DISS. ETH N° 000000

Learning to align single-cell populations

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES

(Dr. sc. ETH Zurich)

presented by

STEFAN STARK

born on March 11, 1992

citizen of United States

accepted on the recommendation of

Prof. Dr. Gunnar Rätsch (ETH Zurich), examiner
Dr. Kjong-van Lehmann (RWATH), examiner
Prof. Dr. Andreas Krause (ETH Zurich), co-examiner
Prof. Dr. Oliver Stegle (EMBL), co-examiner

2024

Contents

1 Introduction

1.0.1 *Scope*

1.0.2 *Publications*

1.0.3 *Collaborators*

1.1 INDIVIDUAL CELLS AS COMPLEX MACHINERY

1.1.1 *central dogma*

1.1.2 *pathways & signaling hierarchies*

1.2 SINGLE-CELL PROFILING TECHNOLOGIES

1.2.1 *scRNA-seq*

1.2.2 *CyTOF*

1.2.3 *4i & Proteomic imaging*

1.3 CONSTRUCTING SINGLE-CELL REPRESENTATIONS

1.3.1 *eg. autoencoders*

1.4 CASTING AS ALIGNMENT

1.4.1 *Holistic view of cells*

2

Multi-modal integration

3 Learning perturbation responses

3.1 RELATED WORKS

3.1.1 *Piece-wise linear approximations*

3.1.2 *Linear shifts in latent space*

3.2 OPTIMAL TRANSPORT

3.2.1 *Primal OT formulation*

3.2.2 *Dual OT formulation*

3.2.3 *Neural optimal transport*

3.2.3.1 *ICNNs*

3.3 PREDICTING PERTURBATION RESPONSES WITH CELLOT

3.3.1 *IID: learning cancer treatment outcomes in two modalities*

3.3.1.1 *utilized metrics*

3.3.1.2 *lack of ground truth & qualitative analyses*

3.3.2 *OOD: generalizing responses*

3.3.2.1 *Cross species*

3.3.2.2 *Lupus patients*

3.3.2.3 *Stem cell development*

4 Clinical applications

4.1 COHORT DESCRIPTION

4.1.1 *challenges, standardization, normalization*

4.1.2 *heterogeneity, overview, etc*

4.2 LEARNING PATIENT RESPONSES

4.2.1 *quantitative metrics*

4.2.2 *clinical associations*

4.3 PREDICTING PATIENT RESPONSES

4.3.1 *quantitative metrics*

4.3.2 *prediction task*

5

Towards single-cell foundation models: interpretable latent spaces

List of Tables

List of Figures

List of Algorithms