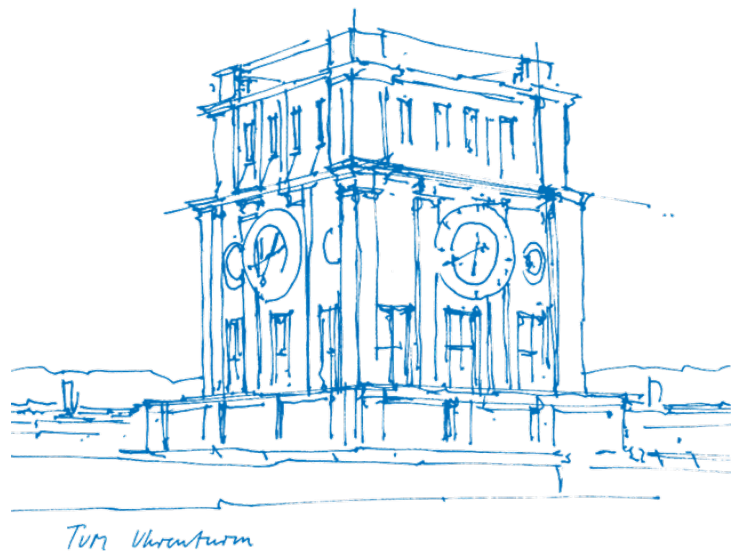


Thesis title

Subtitle of the thesis

Author Name



Thesis title

Subtitle of the thesis

Author Name

Thesis title

Subtitle of the thesis

Author Name

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzende(r):

Prof. Franz X. Gabelsberger

Prüfer der Dissertation:

1. Prof. Dr. Georg Simon Ohm
2. Prof. James Clerk Maxwell

Die Dissertation wurde am 29.04.2016 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 11.07.2016 angenommen.

To Franz X. Gabelsberger, inventor of the street named after him.

Abstract

The abstract of your thesis goes here.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Contents

Abstract	ix
1 Introduction	1
1.1 Pathophysiology of chronic lung diseases	1
1.1.1 Bronchopulmonary Dysplasia (BPD)	1
1.1.2 Asthma	1
1.1.3 Chronic Obstructive Pulmonary Disease (COPD)	1
1.1.4 Idiopathic Pulmonary Fibrosis (IPF)	1
1.2 Computational biology and chronic lung diseases	1
1.2.1 Multi-omics data integration	1
1.2.2 Clinical prediction	1
1.2.3 Systems biology	1
1.2.4 Single-cell transcriptomics	1
1.3 Aims	1
2 Methodology	3
2.1 Data gathering	4
2.1.1 Mice data	4
2.1.2 Human data	4
2.1.3 Public data	4
2.2 Preprocessing	4
2.2.1 Normalization	4
2.2.2 Data imputation	4
2.2.3 Batch-effect detection	4
2.3 Differential expression analysis	4
2.3.1 Limma	4
2.3.2 DESeq2	4
2.4 Enrichment analysis	4
2.4.1 Gene list functional enrichment analysis	4
2.5 Multi-omics factor analysis (MOFA)	4
2.6 Clinical data correlation	5
2.6.1 Linear regression	5
2.6.2 Binomial regression	5
2.6.3 Ordinal regression	5
2.6.4 Multinomial logistic regression	5
2.6.5 Dirichlet regression	5
2.7 Benchmarking of Lasso models dealing with missing values	5
2.7.1 Knowledge guided multi-level network inference	5
2.7.2 Two-steps based models	5
2.7.3 Inverse covariance based methods	5
2.8 Adult data correlation	5
2.8.1 Random forest	5
2.8.2 t-test, manova, log-reg	5

3	Hyperoxia-induced cell cycle arrest drives long-term impairment of lung development and DNA repair in neonates	7
3.1	Summary	8
3.2	Transcriptional characterization of <i>in vitro</i> and <i>in vivo</i> mice under hyperoxia and mechanical ventilation	8
3.2.1	<i>in vitro</i> samples	8
3.2.2	<i>In vivo</i> samples	8
3.3	Transcriptional commonalities between <i>in vitro</i> samples	8
3.3.1	Transcriptional commonalities between AT1/2, EC and MFB	8
3.3.2	Transcriptional commonalities between AT1/2 and MFB	8
3.4	Transcriptional singularities between <i>in vitro</i> samples	8
3.4.1	Alveolar type 1/2 (AT1/2)	8
3.4.2	Myofibroblast (MFB)	8
3.5	Deep transcriptional characterization of cell-cycle arrest under hyperoxia	8
3.5.1	Pre-replication complex hypothesis	8
3.5.2	Experimental validation	8
3.6	Transcriptional comparison between <i>in vitro</i> and <i>in vivo</i> samples	8
3.6.1	Multi-systemic development arrest hypothesis	8
3.7	Discussion	8
3.7.1	Limitations	8
3.7.2	Outlooks	8
4	Multi-omic integration characterizes major endotypes of bronchopulmonary dysplasia linked to adult chronic lung disease	9
4.1	Summary	10
4.2	Feature selection based on differential abundance analysis	10
4.2.1	Proteomics	10
4.2.2	Metabolomics	10
4.3	Enrichment analysis of differentially abundant proteins between BPD and no-BPD samples	10
4.4	Latent factor identification as potential BPD endotypes	10
4.4.1	Factor 1:	10
4.4.2	Factor 2:	10
4.4.3	Factor 3:	10
4.4.4	Factor 4:	10
4.4.5	Factor 5:	10
4.5	Summary of BPD endotypes	11
4.5.1	Acute BPD signature	11
4.5.2	Late/chronic BPD signature	11
4.6	Correlation between BPD endotypes and chronic adult lung diseases	11
4.6.1	Use of imbalanced random forest to separate between COPD and IPF patients using BPD latent factor loadings	11
4.6.2	Characterization of most explanatory features per significant BPD endotype regarding COPD and IPF	11
4.7	Discussion	11
4.7.1	Limitations	11
4.7.2	Outlooks	11
5	Multi-Omics Regulatory Network Inference in the Presence of Missing Data	13
5.1	Summary	13
5.2	Experimental set-up	13
5.2.1	Data preprocessing	13
5.2.2	Parameter selection	13

5.2.3	Missingness simulated scenarios	13
5.2.4	Performance evaluation criteria	13
5.3	Performance evaluation over different missingness scenarios	13
5.3.1	Single-omics random missingness	13
5.3.2	Multi-omics random missingness	13
5.3.3	Block-wise random missingness	13
5.4	General method performance evaluation	13
5.4.1	Performance evaluation over down-sampling	13
5.4.2	Computational runtime evaluation	13
5.5	Discussion	13
5.5.1	Limitations	13
5.5.2	Outlooks	13
6	scNodes: Graphical Interface for Single-Cell Transcriptomics Workflow Development	15
6.1	A section	15
6.2	Another section	15
7	Cytokines Derived from Nasal Epithelial Lining Fluid in Patients with Asthma	17
7.1	A section	17
7.2	Another section	17
8	Discussion	19
A	Appendix	21

1 Introduction

This is the introduction of the thesis.

1.1 Pathophysiology of chronical lung diseases

1.1.1 Bronchopulmonary Dysplasia (BPD)

1.1.2 Asthma

1.1.3 Chronic Obstructive Pulmonary Disease (COPD)

1.1.4 Idiopathic Pulmonary Fibrosis (IPF)

1.2 Cumputational biology and chronic lung diseases

1.2.1 Multi-omics data integration

1.2.2 Clinical prediction

1.2.3 Systems biology

1.2.4 Single-cell transcriptomics

1.3 Aims

2 Methodology

This is the methodology of the thesis.

2.1 Data gathering

2.1.1 Mice data

Transcriptomics

2.1.2 Human data

Transcriptomics

Metabolomics

2.1.3 Public data

Multi-omics bulk data

Neonatal single-cell transcriptomics

2.2 Preprocessing

2.2.1 Normalization

DESeq2

Pareto scaling

Size-effect

2.2.2 Data imputation

Random-forest

knn-Imputation

2.2.3 Batch-effect detection

Principal component analysis (PCA)

Hierarchical clustering

K-BET

2.3 Differential expression analysis

2.3.1 Limma

2.3.2 DESeq2

2.4 Enrichment analysis

2.4.1 Gene list functional enrichment analysis

2.5 Multi-omics factor analysis (MOFA)

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien

est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

2.6 Clinical data correlation

2.6.1 Linear regression

2.6.2 Binomial regression

2.6.3 Ordinal regression

2.6.4 Multinomial logistic regression

2.6.5 Dirichlet regression

2.7 Benchmarking of Lasso models dealing with missing values

2.7.1 Knowledge guided multi-level network inference

2.7.2 Two-steps based models

Grouped adaptive Lasso (GALasso)

Stacked adaptive Lasso (SALasso)

2.7.3 Inverse covariance based methods

Convexed conditioned Lasso (CoCoLasso)

Lasso with high missing rate (HMLasso)

2.8 Adult data correlation

2.8.1 Random forest

Imbalanced random forest

Nested cross-validation in random forest

2.8.2 t-test, manova, log-reg

3 Hyperoxia-induced cell cycle arrest drives long-term impairment of lung development and DNA repair in neonates

This is the first project of the thesis.

3.1 Summary

3.2 Transcriptional characterization of *in vitro* and *in vivo* mice under hyperoxia and mechanical ventilation

3.2.1 *in vitro* samples

Alveolar type 1/2 (AT1/2)

Endothelial cells (EC)

Myofibroblast (MFB)

3.2.2 *In vivo* samples

3.3 Transcriptional commonalities between *in vitro* samples

3.3.1 Transcriptional commonalities between AT1/2, EC and MFB

3.3.2 Transcriptional commonalities between AT1/2 and MFB

3.4 Transcriptional singularities between *in vitro* samples

3.4.1 Alveolar type 1/2 (AT1/2)

3.4.2 Myofibroblast (MFB)

3.5 Deep transcriptional characterization of cell-cycle arrest under hyperoxia

3.5.1 Pre-replication complex hypothesis

3.5.2 Experimental validation

3.6 Transcriptional comparison between *in vitro* and *in vivo* samples

3.6.1 Multi-systemic development arrest hypothesis

3.7 Discussion

3.7.1 Limitations

3.7.2 Outlooks

4 Multi-omic integration characterizes major endotypes of bronchopulmonary dysplasia linked to adult chronic lung disease

This is the second project of the thesis.

4.1 Summary

4.2 Feature selection based on differential abundance analysis

4.2.1 Proteomics

4.2.2 Metabolomics

4.3 Enrichment analysis of differentially abundant proteins between BPD and no-BPD samples

4.4 Latent factor identification as potential BPD endotypes

4.4.1 Factor 1:

Clinical variable correlation

Systems biology of Factor 1

4.4.2 Factor 2:

Clinical variable correlation

Systems biology of Factor 2

4.4.3 Factor 3:

Clinical variable correlation

Systems biology of Factor 3

4.4.4 Factor 4:

Clinical variable correlation

Systems biology of Factor 4

4.4.5 Factor 5:

Clinical variable correlation

Systems biology of Factor 5

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

4.5 Summary of BPD endotypes

4.5.1 Acute BPD signature

4.5.2 Late/chronic BPD signature

4.6 Correlation between BPD endotypes and chronic adult lung diseases

4.6.1 Use of imbalanced random forest to separate between COPD and IPF patients using BPD latent factor loadings

4.6.2 Characterization of most explanatory features per significant BPD endotype regarding COPD and IPF

Factor 1

Factor 4

Factor 5

4.7 Discussion

4.7.1 Limitations

4.7.2 Outlooks

5 Multi-Omics Regulatory Network Inference in the Presence of Missing Data

This is the third project of the thesis.

5.1 Summary

5.2 Experimental set-up

5.2.1 Data preprocessing

5.2.2 Parameter selection

5.2.3 Missingness simulated scenarios

5.2.4 Performance evaluation criteria

5.3 Performance evaluation over different missingness scenarios

5.3.1 Single-omics random missingness

5.3.2 Multi-omics random missingness

5.3.3 Block-wise random missingness

5.4 General method performance evaluation

5.4.1 Performance evaluation over down-sampling

5.4.2 Computational runtime evaluation

5.5 Discussion

5.5.1 Limitations

5.5.2 Outlooks

6 scNodes: Graphical Interface for Single-Cell Transcriptomics Workflow Development

This is the forth project of the thesis.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

6.1 A section

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

6.2 Another section

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

7 Cytokines Derived from Nasal Epithelial Lining Fluid in Patients with Asthma

The fifth project.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

7.1 A section

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

7.2 Another section

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

8 Discussion

This is the discussion of the thesis.

A Appendix

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.