor cells. Int J Mol Sci 2019;20(3) (Article) (In English). DOI: 10.3390/ijms20020477.

36. Andree KC, Mentink A, Zeune LL, et al. Toward a real liquid biopsy in metastatic breast and prostate cancer: Diagnostic LeukApheresis increases CTC yields in a European prospective multicenter study (CTCTrap). Int J Cancer 2018;143(10):2584-2591. (Article) (In English). DOI: 10.1002/ijc.31752.

37. Lambros MB, Seed G, Sumanasuriya S, et al. Single-cell analyses of prostate cancer liquid biopsies acquired by apheresis. Clin Cancer Res 2018;24(22):5635-5644. (Article) (In English). DOI: 10.1158/1078-0432.Ccr-18-0862.

38. Agerbæk MØ, Bang-Christensen SR, Yang MH, et al. The VAR2CSA malaria protein efficiently retrieves circulating tumor cells in an EpCAM-independent manner. Nat Commun 2018;9(1) (Article) (In English). DOI: 10.1038/s41467-018-05793-2.

39. McKay RR, Zukotynski KA, Werner L, et al. Imaging, procedural and clinical variables associated with tumor yield on bone biopsy in metastatic castration-resistant prostate cancer. Prostate Cancer Prostatic Dis 2014;17(4):325-31. (In eng). DOI: pcan201428 [pii]
    10.1038/pcan.2014.28.

40. Lorente D, Omlin A, Zafeiriou Z, et al. Castration-Resistant Prostate Cancer Tissue Acquisition From Bone Metastases for Molecular Analyses. Clin Genitourin Cancer 2016;14(6):485-493. (In eng). DOI: S1558-7673(16)30106-9 [pii] 10.1016/j.clgc.2016.04.016.

41. Ross RW, Halabi S, Ou SS, et al. Predictors of prostate cancer tissue acquisition by an undirected core bone marrow biopsy in metastatic castration-resistant prostate cancer--a Cancer and Leukemia Group B study. Clin Cancer Res 2005;11(22):8109-13. (In eng). DOI: 11/22/8109 [pii] 10.1158/1078-0432.CCR-05-1250.

42. Efstathiou E, Titus M, Wen S, et al. Molecular characterization of enzalutamide-treated bone metastatic castration-resistant prostate cancer. Eur Urol 2015;67(1):53-60. (In eng). DOI: S0302-2838(14)00415-1 [pii] 10.1016/j.eururo.2014.05.005.

43. Spritzer CE, Afonso PD, Vinson EN, et al. Bone marrow biopsy: RNA isolation with expression profiling in men with metastatic castration-resistant prostate cancer--factors affecting diagnostic success. Radiology 2013;269(3):816-23. (In eng). DOI: radiol.13121782 [pii] 10.1148/radiol.13121782.

44. Annala M, Vandekerkhove G, Khalaf D, et al. Circulating tumor DNA genomics correlate with resistance to abiraterone and enzalutamide in prostate cancer. Cancer Discov 2018;8(4):444-457. (Article) (In English). DOI: 10.1158/2159-8290.Cd-17-0937.

45. Wyatt AW, Annala M, Aggarwal R, et al. Concordance of Circulating Tumor DNA and Matched Metastatic Tissue Biopsy in Prostate Cancer. J Natl Cancer Inst 2017;109(12). DOI: 10.1093/jnci/djx118.

46. Sonpavde G, Agarwal N, Pond GR, et al. Circulating tumor DNA alterations in patients with metastatic castration-resistant prostate cancer. Cancer 2019;125(9):1459-1469. (Article) (In English). DOI: 10.1002/cncr.31959.

47. Belic J, Graf R, Bauernhofer T, et al. Genomic alterations in plasma DNA from patients with metastasized prostate cancer receiving abiraterone or enzalutamide. Int J Cancer 2018;143(5):1236-1248. (Article) (In English). DOI: 10.1002/ijc.31397.

48. Vandekerkhove G, Struss WJ, Annala M, et al. Circulating Tumor DNA Abundance and Potential Utility in De Novo Metastatic Prostate Cancer. Eur Urol 2019;75(4):667-675. (Article) (In English). DOI: 10.1016/j.eururo.2018.12.042.

49. Faugeroux V, Lefebvre C, Pailler E, et al. An Accessible and Unique Insight into Metastasis Mutational Content Through Whole-exome Sequencing of Circulating Tumor Cells in Metastatic Prostate Cancer. Eur Urol Oncol 2019 (Article in Press) (In English). DOI: 10.1016/j.euo.2018.12.005.

50. Gupta S, Hovelson DH, Kemeny G, et al. Discordant and heterogeneous clinically relevant genomic alterations in circulating tumor cells vs plasma DNA from men with metastatic castration-resistant prostate cancer. Genes Chromosomes Cancer 2019 (Article in Press) (In English). DOI: 10.1002/gcc.22824.

51. Hodara E, Morrison G, Cunha A, et al. Multiparametric liquid biopsy analysis in metastatic prostate cancer. JCI Insight 2019;4(5) (Article) (In English). DOI: 10.1172/jci.insight.125529.

52. Mayrhofer M, De Laere B, Whittington T, et al. Cell-free DNA profiling of metastatic prostate cancer reveals microsatellite instability, structural rearrangements and clonal hematopoiesis. Genome Med 2018;10(1) (Article) (In English). DOI: 10.1186/s13073-018-0595-5.

53. Ferrarini A, Forcato C, Buson G, 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(3):e0193689. (http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=medc1&AN=29494651).

54. Moreno JG, Miller MC, Gross S, Allard WJ, Gomella LG, Terstappen LW. Circulating tumor cells predict survival in patients with metastatic prostate cancer. Urology 2005;65(4):713-8. DOI: 10.1016/j.urology.2004.11.006.

55. Danila DC, Heller G, Gignac GA, et al. Circulating tumor cell number and prognosis in progressive castration-resistant prostate cancer. Clin Cancer Res 2007;13(23):7053-8. DOI: 10.1158/1078-0432.CCR-07-1506.

56. Choudhury AD, Werner L, Francini E, et al. Tumor fraction in cell-free DNA as a biomarker in prostate cancer. JCI Insight 2018;3(21) (Article) (In English). DOI: 10.1172/jci.insight.122109.

57. Mehra N, Dolling D, Sumanasuriya S, et al. Plasma Cell-free DNA Concentration and Outcomes from Taxane Therapy in Metastatic Castration-resistant Prostate Cancer from Two Phase III Trials (FIRSTANA and PROSELICA). Eur Urol 2018;74(3):283-291. (Article) (In English). DOI: 10.1016/j.eururo.2018.02.013.

58. Torquato S, Pallavajjala A, Goldstein A, et al. Genetic alterations detected in cell-free DNA are associated with enzalutamide and abiraterone resistance in castration-resistant prostate cancer. JCO Precis Oncol 2019;3 (Article) (In English). DOI: 10.1200/po.18.00227.

59. Vlachogiannis G, Hedayat S, Vatsiou A, et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. Science 2018;359(6378):920-926. (In eng). DOI: 359/6378/920 [pii] 10.1126/science.aao2774.

60. Yu M, Bardia A, Aceto N, et al. Cancer therapy. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. Science 2014;345(6193):216-20. (In eng). DOI: 345/6193/216 [pii] 10.1126/science.1253533.

61. Ooft SN, Weeber F, Dijkstra KK, et al. Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. Sci Transl Med 2019;11(513) (In English). DOI: 11/513/eaay2574 [pii] 10.1126/scitranslmed.aay2574.

62. Gao M, Callari M, Beddowes E, et al. Next Generation-Targeted Amplicon Sequencing (NG-TAS): An optimised protocol and computational pipeline for cost-effective profiling of circulating tumor DNA. Genome Med 2019;11(1) (Article) (In English). DOI: 10.1186/s13073-018-0611-9.

63. Taavitsainen S, Annala M, Ledet E, et al. Evaluation of commercial circulating tumor DNA test in metastatic prostate cancer. JCO Precis Oncol 2019;3 (Article) (In English). DOI: 10.1200/po.19.00014.

64. Troll CJ, Kapp J, Rao V, et al. A ligation-based single-stranded library preparation method to analyze cell-free DNA and synthetic oligos. BMC Genomics 2019;20(1) (Article) (In English). DOI: 10.1186/s12864-019-6355-0.

65. Vitale SR, Sieuwerts AM, Beije N, et al. An Optimized Workflow to Evaluate Estrogen Receptor Gene Mutations in Small Amounts of Cell-Free DNA. J Mol Diagn 2019;21(1):123-137. (Article) (In English). DOI: 10.1016/j.jmoldx.2018.08.010.

66. Schweizer MT, Gulati R, Beightol M, et al. Clinical determinants for successful circulating tumor DNA analysis in prostate cancer. Prostate 2019;79(7):701-708. (Article) (In English). DOI: 10.1002/pros.23778.

67. Lolli C, De Lisi D, Conteduca V, et al. Testosterone levels and androgen receptor copy-number variations in castration-resistant prostate cancer treated with abiraterone or enzalutamide. Prostate 2019;79(11):1211-1220. (Article) (In English). DOI: 10.1002/pros.23804.

68. Romanel A, Gasi Tandefelt D, Conteduca V, et al. Plasma AR and abiraterone-resistant prostate cancer. Sci Transl Med 2015;7(312):312re10. DOI: 10.1126/scitranslmed.aac9511.

69. Moses M, Koksal U, Ledet E, et al. Evaluation of the genomic alterations in the androgen receptor gene during treatment with high-dose testosterone for metastatic castrate-resistant prostate cancer. Oncotarget 2020;11(1):15-21. (Article) (In English) (http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L630782915 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6967778/pdf/oncotarget-11-15.pdf).

70. Conteduca V, Jayaram A, Romero-Laorden N, et al. Plasma Androgen Receptor and Docetaxel for Metastatic Castration-resistant Prostate Cancer. Eur Urol 2019;75(3):368-373. (Article) (In English). DOI: 10.1016/j.eururo.2018.09.049.

71. Conteduca V, Castro E, Wetterskog D, et al. Plasma AR status and cabazitaxel in heavily treated metastatic castration-resistant prostate cancer. Eur J Cancer 2019;116:158-168. (Article) (In English). DOI: 10.1016/j.ejca.2019.05.007.

72. Sumiyoshi T, Mizuno K, Yamasaki T, et al. Clinical utility of androgen receptor gene aberrations in circulating cell-free DNA as a biomarker for treatment of castration-resistant prostate cancer. Sci Rep 2019;9(1):4030. (Article) (In English). DOI: 10.1038/s41598-019-40719-y.

73. De Laere B, Oeyen S, Mayrhofer M, et al. TP53 outperforms other androgen receptor biomarkers to predict abiraterone or enzalutamide outcome in metastatic castration-resistant prostate cancer. Clin Cancer Res 2019;25(6):1766-1773. (Article) (In English). DOI: 10.1158/1078-0432.Ccr-18-1943.

74. Goodall J, Mateo J, Yuan W, et al. Circulating Cell-Free DNA to Guide Prostate Cancer Treatment with PARP Inhibition. Cancer Discov 2017;7(9):1006-1017. DOI: 10.1158/2159-8290.CD-17-0261.

75. Quigley D, Alumkal JJ, Wyatt AW, et al. Analysis of Circulating Cell-Free DNA Identifies Multiclonal Heterogeneity of BRCA2 Reversion Mutations Associated with Resistance to PARP Inhibitors. Cancer Discov 2017;7(9):999-1005. DOI: 10.1158/2159-8290.CD-17-0146.

76. Ritch E, Fu SYF, Herberts C, et al. Identification of hypermutation and defective mismatch repair in ctDNA from metastatic prostate cancer. Clin Cancer Res 2019 (Article in Press) (In English). DOI: 10.1158/1078-0432.Ccr-19-1623.

77. Heller G, McCormack R, Kheoh T, et al. Circulating tumor cell number as a response measure of prolonged survival for metastatic castration-resistant prostate cancer: A comparison with prostate-specific antigen across five randomized phase III clinical trials. J Clin Oncol 2018;36(6):572-580. (Conference Paper) (In English). DOI: 10.1200/jco.2017.75.2998.

78. Lorente D, Olmos D, Mateo J, et al. Circulating tumor cell increase as a biomarker of disease progression in metastatic castration-resistant prostate cancer patients with low baseline CTC counts. Ann Oncol 2018;29(7):1554-1560. (Article) (In English). DOI: 10.1093/annonc/mdy172.

79. Paller CJ, Piana D, Eshleman JR, et al. A pilot study of prostate-specific membrane antigen (PSMA) dynamics in men undergoing treatment for advanced prostate cancer. Prostate 2019;79(14):1597-1603. (Article) (In English). DOI: 10.1002/pros.23883.

80. Wyatt AW, Azad AA, Volik SV, et al. Genomic Alterations in Cell-Free DNA and Enzalutamide Resistance in Castration-Resistant Prostate Cancer. JAMA Oncol 2016;2(12):1598-1606. DOI: 10.1001/jamaoncol.2016.0494.

81. Hashimoto K, Tabata H, Shindo T, et al. Serum testosterone level is a useful biomarker for determining the optimal treatment for castration-resistant prostate cancer. Urol Oncol Semin Orig Invest 2019;37(7):485-491. (Article) (In English). DOI: 10.1016/j.urolonc.2019.04.026.

82. Shiota M, Kashiwagi E, Murakami T, et al. Serum testosterone level as possible predictive marker in androgen receptor axis-targeting agents and taxane chemotherapies for castration-resistant prostate cancer. Urol Oncol Semin Orig Invest 2019;37(3):180.e19-180.e24. (Article) (In English). DOI: 10.1016/j.urolonc.2018.10.020.

83. Sakamoto S, Maimaiti M, Xu M, et al. Higher serum testosterone levels associated with favorable prognosis in enzalutamide- and abiraterone-treated castration-resistant prostate cancer. J Clin Med 2019;8(4) (Article) (In English). DOI: 10.3390/jcm8040489.

84. Ando K, Sakamoto S, Takeshita N, et al. Higher serum testosterone levels predict poor prognosis in castration-resistant prostate cancer patients treated with docetaxel. Prostate 2020;80(3):247-255. (Article) (In English). DOI: 10.1002/pros.23938.

85. Ryan CJ, Dutta S, Kelly WK, et al. Androgen decline and survival during docetaxel therapy in metastatic castration-resistant prostate cancer (mCRPC). Prostate Cancer Prostatic Dis 2019 (Article in Press) (In English). DOI: 10.1038/s41391-019-0152-3.

86. Gao X, Dai C, Huang S, et al. Functional Silencing of HSD17B2 in Prostate Cancer Promotes Disease Progression. Clin Cancer Res 2019;25(4):1291-1301. (In eng). DOI: 1078-0432.CCR-18-2392 [pii]
10.1158/1078-0432.CCR-18-2392.

87. Ko HK, Berk M, Chung YM, et al. Loss of an Androgen-Inactivating and Isoform-Specific HSD17B4 Splice Form Enables Emergence of Castration-Resistant Prostate Cancer. Cell Rep 2018;22(3):809-819. (Article) (In English). DOI: 10.1016/j.celrep.2017.12.081.

88. Gent R, du Toit T, Bloem LM, Swart AC. The 11β-hydroxysteroid dehydrogenase isoforms: pivotal catalytic activities yield potent C11-oxy C19 steroids with 11βHSD2 favouring 11-ketotestosterone, 11-ketoandrostenedione and 11-ketoprogesterone biosynthesis. J Steroid Biochem Mol Biol 2019;189:116-126. (Article) (In English). DOI: 10.1016/j.jsbmb.2019.02.013.

Appendices

A

Nederlandse samenvatting

Introductie

Prostaatkanker is de meest voorkomende kanker bij mannen wereldwijd en in Nederland werd in 2018 bij meer dan 12.000 mannen deze diagnose gesteld. Prostaatkanker staat bekend als een niet-agressieve kanker, maar mannen die bij de diagnose uitzaaiingen hebben of deze tijdens het ziektebeloop ontwikkelen vormen een groep met een slechte uitkomst. Gemetastaseerd prostaatkanker kan worden onderverdeeld in twee ziektefasen: hormoongevoelig prostaatkanker die uiteindelijk wordt opgevolgd door hormoon*ongevoelig* ofwel castratieresistent prostaatkanker. Centraal in de behandeling van gemetastaseerd prostaatkanker staat het onderdrukken van de testosteronproductie uit de testes, ook wel androgeen-deprivatietherapie (ADT) genoemd. De meeste patiënten zullen aanvankelijk vaak langere tijd op ADT reageren (de hormoongevoelige fase), maar het is onvermijdelijk dat prostaatkanker cellen uiteindelijk ongevoelig zullen worden voor ADT (castratieresistente fase). Voortschrijding van de ziekte leidt tot aanzienlijke morbiditeit en uiteindelijk tot overlijden. Voor castratieresistente prostaatkanker patiënten zijn er de laatste jaren veel nieuwe behandelingsopties beschikbaar gekomen, waaronder chemotherapie, tweedelijns hormonale therapie, radioactieve middelen, immunotherapie en moleculair-gerichte therapie. Onlangs hebben sommige van deze behandelopties een toepassing gekregen in de hormoongevoelige fase. Hoewel het therapeutisch landschap voor gemetastaseerd prostaatkanker het afgelopen decennium drastisch is veranderd, blijft het een uitdaging om de juiste behandeling, op het juiste moment, in de juiste volgorde, of als combinatie voor een individuele patiënt te selecteren, omdat het ons aan biomarkers ontbreekt die kunnen voorspellen of een patiënt zal reageren en die we kunnen vervolgen om te evalueren of een behandeling aanslaat. Dit is van belang om de beste uitkomst voor een patiënt te realiseren en daarbij de kosten en bijwerkingen te minimaliseren.

Mijn proefschrift beschrijft het genomische landschap van gemetastaseerd prostaatkanker. Dit onderzoek draagt bij aan een beter begrip *wat* deze ziekte drijft en *hoe* we hierop in kunnen grijpen. Met deze kennis heb ik geprobeerd om nieuwe biomarkers te vinden, genaamd 'vloeibare biopten'. Een groot deel van dit onderzoek heeft zich gericht op het standaardiseren van de verzameling en analyses van vloeibare biopten, omdat dit een belangrijk vereiste is voor het toepassen van vloeibare biopten in de kliniek. Daarnaast laat ik in een aantal onderzoeken zien wat de mogelijke toepassingen van vloeibare biopten zijn. Als laatste heb ik ook gekeken naar een andere biomarker die van belang is in de progressie van gemetastaseerd prostaatkanker, namelijk de in de bloedbaan circulerende steroïde hormonen.

Deel I: Het genomische landschap van gemetastaseerd prostaatkanker

De afgelopen jaren hebben uitvoerige analyses aangetoond dat het genomische landschap van gemetastaseerd prostaatkanker erg complex is en meerdere mutaties en structurele varianten bevat, waarbij er veel verschil is tussen de prostaatkanker cellen van patiënten onderling. Een aantal genen hebben vaak afwijkingen bij gemetastaseerd prostaatkanker: *AR, PTEN, TP53* en *RB1*; de genfusie *TMPRSS2-ERG* is de bekendste structurele variant. Structurele varianten zijn grove DNA afwijkingen van ≥ 50 basenparen en ze komen vaak voor in kanker.

AR, het gen dat de DNA code bevat voor de androgeen receptor, is het belangrijkste aangrijpingspunt voor de behandeling van gemetastaseerd prostaatkanker door middel van ADT en tweedelijns hormonale therapie. Zoals hierboven beschreven is, worden prostaatkanker cellen uiteindelijk ongevoelig voor deze behandeling door verschillende resistentiemechanismen. Uit onderzoek is gebleken dat naast de AR signaalroute, ook andere mechanismen van belang zijn in gemetastaseerd prostaatkanker, zoals de PI3K, Wnt en DNA reparatie signaalroutes. Daarnaast lijken ook afwijkingen in niet-coderende delen van het DNA van voorspellende waarde voor het ziektebeloop te zijn en mogelijk hebben deze afwijkingen zelfs een bepaalde functie. In hoeverre we deze mechanismen en afwijkingen kunnen gebruiken als therapeutisch aangrijpingspunt of als marker om te vervolgen om respons op therapie te meten moet verder uitgezocht worden.

In **hoofdstuk 2** van dit proefschrift hebben we bij een grote groep gemetastaseerde castratieresistente prostaatkanker patiënten uitgebreide genomische analyses verricht. We hebben DNA van tumorweefsel, wat verkregen is door een biopt van een metastase te nemen, afgelezen door middel van 'next-generation sequencing' (NGS). Vervolgens hebben we op basis van genomische kenmerken deze patiënten geclassificeerd, waarbij we acht subgroepen of clusters hebben geïdentificeerd:

A) Microsatelliet Instabiliteit (MSI) karakteristieken met hoge 'tumor mutational burden' (TMB) en geassocieerd met mismatch reparatie deficiëntie; 6,6% van ons cohort;

B) Tandem duplicatie (> 100 kbp) fenotype geassocieerd met biallelische *CDK12* inactiviteit; 6,6% van ons cohort;

D) Homologe Recombinatie Deficiëntie (HRD) karakteristieken met veel (> 100 kbp) deleties en geassocieerd met (somatische) mutaties in BRCAness-geassocieerde genen; 11,2% van ons cohort;

F) Tumoren verrijkt voor chromothripsis; 10,2% van ons cohort;

C, E, G, H) Niet-significante genomische kenmerken zonder bekende biologische associatie; 65,4% van ons cohort.

Verschillende clusters bevatten patiënten die op basis van hun genomische kenmerken mogelijk in aanmerking komen voor moleculair-gerichte therapie. Zo bevat cluster A patiënten die

mogelijk in aanmerking komen voor immunotherapie met immuun check-point inhibitoren, zoals pembrolizumab. Van andere kankersoorten die soortgelijke genomische kenmerken hebben, is een hoge gevoeligheid voor immuun check-point inhibitoren beschreven. Twee klinische studies hebben dit verder onderzocht, de KEYNOTE-028 en de KEYNOTE-199, en zij lieten een langdurige respons zien in een subset van gemetastaseerde prostaatkanker patiënten behandeld met immuun check-point inhibitoren. Deze subset van patiënten was voornamelijk geselecteerd op de expressie van PD-L1 in hun tumoren; dit is het aangrijpingspunt van pembrolizumab. Helaas is er (vooralsnog) niets bekend over het genomische landschap van deze tumoren. De patiënten in cluster D komen mogelijk in aanmerking voor behandeling met poly(ADP-ribose) polymerase (PARP) inhibitoren, wat DNA schade reparatie (DDR) signaalroutes blokkeert. Twee klinische studies, de TOPARP-B en de PROfound trial, hebben aangetoond dat de PARP inhibitor olaparib een anti-tumor effect en een verbeterde overleving bewerkstelligt in gemetastaseerde castratieresistente prostaatkanker patiënten met DDR gen afwijkingen, zoals *BRCA2* deficiëntie. Dit zijn veelbelovende resultaten die in toekomstige studies verder gevalideerd moeten worden.

Desalniettemin is er een grote groep van gemetastaseerde castratieresistente prostaatkanker patiënten (cluster C, E, G, H; ongeveer 66% van de patiënten) die vooralsnog niet geclassificeerd kunnen worden op basis van klinisch-relevante of biologische kenmerken. Toekomstig onderzoek moet zich op deze groep patiënten richten om zo te onderzoeken of er therapeutische aangrijpingspunten zijn. Van belang hierbij is dat we recent een cellijn gecreëerd hebben die de genomische kenmerken van cluster E weergeeft, verder beschreven in **hoofdstuk 6**. Deze cellijn kan bijdragen aan verder onderzoek om de biologie achter dit cluster op te helderen.

Deel II: Vloeibare biopten

Met behulp van biomarkers kunnen we (patho)fysiologische processen of reacties op een interventie vastleggen. Dit kan met behulp van klinische metingen, metingen in het laboratorium of door middel van beeldvorming. Er zijn biomarkers met verschillende toepassingen: diagnostisch, prognostisch, predictief en monitoring (van bijvoorbeeld respons op therapie).

In de diagnose en evaluatie van gemetastaseerd prostaatkanker wordt gebruik gemaakt van verschillende laboratorium metingen, waaronder PSA (prostaat-specifiek antigeen), en beeldvorming door middel van botscan, CT-scan, MRI en recent ook de PSMA-PET CT-scan. Echter deze biomarkers hebben hun tekortkomingen. PSA is een prostaat-specifieke marker, maar geen prostaat*kanker*-specifieke marker en PSA staat onder invloed van testosteron. Dit betekent dat er een PSA daling te verwachten is onder ADT en tweedelijns hormonale therapie, maar dat dit niet per definitie betekent dat de prostaatkanker cellen ook stoppen met groeien. Daarnaast komen in gemetastaseerd castratieresistent prostaatkanker PSA level en uitgebreidheid van

ziekte en klinische uitkomsten niet altijd overeen. Beeldvorming kan verandering in tumorgroei vastleggen, maar dit effect is niet direct zichtbaar, waarbij je eigenlijk achter de feiten aan loopt. Verder laat een botscan veranderingen in het botmetabolisme zien en weerspiegelt dus niet de prostaatkanker cellen zelf.

Met de ontdekking van 'vloeibare biopten' is er een nieuwe manier beschikbaar gekomen om tumor DNA te verkrijgen dat gebruikt kan worden als biomarker met verschillende toepassingen. Vloeibare biopten omvatten verschillende entiteiten: circulerend cel-vrij DNA (cfDNA) en RNA (cfRNA), circulerende tumorcellen (CTC's) en exosomen. Het grootste voordeel van vloeibare biopten is dat ze, in tegenstelling tot weefsel biopten, op een minimaal invasieve en veilige manier verkregen kunnen worden en daarom gedurende het ziektebeloop frequent afgenomen kunnen worden. Het gaat namelijk om een afname van bijvoorbeeld bloed, urine of ascites. Een ander belangrijk voordeel is dat alle tumorlocaties in het lichaam deze biomarkers kunnen afgeven waardoor het mogelijk beter de heterogeniteit van de tumor weergeeft dan wanneer je van één enkele plek een biopt afneemt. Aangezien dit proefschrift zich richt op cfDNA en CTC's, zullen deze twee biomarkers in meer detail worden besproken.

Cel-vrij DNA

Cel-vrij DNA zijn kleine fragmenten van nucleïnezuur (lengte van 180–200 bp) die in de bloedcirculatie vrijkomen wanneer weefselcellen en bloedcellen afsterven. In het geval van tumorcellen wordt dit circulerend tumor-DNA (ctDNA) genoemd. De hoeveelheid cfDNA in de bloedcirculatie hangt af van meerdere factoren, zoals (patho)fysiologische omstandigheden (inspanning, ziekte) en pre-analytische condities (opslagtemperatuur, type bloedbuis), maar is meestal laag met ongeveer 10 ng per ml plasma. De fractie van ctDNA hangt af van het tumortype en het ziektestadium en kan variëren van extreem laag (< 0,01%) tot zeer hoog (60%) maar is vaak minder dan 1% van het totale cfDNA.

Circulerende tumorcellen

Circulerende tumorcellen zijn tumorcellen in de bloedcirculatie die afkomstig zijn van de primaire tumor en/of van de metastasen. Aangezien CTC's afkomstig kunnen zijn van alle tumorplaatsen in het lichaam, wordt aangenomen dat ze de genetische en biologische heterogeniteit van de tumor weerspiegelen. Net als ctDNA, komen CTC's weinig voor in het bloed met over het algemeen 1 CTC per ml bloed. Voor het detecteren en kwantificeren van CTC's heeft de FDA (U.S. Food and Drug Administration) het CellSearch systeem goedgekeurd. Dit systeem scheidt CTC's van andere bloedcellen met behulp van met magnetische EpCAM-gecoate ferrofluids (vloeistof met daarin magnetische nanodeeltjes). Binnen gemetastaseerd prostaatkanker zijn CTC's het meest bestudeerd en er is bekend dat CTC's een prognostische en predictieve waarde hebben. Gemetastaseerde prostaatkanker patiënten met ≥ 5 CTC's per 7,5 ml bloed hebben een slechtere uitkomst en toenemende aantallen CTC's in de loop van de tijd hebben ook een prognostische

betekenis. Expressie van AR splice variant 7 (AR-V7) in CTC's, wat een androgeen-onafhankelijke activatie van AR oplevert, is sterk geassocieerd met een slechte respons op tweedelijns hormonale therapie. Daarentegen wordt juist een goede respons op chemotherapie gezien bij patiënten met AR-V7 expressie in CTC's.

Pre-analytische condities

Voor de implementatie van vloeibare biopten in de standaard oncologische zorg is het essentieel om de pre-analytische condities en werkprocessen te standaardiseren. Op deze manier worden betrouwbare en consistente resultaten gewaarborgd.

In **hoofdstuk 3** hebben we verschillende pre-analytische condities onderzocht om cfDNA/ctDNA optimaal te bewaren voordat verdere analyses opgestart worden. We ontdekten dat het gebruik van verschillende bloedafname buizen de detectie van ctDNA niet beïnvloedde, maar dat sommige bloedafname buizen de hoeveelheid cfDNA in de loop van de tijd wel stabiliseerde. Aangezien in veel onderzoeken variant allel frequentie (VAF; fractie ctDNA van cfDNA in %) wordt gerapporteerd, kan dit mogelijk leiden tot een onderschatting van de VAF wanneer een bloedmonster niet op tijd wordt verwerkt. Andere onderzoeken hebben ook verschillende bloedafname buizen getest en vonden net als wij dat, indien het bloedmonster binnen enkele dagen wordt verwerkt, het type bloedbuis triviaal is. Tabel 1 geeft een overzicht van de meest gebruikte bloedafname buizen. Concluderend geldt dat voor RNA-analyses van CTC's een EDTA-buis die binnen 48 uur verwerkt wordt optimaal is, en voor cfDNA/ctDNA analyses een EDTA-buis die binnen zes uur verwerkt wordt óf een cel-stabilisatiebuis die binnen 2-7 dagen verwerkt wordt, even geschikt zijn.

Tabel 1. Meest gebruikte bloedafname buizen voor vloeibare biopten

Type bloedbuis	Monster	Entiteit	Cel stabilisatie
K$_2$EDTA	Plasma	cfDNA CTC	Niet-gestabiliseerd
Cell-Free DNA BCT®	Plasma	cfDNA CTC	Gestabiliseerd
PAXgene Blood ccfDNA	Plasma	cfDNA	Gestabiliseerd
Cell-Free DNA BCT®	Plasma	cfDNA	Gestabiliseerd
CellSave Preservation Tubes	Plasma	cfDNA CTC	Gestabiliseerd
TransFix/EDTA Vacuum BCT	Plasma	CTC	Gestabiliseerd

CTC, circulerende tumor cel
cfDNA, cel-vrij DNA

De volgende stap in de analyse van cfDNA is de isolatie van cfDNA uit plasma; de handmatige 'QIAamp Circulating Nucleic Acid' methode van Qiagen wordt beschouwd als de gouden standaard met de hoogste cfDNA opbrengst. Echter geautomatiseerde methoden hebben de

voorkeur, omdat bij het implementeren van vloeibare biopten in de standaard oncologische zorg er veel samples binnen een korte tijd verwerkt zullen moeten worden en er zo min mogelijk variatie mag optreden tussen de laboranten die de analyses uitvoeren. In **hoofdstuk 4** beschrijven we een directe vergelijking van de QIAamp met twee geautomatiseerde isolatiemethoden, de QIAsymphony (Qiagen) en de Maxwell (Promega). In onze handen presteerde QIAsymphony vergelijkbaar met QIAamp. Hoewel de cfDNA opbrengst het laagst was met Maxwell, was de VAF vergelijkbaar tussen alle methoden. Resultaten van een ander onderzoek toonden aan dat QIAsymphony en Maxwell beiden toereikende geautomatiseerde isolatiemethoden zijn, wat in lijn is met onze resultaten. Verdere optimalisatie van de cfDNA isolatie kan bereikt worden door het plasmavolume te vergroten en het extractievolume te minimaliseren. In **hoofdstuk 4** hebben we het QIAamp protocol verder geoptimaliseerd door het extractievolume driemaal her te gebruiken, waardoor de cfDNA opbrengst toenam.

Om CTC's te isoleren voor verschillende experimentele doeleinden, zoals *ex vivo* celkweken en analyseren van cel heterogeniteit, zijn op filtratie gebaseerde microwells en diagnostische leukaferese interessante methoden. Deze methoden zijn gebaseerd op biofysische kenmerken van CTC's, waaronder celgrootte en celdichtheid. Microwells selecteren individuele CTC's op basis van hun celgrootte en celstijfheid, maar de efficiëntie van de CTC opbrengst is slechts ~ 2%, zoals werd aangetoond in een recente studie. Zoals beschreven in **hoofdstuk 6**, levert het product van diagnostische leukaferese duizenden levensvatbare CTC's op door het gebruik van grote bloedvolumes. In vergelijking met de opbrengst van CTC's vanuit een enkele buis met bloed is dit enorm en dit is in lijn met resultaten uit eerdere studies. Echter, afdoende depletie van witte bloedcellen vormt nog steeds een uitdaging bij het gebruik van de 'RosetteSep depletie' techniek, al dan niet gevolgd door EpCAM-gebaseerde selectie van CTC's, zoals aangetoond in **hoofdstuk 6**, en dit moet verder geoptimaliseerd worden.

Op circulerende tumorcellen-gebaseerde stratificatie met behulp van organoïden

Het gebruik van levende cellen als 'real-life' screeningsmodellen voor de respons op therapie schept een interessant toekomstbeeld om de behandeling van kanker te personaliseren. Dit kan gedaan worden middels een tumorbiopt dat bij een patiënt is afgenomen, wat in het laboratorium wordt gekweekt tot een mini-tumor oftewel een organoïde. Verschillende klinische onderzoeken (bijv. TUMOROID NCT03821870) includeren momenteel patiënten om te onderzoeken of de respons op therapie in organoïden, afkomstig van tumorbiopten, respons op therapie in diezelfde patiënten kan voorspellen. Soortgelijk onderzoek met organoïden van borstkanker patiënten en patiënten met gemetastaseerde gastro-intestinale kankers laten zien dat de respons in organoïden en patiënten overeenkomen.

In **hoofdstuk 6** hebben we ons gericht op het kweken van organoïden afkomstig van CTC's van patiënten met gemetastaseerd prostaatkanker. We verkregen organoïden in 9 uit 18 (50%) monsters

na kortdurende kweek waarin we konden aantonen dat deze CTC-organoïden oorspronkelijk prostaatkanker cellen waren. Bovendien waren we in staat om een organoïde cellijn op te zetten die langdurig in kweek gehouden kon worden (met de genomische kenmerken van cluster E; zie hierboven). Tot nu toe is gebleken dat het erg moeilijk is om prostaatkanker cellijnen tot stand te brengen, dus onze nieuwe cellijn is een waardevolle toevoeging voor het wetenschappelijk onderzoek. In deze organoïde cellijn hebben we de respons op enzalutamide getest, een vorm van tweedelijns hormonale therapie, en dit kwam overeen met de respons van deze patiënt op enzalutamide in de kliniek. Hiermee konden we een 'proof-of-concept' leveren dat patiënt-specifieke organoïden, afkomstig uit CTC's, kunnen dienen als 'real life' screeningsmodel voor de respons op therapie bij gemetastaseerd prostaatkanker. Echter, het huidige slagingspercentage en de groeisnelheid van organoïden afkomstig uit CTC's is momenteel nog te beperkt om dit concept toe te passen in de kliniek.

Responsevaluatie met behulp van circulerend tumor DNA

In gemetastaseerd castratieresistent prostaatkanker is het een uitdaging om de respons op therapie nauwkeurig en tijdig te bepalen met de huidige biomarkers zoals eerder genoemd.

In **hoofdstuk 5** hebben we ons gericht op het detecteren van structurele varianten in ctDNA van patiënten met gemetastaseerd castratieresistent prostaatkanker om de respons op therapie te evalueren. In prostaatkanker zijn structurele varianten betrokken bij biologisch relevante signaalroutes, zoals overexpressie van *AR* en *MYC*, en ETS-fusiegenen. Grote structurele varianten, zoals translocaties tussen chromosomen of binnen chromosomen zijn interessant omdat dit unieke DNA-moleculen zijn die alleen in tumorcellen aanwezig zijn. Dit maakt dat ze een specifieke biomarker voor kanker zijn. Om structurele varianten in ctDNA te detecteren, hebben wij een nieuwe methodologie gebruikt, genaamd 'Oxford Nanopore portable sequencing'. Hiermee zijn eerst structurele varianten geïdentificeerd in tumorcellen verkregen uit weefselbiopten van metastasen. Na een bioinformatische analyse en met behulp van Breakpoint PCR, waren we in staat om structurele varianten in ctDNA te detecteren. Onze eerste resultaten laten zien dat de kwantitatieve meting van structurele varianten in ctDNA overeenkomt met de hoeveelheid kanker in het lichaam en dat deze meting eerder kan wijzen op ziekteprogressie dan een PSA bepaling. Echter, de selectie van structurele varianten moet verder worden geoptimaliseerd, omdat we bij sommige patiënten de correlatie met hoeveelheid kanker en ziekteprogressie niet konden aantonen. Dit kan erop wijzen dat we geen structurele varianten hebben geselecteerd die het meest representatief zijn voor de tumor.

Deel III: 'Steroidomics'

In gemetastaseerde prostaatkanker patiënten die behandeld worden met ADT zijn de circulerende testosteron levels onder castratie niveau. Desondanks kan er ziekteprogressie

plaatsvinden in gemetastaseerde castratieresistente prostaatkanker patiënten door verschillende resistentiemechanismen. Onderzoek toont aan dat de AR signaalroute hierbij van belang blijft en dat resistentie niet alleen afhankelijk is van genomische afwijkingen (bijvoorbeeld overexpressie van *AR*), maar ook van veranderingen in de aanmaak en het metabolisme van steroïden. Steroïde hormonen zijn kleine signaal moleculen die genexpressie reguleren door te binden aan receptoren. Steroïden worden gemaakt uit cholesterol en omvatten vijf groepen van hormonen, waaronder glucocorticoïden en androgenen. Het mannelijk hormoon testosteron behoort tot deze laatste groep. Bepalingen van steroïden worden verricht op bloed-serum monsters met behulp van 'massa spectrometrie' (MS) technieken.

Het effect van ADT wordt geëvalueerd door circulerende testosteron levels in het bloed te bepalen en hoewel dit onder castratie niveau is, is aangetoond dat de hoeveelheid testosteron van voorspellende waarde kan zijn. Patiënten met een relatief hoog testosteron level voor start van therapie hebben een betere uitkomst bij de behandeling met tweedelijns hormonale therapie. Patiënten met een relatief laag testosteron level voor start van therapie hebben een betere uitkomst bij de behandeling met chemotherapie. Verder laat onderzoek een associatie zien bij patiënten met een dalend testosteron gehalte tijdens chemotherapie en PSA respons en ziekteprogressie-vrije overleving.

In **hoofdstuk 7** laten we zien dat CellSave bloedbuizen, die ook gebruikt worden voor CTC en cfDNA analyses, ook gebruikt kunnen worden voor steroïde metingen. Hierdoor kunnen studie samples optimaal gebruikt worden en hoeven er minder verschillende bloedbuizen afgenomen te worden bij patiënten. In **hoofdstuk 8** tonen we aan dat 11-ketotestosteron (11KT), een 11-ketogeen androgeen, de meest voorkomende androgeen is in de bloedcirculatie van gemetastaseerde castratieresistente prostaatkanker patiënten. De circulerende levels van 11KT zijn veel hoger dan die van testosteron en daarnaast is bekend dat 11KT in staat is om de androgeen receptor te activeren. Wij denken dat 11-geoxygeneerd androgenen een belangrijke rol spelen in AR signaalroute activatie na castratie. Hierop aansluitend zagen we dat onze patiënten met een hoog totaal androgeen (11KT + testosteron) level een langere progressievrije overleving hadden. Uit onze resultaten en ander onderzoek blijkt dat ook in castratieresistent prostaatkanker de AR signaalroute een belangrijke speler blijft.

Conclusie

Ik ben ervan overtuigd dat de toepassing van vloeibare biopten in gemetastaseerd prostaatkanker een grote potentie heeft als prognostische, voorspellende en therapie respons biomarker. Een belangrijk voordeel van vloeibare biopten is dat het te verkrijgen is door een simpele bloedafname. Dit is veilig, minimaal invasief en makkelijk te herhalen gedurende de tijd. Bovendien geeft het

een 'real-time' overzicht van de biologische karakteristieken en genomische kenmerken die tumorgroei en ziekteprogressie kunnen beïnvloeden in gemetastaseerd prostaatkanker. Ons onderzoek naar het standaardiseren van pre-analytische condities die van invloed kunnen zijn op de analyse van vloeibare biopten is van groot belang om het gebruik van vloeibare biopten in de klinische praktijk te implementeren. Uitgebreide analyses op ctDNA en CTC's tonen aan dat ze de genomische karakteristieken van de tumor weergeven en dus complementair zijn aan tumor biopten en deze wellicht kunnen vervangen. Het analyseren van genomische afwijkingen, zoals structurele varianten, in ctDNA kan mogelijk therapie respons en resistentiemechanismen voorspellen. Naast de voorspellende waarde van CTC aantallen, kunnen CTC's ook gebruikt worden om buiten het lichaam te kweken tot organoïden. Op deze manier draagt dit bij aan de ontwikkeling van nieuwe ziektemodellen voor gemetastaseerd prostaatkanker en kan het mogelijk toegepast worden om therapie respons te voorspellen. Daarnaast hebben we aangetoond dat 'steroidomics' van belang blijven in gemetastaseerd prostaatkanker met ons onderzoek naar de klinische relevantie van circulerende 11-ketogene androgenen.

Het gebruik van op bloed-gebaseerde biomarkers, zoals vloeibare biopten en circulerende steroïden, als voorspellende markers en therapie response markers moeten verder onderzocht en gevalideerd worden in prospectieve klinische studies. Het is daarom belangrijk om dit standaard mee te nemen in klinische studies naast PSA en radiologische beeldvorming. Daarnaast is het essentieel om meer inzicht te krijgen in de complexe biologie van gemetastaseerd prostaatkanker en om meer te weten te komen over welke genomische afwijkingen belangrijk zijn en mogelijk als aangrijpingspunt kunnen dienen in de grote subgroep van patiënten met gemetastaseerd castratieresistent prostaatkanker zonder duidelijke genomische kenmerken.
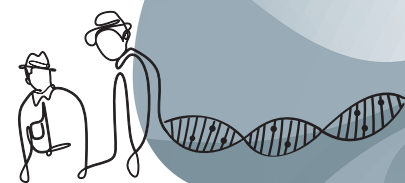
A
Author affiliations

Arlt, Wiebke — Institute of Metabolism and System Research, University of Birmingham, Birmingham, United Kingdom

Beaufort, Corine M. — Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Beije, Nick — Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Bergman, Andries M. — Division on Oncogenomics, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Department of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Blank, de Sam — Department of Genetics, Center for Molecular Medicine, University Medical Center Utrecht and Utrecht University, Utrecht, The Netherlands

Boekhorst, te Peter A.W. — Department of Hematology, Erasmus University Medical Center, Rotterdam, The Netherlands

Boutros, Paul C. — Computational Biology Program, Ontario Institute for Cancer Research, Toronto, Ontario, Canada

Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

Cuppen, Edwin — Center for Molecular Medicine and Oncode Institute, University Medical Center Utrecht, Utrecht, The Netherlands

Hartwig Medical Foundation, Amsterdam, The Netherlands

Erkens-Schulze, Sigrun — Department of Urology, Erasmus University MC, Rotterdam, The Netherlands

Hamberg, Paul — Department of Internal Medicine, Franciscus Gasthuis & Vlietland, Rotterdam, The Netherlands

Heijden, van der Michiel S. — Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Department of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Helmijr, Jean C. — Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Hofland, Johannes — Department of Internal Medicine, Section of Endocrinology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Jansen, Maurice P.H.M. — Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Janssen, Roel — Center for Molecular Medicine and Oncode Institute, University Medical Center Utrecht, Utrecht University, The Netherlands

Jong, de Anouk C. — Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Kloosterman, Wigard P. — Center for Molecular Medicine, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

Kraan, Jaco — Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Livingstone, Julie — Informatics and Biocomputing Program, Ontario Institute for Cancer Research, Toronto, Canada

Lolkema, Martijn P. — Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Look, Maxime P. — Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Martens, John W.M. — Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Mehra, Niven — Department of Medical Oncology, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands

Mout, Lisanne — Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Department of Urology, Erasmus University MC, Rotterdam, The Netherlands

Neves, Rui P.L. — Experimental Surgical Oncology, General, Visceral and Pediatric Surgery, University Hospital and Medical Faculty, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

Oomen-de Hoop, Esther — Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Oort, van Inge M. — Department of Urology, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands

Renkens, Ivo — Department of Genetics, Center for Molecular Medicine, University Medical Center Utrecht and Utrecht University, Utrecht, The Netherlands

Riet, van Job — Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Roosmalen, van Markus J. — Center for Molecular Medicine, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

Sandberg, Yorick — Department of Internal Medicine, Maasstad Hospital Rotterdam, Rotterdam, The Netherlands

†Sieuwerts, Anieta M.
Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Sleijfer, Stefan
Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Smits, Minke
Department of Medical Oncology, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands

Snaterse, Gido
Department of Internal Medicine, Section of Endocrinology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Stangl, Christina
Department of Genetics, Center for Molecular Medicine, University Medical Center Utrecht and Utrecht University, Utrecht, The Netherlands

Division of Molecular Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands

Oncode Institute, Utrecht, The Netherlands

Steeghs, Neeltje
Department of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Stoecklein, Nikolas H.
Experimental Surgical Oncology, General, Visceral and Pediatric Surgery, University Hospital and Medical Faculty, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

Taylor, Angela E.
Institute of Metabolism and System Research, University of Birmingham, Birmingham, United Kingdom

Valle-Inclan, Jose E.
Center for Molecular Medicine, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

Visser, Jenny A.
Department of Internal Medicine, Section of Endocrinology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Vitale, Silvia R.
Department of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Department of Clinical and Experimental Medicine, Center for Experimental Oncology and Hematology, University of Catania, Italy

Vlugt-Daane, van der Michelle
Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Voest, Emile E.
Department of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Oncode Institute, Utrecht, The Netherlands

Weerden, van Wytske M.
Department of Urology, Erasmus Medical Center, Rotterdam, The Netherlands

Werken, van de Harmen J.G
Cancer Computational Biology Center, Erasmus MC Cancer Institute, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands

Wilting, Saskia M.
Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Wit, de Ronald
Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Witte, de Chris J.
Center for Molecular Medicine, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

Woo, Thomas L.C.
Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Yamaguchi, Takafumi N.
Informatics and Biocomputing Program, Ontario Institute for Cancer Research, Toronto, Canada

Zhu, Yanyun
Division on Oncogenomics, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Oncode Institute, Utrecht, The Netherlands

Zwart, Wilbert
Division on Oncogenomics, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Oncode Institute, Utrecht, The Netherlands

A

List of publications

Jose Espejo Valle-Inclan*, Christina Stangl*, Anouk C. de Jong*, **Lisanne F. van Dessel**, Markus J. Van Roosmalen, Jean C. Helmijr, Ivo Renkens, Roel Janssen, Sam de Blank, Chris J. de Witte, John W.M. Martens, Maurice P.H.M. Jansen, Martijn P. Lolkema, Wigard P. Kloosterman.
Optimizing Nanopore sequencing-based detection of structural variants enables individualized circulating tumor DNA-based disease monitoring in cancer patients
Genome Medicine; 2021 in press

Lisanne Mout*, **Lisanne F. van Dessel***, Jaco Kraan, Anouk C. de Jong, Rui P.L. Neves, Sigrun Erkens-Schulze, Corine M. Beaufort, Anieta M. Sieuwerts, Job van Riet, Thomas L.C. Woo, Ronald de Wit, Stefan Sleijfer, Paul Hamberg, Yorick Sandberg, Peter A.W. te Boekhorst, Harmen J.G. van de Werken, John W.M. Martens, Nikolas H. Stoecklein, Wytske M. van Weerden, Martijn P. Lolkema.
Generating human prostate cancer organoids from leukapheresis enriched circulating tumor cells
European Journal of Cancer; 2021 in press

Silvia R. Vitale, Jean C. Helmijr, Marjolein Gerritsen, Hicret Coban, **Lisanne F. van Dessel**, Nick Beije, Michelle van der Vlugt-Daane, Paolo Vigneri, Anieta M. Sieuwerts†, Natasja Dits, Martin E. van Royen, Guido Jenster, Stefan Sleijfer, Martijn P. Lolkema, John W.M. Martens, Maurice P.H.M. Jansen.
Detection of tumor-derived extracellular vesicles in plasma from patients with solid cancer
BMC Cancer; 2021 Mar 24;21(1):315

Gido Snaterse*, **Lisanne F. van Dessel***, Angela E. Taylor, Jenny A. Visser, Wiebke Arlt, Martijn P. Lolkema, Johannes Hofland.
Validation of circulating steroid hormone measurements across different matrices by liquid chromatography-tandem mass spectrometry
Steroids; 2021 Mar;167:108800

Lisanne F. van Dessel, John W.M. Martens, Martijn P. Lolkema.
Fundamentals of liquid biopsies in metastatic prostate cancer: from characterization to stratification
Current Opinion in Oncology; 2020 Sep;32(5):527-534

Lisanne F. van Dessel*, Job van Riet*, Minke Smits, Yanyun Zhu, Paul Hamberg, Michiel S. van der Heijden, Andries M. Bergman, Inge M. van Oort, Ronald de Wit, Emile E. Voest, Neeltje Steeghs, Takafumi N. Yamaguchi, Julie Livingstone, Paul C. Boutros, John W.M. Martens, Stefan Sleijfer, Edwin Cuppen, Wilbert Zwart, Harmen J.G van de Werken, Niven Mehra, Martijn P. Lolkema.
The genomic landscape of metastatic castration-resistant prostate cancers reveals multiple distinct genotypes with potential clinical impact
Nature Communications; 2019 Nov 20;10(1):5251

Lisanne F. van Dessel*, Silvia R. Vitale*, Jean C. Helmijr, Saskia M. Wilting, Michelle van der Vlugt-Daane, Esther Oomen-de Hoop, Stefan Sleijfer, John W.M. Martens, Maurice P.H.M. Jansen, Martijn P. Lolkema.
High-throughput isolation of circulating tumor DNA: a comparison of automated platforms
Molecular Oncology; 2019 Feb;13(2):392-402

Lisanne F. van Dessel, Sarah H.M. Reuvers, Chris H. Bangma, Shafak Aluwini.
Salvage radiotherapy after radical prostatectomy: Long-term results of urinary incontinence, toxicity and treatment outcomes
Clinical and Translational Radiation Oncology; 2018 May 29;11:26-32

Lisanne F. van Dessel, Nick Beije, Jean C. Helmijr, Silvia R. Vitale, Jaco Kraan, Maxime P. Look, Ronald de Wit, Stefan Sleijfer, Maurice P.H.M. Jansen, John W.M. Martens, Martijn P. Lolkema.
Application of circulating tumor DNA in prospective clinical oncology trials - standardization of preanalytical conditions
Molecular Oncology; 2017 Mar;11(3):295-304

Renate K. Hukema, Ronald A.M. Buijsen, Martijn Schonewille, Chris Raske, Lies-Anne W.F.M. Severijnen, Ingeborg Nieuwenhuizen-Bakker, Rob F.M. Verhagen, **Lisanne F. van Dessel**, Alex Maas, Nicolas Charlet-Berguerand, Chris I. de Zeeuw, Paul J. Hagerman, Robert F. Berman, Rob Willemsen.
Reversibility of neuropathology and motor deficits in an inducible mouse model for FXTAS
Human Molecular Genetics; 2015 Sep 1;24(17):4948-57

* These authors contributed equally

A

Curriculum vitae

Lisanne Francisca van Dessel was born on the 29th of January 1989 in Rotterdam, the Netherlands. She completed her secondary school education at Comenius College in Capelle aan den IJssel in 2007. During this period she participated in the Leiden Advanced Pre-University Programme for Top students (LAPP-Top) at the Pre-University College of Leiden University. After graduation, she started her medical training at the Erasmus University Medical Center Rotterdam. In 2009, she commenced with the research master's program Molecular Medicine at the Erasmus University Rotterdam, which she graduated from in 2012. This research master's programme focuses on genetic, molecular, and cellular principles of health and disease and involves two large laboratory research projects. She conducted her senior medical internship at the department of Urology at the Erasmus University Medical Center Rotterdam. During her studies she actively participated in the student rowing club A.R.S.R. Skadi, participated in several committees and the Co-Raad of the Medical Faculty Association Rotterdam (MFVR) and travelled for 6 months through Asia and Australia. In May 2015 she obtained her medical degree after which she started her PhD project on metastatic prostate cancer at the departments of Medical Oncology and Experimental Urology at the Erasmus MC Cancer Institute Rotterdam under the supervision of Prof. dr. R. de Wit, Prof. dr. ir. G. Jenster, Prof. dr. ir. J.W.M. Martens, and Dr. M.P.J.K. Lolkema. Her research projects are described in this thesis and were presented at various (inter)national congresses. During her PhD training she participated in the PhD committee of the Postgraduate School Molecular Medicine and organized multiple (social) activities and research meetings for her laboratory group. Furthermore, she was actively involved in education and supervised several students in extracurricular research and their master thesis. In 2017 she was awarded the Pieter de Mulder Award which enabled her to take on an international research internship at the department of Experimentelle Chirurgische Onkologie at the Universitätsklinikum Düsseldorf in Germany under the supervision of Prof. Dr. Med. N.H. Stoecklein. From December 2018 on, she started working as a medical resident (ANIOS) at the Urology department of the Amphia hospital in Breda and subsequently of the Franciscus Gasthuis & Vlietland hospital in Rotterdam. In 2021 she will start her residency training (AIOS) in Urology in Rotterdam.

Fotografie: Bennie Mentjens

A

PhD portfolio

1. PhD training	Year	Workload
General courses		
- Research Integrity	2015	0.3 ECTS
- BROK ('Basiscursus Regelgeving Klinisch Onderzoek')	2015	1.5 ECTS
- Biostatistical Methods I: Basic Principles	2016	5.7 ECTS
- BKO-training 'Omgaan met groepen'	2016	0.3 ECTS
- Photoshop and Illustrator CS6	2017	0.3 ECTS
- Biomedical English Writing and Communication	2018	3.0 ECTS
- Re-registration BROK	2019	1.0 ECTS
Specific courses (e.g. Research school, Medical Training)		
- Course Biomedical Research Techniques XIV	2015	1.5 ECTS
- NGS in DNA Diagnostics Course	2015	1.0 ECTS
- Course Next Generation Sequencing data analysis	2015	1.4 ECTS
- Course Open Clinica	2015	0.3 ECTS
- Course on R	2016	1.4 ECTS
- Real Time PCR training	2016	0.2 ECTS
- The Galaxy for NGS	2017	1.0 ECTS
Seminars and workshops		
- Erasmus MC Bladder Cancer Research Day	2015	0.3 ECTS
- 4th/5th Daniel den Hoed Day	2016/2017	0.6 ECTS
- Molecular Medicine Day	2016/2017	0.6 ECTS
- Research in mCRPC	2016	0.3 ECTS
- PhD Day	2015/2016	0.6 ECTS
- Illumina 2016 European Genomic Technology Forum	2016	0.3 ECTS
- Symposium Novel Options for Cancer Imaging: Focus on Urological Tumors	2016	0.3 ECTS
- Scientific Meeting Medical Oncology	2015-2017	0.9 ECTS
- Illumina Whole Genome Sequencing Symposium	2016	0.3 ECTS
Presentations		
- Scientific Meeting Medical Oncology – oral	2016	0.2 ECTS
- 20th Molecular Medicine Day – poster	2016	1.0 ECTS
- Research in mCRPC – oral	2016	0.2 ECTS
- Nederlandse Vereniging van Urologie (NVU) Voorjaarsvergadering – oral	2016/2018	0.4 ECTS
- 11th EORTC pathobiology group meeting – oral	2016	0.2 ECTS
- 6th Dutch Uro-Oncology Study group (DUOS) Year Symposium – oral	2016	0.2 ECTS
- Jonge Oncologen avond – oral	2017	0.2 ECTS
- Medical Oncology Research Meeting – oral	2017	0.2 ECTS
- Interlab meeting Twente – oral	2017	0.2 ECTS
- Interlab meeting Antwerpen – oral	2017	0.2 ECTS
- EORTC Gynecological Cancer Group Meeting Porto – oral	2017	0.2 ECTS
- NVMO Oncologiedagen – Pieter de Mulder Award – oral	2017	0.2 ECTS
- AC Urogenitale Tumoren retraite – oral	2018	0.2 ECTS
- CancerID CTC DLA workshop Dusseldorf – oral	2018	0.2 ECTS
- Scholingsdag Nederlandse Vereniging van Oncologie Datamanagers (NVvOD) – oral	2018	0.2 ECTS
- JNI Scientific Lab Meeting – oral	2018	0.2 ECTS
- American Society of Clinical Oncology (ASCO) Annual meeting – poster presentation	2018	1.0 ECTS
- Brigitte and Dr. Konstanze Wegener Seminar / DCC-Net Retreat Krickenbeck – oral	2018	0.2 ECTS
- Tour d'Europe – oral	2018	0.2 ECTS
- Externe refereeravond Urologie – oral	2019	0.2 ECTS

(Inter)national conferences		
- 5th-9th Dutch Uro-Oncology Study group (DUOS) Year Symposium	2015-2019	1.5 ECTS
- CPCT symposium	2015/2016	0.6 ECTS
- CMBD-themadag 'Cell free DNA: grensverleggende innovatie in de moleculaire diagnostiek'	2017	0.3 ECTS
- Nederlandse Vereniging van Urologie (NVU) Voorjaarsvergadering/Najaarsvergadering	2016-2018	0.9 ECTS
- European Association of Urology (EAU) Annual Congress	2018/2019	2.0 ECTS
- 2nd/3rd Brigitte and Dr. Konstanze Wegener Seminar / DCC-Net Retreat Krickenbeck	2018/2019	2.0 ECTS
Other		
- Medical Oncology Journal Club	2015-2018	3.0 ECTS
- Urology Lab Meeting	2015-2018	1.0 ECTS
- Medical Oncology Research Meeting	2015-2018	2.0 ECTS
- JNI Scientific Lab Meeting	2015-2018	3.0 ECTS
- JNI Oncology Lecture	2015-2018	2.0 ECTS

2. Teaching		
Lecturing		
- Klinische les polikliniek 'CIRCUS studie'	2017	0.2 ECTS
Supervising practicals and excursions, Tutoring		
- Urology Skills Training 2nd year medical students	2015-2017	1.0 ECTS
- Tutoring 1st year medical students	2016-2017	1.5 ECTS
- Supervisor "Clinical orientation on the medical profession" for 1st year medicalstudents	2018	0.5 ECTS
Supervising Master's theses		
- Stéphanie Perridon	2017	1.5 ECTS
Other		
- Supervising 'Minor' medical student	2017	0.5 ECTS

3. Other		
- Organizer Medical Oncology Journal Club	2015-2018	1.5 ECTS
- Secretary cfDNA meetings	2016-2018	1.5 ECTS
- Member of PhD committee of the Erasmus Postgraduate School Molecular Medicine	2017-2018	1.0 ECTS
- Research internship Dusseldorf	2018	5.0 ECTS

A

Dankwoord

Voor u ligt misschien wel het belangrijkste, maar in ieder geval het meest gelezen hoofdstuk van een proefschrift. En niet zonder reden, want dit proefschrift was er niet gekomen zonder de hulp, bijdrage en ondersteuning van een hoop mensen.

Allereerst wil ik de patiënten bedanken die hebben deelgenomen aan de klinische studies beschreven in dit proefschrift. Het is bewonderenswaardig dat u zonder enkel belang of voordeel bereid bent geweest materiaal te doneren en uw kostbare tijd beschikbaar te stellen voor het wetenschappelijk onderzoek. De hoed op de cover van dit proefschrift is een blijk van dank; ik zet mijn 'hoed' (pet) voor u af!

Veel dank gaat uit naar de vaste begeleiders van mijn promotietraject de afgelopen jaren. Ik ben begeleid door een gevarieerd team, waarbij ik heb kunnen leren van ieders kwaliteiten. Fijn dat jullie het vertrouwen in mij hebben gesteld om dit traject tot een goed einde te volbrengen.

Prof. dr. De Wit, beste Ronald, naarmate mijn traject vorderde, is ons contact intensiever geworden. Hoewel mijn onderzoeksprojecten af en toe wat technisch waren, zijn jouw expertise en up-to-date kennis van klinische studies op het gebied van prostaatkanker indrukwekkend. Jouw hulp als mijn referent bij mijn sollicitatie voor de opleiding tot uroloog waardeer ik enorm.

Prof. dr. ir. Jenster, beste Guido, via jou ben ik bij dit promotietraject terecht gekomen dus daar wil ik je natuurlijk allereerst voor bedanken! Ik kende je al een beetje via mijn onderzoeksmaster en Geneeskunde studie en het enthousiasme voor het onderzoek dat je daar uitstraalde, heb je na al die jaren nog steeds. Dit werkt enorm motiverend en jouw standaard afsluiting van elke email 'have fun' is daar een weerspiegeling van.

Prof. dr. Martens, beste John, jij hebt mij opgenomen in jouw onderzoekslab en mij wegwijs gemaakt in het meer fundamentele aspect van translationeel onderzoek. Ondanks dat 's ochtends vroeg niet jouw sterkste punt was (waar ik helemaal in kan komen), kon je ineens scherp uit de hoek komen bij de werkbesprekingen. Dat we op een heerlijke zomerdag deze bespreking een keer bij jou in de tuin hebben gedaan, is een van de vele leuke herinneringen. De gezellige sfeer op het lab is ook zeker dankzij jouw Brabantse inborst.

Dr. Lolkema, beste Martijn, jij hebt me onder je hoede genomen gedurende dit traject. Altijd boordevol ideeën en met humor. Je gaf me de ruimte om ook eigen ideeën in te brengen, maar wist altijd een stip op de horizon te stellen. Jouw beeldende metaforen (iets met zand en tennisballen) waren behulpzaam als ik het ook weleens even niet zag zitten. Naast jouw drukke agenda was er ook tijd voor gezelligheid, zoals een bijzonder diner met al jouw PhD'ers bij jou in de tuin of een drankje doen in een soul-bar in Chicago tijdens de ASCO. Ondanks jouw overtuigingskracht heb ik toch voor een vervolg in de urologie gekozen, ik hoop dat je me deze koppigheid vergeeft ;-)

Daarnaast wil ik de leden van mijn kleine commissie, prof. dr. Schalken, prof. dr. Dingemans en prof. dr. Bangma bedanken voor de tijd en moeite die jullie gestoken hebben in het beoordelen van mijn proefschrift.

Prof. dr. Van Moorselaar, prof. dr. Van der Graaf, dr. ir. Van Weerden, prof. dr. Terstappen, prof. dr. De Rijke en prof. dr. Van Laere, bedankt dat jullie bereid zijn om als grote commissie met mij van gedachten te wisselen over mijn proefschrift. Prof. dr. De Bono thank you for the time and effort you have put into reviewing my thesis.

Prof. dr. Sleijfer, beste Stefan, als een ietwat 'buitenbeentje' mocht ik aansluiten bij de Liquid Biopsies Helden bespreking op dinsdagochtend met jouw promovendi. Naast de gezellige noot, heb ik ontzettend veel geleerd hier van alle discussies en kritische vragen. Ook jouw input op mijn manuscripten was zeer nuttig. Ik heb jou niet alleen leren kennen als een deskundig arts, afdelingshoofd en wetenschapper, maar ook als een swingende muzikale encyclopedie ♫

Maurice, zonder jou waren een hoop projecten niet gelukt. Al was het alleen niet voor jouw eindeloze en altruïstische bloeddonaties als 'healthy control'. Jij hebt je vol overgave gestort op de cel-vrij DNA projecten en ik ben blij dat we hierbij altijd zo fijn hebben kunnen samenwerken. Hierbij heb ik veel van jou kunnen leren. Ontzettend leuk ook dat je mij de mogelijkheid gaf om een presentatie over cel-vrij DNA in Porto te geven.

Dr. Kraan, beste Jaco, jouw praktische skills en CTC-kennis maken jou onmisbaar voor het lab. Hoewel jouw relaxte houding een beetje haaks staat op mijn planningsdrang, had je (meestal) gelijk dat het wel goed kwam. Naast de serieuze lab-zaken was er ook tijd over voor de gezellige zaken, zoals een dansje in Schloss Krickenbeck of een ludiek afscheidsfilmpje, jij was overal voor in!

Als ik met een bioinformatisch of computer-gerelateerd probleem zat, wist ik jou, Marcel, altijd te vinden, want jij had het binnen no-time opgelost. Je hebt mijn leven vaak makkelijker gemaakt. En wat heb ik kunnen lachen om jouw photoshop talent en 1 april grappen :-D

Saskia, jouw komst naar het lab is echt een aanwinst. Hoewel het CIRCUS methylatieproject niet goed van de grond kwam, vond ik het erg leuk om met jou hierover te sparren en ik vind het mooi om te zien dat er nu een prachtig manuscript ligt over methylatie-profiling van cfDNA.

Lieve Anieta, ondanks dat je er niet meer bent, verdien jij ook zeker een plek hier. Jouw kennis over PCR was ongeëvenaard. Ik ben dankbaar voor je pipeteerlessen, je PCR lessen en jouw hulp bij het ontwikkelen van ingewikkelde assays. Wat een gemis.

De post-docs Jozien Helleman en Antoinette Hollestelle wil ik ook bedanken. Ondanks dat we niet echt gezamenlijke projecten hadden, was het altijd gezellig tijdens meetings en lunches.

Lieve Joan, onze lab-mama, wat fijn dat je er altijd voor mij was. Wat hebben we gelachen, gekletst, koffietjes gedronken en (gelukkig maar heel af en toe) gehuild. Vraagbaak en rots-in-de-branding voor mij en het lab met jouw ervaring en kennis. Ik weet zeker dat je gemist wordt, maar heerlijk dat je nu van je pensioen en je (klein)kinderen kan genieten, je hebt het verdiend!

Jean, ook jou ben ik veel dank verschuldigd voor jouw hulp en input bij mijn projecten. Wat betreft digital PCRs heb ik veel van je kunnen leren.

Mai, Michelle, Kirsten en Corine, wat hebben jullie mij goed geholpen met het verwerken van de stroom CTC en cfDNA samples. Jullie inzet en precisie zijn voor mij heel waardevol geweest.

Verder wil ik Mieke, Wendy, Vanja, Anita en Renee bedanken voor de praktische hulp op en rondom het lab.

Ook veel dank aan de secretaresses, in het bijzonder Rosita en Gerdien, voor alle hulp bij het inplannen van afspraken in de drukke agenda's van de heren en allerhande praktische zaken waar jullie het antwoord op wisten. Rosita, je hebt jezelf in korte tijd onmisbaar weten te maken door voor alles een oplossing te hebben (iemand een droge broek nodig?) en door jouw altijd opgewekte humeur. Gerdien, wat jammer dat je weg bent gegaan, maar ik hoop dat je het naar je zin hebt bij je nieuwe uitdaging.

Voor hun hulp met de statistische analyses wil ik graag Maxime Look en Esther Oomen-de Hoop bedanken. Als de getallen mij begonnen te duizelen dan waren zij de reddende engelen.

Graag wil ik ook de krachten van het Clinical Trial Center (CTC), Soumia, Nelly, Martin, Susan en Louise bedanken voor hun hulp met het opzetten van de PROuD database, alle studielogistiek en het invullen van de klinische dataformulieren.

Ook wil ik de medische studenten, Thomas en Stephanie, die ik heb mogen begeleiden bedanken voor hun hulp en inzet.

Bij een hoop van mijn onderzoeksprojecten heb ik nauw samengewerkt met verschillende afdelingen en collega PhD'ers, wat ontzettend leuk en leerzaam was.

Silvia, thank you for your help with our cfDNA projects. You are very passionate about your work and life and I hope that this will help you to finish your PhD as well. Arrivederci!

Lisanne M., hoe heeft Martijn zijn PhD'ers zo kunnen kiezen… ;-) ik heb ontzettend fijn met je kunnen samenwerken aan ons CIRCLE project. Je bent een ware 'organoid whisperer' en we kunnen trots zijn op onze mooie publicatie! Ook buiten de projecten om was het altijd gezellig. Ik wens je veel succes met je verdediging en je nieuwe uitdaging in Canada.

Ook wil ik hier Sigrun en Wytske bedanken voor hun ondersteuning en input voor het CIRCLE project, dit was echt een team effort.

Job en Harmen, jullie bioinformatische expertise is van groot belang geweest voor meerdere projecten met als kers op de taart natuurlijk onze publicatie in Nature Communications. Job ik wens je veel succes met het afronden van je PhD, maar dat komt ongetwijfeld goed.

Gido en Hans, jullie hebben mij wegwijs gemaakt in de wereld van LC-MC/MC en steroïden, echt weer een andere tak van sport. Bedankt voor de fijne samenwerking!

Prof. dr. Stoecklein, dear Nick, thank you for welcoming me in your lab those two months. I really enjoyed to 'een kijkje in een andere keuken nemen'. Also, I'm grateful for your hospitality at the DCC-Net Retreat in Schloss Krickenbeck; science and fun are a great combination!

Rui and Guus thank you for your guidance at the lab, I really learned a lot. And off course thank you, and Christiane, Rosa, Elina, Maria and the other colleagues, for all the nice times and tips in Dusseldorf.

Mijn klinische studies waren niet mogelijk geweest zonder de hulp van de oncologen in het Erasmus MC, Franciscus Gasthuis & Vlietland, Haga Ziekenhuis en Maasstad Ziekenhuis. Met name veel dank aan dr. Hamberg, dr. Van der Veldt, dr. Houtsma en dr. Sandberg (MSZ), en aan de research- en oncologie verpleegkundigen Hans, Suzan en Suraya.

Beste dr. Hamberg, jouw inclusiesnelheid is ongekend, waarvoor veel dank. Heel leuk dat ik ook als ANIOS urologie met je heb kunnen samenwerken.

Dr. Te Boekhorst veel dank voor de hulp van u en de hematologie afdeling bij het afnemen van de diagnostische leukaferese samples.

Ook wil ik graag alle co-auteurs bedanken voor hun input en hulp bij het tot stand komen van de manuscripten in dit proefschrift.

Beste Ietje, wat fijn dat ik jou heb leren kennen, want anders had ik niet zo'n mooi proefschrift gehad. Ontzettend bedankt voor je hulp en flexibiliteit bij het vormgeven van mijn proefschrift. De eindsprint hebben we samen gered!

Uiteraard wil ook graag mijn mede-promovendi bedanken voor de gezellige lunches, besprekingen en uitjes: Lindsay, Inge, Marjolein, Nick, Anouk, Pauline, Wendy, Tomasso, Teoman en Manouk.

Het begon met een reisje naar Stockholm en kreeg een vervolg als Texelse beachbabes, Dusselsche dirndels en Volendamse viswijven. Lindsay, Inge en Marjolein, jullie zijn het levende bewijs dat leuke collega's je promotietijd maken. Natuurlijk hebben we veel aan elkaar gehad op en rond het lab, zoals in onze zelfbenoemde 'Coffee Company Be414', tijdens de Kerstbrunches en pre-work ontbijtjes, maar ook daarbuiten tijdens gezellige etentjes, borrels en reisjes!

Marjolein, mijn R-buddy, je hebt je staande weten te houden tussen al die gekke cliniclowns met prikdrang. Jouw biomedische achtergrond en Limburgse gezelligheid was hierop de perfecte aanvulling. Jij begreep mijn euforie als ik eindelijk erin geslaagd was om dat ene streepje in R te programmeren ;-)

Lindsay, een van mijn paranimfen! Jij staat altijd voor iedereen klaar en wat fijn dat je mij op deze spannende dag wilt bijstaan :-D Ik ben blij dat we je weggesleept hebben uit die andere kamer. Als vroege vogel had jij er al een halve werkdag opzitten als iedereen kwam binnendruppelen en het bijkletskwartier (zeg maar gerust een uur) met een Nespresso of Starbucks koffietje begon. Dat jij kan doorpakken blijkt wel uit die superverdiende AIOS-plek en een supermooie promotie op de planning (jammer dat we het toch niet voor elkaar hebben gekregen om er een symposium met Inge van te maken :-P).

Inge, de 'mama' van het halfbroertje van Bintang, in de kattenliefde hadden we elkaar al gauw gevonden ;-) Jouw creativiteit en fotografie-skills kennen geen grenzen en ik vind het knap dat je je eigen weg gevonden hebt. Chicago was maar half zo leuk geweest als we dat niet samen hadden kunnen doen :-D En laten we ook zeker de Fika erin houden!

Nick, ten eerste complimenten dat je het met al die vrouwen in Be-414a hebt weten uit te houden. Jij als senior PhD'er en ik als junior, wist ik gelukkig toch snel je favoriete collega te worden ;-) Bedankt voor de leuke borrels en etentjes en natuurlijk ook voor je hulp toen ik net kwam kijken.

Anouk, wat fijn dat jij mijn studies hebt willen overnemen, want ik weet dat ze in kundige handen zijn. Veel succes met jouw promotietraject!

En natuurlijk ook de promovendi van de 'andere kant' wil ik bedanken: Florence, Bodine, Femke, Koen, Sander, Daan, Melissa, Yarne en Maud.

Florence, wij blijken een hoop gemeenschappelijk te hebben, zoals onze copromotor Martijn en onze liefde voor katten en shoppen in Chicago. Je bent een heerlijke chaoot, maar tegelijk ook enorm gefocust. Veel succes met de laatste loodjes van je proefschrift!

Lieve uro-onderzoekers, ook al werkten we niet direct samen gedurende mijn PhD, ook jullie wil ik bedanken voor de gezellige tijd tijdens de EAU congressen, RUAG borrels en onderwijsmomenten. Hopelijk kunnen we dit snel weer oppakken.

Ook wil ik de urologen bedanken in het Amphia en Franciscus Gasthuis & Vlietland, waar ik heb gewerkt als ANIOS terwijl ik mijn onderzoek nog aan het afronden was. Ik heb veel bij jullie geleerd en wat leuk om mijn onderzoek te presenteren tijdens de onderwijsmomenten en refereeravonden!

Lieve PIT'jes, 'Independent doctors', TSN en mijn andere vriendinnen, bedankt dat jullie zorgden voor een leven buiten mijn promotie :-D Even over iets anders kletsen en lachen tijdens gezellige etentjes en borrels, heerlijk!

Stefanie, wat fijn dat ook jij mij bij wilt staan als paranimf tijdens deze spannende dag. De rust die jij uitstraalt kan ik vast goed gebruiken!

Allerliefste pap en mam, zonder jullie onvoorwaardelijke steun en liefde was ik niet waar ik nu ben. Hoe ontzettend dankbaar ik hiervoor ben, kan ik niet in woorden uitdrukken. Jullie staan altijd voor mij klaar, ik hou van jullie!

Lieve Sab(ine), als mijn grote zus ben jij een voorbeeld voor mij als ik zie hoe je van het leven geniet met je prachtige gezin met Jim en de kids. Onze bijzondere band blijkt wel uit dat we elkaars getuigen zijn op elkaars bruiloft :-D

Ook wil ik mijn lieve schoonouders Zvonko en Monique, mijn zwager Nick, en mijn (schoon) familie bedanken voor alle fijne momenten met elkaar!
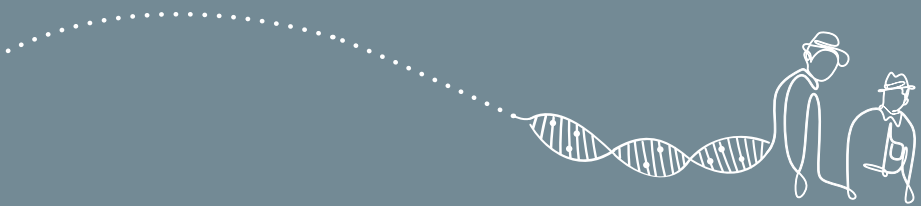
Jelte, soms lijk jij mij beter te kennen dan ik mijzelf. Bij alle hoogte- en dieptepunten was jij er voor mij. Een promotietraject is hoe dan ook enerverend, maar wij besloten ook alle 'life changing events' tijdens deze periode te doen: trouwen, eerste huis kopen, een kind krijgen. Want je hebt gelijk, het komt nooit goed uit en dan blijkt het perfect te kloppen. Met alles ben jij mijn steun en toeverlaat.

Allerliefste Melle, misschien dat mijn proefschrift er eerder ondanks, dan dankzij jou is gekomen, want je weet je moeder wel bezig te houden... ;-) Toch weet je mijn leven ook een stuk simpeler te maken, want met een lach is alles weer goed**.** Ik geniet van alles met jou!

Mijn dankwoord zou niet compleet zijn zonder mijn kat Bintang te noemen. Mijn gehele promotietraject stond hij aan mijn zijde en kon ik als het nodig was de stress van mij af 'aaien'[1].

[1]Allen, K. et al.
Cardiovascular Reactivity and the Presence of Pets, Friends, and Spouses:
The Truth About Cats and Dogs.
Psychosomatic Medicine: September 2002

**Unravelling
the Genomic Landscape
of Metastatic
Prostate Cancer**
a prospect on
patient stratification using
blood-based biomarkers