PrECISE

D1.3

Clonal classification of tumours

Project number:	668858
Project acronym:	PrECISE
Project title:	PrECISE: Personalized Engine for Cancer Integrative Study and Evaluation
Start date of the project:	1st January, 2016
Duration:	36 months
Programme:	H2020-PHC-02-2015

Deliverable type:	OTHER
Deliverable reference number:	PHC-668858 / D1.3 / 1.0
Work package contributing to the deliverable:	WP 1
Due date:	December 2018 – M36
Actual submission date:	20th December, 2018

Responsible organisation:	UZH
Editor:	Dorothea Rutishauser
Dissemination level:	PU
Revision:	1.0

Abstract:	Classification of tumours according to dominant clonal content
Keywords:	copy number alterations, tumour phylogenies, clonal classification of tumours

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 668858.

This work was supported (in part) by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 15.0324-2. The opinions expressed and arguments employed therein do not necessarily reflect the official views of the Swiss Government.

Editor

Dorothea Rutishauser (USZ)

Contributors (ordered according to beneficiary numbers)

Pavel Sumazin (BCM)

Jelena Cuklina (ETH)

Peter Wild (UZH)

Dorothea Rutishauser (UZH)

Matteo Manica (IBM)

Executive Summary

For this deliverable we used clonality interference to select tumours for the follow up by ultra-deep sequencing (done in WP2), and led efforts to improve clonality interference.

Contents

List of Figures

Chapter 1 Introduction

We profiled tumors from 5 CRPC patients at multiple time points and multiple regions per time point for copy number alterations using Affymetrix OncoScan arrays and for mutations using a targeted sequencing panel that evaluates alterations at the loci of 42 genes. Our approach was to introduce a model for the effects of CNAs on mutated-read fractions. We use this model as a basis for simulations with CNA distributions that are compatible with observations from The Cancer Genome Atlas (TCGA)-profiled primary breast, HCCs, PCs, and Wilms' tumors (TCGA, 2017; The Cancer Genome Atlas, 2012; The Cancer Genome Atlas, 2015). We designed Chimaera (Manica et al., 2018) to improve mutation-frequency and CNA estimations from WES of tumours with genetic instability, we developed Chimaera: clonality inference from mutations across biopsies. Chimaera relies on multiple biopsies for the same tumour to, first, approximate CNAs and mutation frequencies; then, identify mutations with similar approximate frequencies and associate them with subclones; and, finally, to estimate the true frequencies of these mutations and the associated subclones. As is the case for estimates made by SCHISM, ABSOLUTE and other methods, Chimaera is not able to produce frequency estimates for all mutations, but compared to existing methods is able to process and determine true frequencies for more variants, exhibiting more power in identifying potentially tumour initiating mutations and disease drivers. Finally, to demonstrate that Chimaera is able to reconstruct subclones from tumour profiles we produced Chimaera-inferred subclones and resulting phylogeny from profiles of ten biopsies taken from a castration-resistant prostate cancer (CRPC) tumour and a set of profiles extracted from five different tumour areas from ten hepatocellular carcinoma (HCC) patients (Lin et al., 2017).

Chapter 2 Patient data overview

From the cohort of CRPC patients, punches have been done for 9 patients (MetaProC) from 1 normal and 5 different tumour areas from each patient. For 42 CRPC patients from ZTMA204 one tumour region was punched from each patient.

Patients have been previously described in D6.2, including selected tumour areas for the different CRPC patients.

Chapter 3 CRPC clonality analysis

In total, 72 samples corresponding to 5 patients were profiled, including a control for each patient. Profiled samples, including profiling time points, Gleason scores, and therapeutic history are given in Table S1[1]. Our analysis was used to infer tumor phylogenies, including the order of the emergence of dominant tumor subclones in these 5 patients (Patient 1-5).

Patient 1 was profiled at 3 time points across 2.3 years, and was assessed mutually disjoint mutations in EP300 and AR (Figure 1A), corresponding to 2 distinct proliferative tumor subclones. The clonal EP300 mutation (p.I997V) is predicted to be

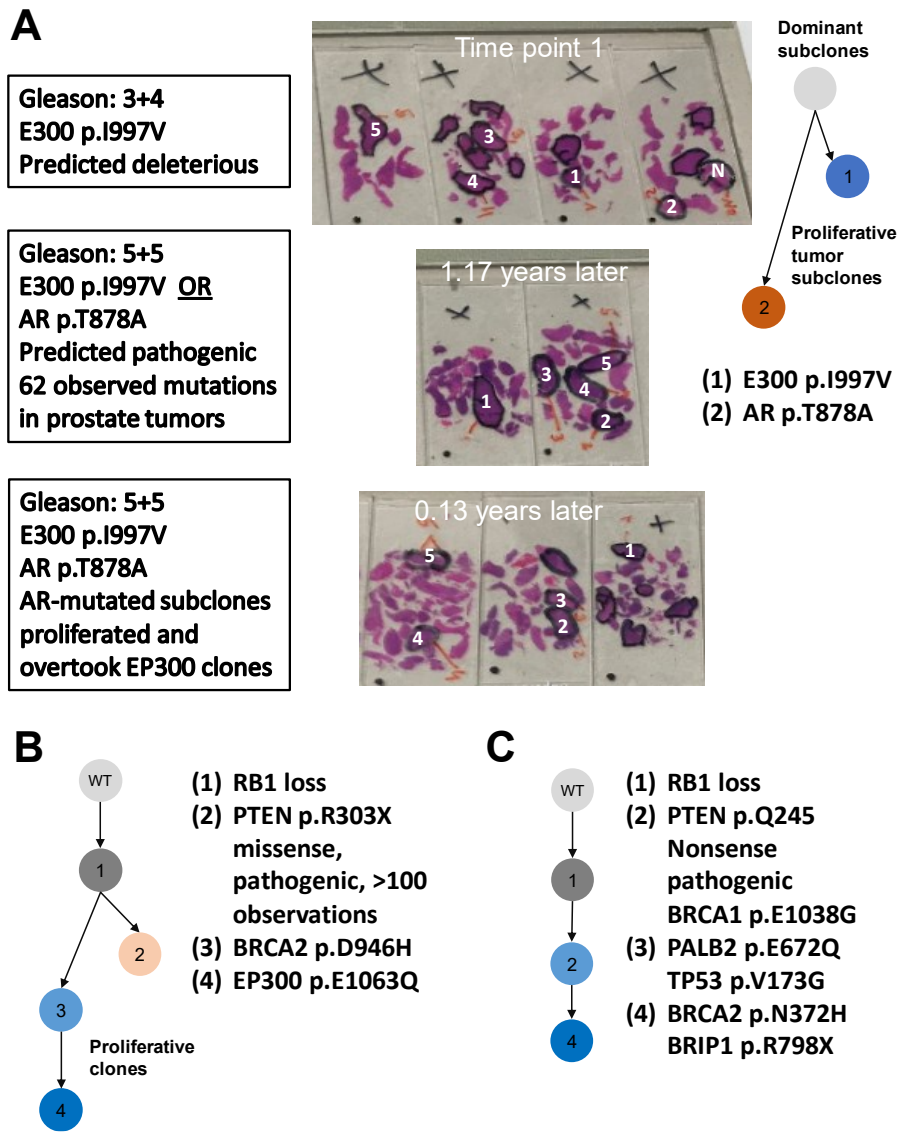

Figure 1. Predicted phylogenies for 3 CRPC patients.

(A) Patient 1 was profiled at 3 time points across 2.3 years following therapy, 5 tumor areas per tumor point. Selected regions are marked as illustrative example. The first profile detected a deleterious mutation in EP300 at Time point 1; Chimaera predicted that a previously observed pathogenic mutation in AR was acquired in an independent tumor subclone at time point 2 and that the proliferation of this tumor subclone coincided with increased Gleason Score. Chimaera analysis also predicted the order of deleterious mutation acquisition in **(B)** Patient 2 and **(C)** Patient 3. The mutations AR p.T878A, PTEN (p.R303X, p.Q245), BRCA1 p.E1038G, and BRCA2 p.N372H have been previously observed in cancers.

[1] Separate confidential document: PrECISE-D1.3-M36-Table-S1-CO.xlsx

deleterious and was inferred to be present in the majority of tumor cells at Time point 1. The well documented pathogenic mutation *AR* p.T878A was detected at Time point 2 and its detection coincided with a loss of the *EP300* mutation and an increase in the tumor's Gleason Score. At Time point 3, regions that were positive for the *AR* mutation tested negative for the *EP300* mutation.

Patient 2 (Figure 1B) was profiled at 5 time points across 1.8 years and was assessed the known pathogenic stop-gain mutation *PTEN* p.R303X at Time point 1 in addition to a heterozygous loss of *RB1*. At Time point 1, *PTEN* p.R303X was inferred in nearly all tumor cells. Following treatment with a luteinizing hormone releasing hormone (LHRH) analogue, this patient gained a mutation in *BRCA2* (p.N372H) and then in *EP300* (p.E1063Q); both were gained in Time point 4 and are predicted to be damaging. This coincided with an increase in Gleason score (5+4 to 5+5) and was followed by the introduction of the combination treatment LHRH and Casodex. Finally, Time point 5 profiles suggest that the *BRCA2-EP300* significantly increase in proliferation and this tumor subclone, which was infrequent at Time point 4, expands and accounts for most of the tumor cells.

Patient 3 (Figure 1C) was profiled at 2 time points and had an increase in Gleason score (5+4 to 5+5). This patient had a castration resistant cancer and his therapy included orchiectomy 7 years prior to the first biopsy. His Chimaera inferred phylogeny suggested a dominant tumor subclone with a heterozygous loss of *RB1* together with previously-observed pathogenic *PTEN* nonsense (p.Q245) and *BRCA1* (p.E1038G) mutations. This subclone later acquired mutations in *PALB2* (p.E672Q) and *TP53* (p.V41G; predicted to be deleterious), followed by *BRCA2* (p.N372H; previously observed) and a stop-gain mutation in *BRIP1* (p.R798X). All of these mutations were present at Time point 1.

Phylogenies inferred for Patients 4 and 5 were simpler. Patient 4's phylogeny included a sequence of 5 intronic and synonymous mutations with unknown significance in *TMPRSS2*. While Patient 5's phylogeny included a predicted initiating mutation in *PIK3CA* (p.Y182H); see Table S1[2] for all mutation frequency data.

Interestingly, analyses of regions that were inferred to be enriched for subclones with specific mutations identified pathway enrichment that supports Chimaera predictions. Specifically:

(1) Comparing profiled regions from tumour samples from patient P6 with low and high frequencies of tumour subclones with a PTEN mutation identified dysregulation of PTEN targets (p<5E-6).
(2) Comparing profiled regions from tumour samples from patient P6 with low and high frequencies of tumour subclones with a TP53 mutation identified dysregulation of TP53 targets (p<2E-17).
(3) Comparing profiled regions from tumour samples from patient P6 with low and high frequencies of tumour subclones with a BRCA1 mutation identified dysregulation of BRCA targets (p<7E-20).
(4) Comparing profiled regions from tumour samples from patient P1 with low and high frequencies of tumour subclones with a AR mutation identified dysregulation of AR targets (p<2E-5).

[2] Separate confidential document: PrECISE-D1.3-M36-Table-S1-CO.xlsx

Chapter 4 Summary and Conclusion

We identified mutations that are known to be involved in cancer development. The unique setup of time-course allowed use to monitor the clonal evolution. The hard part of prostate cancer analysis is that it is not mutation-rich cancer but has a lot of copy number alterations. For this we had to adapt the clonality inference algorithm Chimeara, which was successfully applied to infer clonal composition of 9 patients that have been profiled. However, to determine whether this is typical for prostate cancer, more patients need to be examined. some more words on classification] [proteomics maybe more suitable for prostate cancer patients classification, as not many mutations are underlying prostate cancer] development.

Chapter 5 List of Abbreviations

CRPC	castration-resistant prostate cancer
GS	Gleason Score
LHRH	luteinizing hormone releasing hormone

Chapter 6 Bibliography

Lin, D.-C., Mayakonda, A., Dinh, H. Q., Huang, P., Lin, L., Liu, X., Ding, L.-w., Wang, J., Berman, B. P., and Song, E.-W. (2017). Genomic and epigenomic heterogeneity of hepatocellular carcinoma. Cancer Research *77*, 2255-2265.

Manica, M., Chouvarine, P., Mathis, R., Wagner, U., Oehl, K., Saba, K., Roditi, L. D. V., Pati, A. N., Martínez, M. R., Wild, P. J., and Sumazin, P. (2018). Inferring clonal composition from multiple biopsies of the same tumor. arXiv, 1701.07940.

TCGA (2017). Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. Cell *169*, 1327-1341. e1323.

The Cancer Genome Atlas (2012). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61-70. PMC3465532

The Cancer Genome Atlas (2015). The molecular taxonomy of primary prostate cancer. Cell *163*, 1011-1025. PMC4695400