

---

Establishing a pipeline for analysis of T-cell receptor sequencing data  
and characterizing the impact of the T-cell repertoire on the clinical  
outcome of bladder cancer

---

Master's thesis (30 ECTS)

Asbjørn Kjær | 201709022

Bioinformatics Research Centre (BiRC)

Aarhus University

June 2022



Supervisor:

Nicolai Juul Birkebæk

Associate professor

Department of Molecular Medicine (MOMA)

Aarhus University

## Abstract

T-cells play a crucial role in the fight against cancer, as they can recognize the mutated proteins of the cancer and target the malignant cells for destruction. This contact is mediated through the T-cell receptor (TCR), which has a distinct nucleotide sequence for each T-cell clone. Sequencing of this receptor through T-cell receptor sequencing (TCRseq) can thus be used to characterize the composition of the T-cell repertoire, which can provide insights into the state of the cancer. In this thesis a pipeline for assembling TCRseq data into unique TCR clones was implemented and tested. This pipeline was then used to characterize the impact of the peripheral TCR repertoire on the clinical outcome of bladder cancer in two local cohorts. This analysis revealed that a more diverse TCR repertoire is associated with better prognosis. Patients with good clinical outcome are distinguished by a change in diversity during treatment, while patients with recurrence have a repertoire characterized by constant diversity and a high degree of clonal overlap through time, both when treated with chemotherapy or immunotherapy. This study highlights the importance of the systemic immune system in the fight against cancer and calls for further investigation into using the state of the immune system as a biomarker to guide cancer treatment.

## Abbreviations

Bacillus Calmette-Guérin	(BCG)
Complementarity-determining region	(CDR)
Clonal hematopoiesis of intermediate potential	(CHIP)
High-grade free survival	(HGFS)
Humant leukocyt antigen	(HLA)
Major histocompatibility complex	(MHC)
Muscle invasive bladder cancer	(MIBC)
Neoadjuvant chemotherapy	(NAC)
Non muscle invasive bladder cancer	(NMIBC)
Recurrence free survival	(RFS)
Recover T Cell Receptor	(RTCR)
T-cell receptor	(TCR)

<b>1 INTRODUCTION .....</b>	<b>1</b>
1.1 THE ROLE OF T-CELLS IN THE FIGHT AGAINST CANCER .....	1
1.2 THE MECHANISM BEHIND T-CELL RECEPTOR DEVELOPMENT .....	2
1.3 METHODS FOR T-CELL RECEPTOR SEQUENCING .....	3
1.4 THE CLINICAL VALUE OF THE T-CELL RECEPTOR REPERTOIRE IN CANCER .....	4
1.5 THE TREATMENT OF BLADDER CANCER .....	6
1.6 PRINCIPLES BEHIND BIOINFORMATIC ASSEMBLY OF V(D)J JUNCTIONS .....	7
1.7 BIOINFORMATIC TOOLS FOR V(D)J JUNCTION ASSEMBLY.....	8
1.7.1 <i>Vidjil</i> .....	8
1.7.2 <i>RTCR</i> .....	9
1.7.3 <i>MiXCR</i> .....	10
1.8 AIM OF STUDY.....	10
<b>2 METHODS.....</b>	<b>11</b>
2.1 TCR SEQUENCING PROTOCOL.....	11
2.2 PIPELINES FOR TCR CALLING TOOLS.....	11
2.3 BOOTSTRAP EXPERIMENT .....	11
2.4 TCR OUTPUT ANALYSIS .....	12
2.5 ADDITIONAL BIOINFORMATIC METHODS .....	13
2.5.1 <i>HLA calling and neoantigen prediction</i> .....	13
2.5.2 <i>T-cell fraction prediction</i> .....	13
2.6 COHORT DESCRIPTION .....	14
<b>3 RESULTS .....</b>	<b>15</b>
3.1 EVALUATING PERFORMANCE OF TCR CALLING TOOLS.....	15
3.1.1 <i>Evaluating the output quality of TCR calling tools</i> .....	15
3.1.2 <i>Evaluating the overlap of TCR repertoires between TCR calling tools</i> .....	16
3.1.3 <i>Exploring J and V gene usage of RTCR</i> .....	18
3.1.4 <i>Evaluating the stability of TCR calling tools using bootstrapping</i> .....	18
3.1.5 <i>Evaluating the stability of TCR calling tools using replicate samples</i> .....	20
3.1.6 <i>Final evaluation of TCR calling tools</i> .....	21
3.2 QUALITY CONTROL.....	23
3.2.1 <i>Sample quality</i> .....	23
3.2.2 <i>Distribution of clone sizes</i> .....	23
3.2.3 <i>Defining the optimal filtering threshold</i> .....	24
3.2.4 <i>Evaluating sequencing depth bias</i> .....	26
3.3 INVESTIGATING THE DYNAMICS OF THE TCR REPERTOIRE IN BLADDER CANCER.....	27
3.3.1 <i>Analyzing the TCR repertoire in patients with MIBC</i> .....	27
3.3.1.1 <i>Low baseline TCR repertoire diversity is associated with relapse</i> .....	28
3.3.1.2 <i>TCR repertoire diversity decreases for patients without relapse</i> .....	30
3.3.1.3 <i>Patients without relapse undergo expansion of clones during treatment</i> .....	33
3.3.1.4 <i>A high degree of TCR repertoire overlap is associated with relapse</i> .....	33
3.3.1.5 <i>Correlation between TCR repertoire and other immune system factors</i> .....	36
3.3.2 <i>Analyzing the TCR repertoire in patients with NMIBC</i> .....	38
3.3.2.1 <i>A change in TCR repertoire diversity associates with late recurrence</i> .....	39
3.3.2.2 <i>A high degree of TCR repertoire overlap associates with early recurrence</i> .....	41
3.3.2.3 <i>TCR diversity is not associated with neoantigens or HLA-type</i> .....	43
<b>4 DISCUSSION.....</b>	<b>44</b>
4.1 EVALUATION OF TCR CALLING TOOLS .....	44
4.2 FILTERING OF THE TCR REPERTOIRE.....	45
4.3 COMPARING THE CHARACTERISTICS OF THE TCR REPERTOIRE IN MIBC AND NMIBC .....	46
4.4 CONCLUDING REMARKS .....	51
<b>REFERENCES .....</b>	<b>52</b>
<b>SUPPLEMENTARY .....</b>	<b>57</b>