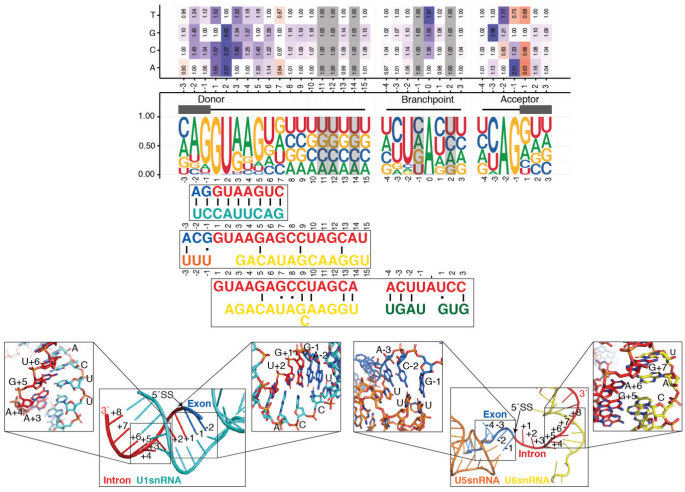
**Selected Topics of Electronics and Micromechatronics**
**Ausgewählte Probleme der Elektronik und Mikromechatronik**

Edited by Gerhard Wachutka, Doris Schmitt-Landsiedel and Gabriele Schrag

Volume 52

Leonhard Wachutka

# Global Donor and Acceptor Splicing Site Kinetics in Human Cells




MUNICH UNIVERSITY OF TECHNOLOGY

Institute for
Physics of Electrotechnology

Institute for
Technical Electronics

Fakultät für Informatik
Technische Universität München

# Global Donor and Acceptor Splicing Site Kinetics in Human Cells

**Leonhard Konrad Friedrich Wachutka**

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

## Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender und Prüfer der Dissertation:
Prof. Dr. Björn Menze
*Technische Universität München*

Prüfende der Dissertation:
1. Prof. Dr. Julien Gagneur
   *Technische Universität München*
2. Prof. Dr. Caroline Friedel
   *Ludwig-Maximilians-Universität München*

Die Dissertation wurde am 27.05.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik 04.08.2020 angenommen.

Selected Topics of Electronics and Micromechatronics
Ausgewählte Probleme der Elektronik und Mikromechatronik

Volume 52

**Leonhard Konrad Friedrich Wachutka**

# Global Donor and Acceptor Splicing Site Kinetics in Human Cells

Shaker Verlag
Düren 2020

**Bibliographic information published by the Deutsche Nationalbibliothek**
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Zugl.: München, Techn. Univ., Diss., 2020



Printed in Germany.

ISBN 978-3-8440-7710-0
ISSN 1618-7539

Shaker Verlag GmbH • Am Langen Graben 15a • 52353 Düren
Phone: 0049/2421/99011-0 • Telefax: 0049/2421/99011-9
Internet: www.shaker.de • e-mail: info@shaker.de

# I. Acknowledgments

So finally this thesis comes to an end. If someone had told me five years ago that this ever would happen, I would not have believed any word. I would like to thank the countless people who made this possible. The people, who gave me hope, when I had no more. The people, who made impossible things possible. The people, who supported me to do it 'my way', although I know it was not the easiest. And all the people I met during the past five years, who inspired me never to give up. I want to say thank you to every single one of these many people, because nothing would have worked out without you.

In addition, I want to say special thanks to Julien Gagneur, my supervisor, who didn't let the faith go in me. He has been a constant source of new ideas and the best boss one can think of. And I thank Livia Caizzi: without her help and her endless patience during the experiments nothing of this would have become possible. I thank Patrick Cramer for all the useful input and support. I thank Carina Demel, who inspired me during my master thesis and who helped me with bioinformatics. I thank Jun Cheng, who jump started the bioinformatics part of this thesis and who always had good ideas in his mind. I thank Christian Mertes and Žiga Avsec, for countless hours of fruitful discussions and their continuous help. And I thank all people of the past and present Gagneurlab. You folks are great!

I want to express my special thanks to my whole family for their unbelievable support, my father who opened up my mind for natural science, my mother who always grounded me, and my lovely brother, who makes my life happier every day. And I want to express my particular thanks to Anna, Tins and Lexi for their endless love. Without love this live wouldn't be worth living.

# II. Abstract

Extensive *in vitro* studies have strongly advanced our understanding of the splicing process, but the kinetics of the successive catalytic steps of splicing *in vivo* remains far less understood. Most existing studies fall short of a sound theoretical interpretation of measured data because of the entangled nature of RNA synthesis, splicing and degradation in standard RNA-Seq protocols.

TT-seq is a labeled RNA sequencing protocol combined with an early RNA fragmentation step that makes it possible to quantify the human RNA metabolism at the resolution of individual phosphodiester bonds. Using Transient Transcriptome sequencing (TT-seq) enabled, for the first time, the genome-wide quantification of the half-lives of individual donor site and acceptor site bonds as well as the rate of formation of splice junctions.

A careful analysis of the kinetic times provided the basis for a systematic search for genetic predictors of fast and slow splicing processes. It turned out that the donor site cleavage time is limited by polymerase elongation and allowed for estimating the interaction effects between several spliceosome components and the RNA at single nucleotide resolution. These findings agree very well with published structural data.

During this study it was necessary to set up solid mathematical models of single bond splicing kinetics to overcome the experimental ambiguities caused by alternative splicing, which are common to massive parallel sequencing methods. This approach, for the first time, allows for introducing and measuring the splicing yield, which is the proportion of precursor RNA successfully converted into spliced RNA.

With a view to facilitating similar analyses for other researchers, the software package rCube has been developed, which opens an easy access to the analysis of TT-seq data to estimate RNA kinetics. It is not only applicable to infer splicing kinetics, but also allows for the high precision measurement of RNA turnover rates.

In summary, this thesis provides conceptual advances in the understanding of RNA kinetics, demonstrates the power of the approach by finding de-novo known and unknown genetic determinants of splicing, and offers a rich source of data for further analysis.

# III. Publications

## Global donor and acceptor splicing site kinetics in human cells

**Leonhard Wachutka**[1]†, Livia Caizzi[2]†, Julien Gagneur[1], Patrick Cramer[2]

[1]Department of Informatics, Technical University of Munich, Garching, Germany; [2]Department of Molecular Biology, Max-Planck-Institute for Biophysical Chemistry, Göttingen, Germany

†These authors contributed equally to this work

(2019) eLife. DOI: 10.7554/eLife.45056. (Wachutka *et al.*, 2019)

Author contributions: **Leonhard Wachutka**, conceptualization, data curation, software, formal analysis, validation, investigation, visualization, methodology, writing - original draft, writing - review and editing, designed and carried out the bioinformatics analysis; Livia Caizzi, conceptualization, data curation, validation, investigation, visualization, methodology, writing - original draft, writing - review and editing, optimized and carried out TT-seq experiments, contributed to the design of the bioinformatics analysis, and used molecular modeling to interpret results; Julien Gagneur, conceptualization, resources, software, supervision, funding acquisition, investigation, visualization, methodology, writing - original draft, project administration, Writing - review and editing; Patrick Cramer, conceptualization, resources, supervision, funding acquisition, investigation, visualization, methodology, writing - original draft, project administration, writing - review and editing author

## Measures of RNA metabolism rates: Toward a definition at the level of single bonds

**Leonhard Wachutka** and Julien Gagneur

Department of Informatics, Technical University of Munich, Garching, Germany

(2017) Transcription. DOI: 10.1080/21541264.2016.1257972. (Wachutka and Gagneur, 2017)

Author contributions: **Leonhard Wachutka**, conceptualization, visualization, writing - original draft, writing - review and editing; Julien Gagneur, conceptualization, resources, supervision, funding acquisition, writing - original draft, project administration, writing - review and editing

## Transient transcriptome sequencing: computational pipeline to quantify genome-wide RNA kinetic parameters and transcriptional enhancer activity

Gabriel Villamil[1]†, **Leonhard Wachutka[2]†**, Patrick Cramer[1], Julien Gagneur[2], Björn Schwalb[1]†

[1]Department of Molecular Biology, Max-Planck-Institute for Biophysical Chemistry, Göttingen, Germany
[2]Department of Informatics, Technical University of Munich, Garching, Germany;

†These authors contributed equally to this work

(2019) bioRxiv. DOI: 10.1101/659912. (Villamil *et al.*, 2019)

Author contributions: TT-seq protocol description, writing: Gabriel Villamil, Patrick Cramer, Björn Schwalb. rCube analysis description, writing: Leonhard Wachutka, Julien Gagneur

## OCR-Stats: Robust estimation and statistical testing of mitochondrial respiration activities using Seahorse XF Analyzer

Vicente A. Yépez[1,2], Laura S. Kremer[3,4], Arcangela Iuso[3,4], Mirjana Gusic[3,4], Robert Kopajtich[3,4], Eliska Koňařıková[3,4], Agnieszka Nadel[3,4], **Leonhard Wachutka**[1], Holger Prokisch[3,4], Julien Gagneur[1,2]

[1] Department of Informatics, Technical University of Munich, Garching, Germany, [2] Quantitative Biosciences Munich, Gene Center, Department of Biochemistry, Ludwig-Maximilians Universität München, Munich, Germany, [3] Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany, [4] Institute of Human Genetics, Klinikum Rechts der Isar, Technical University of Munich, Munich, Germany

(2018) PLoS ONE. DOI: 10.1371/journal.pone.0199938 (Yépez *et al.*, 2018)

Author Contributions:
Conceptualization: Vicente A. Yépez, Laura S. Kremer, Arcangela Iuso, Mirjana Gusic, Robert Kopajtich, Eliska Koňařıková, Agnieszka Nadel, Holger Prokisch, Julien Gagneur.
Data curation: Laura S. Kremer, Arcangela Iuso, Mirjana Gusic, Robert Kopajtich, Eliska Koňařıková, Agnieszka Nadel.
Formal analysis: Vicente A. Yépez, Holger Prokisch, Julien Gagneur.
Investigation: Vicente A. Yépez, Julien Gagneur. Software: Vicente A. Yépez, **Leonhard Wachutka.** Supervision: Holger Prokisch, Julien Gagneur.
Visualization: Vicente A. Yépez, Julien Gagneur. Writing – original draft: Vicente A. Yépez, Holger Prokisch, Julien Gagneur.

*Writing – review & editing*: Vicente A. Yépez, Laura S. Kremer, Arcangela Iuso, Mirjana Gusic, Robert Kopajtich, Eliska Koňařıková, Agnieszka Nadel, **Leonhard Wachutka**, Holger Prokisch, Julien Gagneur.

# IV. Content

*Parts of this thesis have already been published. The respective publications and the contributions of other co-authors are clearly indicated at the beginning of each chapter. The first person plural form 'we' is used throughout the thesis to avoid unnecessary switching between forms 'I' and 'we'.*