

UNIVERSITÀ DEGLI STUDI DI SALERNO  
DOTTORATO IN MANAGEMENT & INFORMATION TECHNOLOGY



CURRICULUM: INFORMATION SECURITY & INNOVATION SYSTEMS

COORDINATORE: Ch.mo. Prof. Antonelli Valerio

Ciclo XVII N.S.

Novel tools for reproducible  
Next Generation Sequencing data analysis and integration

**Relatori**

Ch.mo. Prof. Tagliaferri Roberto

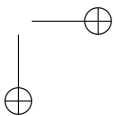
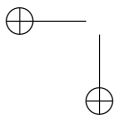
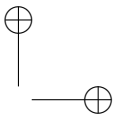
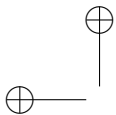
Ch.mo. Prof. Angelini Claudia

**Candidato**

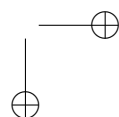
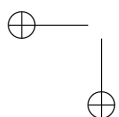
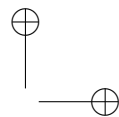
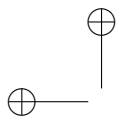
Righelli Dario

Matr. 8800800010

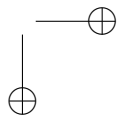
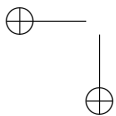
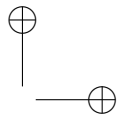
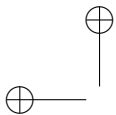
ANNO ACCADEMICO 2017/2018



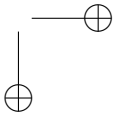
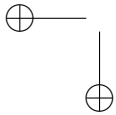
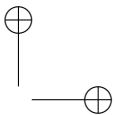
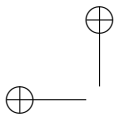
*How to reach a goal?*  
*Without haste but without rest*  
*Goethe*



Add acknowledgements here



Write your abstract here





# Contents

<b>Acknowledgements</b>	<b>5</b>
<b>Abstract</b>	<b>7</b>
<b>1 Introduction</b>	<b>11</b>
1.1 Biological Background . . . . .	11
1.2 Sequencing Techniques . . . . .	11
1.3 Computational Aspects . . . . .	11
<b>2 TiCoRSe - Time Course RNA-Seq data analysis</b>	<b>13</b>
2.1 Introduction . . . . .	13
2.1.1 Time Course RNA-Seq . . . . .	13
2.2 Methods . . . . .	13
2.2.1 General Approach . . . . .	13
2.2.2 Time Course Methods . . . . .	13
2.2.3 Other Methods . . . . .	13
2.2.4 Additional Features . . . . .	13
2.3 Results . . . . .	13

<b>3</b>	<b>DEScan2 - Differential Enriched Scan 2</b>	<b>15</b>
3.1	Introduction . . . . .	15
3.2	Methods . . . . .	15
3.2.1	Peak Caller . . . . .	15
3.2.2	Peak Filtering and Alignment . . . . .	15
3.2.3	Peak Counts . . . . .	15
3.2.4	Additional Features . . . . .	15
3.3	Results . . . . .	15
<b>4</b>	<b>IntegrHO - Integration of High-Throughput Omics data</b>	<b>17</b>
4.1	Introduction . . . . .	18
4.2	Methods . . . . .	18
4.2.1	Single Omic Approach . . . . .	18
4.2.2	Multi Omic Approach . . . . .	18
4.3	Implementation Aspects . . . . .	18
4.4	Reproducible Computational Research . . . . .	18
4.5	Results . . . . .	18
<b>5</b>	<b>Conclusions &amp; Future Works</b>	<b>19</b>
	<b>Appendices</b>	<b>21</b>
.1	R Language . . . . .	23
.2	Rmarkdown Language . . . . .	23
<b>6</b>	<b>Bibliography</b>	<b>25</b>
	<b>List of Figures</b>	<b>27</b>
	<b>List of Tables</b>	<b>29</b>

# Chapter 1

## Introduction

### 1.1 Biological Background

### 1.2 Sequencing Techniques

### 1.3 Computational Aspects



# Chapter 2

## TiCoRSe - Time Course RNA-Seq data analysis

### 2.1 Introduction

#### 2.1.1 Time Course RNA-Seq

### 2.2 Methods

#### 2.2.1 General Approach

#### 2.2.2 Time Course Methods

#### 2.2.3 Other Methods

#### 2.2.4 Additional Features

### 2.3 Results



# Chapter 3

## DEScan2 - Differential Enriched Scan 2

### 3.1 Introduction

### 3.2 Methods

#### 3.2.1 Peak Caller

#### 3.2.2 Peak Filtering and Alignment

#### 3.2.3 Peak Counts

#### 3.2.4 Additional Features

### 3.3 Results





# Chapter 4

# IntegrHO - Integration of High-Throughput Omics data

## 4.1 Introduction

## 4.2 Methods

### 4.2.1 Single Omic Approach

### 4.2.2 Multi Omic Approach

Low Level Itegration

High Level Itegration

## 4.3 Implementation Aspects

## 4.4 Reproducible Computational Research

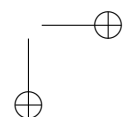
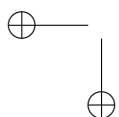
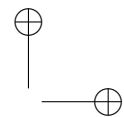
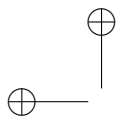
## 4.5 Results

# Chapter 5

## Conclusions & Future Works



# Appendices



## **.1. R LANGUAGE**

---

**23**

### **.1 R Language**

### **.2 Rmarkdown Language**





# Chapter 6

## Bibliography

1. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics* **10**, 57–63. ISSN: 1471-0064 (2009).



## List of Figures



## List of Tables