

DISS. ETH NO. 26736

Systems analysis of clonal selection and evolution in adaptive immune repertoires

A thesis submitted to attain the degree of  
DOCTOR OF SCIENCES of ETH ZURICH  
(Dr. sc. ETH Zurich)

Presented by  
ALEXANDER DIMITRI YERMANOS

Master of Science in Bioinformatics and Computational Biology, ETH Zurich &  
University of Zurich

born on 07.03.1992  
citizen of the United States of America

accepted on the recommendation of:

Prof. Dr. Sai Reddy, examiner  
Prof. Dr. Annette Oxenius, co-examiner  
Prof. Dr. Doron Merkler, co-examiner  
Prof. Dr. David Gfeller, co-examiner

Zurich, 2020

## SUMMARY

Recent developments in deep sequencing and bioinformatics have catalyzed a systems level of analysis adaptive immunity. Fundamental questions investigating the interplay between clonal selection, expansion, and evolution can now be addressed through adaptive immune (B cell receptors, secreted antibodies, and T cell receptors) repertoire sequencing and bioinformatic analysis. These inquiries inform both how the immune system overcomes infection and how repertoire analysis can instruct the discovery and development of therapeutics. In this thesis, I have leveraged an array of deep sequencing and bioinformatic analysis methods to profile B cell receptor (BCR) and T cell receptor (TCR) repertoires in the context of immunization and viral infection. In concert with this, we developed and validated a computational framework to benchmark antibody repertoire evolution. Combining *in silico* and *in vivo* repertoire sequencing experiments, we explored the robustness of various bioinformatics processing in determining evolutionary parameters such as mutation rates, phylogenetic tree topology, and clonal selection. Utilizing the murine lymphocytic choriomeningitis virus (LCMV) and murine cytomegalovirus (MCMV) models, we next investigated the B and T cell repertoires following chronic, acute, and persistent infection. Additionally, we leveraged single cell repertoire sequencing technologies to discover monoclonal antibodies following both protein immunizations and chronic viral infection. Finally, we developed a bioinformatics framework to quantify genome-encoded sequence diversity of the antibody locus across several commonly-studied species. Combining computational and experimental sequencing experiments, we molecularly quantified the adaptive immune response following infections and immunizations and explored the resistance of our conclusions to bioinformatic methodology.

## ZUSAMMENFASSUNG

Aktuelle Entwicklungen im Deep Sequencing sowie in der Bioinformatik haben eine systemische Analyse der adaptiven Immunabwehr ermöglicht. Grundlegende Fragen bezüglich des Zusammenspiels von klonaler Selektion, Expansion und Evolution können nun durch die Sequenzierung des adaptiven Immunrepertoires (B-Zell-Rezeptoren, sezernierte Antikörper und T-Zell-Rezeptoren) und mittels bioinformatischer Analysen angegangen werden. Daraus lassen sich Rückschlüsse auf Mechanismen ziehen, mit denen das Immunsystem Krankheiten überwindet, und Anleitungen bezüglich der Wirkstoffentdeckung und -entwicklung folgern. In der vorliegenden Arbeit habe ich eine Reihe analytischer Methoden aus den Bereichen des Deep Sequencing und der Bioinformatik eingesetzt, um die Repertoires von B- und T-Zell-Rezeptoren im Kontext von Immunisierung und vitaler Infektion zu beschreiben. Dafür haben wir ein Framework entwickelt und validiert, mit dem die Evolution von Antikörpern verglichen werden kann. Indem wir *in silico* und *in vivo* Sequenzierungsexperimente von Immunrepertoires kombiniert haben, waren wir in der Lage, die Robustheit verschiedener Methoden der Bioinformatik im Hinblick auf Parameterschätzungen zu bestimmen. Zu den betrachteten Parametern gehörten unter anderem die Mutationsrate, Stammbaumtypologie sowie klonale Selektion. Unter Zuhilfenahme von Modellen basierend auf murinem lymphozytärem Chriomeningitis-Virus (LCMV) und murinem Zytomegalovirus untersuchten wir B- und T- Zell-Repertoires nach akuter, chronischer und anhaltender Infektion. Ausserdem nutzen wir Einzelzell-Sequenzierungstechnologien, um monoklonale Antikörper sowohl nach Immunisierung als auch nach chronischer viraler Infektion zu entdecken. Abschliessend haben wir ein bioinformatisches Framework entwickelt, um in einigen oft untersuchten Spezies die im jeweiligen Antikörper-Genlocus kodierte Sequenzdiversität zu ermitteln. Durch die Kombination von *in silico* und *in vivo* Sequenzierungsexperimenten war es uns möglich, die adaptive Immunantwort nach Infektionen sowie nach Immunisierungen zu quantifizieren und die Beständigkeit unserer Schlussfolgerungen in Abhängigkeit von bioinformatischer Methodik zu ergründen.