

Bayerische Julius–Maximilians–Universität Würzburg



**Systembiologische Analysen von Interaktionen: Zytokine  
(Pflanzenpathogene), 3D-Zellkulturen (Krebstherapie)  
und Drugtargets**

Dissertation zur Erlangung des  
naturwissenschaftlichen Doktorgrades  
der Julius-Maximilians-Universität Würzburg

vorgelegt von

Meik Kunz

aus Rodewisch

Würzburg 2016

Eingereicht am: .....

Mitglieder der Promotionskommission:

Vorsitzender: .....

1. Gutachter: Prof. Dr. Thomas Dandekar

2. Gutachter: Prof. Dr. Heike Walles

Tag des Promotionskolloquiums: .....

Doktorurkunde ausgehändigt am: .....

Wesentliche Teile dieser Arbeit wurden veröffentlicht in:

### **Teil: Die pflanzliche Immunabwehr**

1. Naseem M\*, **Kunz M\***, Dandekar T. Probing the unknowns in cytokinin-mediated immune defense in Arabidopsis with systems biology approaches. *Bioinform Biol Insights*. 2014 Feb 13;8:35-44. doi: 10.4137/BBI.S13462.
2. **Kunz M**, Ahmed N, Dandekar T, and Naseem M. Hormone Signaling Networks Open Multiple Routes for Immunity and Disease in Plants. *Biohelikon: Immunity and Diseases*, 2013 1:2. **Review**.
3. **Kunz M**, Dandekar T and Naseem M. A systems biology methodology combining transcriptome and interactome datasets assess the implications of cytokinin signaling for plant immune networks. *Methods Mol Biol*. 2016, accepted manuscript.

### **Teil: Dynamische Modellierungen im *in silico/in vitro*-Lungentumormodell**

4. Göttlich C\*, Müller LC\*, **Kunz M\***, Schmitt F, Walles H, Walles T, Dandekar T, Dandekar G, Nietzer SL. A Combined 3D Tissue Engineered In Vitro/In Silico Lung Tumor Model for Predicting Drug Effectiveness in Specific Mutational Backgrounds. *J Vis Exp*. 2016 Apr 6;(110). doi: 10.3791/53885.

### **Teil: RNA- und Protein-Interaktionen**

5. **Kunz M**, Xiao K, Liang C, Viereck J, Pachel C, Frantz S, Thum T, Dandekar T. Bioinformatics of cardiovascular miRNA biology. *J Mol Cell Cardiol*. 2015 Dec;89(Pt A):3-10. doi: 10.1016/j.yjmcc.2014.11.027. **Review**.
6. Fiedler J, Breckwoldt K, Remmele CW, Hartmann D, Dittrich M, Pfanne A, Just A, Xiao K, **Kunz M**, Müller T, Hansen A, Geffers R, Dandekar T, Eschenhagen T, Thum T. Development of Long Noncoding RNA-Based Strategies to Modulate Tissue Vascularization. *J Am Coll Cardiol*. 2015 Nov 3;66(18):2005-15. doi:10.1016/j.jacc.2015.07.081.
7. Viereck J, Kumarswamy R, Foinquinos A, Xiao K, Avramopoulos P, **Kunz M**, Dittrich M, Maetzig T, Zimmer K, Remke J, Just A, Fendrich J, Scherf K, Bolesani E, Schambach A, Weidemann F, Zweigerdt R, de Windt LJ, Engelhardt S, Dandekar T, Batkai S, Thum T. Long noncoding RNA Chast promotes cardiac remodeling. *Sci. Transl Med*. 2016 Feb 17;8(326):326ra22.
8. **Kunz M\***, Liang C\*, Nilla S, Cecil A, Dandekar T. The drug-minded protein interaction database (DrumPID) for efficient target analysis and drug development. *Database (Oxford)*. 2016 Apr 7;2016. pii: baw041. doi: 10.1093/database/baw041.

\* authors contributed equally

Weitere im Rahmen der Doktorarbeit erschienene Publikationen:

Wiegering A, **Kunz M**, Hussein M, Klein I, Wiegering V, Uthe FW, Dietz UA, Jurowich C, Bley T, Dandekar T, Germer CT, Ritter C. Diagnostic value of preoperative CT scan to stratify colon cancer for neoadjuvant therapy. *Int J Colorectal Dis.* 2015 Aug; 30(8):1067-73. doi: 10.1007/s00384-015-2265-z.

“I am a firm believer, that without speculation there is no good & original observation.”

Charles Darwin: Letter to A. R. Wallace, 22 December 1857. In F. Burkhardt and S. Smith (eds.), *The Correspondence of Charles Darwin 1844-1846* (1987), Vol. 6, 514.

## Inhaltsverzeichnis

<b>1</b>	<b>Zusammenfassung</b> .....	<b>9</b>
<b>2</b>	<b>Abstract</b> .....	<b>13</b>
<b>3</b>	<b>Einleitung</b> .....	<b>17</b>
3.1	Überblick und Motivation .....	17
3.2	Die pflanzliche Immunabwehr am Beispiel von <i>A. thaliana</i> .....	18
3.3	Lungenkrebs - Allgemeine Grundlagen, Gene und Signalwege .....	21
3.4	Nicht-kodierende RNAs .....	26
3.5	Systembiologische Methoden zur Modellierung und Simulation biologischer Daten .....	32
3.6	Die DrumPID-Datenbank .....	35
3.7	Ziele dieser Arbeit .....	35
<b>4</b>	<b>Material und Methoden</b> .....	<b>37</b>
4.1	Integrierte Netzwerkanalysen zum pflanzlichen Immunsystem von <i>A. thaliana</i> .....	37
4.2	Integrierte Analysen zu Lungenkrebs .....	39
4.2.1	Experimentelle Arraydaten der 17AAG-Behandlung eines <i>in vitro</i> -Experiments .....	39
4.2.2	Systembiologische Analyse der Arraydaten .....	40
4.2.3	<i>In silico</i> -Analyse von Resistenzmechanismen nach einer Gefitinib-Behandlung .....	44
4.3	Integrierte Analysen von lncRNAs .....	45
4.3.1	Systembiologische Untersuchungen zur hypertrophie-assoziierten lncRNA Chast .....	45
4.3.2	Systembiologische Untersuchungen zu den hypoxie-induzierten lncRNAs LINC00323-003 und MIR503HG .....	48
4.4	Analysen zur DrumPID-Datenbank .....	49
<b>5</b>	<b>Ergebnisse</b> .....	<b>52</b>
5.1	Integrierte Netzwerkanalysen am Beispiel von <i>A. thaliana</i> infiziert mit <i>Pst</i> DC3000 .....	52
5.2	Dynamische Modellierung am Beispiel des <i>in silico/in vitro</i> - Lungentumormodells .....	62
5.2.1	Integrierte systembiologische Analysen der experimentellen Arraydaten .....	62
5.2.2	<i>In silico</i> -Analysen der Resistenzmechanismen nach Gefitinib-Behandlung .....	76
5.3	Identifizierung von RNA- und Protein-Interaktionen .....	87
5.3.1	Ergebnisse der Analysen von hypertrophie- und hypoxie-assoziierten lncRNAs .....	87

	5.3.1.1	Herzhypertrophie-assoziierte lncRNA Chast .....	87
	5.3.1.2	Hypoxie-induzierte lncRNAs LINC00323-003 und MIR503HG .....	97
	5.3.2	Die DrumPID-Datenbank .....	100
<b>6</b>	<b>Diskussion</b> .....		<b>109</b>
	6.1	Eine grundlegende systembiologische Methode für verschiedene wissenschaftliche Fragestellungen und biologische Aspekte .....	109
	6.2	Einfluss verschiedener Virulenzfaktoren auf die Pathogenabwehr von <i>A. thaliana</i> .....	110
	6.2.1	<i>Pst</i> DC3000 beeinflusst das zelluläre Wirtsnetzwerk und zeigt, wie sich die zwei Phasen der Immunantwort (PTI und ETI) in <i>A. thaliana</i> überlappen .....	111
	6.2.2	<i>Pst</i> DC3000 reguliert wichtige Hubproteine für die Immunabwehr in <i>A. thaliana</i> .....	114
	6.3	Systembiologische Analysen am Beispiel eines <i>in silico/in vitro</i> - Lungentumormodells .....	118
	6.3.1	Systematische Analyse einer HSP90-Inhibierung .....	118
	6.3.1.1	P53 und STAT zeigen eine starke Vernetzung und haben einen Einfluss auf das unterschiedliche Behandlungsansprechen im 2D- und 3D-System .....	119
	6.3.1.2	Die dynamische Modellierung reflektiert die experimentellen Daten der HSP90-Inhibierung im 3D- System .....	121
	6.3.1.3	Die systembiologischen Analysen zeigen eine individuelle Signatur möglicher prognostischer Marker und Drugtargets für die KRAS-mutierten Zelllinien A549 und H441 .....	121
	6.3.1.4	HIF1A und AMPK als neue therapeutische Targets für die KRAS-mutierten Zelllinien A549 und H441 .....	125
	6.3.2	Systematische Analysen von Resistenzmechanismen gegenüber einer Gefitinib-Behandlung .....	127
	6.3.2.1	Eine kombinierte Inhibierung von PI3K und MEK als möglicher Therapieansatz .....	127
	6.3.2.2	Die Analysen zeigen eine Inhibierung der miRNA-21 und resultierende Aktivierung von PTEN als möglichen Therapieansatz und prognostischen Marker .....	129
	6.3.3	Schlussbetrachtung zu den Analysen des Lungentumormodells .....	131
	6.4	Ein systembiologisches Analyseprotokoll zur systematischen Charakterisierung von lncRNAs .....	132
	6.5	Die DrumPID-Datenbank als nützliche Datenbank für den täglichen wissenschaftlichen Gebrauch auf dem Gebiet der Analyse von Protein- Pharmaka-Interaktionen .....	136

<b>7</b>	<b>Schlussfolgerung und Ausblick .....</b>	<b>139</b>
<b>8</b>	<b>Literaturverzeichnis .....</b>	<b>142</b>
<b>9</b>	<b>Anhang .....</b>	<b>159</b>
9.1	Abkürzungsverzeichnis .....	159
9.2	Verzeichnis der Abbildungen und Tabellen .....	163
9.2.1	Abbildungsverzeichnis .....	163
9.2.2	Tabellenverzeichnis .....	164
9.3	Anhang der Abbildungen und Tabellen .....	165
9.4	Danksagung .....	177
9.5	Ehrenwörtliche Erklärung .....	178



## 1 Zusammenfassung

Der Einsatz von computergestützten Analysen hat sich zu einem festen Bestandteil der biowissenschaftlichen Forschung etabliert. Im Rahmen dieser vorliegenden Arbeit wurden systembiologische Untersuchungen auf verschiedene biologische Themengebiete und Organismen angewendet. In diesem Zusammenhang liefert die Arbeit einen innovativen und interdisziplinären methodischen Ansatz. Die grundlegende Frage lautet: Wie verstehe und beschreibe ich Signalwege und wie kann ich sie beeinflussen? Der Ansatz verknüpft verschiedene biologische Datensätze und Datenebenen miteinander, beginnend vom Genom und Interaktionskontext über semiquantitative Simulationen hin zu neuen Interventionen und Experimenten, welche therapeutisch und biotechnologisch genutzt werden können. Die Analysen können auf diese Weise

- zu einem besseren Verständnis experimenteller Daten und biologischer Fragestellungen beitragen und ermöglichen ein systematisches Verständnis der zugrunde liegenden Signalwege und Netzwerkeffekte (z.B. in Pflanzen).

- Darüber hinaus ermöglichen sie die Identifizierung wichtiger funktioneller Hubproteine und die Entwicklung neuer therapeutischer Strategien für weitere experimentelle Testungen (z.B. Tumormodelle),

- stellen zudem einen hilfreichen Schritt auf dem Weg zur personalisierten Medizin (z.B. lncRNAs und Tumormodelle) und Medikamentenentwicklung (z.B. Datenbank DrumPID) dar.

(i) Als Grundlage wurde hierzu eine integrierte systembiologische Methode entwickelt, welche experimentelle Daten (z.B. Transkriptomdaten) hinsichtlich ihrer biologischen Funktionen untersucht und die Identifizierung relevanter funktioneller Cluster und Hubproteine ermöglicht. In einem **ersten Teil** wurden Analysen zum **pflanzlichen Immunsystem** durchgeführt. Mithilfe der entwickelten Methode wurden Genexpressionsdatensätze von *A. thaliana*, die mit dem Pathogen *Pst* DC3000 infiziert wurden, untersucht, um den Einfluss verschiedener Virulenzfaktoren auf das Interaktom der Wirtspflanze zu untersuchen und neue Modulatoren einer CK-vermittelten Immunabwehr zu finden. In diesem Zusammenhang konnte gezeigt werden, dass die von *Pst* DC3000 sekretierten Abwehrstoffe wichtige pflanzliche Hormonsignalwege für die Immunabwehr in *A. thaliana* beeinflussen. Die Ergebnisse zeigen zudem, dass sich der Einfluss auf das Netzwerkverhalten der Effektorproteine und COR-Phytotoxine von dem der PAMPs unterscheidet, sich jedoch auch eine Regulierung

gemeinsamer Signalwege und eine Überlappung der beiden Phasen der Immunantwort (PTI und ETI) in *A. thaliana* finden lassen. Die komplexe Immunantwort auf eine Infektion spiegelt sich zudem in einer höheren Anzahl an funktionellen Clustern und Hubproteinen in *Pst* DC3000 gegenüber den beiden untersuchten Mutanten wider, wobei sich für *Pst* DC3000 insbesondere ein stark vernetztes immunrelevantes Cluster um den JA-Signalweg zeigt. Weiterhin wurden anhand der entwickelten Methode wichtige Hubproteine für die Immunabwehr identifiziert. Als bedeutende Vertreter sind AHK2 und AAR14 zu nennen, welche Teil des Zweikomponentensystems der Signalübertragung von CK sind und hierbei wichtige Modulatoren für eine CK-vermittelte Immunabwehr darstellen.

(ii) Im **zweiten Teil** der Arbeit schließen sich Untersuchungen an einem *in vitro*-Experiment einer **2D- und 3D-Zellkultur** einer HSP90-Behandlung in einem **Lungentumormodell** an. In diesem Zusammenhang wurden mithilfe der entwickelten Methode Unterschiede zwischen den beiden Zellkultursystemen gefunden, die das unterschiedliche Behandlungsansprechen erklären, und für die beiden KRAS-mutierten Zelllinien A549 und H441 des 3D-Testsystems neue prognostische und therapeutische Kandidaten identifiziert. Hierbei haben die durchgeführten Analysen zwei funktionelle Cluster von Protein-Interaktionen um p53 und die STAT-Familie gefunden, welche eine Verbindung zu HSP90 haben und die entsprechenden Behandlungsunterschiede nach einer HSP90-Inhibierung zwischen den beiden Zellkultursystemen erklären können. Unter Berücksichtigung des zelllinien-spezifischen Mutationshintergrunds wurde eine prognostische Markersignatur und daraus abgeleitet HIF1A für die H441-Zelllinie und AMPK für die A549-Zelllinie als neue therapeutische Targets gefunden, wobei die anschließend durchgeführten *in silico*-Simulationen einen potentiellen therapeutischen Effekt aufzeigen konnten. Weiterhin wurden wichtige experimentelle Readout-Parameter in ein *in silico*-Lungentumormodell integriert, wobei unter Einbeziehung des Mutationshintergrunds für die verwendeten Zelllinien die HSP90-Behandlung des 3D-Testsystems computergestützt abgebildet werden konnte. Im weiteren Verlauf wurden im *in silico*-Lungentumormodell Resistenzmechanismen nach einer Gefitinib-Behandlung mit bekanntem Mutationsstatus für die Zelllinien HCC827 und A549 untersucht und daraus folgend neue Therapieansätze abgeleitet, die von potentieller klinischer Bedeutung sein können. Die durchgeführten *in silico*-Simulationen für HCC827 konnten hierbei zeigen, dass eine EGFR- und c-MET-Koaktivierung zu einer Gefitinib-Resistenz führen kann, wohingegen bei den A549 eine Komutation von KRAS und IGF-1R zu einem geringen Behandlungsansprechen beiträgt.

Die Simulationen lassen zudem erkennen, dass eine direkte Inhibierung der an der Resistenzentwicklung beteiligten Rezeptoren c-MET und IGF-1R in beiden Fällen nicht die bestmögliche Therapiestrategie darstellt. In beiden Zelllinien konnte gezeigt werden, dass eine kombinierte Inhibierung von PI3K und MEK den bestmöglichen therapeutischen Effekt liefert, was demnach einen vielversprechenden Therapieansatz bei Gefitinib-resistenten Lungentumorpatienten darstellt. In einem weiteren Schritt wurde das therapeutische Potential der miRNA-21 im *in silico*-Modell für die HCC827-Zelllinie untersucht. Die durchgeführten Simulationen zeigen, dass eine miRNA-21-Überexpression zu einer Resistenzentwicklung nach Gefitinib-Behandlung beitragen kann, wobei eine Inhibierung der miRNA-21 diesen Effekt umkehren kann. Die Ergebnisse lassen zudem erkennen, dass eine PTEN-Aktivierung als potentieller Marker einer erfolgreichen therapeutischen Inhibierung der miRNA-21 fungieren kann, wohingegen eine reduzierte miRNA-21-Expression als möglicher Marker für eine erfolgreiche Gefitinib-Behandlung dienen kann.

(iii) Im **dritten Teil** der Arbeit wurden systematisch **RNA- und Protein-Interaktionen** untersucht. Hierzu wurden integrierte systembiologische Analysen an neu identifizierten und funktionell bislang unbekanntem lncRNAs durchgeführt. Die Analysen für die infolge einer Herzhypertrophie hochregulierte **lncRNA Chast** haben umfassend gezeigt, dass diese Proteine und Transkriptionsfaktoren regulieren und binden kann, welche die Signalübertragung und Genexpression regulieren, aber auch eine Verbindung zum kardiovaskulären System und stressinduzierter Herzhypertrophie besitzt. Anhand der Ergebnisse lässt sich schlussfolgern, dass Chast direkt und indirekt (a) Proteine binden und die Translation beeinflussen kann, zudem eine Chromatin-modifizierende Funktion besitzt und so die Transkription, z.B. für herz- und stress-assoziierte Gene, reguliert, und/oder (b) in einem negativen Feedbackloop seine eigene Transkription reguliert. Obwohl lncRNAs meist eine geringe Konservierung aufweisen, konnten die durchgeführten Analysen für Chast eine Sequenz-Struktur-Konservierung in Säugetieren aufzeigen. Weiterhin haben die Untersuchungen an zwei hypoxie-induzierten lncRNAs in Endothelzellen gezeigt, dass die **lncRNA MIR503HG** eine hohe Sequenz-Struktur-Konservierung in Säugetieren besitzt, wohingegen die **LINC00323-003** eine geringe Konservierung aufweist. Dies untermauert die Tatsache, dass lncRNAs häufig eine geringe Konservierung aufweisen, was Untersuchungen in Modellorganismen hinsichtlich einer therapeutischen Nutzung schwierig machen.

Da sich zahlreiche Untersuchungen auf Interaktionen und Signalwege konzentriert haben, wurde abschließend eine Datenbank entwickelt, welche Analysen von Protein-Interaktionen und Signalwegen nachhaltig voranbringt. Die entwickelte **DrumPID-Datenbank** stellt insbesondere die **Interaktion zwischen einem Medikament und seinem Target** in den Fokus und ermöglicht Analysen einzelner Interaktionen und beteiligter Signalwege, bietet zusätzlich aber auch verschiedene Links zu anderen Datenbanken für individuelle weiterführende Analysen. DrumPID ermöglicht ein geeignetes Medikament u. a. für ein vorgegebenes Zielprotein zu finden und dessen Wirkmechanismus und Interaktionskontext zu untersuchen, was zu einem besseren experimentellen Verständnis beitragen kann. Zudem erlaubt DrumPID eine potentielle chemische Leitstruktur für ein Zielprotein zu entwickeln, was z.B. spezifisch ein parasitisches Protein inhibiert, ohne dabei einen toxischen Effekt im Menschen zu haben. Zahlreiche weitere Pharmakabeispiele belegen, dass DrumPID für den täglichen wissenschaftlichen Gebrauch auf dem Gebiet der Analyse von Protein-Pharmaka-Interaktionen und der Medikamentenentwicklung geeignet ist.

Die beschriebenen Ergebnisse der Promotionsarbeit wurden in fünf Originalarbeiten, zwei Übersichtsartikeln und einem Buchteil, u. a. in Science Translational Medicine, veröffentlicht, sechs dieser Publikationen erfolgten im Rahmen von Erstautorschaften.

## 2 Abstract

The use of computer-based analysis has become an integral part of life science research. Within this thesis, systems biology investigations have been applied to various biological topics and organisms which provides an innovative and interdisciplinary methodological approach. The basic question was: How do I understand and describe signaling pathways and how can I influence them? The approach combines various biological data sets and data levels starting from the genome and interaction context over semiquantitative simulations towards new interventions and experiments which can be used therapeutically and biotechnologically. The analysis can contribute to

- a better understanding of experimental data and biological questions and enables a systematic understanding of the signaling pathways and network effects (e.g. in plants).
- They enable the identification of important functional hub nodes as well as the development of new therapeutic strategies for further experimental testing (e.g. tumor models),
- also representing a helpful step on the path to personalized medicine (e.g. lncRNAs and tumor models) and drug development (e.g. database DrumPID).

(i) As a basis, an integrated systems biology methodology was developed which examines experimental data sets (e.g. transcriptome data) with respect to their biological functions and enables the identification of relevant functional clusters and hub nodes.

In the **first part** of the thesis, analyzes regarding the **plant immune system** were accomplished. Using the developed methodology, gene expression datasets of *A. thaliana* infected with the pathogen *Pst* DC3000 were analyzed in order to investigate the influence of different virulence factors on the host interactome, and to find new modulators of CK-mediated immune defense. In this context, the analysis could show that the secreted defense compounds of *Pst* DC3000 influence important plant hormone signaling pathways for the immune defense in *A. thaliana*. Moreover, the results show that the impact on the network behavior of the effector proteins and COR phytoalexins differ from the PAMPs, but there also exists an overlap in common regulated signal pathways as well as an overlap between the two phases of immune response (PTI and ETI) in *A. thaliana*. In addition, the complex immune response to an infection is also reflected by a higher number of functional clusters and hub nodes in *Pst* DC3000 compared to the two studied mutants, whereby for *Pst* DC3000 a highly connected immune-relevant cluster around the JA pathway has been found. Furthermore, using the developed methodology several

important hub nodes for the immune defense have been identified. As most important candidates, AHK2 and AAR14 have to be highlighted which are part of the two-component-system of signal transduction of CK and represent in this context important modulators for a CK mediated immune defense.

(ii) In the **second part** of the thesis, analyzes of a HSP90 treatment in **lung cancer** in an *in vitro* experiment in **2D and 3D cell cultures** were accomplished. In this context using the developed methodology, differences between the two cell cultures explaining the differences in treatment responses were found, and for the two KRAS mutated cell lines A549 and H441 of the 3D test system new prognostic marker and therapeutic drug candidates were identified. However, the analyzes found two functional clusters of protein interactions around p53 and the STAT family which have a connection to HSP90 and might explain the observed treatment differences for the HSP90 inhibition between the two cell culture systems. Considering the mutational background of the cell lines, a prognostic marker signature were found and derived from it HIF1A for the H441 cell line and AMPK for the A549 cell line as new therapeutic drug targets. Moreover, the subsequently performed *in silico* simulations could show a potential therapeutic effect of the identified drug targets. Furthermore, important experimental read-out parameters were integrated into the *in silico* lung tumor model, and by considering the mutation background of the used cell lines the HSP90 treatment of the 3D test system could be *in silico* simulated. In the further course of the thesis, resistance mechanisms after gefitinib treatment with known mutation status for the HCC827 and A549 cell lines were investigated in the *in silico* lung tumor model and consequently new therapeutic approaches were derived which may be of potential clinical relevance. Here, the *in silico* simulations for HCC827 cells show that a co-activation of EGFR and c-MET can lead to a gefitinib resistance, whereas in the A549 a co-mutation of KRAS and IGF-1R can contribute to the reduced treatment response. In addition, the simulations reveal that a direct inhibition of the resistance contributing receptors c-MET and IGF-1R reflect not the best treatment strategy in both cases. However, in both cell lines a combined inhibition of PI3K and MEK provides the best therapeutic effect, thus representing a promising new therapeutic approach in gefitinib resistant lung cancer patients. In a further step, the therapeutic potential of the miRNA-21 was examined in the *in silico* model for the HCC827 cells. The simulations show that an overexpression of the miRNA-21 can contribute to a resistance development after gefitinib treatment, in which an inhibition of the miRNA-21 reverses this effect. Moreover, the results show that a PTEN activation can function as a

potential marker of therapeutic success of miRNA-21 inhibition whereas a reduced miRNA-21 expression may serve as a potential marker for a successful gefitinib treatment.

(iii) In the **third part** of the thesis, systematic **RNA and protein interactions** were investigated. For this, integrated systems biology analyzes were carried out on new identified and previously functional unknown lncRNAs. The analyzes of the cardiac hypertrophy caused upregulated **lncRNA Chast** have extensively demonstrated that Chast can regulate and bind proteins and transcription factors which regulate signal transduction and gene expression, but it has also a connection to the cardiovascular system and stress-induced cardiac hypertrophy. Based on the results, it can be concluded that Chast can directly and indirectly (a) bind proteins and influence the translation but also possess a chromatin-modifying function and regulate transcription e.g. for cardiac and stress-associated genes, and/or (b) regulate its own transcription in a negative feedback loop. Although lncRNAs often have a low conservation the analysis could show a sequence-structure-conservation for Chast in mammals. Furthermore, the investigations for two hypoxia induced endothelial lncRNAs have shown that the **lncRNA MIR503HG** represents a high sequence-structure-conservation in mammals, whereas the **LINC00323-003** shows a low conservation. This underscores the fact that lncRNAs often have a low conservation thereby making studies regarding the therapeutic potential in model organisms difficult.

Finally, as numerous analyzes in this thesis have focused on interactions and signaling pathways, a database was developed which brings a sustainable progress in analysis of protein interactions and signaling pathways. The developed **DrumPID database** puts especially the **interaction between a drug and its target** into its focus and allows analysis of individual interactions and involved signaling pathways but, additionally, provides various crosslinks to other databases for individual further analysis. DrumPID enables to find a suitable drug, e.g. for a given target protein, and to analyze its mechanism of action as well as interaction context which can contribute to a better understanding of experimental data. Moreover, DrumPID allows to develop a potential chemical lead structure for a target protein which e.g. specifically inhibits a parasitic protein but has no toxic effect in humans. Numerous additional pharmaceutical examples verify that DrumPID is suitable for the daily scientific usage in the field of analysis of protein-drug-interactions and drug development.

The described results of the doctoral thesis were published in five research papers, two review articles and a book chapter, e.g. in *Science Translational Medicine*, including six first authorships.



### **3 Einleitung**

#### **3.1 Überblick und Motivation**

Der Fortschritt der biowissenschaftlichen Forschung und die immer besser werdenden experimentellen Technologien produzieren täglich riesige Mengen an biologischen Daten, die es gilt, wissenschaftlich auszuwerten, um zu bislang unbekanntem Erkenntnissen zu gelangen. Die Systembiologie hat sich in diesem Zusammenhang zu einem nicht mehr wegzudenkenden Gebiet der heutigen biowissenschaftlichen Forschung etabliert. Sie versucht biologische Fragestellungen computergestützt zu lösen und geeignete Analysemethoden bereitzustellen, um zelluläre Signalkaskaden und biologische Systeme und letztendlich einen Organismus ganzheitlich zu verstehen.

Die vorliegende Arbeit soll einen Einblick und zugleich Ansatzpunkte über die Möglichkeiten liefern, wie systembiologische Analysen interdisziplinär auf verschiedene Fragestellungen und Organismen angewendet werden können. Hierbei werden Analysen an verschiedenen Themengebieten mit hoher biologischer Relevanz durchgeführt. Es finden Untersuchungen zur pflanzlichen Pathogenabwehr in *A. thaliana*, an einem Lungentumormodell und lncRNAs statt, was u. a. relevant hinsichtlich dem Pflanzenschutz und der Nahrungsmittelsicherung, aber auch einer therapeutischen Anwendung im Menschen ist. Dabei gibt es im Wesentlichen drei methodische Schritte: (i) Datenbanken und Server sammeln, integrieren und stellen biologische Daten bereit, welche entsprechend als Datengrundlage genutzt werden. Zusätzlich wurde eine eigene Datenbank entwickelt. (ii) Anhand verschiedener Programme und Software werden Datensätze und Moleküle untersucht, etwa hinsichtlich ihrer Sequenzkonservierung und Interaktionen. (iii) Darauf aufbauend wird ein geeignetes biologisches bzw. systembiologisches Modell entwickelt, was für individuelle wissenschaftliche Fragestellungen genutzt wird.

Die dargestellten systembiologischen Analysen zielen darauf ab, durch die Kombination verschiedener biologischer Prozesse und Datenebenen einen integrierten biologischen Blick zu erhalten, um wichtige zelluläre Signalwege und beteiligte Interaktionen im Detail zu analysieren und zu neuen wissenschaftlichen Einblicken zu gelangen. Hierbei sollen nicht nur Signalnetzwerke besser verstanden, sondern auch neue funktionelle Komponenten und Drugtargets detektiert werden, etwa für potentielle therapeutische Interventionen.

### 3.2 Die pflanzliche Immunabwehr am Beispiel von *A. thaliana*

Pflanzen sind aufgrund ihrer sessilen Lebensweise auf ein ausgeklügeltes und gut funktionierendes Immunsystem angewiesen. Eine allseits anerkannte Beschreibung der pflanzlichen Immunabwehr ist das ZigZag-Modell, welches auf zwei Mechanismen erfolgen kann (Jones and Dangl 2006). In einer ersten Phase erkennen Pflanzen Pathogen-/Mikroben-assoziierte Molekülmuster, sogenannte MAMPs/PAMPs (englisch: Microbial/Pathogen-associated molecular patterns) (Jones and Dangl 2006). Diese werden über entsprechende Rezeptoren (PRRs, englisch: Pattern recognition receptors) in der Pflanze wahrgenommen und lösen eine PAMP-vermittelte Immunabwehr aus (PTI, englisch: PAMP-triggered immunity) (Monaghan and Zipfel 2012) (Zipfel, Kunze et al. 2006) (Jones and Dangl 2006) (Pieterse, Van der Does et al. 2012). Bekannte PAMPs sind Flagellin, Chitin, Glykoproteine und Lipopolysaccharide, wodurch Pflanzen z.B. durch eine schnelle Produktion reaktiver Sauerstoffspezies (englisch: reactive oxygen species), Membrandepolarisation, Calciumsignaturen, Aktivierung von Mitogen-aktivierten Proteinkinasen (MAPK) und der Induktion von Abwehrgenen, wie etwa WRKY29, WRKY30, FRK1 und PR1 (englisch: Pathogenesis-related protein 1), eine PTI auslösen (Monaghan and Zipfel 2012) (Zipfel, Kunze et al. 2006) (Asai, Tena et al. 2002) (Chinchilla, Zipfel et al. 2007). Die PTI stellt eine erste basale pflanzliche Immunabwehr dar, die kurzfristig und effektiv Pathogene abwehrt (Asai, Tena et al. 2002) (Chinchilla, Zipfel et al. 2007).

Um Pflanzen trotz alledem erfolgreich zu infizieren, haben Pathogene entsprechende Strategien entwickelt (Gohre and Robatzek 2008) (Kazan and Lyons 2014). Sie induzieren Effektorproteine in die Wirtszelle, um die pflanzliche Immunabwehr zu überwinden, was auch als ETS (englisch: Effector-triggered susceptibility) bezeichnet wird (Jones and Dangl 2006) (Gohre and Robatzek 2008) (Sarris, Duxbury et al. 2015). Bekannte Effektoren sind z.B. AvrE und HopM1, welche direkt die Salicylsäure (SA)-vermittelte Immunabwehr unterdrücken (DebRoy, Thilmony et al. 2004) (Kazan and Lyons 2014). Diese Effektorproteine haben sich Pflanzen ihrerseits zu Nutze gemacht. Intrazelluläre Rezeptoren (NB-LRR) erkennen die Effektoren und lösen eine hypersensitive Reaktion aus, die zum Zelltod und zur Pathogenabwehr führt, wobei man in diesem Zusammenhang von einer Effektor-vermittelten Immunabwehr (ETI, englisch: Effector-triggered immunity) spricht (Jones and Dangl 2006) (Gohre and Robatzek 2008) (Pieterse, Van der Does et al. 2012) (Sarris, Duxbury et al. 2015) (Cui, Tsuda et al. 2015). Während die PTI eine basale Immunantwort gegenüber einer Vielzahl von Molekülen in verschiedenen Mikroorganismen und auch Nicht-Pathogenen ist, stellt die

ETI eine robuste Immunreaktion infolge pathogener Virulenzfaktoren dar (Gohre and Robatzek 2008) (Cui, Tsuda et al. 2015) (Jones and Dangl 2006).

Das ZigZag-Modell liefert eine konzeptionelle Darstellung der gegenseitigen Erkennung zwischen Wirt und Pathogen und beschreibt weitestgehend die Abwehrmechanismen, die um Überleben oder Tod der Pflanze entscheiden. Es liefert allerdings keine vollständigen Aussagen über beteiligte Interaktionen, gemeinsame Überlappungen zwischen den beiden Phasen (PTI und ETI) und erlaubt auch keine Auskunft über bislang unbekannte Vorgänge während der pflanzlichen Immunabwehr (Cui, Tsuda et al. 2015) (Dong, Jiang et al. 2015). Die pflanzliche Immunabwehr ist zudem ein dynamischer Prozess, wobei die beteiligten Signalwege über mehrere Stufen zusammenarbeiten, etwa Transkriptom und Metabolom (Jones and Dangl 2006) (Sarris, Duxbury et al. 2015) (Pieterse, Leon-Reyes et al. 2009) (Naseem, Philippi et al. 2012). Es konnte ebenfalls gezeigt werden, dass Effektorproteine mit zahlreichen Proteinen interagieren und einen direkten Einfluss auf das pflanzliche Interaktionsnetzwerk haben (Tully, Hill et al. 2014) (Mukhtar, Carvunis et al. 2011). Dies verdeutlicht ein komplexes Zusammenspiel einzelner Komponenten und Signalwege für eine gut funktionierende pflanzliche Immunabwehr.

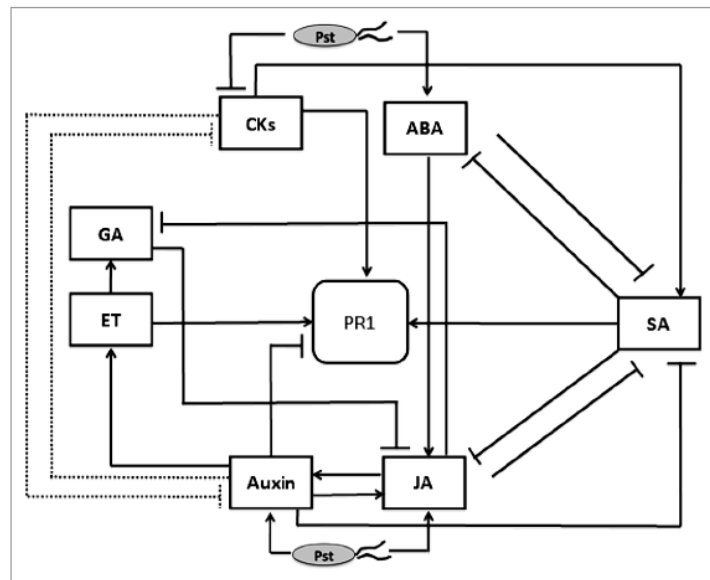
Hierfür verantwortlich sind Hormone und deren Interaktionen (Abb. 3.2). Hormone spielen in Pflanzen neben einer Immunabwehr eine essentielle Rolle in zahlreichen Funktionen, etwa für Wachstum, Entwicklung und Reproduktion (Hwang, Sheen et al. 2012) (Kunz M 2013). Die Grundlage der pflanzlichen Immunabwehr ist hierbei eine antagonistische Interaktion zwischen SA und Jasmonsäure/Ethylen (JA/ET), wobei SA über PR1 eine entsprechende Immunantwort induziert (Robert-Seilaniantz, Grant et al. 2011) (Grant and Jones 2009) (Caarls, Pieterse et al. 2015) (Abb. 3.2). Weitere an der pflanzlichen Immunabwehr beteiligte Hormone sind Auxin (AUX), Abscisinsäure (ABA) und Gibberellinsäure (GA). AUX inhibiert PR1 (Kazan and Manners 2009) (Robert-Seilaniantz, Grant et al. 2011) (Grant and Jones 2009), des Weiteren wirken AUX und ABA inhibierend auf SA und aktivierend auf JA, wobei letzteres wiederum von GA inhibiert wird (Anderson, Badruzsaufari et al. 2004) (Mohr and Cahill 2007) (Yang, Yao et al. 2012) (Naseem, Kunz et al. 2013) (Abb. 3.2). Eine weitere Schlüsselrolle kommt dem Hormon Zytokinin (im Folgenden als Cytokinin (CK) bezeichnet) zu. CKe beeinflussen die pflanzliche Immunabwehr über eine Verstärkung von SA und PR1 und können zu einer erhöhten Resistenz beitragen (Naseem, Kunz et al. 2014). Weiterhin wird davon ausgegangen, dass eine antagonistische Interaktion zwischen CK und AUX mit einem gegensätzlichen Effekt auf die pflanzliche Immunabwehr nach einer Infektion mit *Pseudomonas syringae* pv. *tomato*

DC3000 (*Pst* DC3000) besteht (Naseem, Kunz et al. 2013) (Naseem, Philippi et al. 2012) (Abb. 3.2). Während die Funktionen von JA und SA in der Immunabwehr umfangreich bekannt ist, wurde die Rolle von CK, etwa hinsichtlich der Interaktion mit JA und AUX, bislang noch nicht vollständig geklärt (Naseem, Kunz et al. 2014) (Kunz M 2013).

Die pflanzliche Signalübertragung von CK erfolgt über ein Zweikomponentensystem. Histidinkinasen (AHKs, englisch: Arabidopsis Histidine Kinases) erkennen und phosphorylieren CK, entsprechende Histidin-Phosphotransfer-Proteine (AHPs, englisch: Arabidopsis Histidine Phosphotransfer Proteins) übertragen dieses anschließend an B-type-ARRs (englisch: Arabidopsis Response Regulators), welche ihrerseits A-type-ARRs aktivieren. B-type-ARRs fungieren dabei als positive CK-Regulatoren, wohingegen A-type-ARRs das CK-Signal reprimieren (Naseem, Kunz et al. 2013) (Naseem and Dandekar 2012) (Kunz M 2013). In diesem Zusammenhang konnte gezeigt werden, dass ARR2, ein B-type-ARRs, über TGA1A-related gene 3 (TGA3) mit dem SA-Signalweg interagiert und zu einer erhöhten Resistenz gegenüber *Pst* DC3000 beitragen kann, wohingegen A-type-ARRs eine entgegengesetzte Regulierung zur SA-vermittelten Immunabwehr aufzeigen (Choi, Huh et al. 2010) (Choi, Choi et al. 2011) (Naseem, Philippi et al. 2012).

Pflanzliche Pathogene werden anhand ihrer Lebensweise in drei Gruppen unterteilt: Nekrotrophe Pathogene töten ihre Wirtszelle und ernähren sich von dieser, wohingegen biotrophe Pathogene ihre Nährstoffe von lebenden Zellen beziehen. Daneben existieren noch hemibiotrophe Pathogene, welche nach einer anfänglichen biotrophen Lebensweise letztendlich die befallenen Zellen töten und sich hiervon ernähren. Somit werden hemibiotrophe Pathogene als eine Zwischenform beider Gruppen angesehen (Glazebrook 2005) (Pieterse, Leon-Reyes et al. 2009) (Pieterse, Van der Does et al. 2012) (Kunz M 2013). Ein gut untersuchtes hemibiotropes Pathogen in *Arabidopsis thaliana* (*A. thaliana*) und Tomaten ist *Pst* DC3000 (Xin and He 2013) (Zwiesler-Vollick, Plovanich-Jones et al. 2002), auf dessen experimenteller Untersuchung auch das ZigZag-Modell beruht (Jones and Dangl 2006). *Pst* DC3000 infiziert seine Wirtspflanze u. a. durch ein aus mehreren Effektorproteinen bestehendes Proteinsekretionssystem (Typ III), welches von den Genen *HrpS* und *HrpL* (hypersensitive response and pathogenicity hrp) reguliert wird (Zwiesler-Vollick, Plovanich-Jones et al. 2002) (Roine, Wei et al. 1997). Neben Effektorproteinen sekretiert *Pst* DC3000 auch Phytotoxine, etwa Coronatin (COR; MAMPs), welches die Wirkung von JA nachahmt. COR interagiert mit dem JA-Rezeptor Coronatine-insensitive 1 (COI1) und sorgt so für eine Aktivierung von JA in der Wirtszelle, was die SA-vermittelte Immunantwort unterdrückt und die Anfälligkeit der

Wirtspflanze für eine Infektion erhöht (Geng, Cheng et al. 2012) (Kazan and Lyons 2014) (Kazan and Manners 2013). Aufgrund ihrer Fähigkeiten, etwa Hormone der Wirtszelle zu imitieren oder aber Signalwege und den Stoffwechsel zu beeinflussen, spielen die pathogenen Abwehrstoffe PAMPs, Effektorproteine und Phytotoxine eine wichtige Rolle in der pflanzlichen Immunabwehr (Kazan and Lyons 2014) (Attaran, Major et al. 2014).



**Abbildung 3.2: Überblick über die pflanzliche Immunabwehr.** Hormon-Interaktionen regulieren die pflanzliche Immunabwehr nach einer pathogenen Infektion, hier exemplarisch dargestellt für *Pst* DC3000. Eine antagonistische Interaktion zwischen SA und JA/ET stellt das Grundgerüst der Immunabwehr dar, wobei PR1 eine entsprechende Immunreaktion induziert. Weitere wichtige Hormone sind AUX, ABA und GA, die sich inhibierend und/oder aktivierend auf SA und JA auswirken. Eine weitere Schlüsselrolle spielt CK, welches die pflanzliche Immunabwehr und Resistenz gegenüber Pathogenen erhöhen kann. (→) Aktivierung, (–) Inhibierung. (Abbildung wurde in Naseem, Kunz et al. (2013) veröffentlicht.)

### 3.3 Lungenkrebs - Allgemeine Grundlagen, Gene und Signalwege

Lungenkrebs hat mit weltweit jährlich 1,59 Millionen Todesfällen (353.000 in Europa; Stand 2012) die höchste Sterblichkeitsrate aller Tumorerkrankungen (Ferlay, Steliarova-Foucher et al. 2013) (Ferlay J Available from: <http://globocan.iarc.fr>) (Stahlhut and Slack 2015) (Herbst, Heymach et al. 2008) (Boeri, Pastorino et al. 2012). Lungenkrebs wird in kleinzellige (SCLC, englisch: Small Cell Lung Cancer) und nicht-kleinzellige Lungentumore (NSCLC, englisch: Non-Small Cell Lung Cancer) klassifiziert. NSCLC ist mit 85% die häufiger vorkommende Form und wird zusätzlich in Plattenepithelkarzinome (englisch: squamous cell carcinoma),

großzellige Karzinome (englisch: large cell carcinoma) und Adenokarzinome (englisch: adenocarcinoma) unterteilt, wobei letztere die dominante NSCLC-Form ist (Stahlhut and Slack 2015) (Herbst, Heymach et al. 2008) (Minna, Roth et al. 2002) (Sekido, Fong et al. 2003) (Shigematsu, Lin et al. 2005) (Heegaard, Schetter et al. 2012) (Ramalingam, Owonikoko et al. 2011) (Landi, Zhao et al. 2010) (Tang, Shigematsu et al. 2005) (WD. Travis 2004). Die Fünf-Jahres-Überlebensrate nach Diagnose liegt bei unter 15%, was hierbei primär an einer zu späten Diagnose und resultierendem fortgeschrittenem Tumorstadium liegt. Zum Vergleich hierzu liegt die Fünf-Jahres-Überlebensrate bei einer Diagnose in Stadium 2 bei ca. 33% und in Stadium 1 bereits bei 60 - 70% (Stahlhut and Slack 2015) (Herbst, Heymach et al. 2008) (Minna, Roth et al. 2002) (Heegaard, Schetter et al. 2012) (Landi, Zhao et al. 2010) (Tang, Shigematsu et al. 2005) (van Rens, de la Riviere et al. 2000).

Die aktuellen diagnostischen Verfahren beruhen auf einer nicht-invasiven Computertomographie oder Positronen-Emissions-Tomographie/Computertomographie, gefolgt von einer invasiven Bronchoskopie und Biopsie (Heegaard, Schetter et al. 2012) (Coate, John et al. 2009) (Boeri, Pastorino et al. 2012). Die hohe Falsch-Positiv-Rate und die damit verbundenen hohen Kosten machen Computertomographie-Untersuchungen als alleinige nicht-invasive diagnostische Tests unbrauchbar (Sozzi, Boeri et al. 2014) (Montani, Marzi et al. 2015). Neue diagnostische Verfahren konzentrieren sich auf nicht-invasive Bluttests, wobei erste Tests basierend auf microRNA-Signaturen als ergänzender diagnostischer Test zu den bildgebenden Computertomographie-Verfahren bereits erfolgversprechende Ergebnisse aufzeigen (Sozzi, Boeri et al. 2014) (Bianchi, Nicassio et al. 2011) (Montani, Marzi et al. 2015). Trotz intensiver Bemühungen befindet sich jedoch aktuell kein geeigneter Bluttest in klinischer Anwendung.

Die Standardbehandlung von Lungentumoren beruht auf einer Platin-basierten Chemotherapie oder zielgerichteten Therapien, etwa Gefitinib (EGFR-Inhibitor), wobei diese aber oftmals einen geringen klinischen Erfolg und teils auch starke toxische Nebenwirkungen in Patienten zeigen (Coate, John et al. 2009). Erschwerend kommt hinzu, dass sich aufgrund von Begleitmutationen (auch als Komutationen bezeichnet) entsprechende Resistenzen gegenüber aktuell verfügbaren Behandlungen entwickeln.

### ***in silico/in vitro*-Lungentumormodell**

Wie bereits erwähnt, besitzt Lungenkrebs eine hohe Sterblichkeitsrate, was häufig auf eine zu späte Diagnose und Behandlungsresistenzen zurückzuführen ist (Ferlay J Available from: <http://globocan.iarc.fr>, accessed on 13/08/2015.) (Stahlhut and Slack 2015) (Herbst, Heymach et al. 2008). Die Entwicklung einer neuen klinisch zugelassenen Therapie ist ein langwieriger und mit enormen Kosten verbundener Prozess. Ein generelles Problem hinsichtlich neuer Therapieoptionen ist, dass sich bei neu entwickelten Tumorbehandlungen oftmals nach erfolgreicher vorklinischer Testung (z.B. in Zellkultur) eine geringe Wirksamkeit in den anschließenden klinischen Studien zeigt, was zu einer Untersagung der klinischen Zulassung führt (Bhattacharjee 2012) (Kola and Landis 2004) (Morgan, Van Der Graaf et al. 2012) (Arrowsmith and Miller 2013) (Gottlich, Muller et al. 2016) (Stratmann, Fecher et al. 2014) (Arrowsmith 2011). Ein Grund hierfür sind häufig mangelnde vorklinische Testsysteme, welche meist nur ungenau die tatsächliche Situation und Wirksamkeit im Patienten widerspiegeln. So zeigen 2D-Zellkultursysteme nicht die komplexen Tumorumgebungen und -gewebe, wohingegen Tiermodelle dahingehend zwar komplexer organisiert sind, sich aber andererseits oftmals auch Behandlungsunterschiede zwischen Mensch und Maus zeigen (Tursz, Andre et al. 2011) (Gottlich, Muller et al. 2016) (Nietzer, Baur et al. 2016) (Moll, Reboredo et al. 2013) (Pampaloni, Reynaud et al. 2007) (Hartung 2009). Demnach repräsentieren diese Testsysteme teils nur ungenau die Situation in Patienten, was neue Test- und Analyseverfahren notwendig macht. Um die Lücke zwischen 2D-Zellkultursystemen und komplexen *in vivo*-Tumoren zu schließen, wurden in den letzten Jahren 3D-Tumortestsysteme (Tissue Engineering) entwickelt, welche auf unterschiedlichen Ansätzen beruhen (Schanz, Pusch et al. 2010) (Stratmann, Fecher et al. 2014) (Moll, Reboredo et al. 2013) (Gottlich, Muller et al. 2016) (Nietzer, Baur et al. 2016) (Pampaloni, Reynaud et al. 2007). Ein solches neuartiges *in vitro*-3D-Zellkultursystem basiert beispielsweise auf einer dezellularisierten Gewebematrix aus porcinem Jejunum, Schweinedünndarm (BioVaSc®, Biological Vascularized Scaffold), welcher zur Kultivierung von Zellen dient, der sogenannten SISmuc (Small Intestinal Submucosa mit erhaltener Mucosa) (Linke, Schanz et al. 2007). Diese *in vitro*-3D-Modelle verfügen u. a. über Gewebestrukturen und Komponenten der extrazellulären Matrix (z.B. Kollagene und Basalmembranstrukturen), was eine bessere Darstellung des Tumorgewebes und Abbildung therapeutischer Behandlungen erlaubt (Linke, Schanz et al. 2007). Es konnte in diesem Zusammenhang bereits gezeigt werden, dass ein auf dieser Methode basierendes 3D-

Lungentumormodell z.B. eine geringere Proliferationsrate, ein wichtiger Parameter zur Beurteilung der Wirksamkeit von Medikamenten, aber auch weitere Signalkomponenten (z.B. Rezeptoren) gegenüber herkömmlichen 2D-Zellkultursystemen zeigt, was dabei Medikamententestungen ähnlich den *in vivo*-Tumorbedingungen möglich macht (Stratmann, Fecher et al. 2014) (Gottlich, Muller et al. 2016). In Kombination mit einem *in silico*-Tumormodell, welches wichtige Tumorsignalwege und deren Verschaltung abbildet, kann hier zusätzlich ein schnelles und effizientes Medikamentenscreening sowie die Entwicklung und Vorhersage von neuen Therapiestrategien, etwa nach Resistenzentwicklung, erreicht werden, was bereits anhand weniger experimenteller Daten erfolgen kann (Stratmann, Fecher et al. 2014) (Gottlich, Muller et al. 2016). Aus diesem Grund wurden in dieser Arbeit Untersuchungen an einem kombinierten *in silico/in vitro*-Lungentumormodell durchgeführt, wobei Zelllinien mit einer Deregulierung von KRAS und EGFR verwendet wurden. Im Folgenden sollen wichtige Gene und Signalwege bei Lungenkrebs vorgestellt werden.

### **Wichtige Gene und Signalwege bei Lungenkrebs**

Lungentumore zeigen eine Deregulierung zahlreicher Gene, welche mit wichtigen biologischen Signalwegen, wie etwa für Tumorwachstum, Metastasierung oder Überleben, assoziiert sind (Herbst, Heymach et al. 2008) (Minna, Roth et al. 2002) (Sekido, Fong et al. 2003) (Sun, Schiller et al. 2007). Im Folgenden soll auf die drei wichtigen Gene *KRAS*, *EGFR* und *p53* bei NSCLC-Patienten näher eingegangen werden.

Das Tumor-Protein P53 (p53, auch TP53 genannt) ist in 50 - 70% aller NSCLC-Patienten dereguliert. Dieses Tumorsuppressorgen reguliert u. a. p21 und BAX und spielt eine wichtige Rolle im Zellzyklus und der Apoptose. Es zeigt sich zudem, dass eine Deregulierung von p53 mit einer geringen Apoptose und damit verbundenen erhöhten Tumorüberlebensrate korreliert (Stahlhut and Slack 2015) (Herbst, Heymach et al. 2008) (Minna, Roth et al. 2002) (Sekido, Fong et al. 2003) (Sun, Schiller et al. 2007) (Fong, Sekido et al. 2003).

Das Kirsten Rat Sarcoma Viral Oncogene Homolog (KRAS, zu 15 - 20% mutiert in NSCLC) und der Epidermal Growth Factor Receptor (EGFR, zu 15 - 30% dereguliert in NSCLC) sind beide spezifisch dereguliert in NSCLC und beeinflussen wichtige Signalwege für die Proliferation und Apoptose. KRAS wird aktiviert durch EGFR und reguliert den RAF/MEK/ERK-Signalweg, welcher mit Proliferation, Tumortransformation und -progression assoziiert ist. Es zeigt sich, dass *KRAS* häufig in frühen Stadien von NSCLC mutiert und dabei



verantwortlich für Behandlungsresistenzen ist, etwa für den EGFR-Antikörper Cetuximab und Chemotherapeutika (Stahlhut and Slack 2015) (Herbst, Heymach et al. 2008) (Minna, Roth et al. 2002) (Pao, Wang et al. 2005) (Eberhard, Johnson et al. 2005) (Fong, Sekido et al. 2003) (Riely, Kris et al. 2008). KRAS ist diesbezüglich ein interessantes therapeutisches Target. Entsprechend entwickelte Therapeutika, wie etwa der Farnesyltransferase-Inhibitor oder das Ras-Umwandlungs-Enzym 1 (RCE1, englisch: Ras-converting enzyme-1), zeigen aber allesamt eine geringe klinische Wirksamkeit (Fong, Sekido et al. 2003) (Baines, Xu et al. 2011) (Mortazavi, Lu et al. 2015) (Adjei 2001).

Wie bereits erwähnt, zeigen Lungentumore häufig eine Überexpression der Rezeptor-Tyrosinkinase (RTK, englisch: Receptor Tyrosine Kinase) EGFR. EGFR steuert die Signalwege KRAS/RAF/MEK/ERK und PI3K/AKT, die wichtige Prozesse wie Apoptose und Proliferation regulieren, aber auch eine Tumorentstehung, Angiogenese und Metastasierung begünstigen. Es ist zudem mit einer verminderten Überlebensrate und Behandlungsresistenzen, etwa gegenüber Chemotherapeutika, assoziiert (Stahlhut and Slack 2015) (Herbst, Heymach et al. 2008) (Sekido, Fong et al. 2003) (Shigematsu, Lin et al. 2005) (Tang, Shigematsu et al. 2005) (Marchetti, Martella et al. 2005) (Woodburn 1999) (Wakeling, Guy et al. 2002) (Nicholson, Gee et al. 2001). Verschiedene Therapieansätze fokussieren sich auf EGFR, etwa die Tyrosin-Kinase-Inhibitoren (TKIen) Gefitinib und Erlotinib. Diese werden als Ersttherapeutikum bei einer EGFR-Überexpression oder aber als Zweit- oder Dritttherapeutikum bei Patienten mit entsprechenden Behandlungsresistenzen gegenüber Chemotherapeutika verwendet (Stahlhut and Slack 2015) (Suzuki, Nakagawa et al. 2003) (Fukuoka, Yano et al. 2003) (Kris, Natale et al. 2003) (Herbst and Sandler 2008). Jedoch zeigt sich oftmals nach einem anfänglichen Therapieansprechen nach gewissem Zeitablauf eine Behandlungsresistenz aufgrund zusätzlich existierender Begleitmutationen, etwa KRAS, MET Proto-Oncogene (c-MET) oder Insulin-Like Growth Factor 1 Receptor (IGF-1R) (Stahlhut and Slack 2015) (Suzuki, Nakagawa et al. 2003) (Engelman, Zejnullahu et al. 2007) (Morgillo, Kim et al. 2007).

Dies verdeutlicht die Notwendigkeit neuer diagnostischer und therapeutischer Verfahren zur Behandlung von Lungenkrebs, die sowohl eine Früherkennung als auch bessere Behandlungsaussichten ermöglichen. Aus diesem Grund wird ein enormer Forschungsaufwand betrieben, um geeignete diagnostische und prognostische Marker sowie neue Therapieansätze zu finden. Auf eine dieser neuen Möglichkeiten soll im Folgenden eingegangen werden.

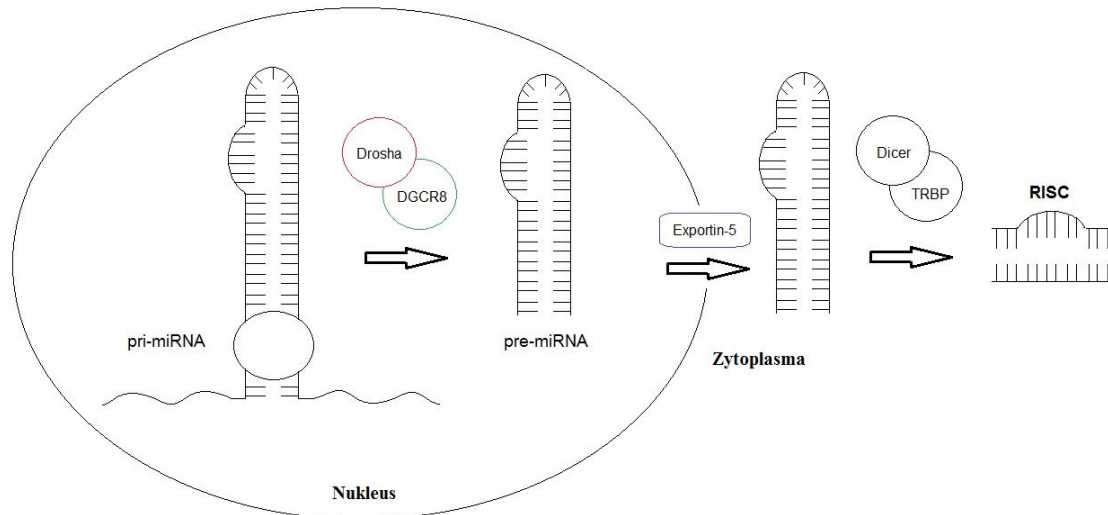
### 3.4 Nicht-kodierende RNAs

Die Identifizierung und Charakterisierung funktioneller nicht-kodierender Ribonukleinsäuren (ncRNAs, englisch: noncoding Ribonucleic Acids) hat zu einem enormen Fortschritt auf dem Gebiet der klinischen Forschung beigetragen (Fiedler, Breckwoldt et al. 2015). Sie sind an der Regulation wichtiger biologischer Prozesse und Signalwege beteiligt, etwa für Zellwachstum und Pathogenese, und mit verschiedenen Krankheiten, z.B. Tumor- und Herzerkrankungen, assoziiert. Diesbezüglich haben zahlreiche Studien gezeigt, dass sich ncRNAs aufgrund ihrer Eigenschaften und genregulatorischen Funktionen besonders als diagnostisches und therapeutisches Werkzeug eignen. ncRNAs werden anhand ihrer Transkriptlänge in kleine (<200 Nukleotide, Nte) und lange (>200 Nte) Ribonukleinsäuren (RNAs, englisch: Ribonucleic Acids) unterteilt (Bell, Long et al. 2014). Im Folgenden soll auf miRNAs und lncRNAs eingegangen werden, worauf sich auch ein Teil der vorliegenden Arbeit konzentriert (Resultate Punkt 5.2.2 und 5.3.1).

#### **Kleine nicht-kodierende microRNAs**

MicroRNAs (miRNAs) zeigen eine Sequenzlänge von 22 - 25 Nten und sind hoch konserviert zwischen verschiedenen Organismen. Sie liegen im Genom in intronischen (Introns) und intergenischen Regionen kodierender und nicht-kodierender Gene und werden einzeln oder in Clustern exprimiert (Kunz, Xiao et al. 2014) (Kayani, Kayani et al. 2011) (Erson and Petty 2009) (Jansson and Lund 2012) (Le Quesne and Caldas 2010). miRNAs werden im Zellkern (Nukleus) von einer RNA-Polymerase II als primäre miRNA (pri-miRNA, englisch: primary-miRNA) transkribiert (Abb. 3.4). Die pri-miRNA wird anschließend von dem Ribonuklease (RNase)-III-Enzym Drosha und dessen Kofaktor DiGeorge Syndrome Critical Region 8 (DGCR8) in eine ca. 70 Nte lange Vorläufer miRNA (pre-miRNA, englisch: precursor-miRNA) geschnitten, welche eine charakteristische Haarnadelstruktur (englisch: Stem-loop) besitzen (Abb. 3.4) (Kunz, Xiao et al. 2014) (Erson and Petty 2009) (Jansson and Lund 2012). Die pre-miRNA wird durch das Exportprotein Exportin-5 in das Zytoplasma transportiert, wo es von dem Enzym Dicer und dessen RNA-Bindungspartner TRBP (Trans-Activation-Responsive RNA-Binding Protein) in die ca. 22 Nte lange doppelsträngige miRNA, bestehend aus einem Führungs- und Passagierstrang (englisch: guide strand, passenger strand), geschnitten und zur reifen miRNA (guide strand) weiter prozessiert wird (Abb. 3.4). Hierbei wird in der Regel der thermodynamisch stabilere Passagierstrang (passenger strand) entfernt,

wohingegen der weniger stabile Strang an der 5'-untranslatierten-Region (5'-UTR) von dem Protein Argonaute 2 (AGO2) gebunden wird und einen Komplex mit RISC (englisch: RNA-inducible silencing complex) bildet, welcher letztendlich die genregulatorische Funktion ausübt (Kayani, Kayani et al. 2011) (Erson and Petty 2009) (Jansson and Lund 2012) (Le Quesne and Caldas 2010) (Kunz, Xiao et al. 2014) (Abb. 3.4).



**Abbildung 3.4: Biogenese von miRNAs.** miRNAs werden im Zellkern als pri-miRNA (links im Nukleus) transkribiert und von dem RNase-III-Enzym Drosha (rot umrandet) und dessen Kofaktor DGCR8 (grün umrandet) in eine ca. 70 Nte lange pre-miRNA (rechts im Nukleus) geschnitten. Die pre-miRNA mit der charakteristischen Haarnadelstruktur wird anschließend durch Exportin-5 (schwarzes Rechteck) in das Zytoplasma transportiert (links im Zytoplasma) und von dem Enzym Dicer und dessen Bindungspartner TRBP (jeweils schwarz umrandet im Zytoplasma) in die ca. 22 Nte lange doppelsträngige miRNA geschnitten, welche weiter zur reifen miRNA prozessiert wird (hier nicht gezeigt). Die miRNA wird von AGO2 gebunden (hier nicht gezeigt) und bildet einen Komplex mit RISC (rechts im Zytoplasma), welcher funktionell an mRNAs bindet. (Details siehe Text.)

miRNAs binden mit ihrer konservierten seed-Region (ca. sieben Nte lang) im Zytoplasma an die 3'-UTR komplementäre messenger-RNA (mRNA) und regulieren wichtige zelluläre Signalwege (Kunz, Xiao et al. 2014). Es zeigt sich dabei, dass eine unvollständige Sequenzübereinstimmung zu einer mRNA-Inhibierung und Herunterregulierung der Translation führt, wohingegen eine vollständige Sequenzübereinstimmung in einem mRNA-Abbau und einer Translationsinhibierung resultiert (Landi, Zhao et al. 2010) (Kunz, Xiao et al. 2014) (Erson and Petty 2009) (Le Quesne and Caldas 2010) (Kayani, Kayani et al. 2011) (Jansson and Lund 2012) (Lewis, Shih et al. 2003). miRNAs können meist an mehrere mRNAs binden (unvollständige Sequenzübereinstimmung teils über 100 Targets) und regulieren so die

Genexpression einer Vielzahl menschlicher Gene (ca. 30 - 60% eukaryotischer Gene) (Landi, Zhao et al. 2010) (Kunz, Xiao et al. 2014) (Erson and Petty 2009) (Le Quesne and Caldas 2010) (Kayani, Kayani et al. 2011) (Jansson and Lund 2012) (Friedman, Farh et al. 2009). Im Hinblick auf Lungenkrebs konnte gezeigt werden, dass miRNAs entweder als Tumorsuppressoren oder aber als Onkogene fungieren und wichtige Prozesse und Signalwege, z.B. für die Tumorentstehung, Metastasierung, Invasion und Resistenzentwicklung, beeinflussen, aber auch mit dem Tumorstadium, Behandlungserfolg und der Überlebensrate korrelieren (Bjaanæs, Halvorsen et al. 2014) (Yu, Chen et al. 2008) (Yanaihara, Caplen et al. 2006) (Takamizawa, Konishi et al. 2004) (Tang, Shen et al. 2013) (Charkiewicz, Pilz et al. 2015) (Stahlhut and Slack 2015) (Kasinski and Slack 2012) (Dong, Ren et al. 2015) (Okayama, Saito et al. 2013). Dies macht sie zu einem nützlichen therapeutischen Hilfsmittel der posttranskriptionellen Beeinflussung der Genregulierung, aber auch zu diagnostischen und prognostischen Markern (Garofalo and Croce 2011) (Esquela-Kerscher and Slack 2006) (Xia and Hui 2012) (Kunz, Xiao et al. 2014) (Boeri, Pastorino et al. 2012). Diesbezüglich haben verschiedene Studien bereits erste therapeutische Erfolge in Zellkultur und Tiermodellen gezeigt sowie vielversprechende Ergebnisse als nicht-invasive diagnostische Verfahren geliefert (Sozzi, Boeri et al. 2014) (Bianchi, Nicassio et al. 2011) (Montani, Marzi et al. 2015).

Im Folgenden sollen wichtige miRNAs in Lungenkrebs, die eine Verbindung zu KRAS, EGFR und p53 besitzen, kurz vorgestellt werden. Die Auswahl basiert auf eigenen wissenschaftlichen Arbeiten und Interessen bezüglich dieser miRNAs, daneben gibt es selbstverständlich zahlreiche weitere miRNAs mit einer wichtigen Rolle in NSCLC, z.B. miRNA-155, miRNA-221/222 und miRNA-210.

Eine bei vielen Tumoren überexprimierte miRNA ist die miRNA-21. Die miRNA-21 wird durch den EGFR-Signalweg aktiviert, begünstigt Proliferation sowie Metastasierung und ist mit einer niedrigen Überlebensrate in NSCLC assoziiert (Dong, Ren et al. 2015) (Qi and Mu 2012) (Yang, Shen et al. 2013) (Hatley, Patrick et al. 2010). Die miRNA-21 reguliert u. a. Programmed cell death protein 4 (PDCD4), Phosphatase and Tensin homolog (PTEN) und Sprouty RTK Signaling Antagonist 1 (Spry1) und übernimmt eine bedeutende Aufgabe in der Regulation der RAS/MEK/ERK- und RAS/PI3K/AKT-Signalwege (Gomez, Wykosky et al. 2013). PTEN ist ein wichtiger negativer Regulator von Phosphoinositide-3-Kinase (PI3K) und V-Akt Murine Thymoma Viral Oncogene Homolog (AKT) und spielt eine wesentliche Rolle in der Regulation von Zellzyklus, Apoptose und Überleben (Kayani, Kayani et al. 2011)

(Jansson and Lund 2012) (Gomez, Wykosky et al. 2013) (Frezza, Menna et al. 2011) (Leever, Vanhaesebroeck et al. 1999) (Frankel, Christoffersen et al. 2008) (Ramaswamy, Nakamura et al. 1999) (Brighenti 2015). Interessanterweise lassen sich vor allem auch Unterschiede in den Interaktionspartnern (IPn; auch als Targets und Zielprotein bezeichnet) feststellen. Während sich bei Lungen- und Brusttumoren z.B. ein verringertes PTEN-Level infolge einer miRNA-21-Überexpression zeigt, lässt sich dieser Effekt bei Darmtumoren nicht erkennen (Tokarz and Blasiak 2012) (Vickers, Bar et al. 2012). Des Weiteren steht die miRNA-21 in engem Zusammenhang mit einer Resistenzentwicklung gegenüber Medikamenten, wobei sich erste gezielte experimentelle Modulationen der miRNA-21 bei Gefitinib-resistenten NSCLC-Zelllinien bereits als therapeutisch wirksam erwiesen haben (Shen, Zhu et al. 2014) (Li, Ren et al. 2014).

Bekannte herunterregulierte miRNAs in Lungentumoren sind let-7 (let-7), die miRNA-34 und miRNA-143/145. Die miRNA-Familie let-7 bindet z.B. KRAS und c-Myc und ist mit der Tumorentstehung und einer niedrigen Überlebensrate assoziiert (Takamizawa, Konishi et al. 2004) (Gomez, Wykosky et al. 2013) (Johnson, Grosshans et al. 2005) (Sampson, Rong et al. 2007) (Lee and Dutta 2007) (Kasinski, Kelnar et al. 2015). Eine experimentelle Erhöhung der let-7-Expression hat ebenfalls erste therapeutische Effekte in Lungenkrebszellen gezeigt (Takamizawa, Konishi et al. 2004) (Kumar, Erkland et al. 2008). Die miRNA-34-Familie (miRNA-34a, -34b und -34c) wird durch p53 aktiviert, beeinflusst Proliferation und Tumorgenese und ist mit einer geringen Überlebensprognose in NSCLC assoziiert (Jansson and Lund 2012) (He, He et al. 2007) (Gomez, Wykosky et al. 2013). Bekannte Targets der miRNA-34 sind z.B. c-Myc und c-MET (Bandi and Vassella 2011) (Kim, Kim et al. 2011) (Jansson and Lund 2012) (Siemens, Jackstadt et al. 2011) (Kasinski and Slack 2012) (Gomez, Wykosky et al. 2013) (Corney, Flesken-Nikitin et al. 2007) (Zhou, Chen et al. 2014) (Bommer, Gerin et al. 2007) (Tazawa, Tsuchiya et al. 2007). Im Hinblick auf Lungentumore konnte eine experimentelle Induktion der miRNA-34a einen therapeutischen Effekt auf die Apoptose und das Tumorwachstum zeigen (Kasinski and Slack 2012) (He, He et al. 2007) (Chang, Wentzel et al. 2007) (Bandi and Vassella 2011) (Kim, Kim et al. 2011) (Siemens, Jackstadt et al. 2011), dagegen hat eine kombinierte Beeinflussung der miRNA-34a und let-7 die Wirkung von Erlotinib in NSCLC-Zellen verstärkt (Stahlhut and Slack 2015). Die beiden kotranskribierten miRNAs miRNA-143/145 sind bekannte Tumorsuppressoren und assoziieren mit Proliferation und Tumorgenese in verschiedenen Tumorentitäten, z.B. Lungen-, Brust- und Darmkrebs (Qi and Mu 2012) (Xin, Small et al. 2009) (Izzotti, Calin et al. 2009) (Cho, Chow et al. 2009) (Gao,

Yu et al. 2010) (Fang, Xiao et al. 2012). Die miRNA-143 hat zudem einen Einfluss auf Invasion und Metabolismus, z.B. durch Inhibierung des Lungenstammzellmarkers CD44v3 und Bindung von Hexokinase 2 (Ma, Jiang et al. 2013) (Templeton, Miyamoto et al. 2014) (Fang, Xiao et al. 2012), wohingegen die miRNA-145 Proliferation, Metastasierung und Zellzyklusprozesse über eine Inhibierung von beispielsweise c-Myc, MUC1 und EGFR reguliert (Qi and Mu 2012) (Chen, Zeng et al. 2010) (Cho, Chow et al. 2011) (Yin, Zhang et al. 2011) (Sachdeva and Mo 2010).

### **Lange nicht-kodierende RNAs**

Eine erst seit Neuestem durch die großen Fortschritte in der Transkriptomsequenzierung bekannte Klasse von ncRNAs sind lange nicht-kodierende Ribonukleinsäuren (lncRNAs, englisch: long noncoding RNAs). lncRNAs (Sequenzlänge über 200 Nte) können im Genom entweder in gleicher (sense) oder entgegengesetzter (antisense) Position zu proteinkodierenden Genen oder auch zwischen (intergenisch) diesen liegen (Fiedler, Breckwoldt et al. 2015). Im Gegensatz zu miRNAs besitzen lncRNAs oftmals eine geringe Sequenz-Struktur-Konservierung und haben verschiedene Wirkmechanismen. Sie binden im Zellkern z.B. Transkriptionsfaktoren (TFen) und Chromatin-modifizierende Proteine (Chromatin modifier) und haben einen Einfluss auf die Transkription (nukleare lncRNAs) und epigenetische Prozesse. Zusätzlich binden sie im Zytoplasma mRNAs, miRNAs und Proteine und beeinflussen die Translation (zytoplasmatische lncRNAs) (Lee 2012) (Ulitsky and Bartel 2013, Bell, Long et al. 2014) (Boon, Jaé et al. 2016) (Paralkar and Weiss 2013). Je nach ihrer Expression im Zytoplasma oder Zellkern üben lncRNAs eine komplexe regulatorische Funktion aus. Erste Studien zeigen bereits lncRNAs mit krankheitsassoziierten Funktionen. So zeigt die lncRNA Myosin heavy-chain-associated RNA transcript (Myheart) eine herzspezifische Expression und spielt über eine Bindung mit Brg1 (Smarca4; SWI/SNF-Related Matrix-Associated Actin-Dependent Regulator Of Chromatin Subfamily A Member 4) eine wichtige Rolle in der Herzfunktion, wodurch sie sich als therapeutisches Target für eine Herzhypertrophie eignet (Han, Li et al. 2014) (Fiedler, Breckwoldt et al. 2015). Eine Herzhypertrophie, z.B. durch Stress ausgelöst, ist eine krankheitsbedingte Größenzunahme des Herzens, die zu einer Herzmuskelschwäche führt (Viereck, Bang et al. 2014).

Angiogenese und Vaskularisation übernehmen eine entscheidende Aufgabe in der Entstehung und Bildung von Blutgefäßen und sind an der zellulären Differenzierung beteiligt (Fiedler, Breckwoldt et al. 2015) (Bell, Long et al. 2014). Hierfür wichtige lncRNAs sind Smooth muscle and Endothelial cell-enriched migration/differentiation-associated long Non-Coding RNA (SENCR) und Myocardial infarction-associated transcript (MIAT) (Bell, Long et al. 2014) (Yan, Yao et al. 2015) (Fiedler, Breckwoldt et al. 2015). SENCR zeigt eine spezifische Expression in vaskulären Zellen und ist mit Migration und Differenzierung assoziiert (Bell, Long et al. 2014) (Fiedler, Breckwoldt et al. 2015). Die lncRNA MIAT hingegen besitzt neben einer Rolle bei einem Myokardinfarkt auch eine Verbindung zu Diabeteserkrankungen, wobei sich hinsichtlich einer therapeutischen Intervention bereits in einem *in vivo*-Modell eine verbesserte mikrovaskuläre Dysfunktion durch Bindung von MIAT an Vascular Endothelial Growth Factor (VEGF) und die miRNA-150 gezeigt hat (Yan, Yao et al. 2015). Auch haben erste Studien begonnen, die Funktion von lncRNAs als potentielle Biomarker zu untersuchen, etwa bei Herzerkrankungen. So haben Kumarswamy et al. (2014) in einer Transkriptomanalyse von 788 Patienten gezeigt, dass die mitochondriale lncRNA LIPCAR nach einem Myokardinfarkt dereguliert ist und Aussagen über die Überlebensrate erlaubt (Kumarswamy, Bauters et al. 2014).

Aufgrund ihres genregulatorischen Einflusses und des enormen diagnostischen und therapeutischen Potentials wächst die Zahl der neu annotierten lncRNAs rasant (Bell, Long et al. 2014) (Cabili, Trapnell et al. 2011). Trotz alledem sind bislang nur eine geringe Zahl an lncRNAs hinsichtlich ihrer Funktionsweise vollständig charakterisiert, was vor allem durch die teils geringe Konservierung und komplexen Wirkmechanismen einen komplizierten und aufwendigen experimentellen Prozess darstellt. In diesem Zusammenhang kann eine integrierte systembiologische Analyse zu einem besseren funktionellen Verständnis beitragen und die experimentelle Charakterisierung unterstützen. Im Folgenden soll auf systembiologische Methoden zur Analyse von Sequenz, Struktur und Signalnetzwerken, die in dieser Arbeit hauptsächlich Anwendung fanden, eingegangen werden.

### 3.5 Systembiologische Methoden zur Modellierung und Simulation biologischer Daten

#### Sequenz-Struktur-Analyse

Eine Analyse von miRNAs und lncRNAs sollte immer hinsichtlich einer Sequenz-Struktur-Konservierung untersuchen. Dies kann neben homologen Sequenzen in anderen Organismen und konservierten Sekundärstrukturelementen, z.B. Haarnadelstruktur, vor allem mögliche Aussagen über die Funktion liefern. Die RNA-Sekundärstrukturfaltung ist ein komplexer Prozess, da neben einer komplementären Sequenz auch die Faltungsenergie betrachtet werden muss. RNA-Sekundärstrukturen können bioinformatisch vorhergesagt werden. Dynamische Programmieralgorithmen, z.B. der Zuker-Algorithmus (RNAfold- und Mfold-Programme), berechnen für eine Sequenz die thermodynamisch optimale Sekundärstruktur basierend auf einer Energieminimierung (minimale freie Faltungsenergie, englisch: minimum free energy). Diese Algorithmen berechnen sehr akkurat die optimale Sekundärstruktur, was sie aber für eine Analyse mit mehreren Sequenzen aufgrund der hohen Rechenzeit meist unbrauchbar macht (Hofacker 2003) (Zuker 2003) (Reeder and Giegerich 2005) (Dalli, Wilm et al. 2006) (Bernhart, Hofacker et al. 2008) (Gomes, Cho et al. 2013). Hierfür sind effektivere Algorithmen notwendig, wie etwa der Sankoff-Algorithmus, die gleichzeitig mehrere Sequenzen alignieren und falten können (Sankoff 1985) (Reeder and Giegerich 2005) (Dalli, Wilm et al. 2006) (Hofacker, Fekete et al. 2002) (Bernhart, Hofacker et al. 2008). Programme, welche auf dem Sankoff-Algorithmus basieren, sind z.B. RNAalifold (Bernhart, Hofacker et al. 2008), FOLDALIGN (Havgaard, Lyngsø et al. 2005) und LocARNA (Smith, Heyne et al. 2010). Hierbei wird ein paarweises Sequenzalignment generiert, welches anschließend simultan gefaltet und aligniert wird, um die konservierte Sekundärstruktur zu berechnen (Sankoff 1985) (Havgaard, Lyngsø et al. 2005) (Bernhart, Hofacker et al. 2008) (Smith, Heyne et al. 2010). Darüber hinaus gibt es noch optimiertere und effizientere Ansätze. Dabei handelt es sich um Weiterentwicklungen des Sankoff-Algorithmus, welche eine noch schnellere Rechenzeit ermöglichen, etwa das RNashapes-Programm basierend auf einem nicht-heuristischen Ansatz (lineare Rechenzeit, Sankoff-Algorithmus hingegen exponentiell) (Reeder and Giegerich 2005) (Steffen, Voss et al. 2006).

Neben einer Sequenz-Struktur-Konservierung sollte eine systembiologische Analyse von miRNAs und lncRNAs insbesondere den genomischen Kontext (sense, antisense), Promotor, IP und assoziierte biologische Prozesse und Signalwege berücksichtigen. Dies kann zu einem



umfassenden funktionellen Verständnis beitragen und Aussagen über eine mögliche gewebe- und zelltypspezifische Regulation aufzeigen. Auf eine Beschreibung soll an dieser Stelle verzichtet werden, zusätzliche Informationen sind in Kunz, Xiao et al. (2014) zu finden.

### **Analyse und Modellierung biologischer Signalwege**

Dies kann dabei wie nachstehend aufgeführt erfolgen: (i) Signalnetzwerkrekonstruktion, (ii) Signalnetzwerkanalyse, (iii) *in silico*-Modellierung und zeitliche Analyse eines Signalnetzwerkes. Anhand experimenteller Daten und Expertenwissen können hierbei individuelle Netzwerkmodelle, etwa für Lungenkrebs, entwickelt werden, um daraus computergestützt biologische Fragestellungen zu beantworten und zu neuen Erkenntnissen zu gelangen. Hieraus kann dann letztendlich ein systembiologisches Modell entwickelt werden. Dieses kann, je nach Fragestellung, verschiedene Antworten geben, z.B. welches logische Schaltverhalten die Netzwerkkomponenten aufweisen, welche Komponenten miteinander interagieren oder das resultierende Netzwerkverhalten im Zeitablauf vorhersagen. Meist interessiert insbesondere die Funktion des Signalnetzwerks, deren gewebe- und zelltypspezifischen Eigenschaften und vor allem deren therapeutische Intervention.

(i) Um einen Signalweg modellieren zu können, muss zuerst ein Signalnetzwerk anhand experimenteller Daten (eigene Experimente oder aus Datenbanken, z.B. Gene Expression Omnibus (GEO)) und Literaturdaten rekonstruiert werden, z.B. Rezeptoren und Proteine, etwa mit den Datenbanken HPRD (Mensch), AINM (Pflanzen) oder STRING (verschiedene Organismen) (Naseem, Kunz et al. 2014) (Kunz, Dandekar et al. 2016).

(ii) Das erstellte Netzwerkmodell kann anschließend systembiologisch untersucht werden, z.B. mit Cytoscape. Cytoscape ist eine systembiologische Software zur Visualisierung und Analyse regulatorischer Netzwerke (Shannon, Markiel et al. 2003). Es verfügt über zahlreiche Plug-ins, die verschiedene Analysen, etwa nach biologischen Funktionen und Signalwegen (z.B. BiNGO und ClueGO) oder hinsichtlich der Netzwerkeigenschaften/-topologie (NetworkAnalyzer), ermöglichen (Naseem, Kunz et al. 2014) (Kunz, Dandekar et al. 2016).

(iii) Das Netzwerkmodell kann in einem weiteren Schritt mithilfe mathematischer Methoden computergestützt abgebildet werden, z.B. die Änderung der Aktivität der Signalwege nach Medikamentengabe anhand von Differentialgleichungen sowohl mit als auch ohne experimentelle Daten.

Boolesche Modelle beschreiben die Knoten (z.B. Proteine) eines Netzwerkes anhand logischer

Operatoren (AND, OR, NOT; boolesche Funktion). Knoten sind demnach inaktiv (OFF/nicht exprimiert=0) oder aktiv (ON/exprimiert=1) und erhalten entsprechend ihrem Aktivierungszustand entweder den Wert 0 oder 1 zugewiesen. Diese Methode gibt eine qualitative Aussage über das logische Schaltverhalten der Netzwerkkomponenten und Systemzustände wieder und ist z.B. für genregulatorische Netzwerke geeignet (Schlatter, Philippi et al. 2012) (Thakar and Albert 2010).

Eine weitere Möglichkeit der Abbildung eines Netzwerkmodells ist die dynamische Modellierung. Diese Methode geht davon aus, dass jeder Knoten im Netzwerk mit einer gewissen Konzentration interagiert, deren Modellierung anhand gewöhnlicher Differentialgleichungen erfolgen kann. Dies verlangt exakte Kenntnisse über die Konzentrationslevel beteiligter Komponenten, wofür sehr gute und zeitaufgelöste experimentelle Daten notwendig sind. Dynamische Modellierungen erlauben eine quantitative Beschreibung der Netzwerkkomponenten im Zeitablauf und sind zur Modellierung kinetischer Daten hilfreich, wie etwa Signaltransduktion oder Enzymreaktionen (z.B. Michaelis-Menten-Kinetik) (Maiwald and Timmer 2008; Wangorsch, Butt et al. 2011) (Schlatter, Philippi et al. 2012).

Um auch ohne exakte kinetische Daten entsprechende Aussagen über das Netzwerkverhalten im Zeitverlauf zu erhalten, ist die semiquantitative Modellierung eine geeignete Methode. Diese kombiniert dabei die boolesche mit der dynamischen Modellierung durch eine mathematische Transformation. Eine semiquantitative Modellierung ermöglicht somit eine intuitive Betrachtung der Systemzustände und eignet sich sehr gut, um einen Überblick über kinetische Zusammenhänge zu bekommen, was auch ohne Wissen über kinetische Parameter möglich ist (Schlatter, Philippi et al. 2012) (Philippi, Walter et al. 2009) (Garg, Di Cara et al. 2008) (Di Cara, Garg et al. 2007).

Aus diesem Grund wurde in der vorliegenden Arbeit eine semiquantitative Modellierung mittels der Software SQUAD durchgeführt (Di Cara, Garg et al. 2007). SQUAD betrachtet den Systemzustand eines Netzwerkes zunächst als diskretes System (boolesches System) und berechnet darauf aufbauend alle Steady-state-Zustände. Diese Steady-states beschreiben stabile Gleichgewichtszustände, in welche ein Netzwerk auch durch Störungen bzw. Änderungen immer wieder zurückkehrt. Anschließend erfolgt eine Transformation in ein dynamisches System mittels Differentialgleichungen (Exponentialfunktion). SQUAD ist somit in der Lage, die einzelnen Netzwerkkomponenten im Intervall zwischen 0 und 1 zu betrachten, was

Rückschlüsse über den Aktivierungsverlauf der Netzwerkkomponenten im Zeitverlauf ermöglicht (Di Cara, Garg et al. 2007).

### 3.6 Die DrumPID-Datenbank

Der Fortschritt in der Entwicklung neuer experimenteller Technologien hat zu einer Ansammlung riesiger Datenmengen beigetragen, die mithilfe von Datenbanken gesammelt und gespeichert werden müssen, aber auch um weitere Untersuchungen zu ermöglichen. Ein Ziel ist dabei, die Datenbanken den individuellen Bedürfnissen der Nutzer anzupassen. Da sich ein Großteil der Untersuchungen in dieser Arbeit u. a. auf Signalnetzwerkanalysen konzentriert haben, sollte hierfür eine Datenbank bereitgestellt werden. Die DrumPID- (Drug-minded Protein Interaction Database) Datenbank integriert verschiedene Datenbanken und stellt die Interaktion zwischen einem Medikament (auch als Pharmakon bezeichnet) und seinem Target in den Fokus.

### 3.7 Ziele dieser Arbeit

Systembiologische Analysen ermöglichen fundierte Aussagen über grundlegende Mechanismen und Funktionsweisen und können zu einem besseren Verständnis biologischer Daten beitragen. Das grundlegende Ziel dieser Arbeit ist die Entwicklung eines systematischen Ansatzes für systembiologische Analysen, um diesen individuell an verschiedenen biologischen Fragestellungen und Organismen anwenden zu können und zu neuen Einsichten und Interventionen zu gelangen.

Hierzu soll (i) eine integrierte systembiologische Methode entwickelt werden, welche verschiedene Datenebenen integriert, um z.B. experimentelle Transkriptomdaten funktionell zu untersuchen und relevante Hubproteine zu identifizieren. Im ersten Teil werden Analysen zum pflanzlichen Immunsystem durchgeführt. Diese sollen insbesondere konzeptionell das ZigZag-Modell in Bezug auf den Einfluss von *Pst* DC3000 auf das zelluläre Interaktom von *A. thaliana* untersuchen und neue Modulatoren einer CK-vermittelten Immunabwehr identifizieren.

(ii) Darauf aufbauend schließen sich dynamische Modellierungen der Signalverarbeitung und -modulation an. Als Beispiel dient ein *in silico/in vitro*-Lungentumormodell einer HSP90-Therapie, um das unterschiedliche Behandlungsansprechen erklären zu können sowie neue

prognostische Marker und therapeutische Targets mit potentieller klinischer Relevanz zu identifizieren. Zusätzlich sollen eine Gefitinib-Behandlung anhand bekannter Mutationen und der miRNA-21 computergestützt nachgebildet und die bestmögliche Therapiestrategie nach Resistenzentwicklung herausgearbeitet werden.

(iii) Im dritten Teil der Arbeit werden systematisch RNA- und Protein-Interaktionen analysiert. Als Beispiel dienen a) neu detektierte hypertrophie- und hypoxie-assoziierte lncRNAs. Hierzu werden Untersuchungen beginnend mit einer Sequenz-Struktur-Konservierung über Promotoranalysen bis hin zu funktionellen Analysen durchgeführt. Ein weiteres Beispiel stellen b) Protein-Pharmaka-Interaktionen dar. Hierfür soll eine neue Datenbank entwickelt werden, die Drug-minded Protein Interaction Database (DrumPID), welche Analysen zu einem Medikament, dessen Interaktionen und beteiligter Signalwege sowie resultierender Netzwerkeffekte ermöglicht.

## 4. Material und Methoden

### 4.1 Integrierte Netzwerkanalysen zum pflanzlichen Immunsystem von *A. thaliana*

Um relevante funktionelle Hubproteine in experimentellen Genexpressionsdatensätzen hinsichtlich ihres Interaktionskontextes und biologischen Funktion zu detektieren, wurde eine integrierte systembiologische Methode entwickelt (Naseem, Kunz et al. 2014) (Kunz, Dandekar et al. 2016). Entsprechende Vorarbeiten an einem Kontrolldatensatz wurden in meiner Masterarbeit begonnen (hier nicht gezeigt). Die entwickelte Methode wird im Folgenden für einen neuen komplexeren Datensatz beschrieben, dessen Ergebnisse in dieser Arbeit gezeigt werden.

#### Einfluss verschiedener Virulenzfaktoren auf das zelluläre Wirtsnetzwerk

Die durchgeführten Analysen wurden entsprechend Naseem, Kunz et al. (2014) und Kunz, Dandekar et al. (2016) durchgeführt. Interaktomdaten von *A. thaliana* wurden aus der Plant Interactome Database (*Arabidopsis* Interactome Network Map, [http://interactome.dfc.harvard.edu/A\\_thaliana/](http://interactome.dfc.harvard.edu/A_thaliana/)) verwendet (experimentelle binäre Yeast Two-Hybrid (Y2H) Protein-Protein-Interaktionen (PPIen) und Literaturdaten; 4866 Proteine (Knoten), 11374 Interaktionen (Kanten); Stand 04/2016). Als Transkriptomdaten wurde ein Microarray von *A. thaliana* infiziert mit dem Pathogen *Pst* DC3000 aus der GEO-Datenbank verwendet (Nummer: GSE5520). Datensatz 1 (Normalinfektion): *Pst* DC3000 gegen (vs.) Mock (3 Replikate, jeweils 24h behandelt; GSM128654, GSM128655, GSM128656); Datensatz 2 (Mutante): *Pst* DC3000 COR<sup>-</sup> vs. Mock (3 Replikate, jeweils 24h behandelt; GSM128648, GSM128649, GSM128650); Datensatz 3 (Doppelmutante): *Pst* DC3000 COR<sup>-</sup> *hrpS* vs. Mock (3 Replikate, jeweils 24h behandelt; GSM128651, GSM128652, GSM128653) (in jedem Datensatz insgesamt jeweils 22810 Gene). Als Mock (Kontrollbehandlung) wurde für alle drei Datensätze GSM128657, GSM128658 und GSM128659 (24h) gewählt (Stand 04/2016). Unbekannte Annotationen und Gennamen zu den Affimetrix-Identifizier wurden aus der Datenbank EnsemblPlants BioMart (<http://plants.ensembl.org/biomart/martview/9bdaa7174df36f0499025c8de4f750fd>; Datensatz: *Arabidopsis thaliana* genes, TAIR10 (2010-09-TAIR10)) entnommen. Eine Normalisierung der Daten wurde mit GEO-2R durchgeführt (Standardeinstellungen verwendet). Als signifikant differentiell-exprimierte Gene (DEGe) wurden Gene mit einem  $\log_{2}FC \geq +0,5 / \leq -0,5$  und  $p\text{-Value} < 0,05$  ausgewählt und grafisch in einem Plot mit dem Statistikprogramm R (Version

3.1.2) dargestellt (The R Development Core Team. Vienna, Austria, 2010. R Foundation for Statistical Computing, ISBN 3-900051-07-0). Als erster Schritt wurden die DEGe für jeden Datensatz gegen das Interaktom gemappt (Transkriptom-Interaktom-Mapping), um so die entsprechenden überlappenden Knoten zwischen Transkriptom und Interaktom zu erhalten (Mapping wurde in phpMyAdmin [http://www.phpmyadmin.net/home\\_page/index.php](http://www.phpmyadmin.net/home_page/index.php) durchgeführt). Zusätzlich wurden die beiden Datensätze der Mutanten (Datensatz 2 und Datensatz 3) gegen Datensatz 1 (*Pst* DC3000) gemappt, um überlappende Gene zu identifizieren (ebenfalls in phpMyAdmin; grafische Darstellung als Volcano-Plot unter Verwendung des ggplot2-Pakets aus dem Statistikprogramm R (Version 3.1.2) (The R Development Core Team. Vienna, Austria, 2010. R Foundation for Statistical Computing, ISBN 3-900051-07-0). Anschließend wurde anhand der DEGe und unter Verwendung der direkten IP aus dem Interaktom für jeden Datensatz ein expressions-basiertes Netzwerk (EBN) erstellt und in Cytoscape 2.8.3 (Shannon, Markiel et al. 2003) visualisiert. Die EBNe wurden im Anschluss auf ihre Netzwerktopologie mit dem Cytoscape-Plug-in NetworkAnalyzer (Assenov, Ramirez et al. 2008) untersucht (Eingabeparameter: Treat the network as undirected), wobei die Parameter „durchschnittliche Anzahl an Interaktionsnachbarn“, „Netzwerkzentralität“ und „Netzwerkheterogenität“ verwendet wurden. Des Weiteren wurden die EBNe für jeden Datensatz auf funktionelle Module und Proteinkomplexe (Cluster bzw. gut vernetzte Subgraphen) mit dem Cytoscape-Plug-in AllegroMCODE untersucht, welches auf dem Molecular Complex Detection (MCODE) Algorithmus basiert (Bader and Hogue 2003) (Saito, Smoot et al. 2012). Hierbei wurden Cluster mit einem Score>1 ausgewählt. Die Knoten der EBNe wurden ebenfalls gegen bekannte 137 Effektorproteine in *A. thaliana* gemappt, welche aus der Publikation Mukhtar, Carvunis et al. (2011) erwendet wurden.

### **Neue Modulatoren einer CK-vermittelten Immunabwehr**

Um für jeden Datensatz funktionelle Hubproteine zu identifizieren, wurden die DEGe hinsichtlich ihrer biologischen Funktion mit dem Cytoscape-Plug-in Biological Networks Gene Ontology (BiNGO (Maere, Heymans et al. 2005), Suchparameter „biological process“ in *A. thaliana*; p-Value<0,05; Standardparameter) untersucht, wobei lediglich biologische Prozesse assoziiert mit dem pflanzlichen Immunsystem für die weitere Analyse ausgewählt wurden. BiNGO identifiziert signifikant überrepräsentierte biologische Prozesse entsprechend ihrer Gene-Ontology (GO (Ashburner, Ball et al. 2000))-Kategorie. Mithilfe der für das Immunsystem relevanten ausgewählten Proteine wurde analog der beschriebenen Schritte für

die EBNe für jeden Datensatz ein Immunnetzwerk erstellt und auf ihre Netzwerktopologie untersucht. Abschließend wurden in den Immunnetzwerken unter Verwendung einer hohen Netzwerkvernetzung (hohe Anzahl an IP) und Immunfunktion funktionelle Hubproteine für die pflanzliche Immunabwehr detektiert (Auswahl erfolgte unter Verwendung der bekannten Literaturdatenbanken TAIR ([www.arabidopsis.org](http://www.arabidopsis.org)) und PubMed (<http://www.ncbi.nlm.nih.gov/pubmed/>)).

## 4.2 Integrierte Analysen zu Lungenkrebs

### 4.2.1 Experimentelle Arraydaten der 17AAG-Behandlung eines *in vitro*-Experiments

Um eine Korrelation eines 3D-Tumortestsystems basierend auf der SISmuc mit herkömmlichen 2D-Tumortestsystemen und klinischen Beobachtungen zu untersuchen, wurden drei verschiedene NSCLC-Zelllinien mit dem HSP90-Inhibitor 17AAG (17-allylamino-geldanamycin, Tanespimycin) behandelt. Für das 3D-Lungentumormodell wurden die drei Zelllinien HCC827 (EGFR mutiert, KRAS wildtypisch), A549 und H441 (KRAS mutiert, EGFR wildtypisch) auf der SISmuc-Matrix entsprechend Stratmann, Fecher et al. (2014) kultiviert, ebenfalls wurden diese Zelllinien unter 2D-Zellkulturbedingungen kultiviert. Nach einer 72-stündigen Behandlung mit dem HSP90-Inhibitor 17AAG wurden die Zellen lysiert und anschließend hinsichtlich ihrer Signalaktivität mittels RTK- und Phosphokinase (PK)-Arraydaten untersucht. Der RTK-Array besteht aus 15 Rezeptorproteinen und der PK-Array aus 45 Proteinen. Die experimentellen Ergebnisse zeigen zwischen dem 2D- und 3D-Tumormodell in den verwendeten Zelllinien einen ähnlichen Effekt in der Apoptose nach einer 17AAG-Behandlung, allerdings zeigt sich im 3D-Modell für die A549-Zelllinie ein deutlich geringerer Effekt in der Behandlungsstärke (Tab. 4.2.1). Hinsichtlich der Proliferation lässt sich im 2D-Modell in allen drei Zelllinien ein therapeutischer Effekt erkennen, wohingegen sich dies lediglich im 3D-Modell für die HCC827-Zellen beobachten lässt (Tab. 4.2.1). Zudem lassen sich teils vollständig konträre Unterschiede in der Regulierung der experimentellen RTK- und PK-Arraydaten erkennen. Die experimentellen Arbeiten wurden von Claudia Göttlich aus der AG Tumor-Testsysteme (Dr. Gudrun Dandekar) am Lehrstuhl für Tissue Engineering und Regenerative Medizin (Leitung: Prof. Dr. Heike Walles, Universitätsklinikum Würzburg und Translationszentrum „Regenerative Therapien für Krebs- und

Muskuloskelettale Erkrankungen“) durchgeführt. Die Ergebnisse wurden in einer Excel-Tabelle zusammengefasst und für die systembiologischen Analysen zur Verfügung gestellt.

**Tabelle 4.2.1: Übersicht der Änderungen auf Apoptose und Proliferation nach 17AAG-Behandlung.** Die Tabelle fasst die Änderungen auf Apoptose und Proliferation zwischen dem 2D- und 3D-Zellkultursystem zusammen. Es lässt sich ein ähnlicher Effekt in beiden Systemen in der Apoptose, nicht aber Proliferation, erkennen, allerdings zeigt das 3D-System eine geringere Ansprechrate nach 17AAG-Behandlung (Proliferation: Werte repräsentieren Differenz aus behandelt minus unbehandelt; Apoptose: Werte repräsentieren die x-fache Erhöhung nach Behandlung gegenüber unbehandelt; signifikant erhöhte Werte in grün, in rot entsprechend erniedrigte Werte nach 17AAG-Behandlung).

	2D			3D		
	A549	HCC827	H441	A549	HCC827	H441
Proliferation	-71,21%	-12,73%	-61,47%	-1,95%	-10,80%	-2,93%
Apoptose	4,292	1,586	0,785	1,481	1,647	0,938

#### 4.2.2 Systembiologische Analyse der Arraydaten

Für meine folgenden systembiologischen Analysen habe ich als erstes anhand der Arraydaten für das 2D- und 3D-System die Differenz zwischen behandelten und unbehandelten Zellen berechnet (Differenz: +17AAG minus -17AAG). Die Werte reflektieren das entsprechende Ansprechen der Zellen auf die 17AAG-Behandlung. Ein positiver Wert bedeutet eine höhere Signalaktivierung, wohingegen ein negativer Wert eine reduzierte Aktivierung nach Behandlung aufzeigt. Anhand dieser Berechnung haben sich die systembiologischen Analysen auf die Unterschiede zwischen dem 2D- und 3D-System sowie auf die beiden KRAS-mutierten Zelllinien im 3D-System konzentriert.

#### *In silico*-Analyse hinsichtlich der Unterschiede zwischen dem 2D- und dem 3D-System

Um Unterschiede zwischen 2D und 3D zu finden, wurden die berechneten Werte miteinander verglichen. Da die Aktivierungswerte in 2D geringer als in 3D waren, wurde für das 2D-System eine Differenz von  $\pm 0,5\%$  und für das 3D-System von  $\pm 5\%$  als signifikant angesehen. Es wurden jeweils diejenigen Proteine ausgewählt, die entweder signifikant reguliert in 2D oder 3D bzw. signifikant reziprok reguliert in mindestens einer Zelllinie zwischen 2D und 3D sind und entsprechend als differentiell-reguliertes Protein (DRP) bezeichnet. Die gefundenen DRPe wurden anschließend entsprechend der Methode aus Naseem, Kunz et al. (2014) und Kunz, Dandekar et al. (2016) analysiert. Als Interaktomdaten dienten alle humanen direkten binären



Y2H PPIen aus der HPRD-Datenbank (9617 Knoten und 39184 Kanten, release 9 vom 13. April 2010). Die DRPe wurden auf das Interaktom gemappt und unter Verwendung der direkten IP ein 2D-3D-Zellkulturnetzwerk erstellt (Mapping erfolgte in phpMyAdmin ([http://www.phpmyadmin.net/home\\_page/index.php](http://www.phpmyadmin.net/home_page/index.php))). Für die Visualisierung des Netzwerks wurde Cytoscape 2.8.3 verwendet (Shannon, Markiel et al. 2003). Im Anschluss wurde das Netzwerk hinsichtlich funktioneller Module mit dem Plug-in AllegroMCODE (Score>1) untersucht (Bader and Hogue 2003) (Saito, Smoot et al. 2012).

### **Dynamische Simulation der 17AAG-Behandlung für die HCC827-Zelllinie für das 3D-System**

Für das 3D-System wurde für die HCC827-Zelllinie eine Simulation der 17AAG-Behandlung durchgeführt. Hierfür wurden die experimentellen Daten und der spezifische Mutationshintergrund der Zelllinie in ein *in silico*-Modell (Stratmann, Fecher et al. 2014) (Gottlich, Muller et al. 2016) unter Verwendung der Software CellDesigner 3.5.1 (Funahashi, Morohashi et al. 2003) integriert und ein individuelles HCC827-Netzwerk erstellt. Das *in silico*-Modell integriert die in der Klinik verwendeten prädiktiven Marker für eine anti-EGFR-Therapie (aktivierende Mutation im EGFR) sowie beteiligte Signalwege der für den therapeutischen Effekt wichtigen Parameter Apoptose (PI3K/AKT-Signalweg) und Proliferation (RAS/RAF/ERK-Signalweg). Die Mutationen und entsprechenden Interaktionen wurden aus Literaturrecherche und bekannten Interaktionsdatenbanken, z.B. PubMed (<https://www.ncbi.nlm.nih.gov/pubmed>), HPRD (Keshava Prasad, Goel et al. 2009), KEGG (Kanehisa, Sato et al. 2016) und QIAGEN (GeneGlobe; <https://www.qiagen.com/us/geneglobe/>) verwendet. Das Netzwerk wurde anschließend in die Software SQUAD (Di Cara, Garg et al. 2007) eingelesen. SQUAD repräsentiert die Netzwerktopologie (Aktivierung, Inhibierung) mittels logischer boolescher Operatoren (AND, OR, NOT) und interpoliert diese anschließend mittels mathematischer Exponentialfunktionen (Di Cara, Garg et al. 2007). Das Netzwerkverhalten wird graphisch als Veränderungen der Netzwerkknoten (y-Achse) im Zeitverlauf (x-Achse) dargestellt, was eine *in silico*-Simulation verschiedener Netzwerkszenarien, etwa Medikamentengabe, erlaubt. Für die Simulationen wurden jeweils Simulationsprotokolle über die Perturbator-Funktion in SQUAD geschrieben. Die Netzwerkknoten wurden basierend auf dem bekannten Mutationsstatus und den experimentellen Arraydaten gesetzt, wobei ein Wert (Value) von 1 eine maximale Aktivierung

und 0 eine Inaktivierung reflektiert. Hierbei wurde folgendes Skript für die HCC827-Zelllinie verwendet:

1. Unbehandelte Zellen: initialstate=State1; constant node pulse (parameter=state): low\_glucose: time=0, value=0,4; \*(EGF-EGFR): time=0, value=0,21; (c-MET): time=0, value=0,21; Stress: time=0, value=0,7; (c-RET): time=0, value=0,05; (Erb2/3): time=0, value=0,05; active node pulse: alle Knoten state=0.

2. 17AAG-Behandlung: initialstate=State1; constant node pulse (parameter=state): 17AAG: time=0, value=1; low\_glucose: time=0, value=0,4; \*(EGF-EGFR): time=0, value=0,17; (c-MET): time=0, value=0,05; Stress: time=0, value=0,7; active node pulse: alle Knoten state=0.

### **Identifizierung neuer prognostischer Marker und therapeutischer Targets sowie dynamische Simulation des therapeutischen Effekts für die beiden KRAS-mutierten Zelllinien A549 und H441**

Beide KRAS-mutierten Zelllinien des 3D-Systems wurden auf ihre Behandlungsunterschiede und neuen prognostischen und therapeutischen Targets entsprechend der Methode aus Naseem, Kunz et al. (2014) und Kunz, Dandekar et al. (2016) untersucht. Hierfür wurden anhand der berechneten Werte aus den experimentellen Daten (Differenz: +17AAG minus -17AAG; Signifikanzniveau  $\pm 5\%$ ) DRPe ausgewählt, die entweder signifikant reguliert in mindestens einer der beiden Zelllinien oder signifikant reziprok reguliert waren. Die DRPe wurden auf das Interaktom gemappt und unter Verwendung der direkten IP ein KRAS-Netzwerk erstellt (Mapping erfolgte in phpMyAdmin ([http://www.phpmyadmin.net/home\\_page/index.php](http://www.phpmyadmin.net/home_page/index.php))). Die Visualisierung wurde in Cytoscape 2.8.3 (Shannon, Markiel et al. 2003) durchgeführt. Zusätzlich wurde bei dieser Analyse noch der Mutationshintergrund beider Zelllinien einbezogen. Hierzu wurden Daten aus der COSMIC-Datenbank ([http://cancer.sanger.ac.uk/cell\\_lines](http://cancer.sanger.ac.uk/cell_lines); Option Zelllinienprojekt) verwendet: A549 Zelllinie: Sample Name: A549, Sample ID: 905949, analysierte Gene: 1126, Mutationen: 651; H441 Zelllinie: Sample Name: NCI-H441, Sample ID: 908460, analysierte Gene: 1646, Mutationen: 995; Stand 25.04.2016. Die jeweiligen Mutationen jeder Zelllinie wurden im Anschluss mithilfe von Venny (Oliveros (2007-2015)) auf das erstellte KRAS-Netzwerk gemappt, um so die spezifischen Überlappungen beider Zelllinien mit dem Netzwerk zu erhalten. Unter Verwendung der überlappenden Knoten zwischen dem KRAS-Netzwerk und der jeweiligen Zelllinie, der beiden Knoten HSP90AA1 und HSP90AB1 und den direkten IPn aus dem Interaktom wurde ein individuelles zelllinien-spezifisches Netzwerk erstellt (A549-Netzwerk

und H441-Netzwerk; Visualisierung in Cytoscape 2.8.3 (Shannon, Markiel et al. 2003)). Die weitere Analyse hinsichtlich funktioneller Module und Proteinkomplexe erfolgte mit dem Plugin AllegroMCODE (Score>1) (Bader and Hogue 2003) (Saito, Smoot et al. 2012). Darauf basierend wurden neue prognostische Marker und Drugtargets für die beiden KRAS-mutierten Zelllinien identifiziert, wobei die zwei neuen Targets Hypoxia-inducible factor 1-alpha (HIF1A) und AMP-aktivierte Proteinkinase (AMPK) in einem letzten Schritt auf ihren potentiellen therapeutischen Effekt mittels einer *in silico*-Simulation untersucht wurden. Das Erstellen des Netzwerks und die anschließende Simulation wurden analog der vorher beschriebenen Schritte für das HCC827-Netzwerk durchgeführt. Hierbei wurden die Netzwerkparameter Apoptose und Proliferation als Indikatoren für den therapeutischen Effekt verwendet (HIF1A-Inhibitor wurde aus Selleckchem (<http://www.selleckchem.com/>) bzw. AMPK-Aktivator 5-Aminoimidazole-4-carboxamide ribonucleotide (AICAR) aus Analysen der DrugPID-Datenbank (Kunz, Liang et al. 2016) ausgewählt). Zusätzlich wurde die 17AAG-Behandlung für die beiden Zelllinien simuliert. Hierbei wurden folgende Skripte verwendet:

#### A549:

1. Unbehandelte Zellen: initialstate=State1; constant node pulse (parameter=state): Kras: time=0, value=0,4; low\_glucose: time=0, value=0,4; \*(EGF-EGFR): time=0, value=0,15; Akt\*: time=0, value=0,2; Stress: time=0, value=0,7; active node pulse: alle Knoten state=0.

2. AICAR-Behandlung: initialstate= State1; constant node pulse (parameter=state): Kras: time=0, value=0,4; low\_glucose: time=0, value=0,4; \*(EGF-EGFR): time=0, value=0,15; Akt\*: time=0, value=0,2; Stress: time=0, value=0,7; AICAR: time=0, value=1; active node pulse: alle Knoten state=0.

3. 17AAG-Behandlung: initialstate= State1; constant node pulse (parameter=state): Kras: time=0, value=0,4; low\_glucose: time=0, value=0,4; \*(EGF-EGFR): time=0, value=0,15; Akt\*: time=0, value=0,2; Stress: time=0, value=0,7; 17AAG: time=0, value=1; active node pulse: alle Knoten state=0.

#### H441:

1. Unbehandelte Zellen: initialstate=State1; constant node pulse (parameter=state): Kras: time=0, value=0,4; \*(EGF-EGFR): time=0, value=0,25; (c-MET): time=0, value=0,25; Stress: time=0, value=0,22; active node pulse: alle Knoten state=0.

2. HIF1A-Inhibierung: initialstate= State1; constant node pulse (parameter=state): Kras: time=0, value=0,4; \*(EGF-EGFR): time=0, value=0,25; (c-MET): time=0, value=0,25; Stress: time=0, value=0,22; HIF1a-Inhibitor: time=0, value=1; active node pulse: alle Knoten state=0.

3. 17AAG-Behandlung: initialstate= State1; constant node pulse (parameter=state): Kras: time=0, value=0,4; \*(EGF-EGFR): time=0, value=0,25; (c-MET): time=0, value=0,25; Stress: time=0, value=0,22; 17AAG: time=0, value=1; active node pulse: alle Knoten state=0.

#### 4.2.3 *In silico*-Analyse von Resistenzmechanismen nach einer Gefitinib-Behandlung

Das *in silico*-Modell sollte in einem weiteren Schritt verschiedene Resistenzmechanismen mit bekanntem Mutationsstatus abbilden und hieran anschließend geeignete Behandlungsstrategien entwickelt werden.

##### **HCC827- und A549-Zelllinie**

In der HCC827-Zelllinie wurde für eine Therapieresistenz gegenüber einer Gefitinib-Behandlung eine Komutation des c-MET-Rezeptors und für die A549-Zelllinie des IGF-1R in das *in silico*-Modell (Stratmann, Fecher et al. 2014) (Gottlich, Muller et al. 2016) unter Verwendung von CellDesigner 3.5.1 (Funahashi, Morohashi et al. 2003) integriert. Anschließend wurde das erstellte Netzwerk in SQUAD (Di Cara, Garg et al. 2007) simuliert. Hierbei wurden folgende Skripte verwendet:

##### HCC827:

1. Unbehandelte Zellen: initialstate=steady state 4 (SS-4); constant node pulse (parameter=state): FLIP: time=0, value=0,4; active node pulse: (EGF-EGFR): state=0,8; alle anderen Knoten state=0.
2. Gefitinib-Behandlung: initialstate=steady state 4 (SS-4); constant node pulse (parameter=state): FLIP: time=0, value=0,4; gefitinib: time=0, value=1; active node pulse: (EGF-EGFR): state=0,8; alle anderen Knoten state=0.
3. Gefitinib-Resistenz: initialstate=steady state 4 (SS-4); constant node pulse (parameter=state): FLIP: time=0, value=0,4; c-MET: time=0, value=0,4; gefitinib: time=0, value=1; active node pulse: (EGF-EGFR): state=0,8; alle anderen Knoten state=0.
4. Alleinige und kombinierte Inhibierung von PI3K, MEK und c-MET: initialstate=steady state 4 (SS-4); constant node pulse (parameter=state): FLIP: time=0, value=0,4; c-MET: time=0, value=0,4; active node pulse: (EGF-EGFR): state=0,8; alle anderen Knoten state=0; je nachdem, alleinige oder kombinierte Inhibierung, Parameter PI3K-Inh, MEK-Inh, MET-Inh als: constant node pulse (parameter=state), time=0, value=1.

A549:

1. Unbehandelte Zellen: initialstate=steady state 4 (SS-4); constant node pulse (parameter=state): FLIP: time=0, value=0,4; Kras: time=0, value=0,4; IGF-1R: time=0, value=0,4; active node pulse: alle Knoten state=0.
2. Gefitinib-Resistenz: initialstate=steady state 4 (SS-4); constant node pulse (parameter=state): FLIP: time=0, value=0,4; Kras: time=0, value=0,4; IGF-1R: time=0, value=0,4; gefitinib: time=0, value=1; active node pulse: alle Knoten state=0.
3. Alleinige und kombinierte Inhibierung von PI3K, MEK und IGF-1R: initialstate=steady state 4 (SS-4); constant node pulse (parameter=state): FLIP: time=0, value=0,4; Kras: time=0, value=0,4; IGF-1R: time=0, value=0,4; active node pulse: alle Knoten state=0; je nachdem, alleinige oder kombinierte Inhibierung, Parameter PI3K-Inh, MEK-Inh, IGF-1R-Inh als: constant node pulse (parameter=state), time=0, value=1.

**Effekt der miRNA-21 in der HCC827-Zelllinie**

Des Weiteren sollte der Einfluss und das therapeutische Potential der miRNA-21 untersucht werden. Das *in silico*-Netzwerk wurde ebenfalls in CellDesigner 3.5.1 (Funahashi, Morohashi et al. 2003) erstellt und anschließend in SQUAD (Di Cara, Garg et al. 2007) simuliert. Hierbei wurde folgendes Skript verwendet:

1. Unbehandelte Zellen: initialstate=steady state 4 (SS-4); constant node pulse (parameter=state): FLIP: time=0, value=0,4; c-MET: time=0, value=0,4; active node pulse: (EGF-EGFR): state=0,8; alle anderen Knoten state=0.
2. Gefitinib-Resistenz: initialstate=steady state 4 (SS-4); constant node pulse (parameter=state): FLIP: time=0, value=0,4; c-MET: time=0, value=0,4; gefitinib: time=0, value=1; active node pulse: (EGF-EGFR): state=0,8; alle anderen Knoten state=0.
3. miRNA-21-Inhibierung: initialstate=steady state 4 (SS-4); constant node pulse (parameter=state): FLIP: time=0, value=0,4; c-MET: time=0, value=0,4; miRNA-21-Inh: time=0, value=1; active node pulse: (EGF-EGFR): state=0,8; alle anderen Knoten state=0.

**4.3 Integrierte Analysen von lncRNAs****4.3.1 Systembiologische Untersuchungen zur hypertrophie-assoziierten lncRNA Chast**

Die experimentellen Arbeiten haben die lncRNA Gm11641 (ENSMUST00000130556) als signifikant hochregulierter (3,85-fach erhöht) Kandidat infolge einer Herzhypertrophie gezeigt (hierbei wurden Mausherzen sechs Wochen dem Quer-Aorta-Verengung-Verfahren (TAC) ausgesetzt (deAlmeida, van Oort et al. 2010)). Hieran anschließende experimentelle

Untersuchungen haben herausgefunden, dass die lncRNA eine pro-hypertrophe Funktion besitzt, sodass sie als cardiac hypertrophy-associated transcript (Chast) bezeichnet wurde. Ebenso konnte gezeigt werden, dass die lncRNA Chast neben dem Herzen (in Kardiomyozyten, Fibroblasten und Endothelzellen) auch in verschiedenen anderen Geweben und Organen, etwa Blutplasma, Lunge und Leber, exprimiert ist und keine spezifische Expression im Herzen besitzt, sondern sowohl im Zellkern, Zytoplasma als auch Chromatin vorkommt. Die vorgenannten experimentellen Arbeiten wurden von Dr. Janika Viereck am Institut für Molekulare und Translationale Therapiestrategien (Leitung: Prof. Dr. Dr. med. Thomas Thum, Medizinische Hochschule Hannover) durchgeführt, für weitere Details zu den experimentellen Arbeiten siehe Viereck, Kumarswamy et al. (2016).

Die systembiologischen Analysen haben sich auf das proteinkodierende Potential, eine Sequenz-Struktur-Konservierung, den Interaktionskontext (mRNA, miRNA und Protein), eine funktionelle COG/KOGs-Annotation sowie Promotoranalyse konzentriert.

### **Proteinkodierendes Potential**

Das proteinkodierende Potential der cDNA-Sequenz wurde mit dem Coding Potential Calculator (Kong, Zhang et al. 2007) untersucht (Standardparameter gewählt).

### **Sequenz-Struktur-Konservierung**

Die cDNA-Sequenz wurde nach homologen Genen in verschiedenen Säugetieren mit dem BLAST-Algorithmus (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) analysiert (blastn, Standardparameter, E-Value $\leq$ 0,05). Die identifizierten homologen Sequenzen ausgewählter Säugetiere wurden anschließend auf ihre Sequenz-Struktur-Konservierung mit LocARNA (Smith, Heyne et al. 2010) untersucht (Parameter: Alignmenttyp: Global, Alignment Mode: Standard).

### **Interaktionspartner**

Potentielle Protein-Interaktions-Partner (PIP) wurden mit der Software catRAPID identifiziert (catRAPID omics; Parameter: Mus musculus, full-length proteins, RNA- and DNA-binding, no disordered proteins), wobei das Programm als Datengrundlage den genomweiten Proteomdatensatz der Maus (UniprotKB database, version 2012\_11) verwendet (Agostini, Zanzoni et al. 2013). CatRAPID identifiziert potentielle Protein-Interaktionen (PIen) anhand eines statistischen Hidden-Markov-Modells (HMMscan) (Agostini, Zanzoni et al. 2013). Die

gefundenen PIP wurden entsprechend dem von catRAPID empfohlenen Parameter „discriminative power“ (>75%) ausgewählt, der anhand den von catRAPID verwendeten Trainingsdatensätzen auf eine hohe Interaktionswahrscheinlichkeit hinweist (Agostini, Zanzoni et al. 2013). Die gefundenen PIP wurden ebenfalls nach ihrer biologischen Funktion und beteiligter Signalwege mit dem Cytoscape-Plug-in ClueGO untersucht (Bindea, Mlecnik et al. 2009) (Organismus Maus, p-Value<0,005, Standardparameter gewählt).

Potentielle miRNA-IP wurden mithilfe des miRanda-Algorithmus (Enright, John et al. 2003) vorhergesagt (lokal heruntergeladen; Parameter: miRNA-Suche mit Maus miRNA release 17; Gap Open Penalty: -2, Gap Extend: -8, Score Threshold: 80, Energy Threshold: -21 kcal/mol, Scaling Parameter: 2). Die identifizierten miRNAs wurden anschließend hinsichtlich relevanter miRNAs mit einer generellen Funktion im Herzen und/oder bei kardiovaskulärer Hypertrophie untersucht (miRNA-Referenzkatalog wurde mithilfe der Datenbank Wikipathways (<http://www.wikipathways.org/index.php/WikiPathways>) und Literaturrecherche zusammengestellt).

Zur Identifizierung von mRNA-IPn wurde nach antisense liegenden Genen für die lncRNA in Ensemble (<http://www.ensembl.org/index.html>) gesucht und anschließend auf eine potentielle Interaktion mit der Software IntaRNA (Smith, Heyne et al. 2010) untersucht. Hierzu wurde die Software lokal auf dem Rechner installiert und folgende Linux-Kommandozeile verwendet:

a) IntaRNA -t intarna\_target\_Chast.fasta -m NM\_183034.1\_Plekhm1.fasta -o > NM\_183034.1\_Plekhm1.result.txt und b) IntaRNA -t intarna\_target\_Chast.fasta -m NM\_183288.3\_Arhgap27.fasta -o > NM\_183288.3\_Arhgap27.result.txt.

Es wurden entsprechend der Annotation aus Ensemble die beiden proteinkodierenden Transkripte Plekhm1\_001 (ENSMUST00000041272; RefSeq: NM\_183034.1) und Arhgap27\_002 (ENSMUST00000092557; RefSeq: NM\_183288.3) verwendet, die zugehörigen RefSeq mRNA-Sequenzen wurden aus NCBI heruntergeladen (Stand 04/2016).

Potentielle regulatorische Elemente wurden mit der Software Nsite (Solovyev, Shahmuradov et al. 2010) vorhergesagt, wobei die Standardparameter verwendet wurden.

### **Funktionelle orthologe Gruppen**

Anhand einer Suche nach orthologen Gruppen (Clusters of Orthologous Groups/Eukaryotic Orthologous Groups (COG/KOGs (Tatusov, Fedorova et al. 2003)) mit dem JANE COGmaster (Liang, Schmid et al. 2009) wurde auf potentielle Funktionen hin untersucht. Entsprechende

Proteine für die orthologen Gruppen wurden aus der STRING- und EggNOG-Datenbank bezogen (Szklarczyk, Franceschini et al. 2015) (Huerta-Cepas, Szklarczyk et al. 2016), funktionelle orthologe Interaktionsgruppen wurden über die STRING-Datenbank identifiziert.

### **Promotoranalyse**

Als Promotorsequenz wurde der Sequenzbereich -499 bis +100 bp (mm10\_chr11: 103,362,714 – 103,363,313) relativ zur Transkriptionsstartstelle (TSS) ausgewählt. Die Sequenz wurde anschließend auf potentielle Transkriptionsfaktorbindestellen (TFBS) mit der Software PROMO (Alggen Server, (Messeguer, Escudero et al. 2002)) analysiert (Parameter: „Maximum matrix dissimilarity rate“ less or equal 15%). PROMO verwendet als Datengrundlage TFBS aus der TRANSFAC-Datenbank (release version 8.3) (repräsentiert als positionsspezifische Gewichtsmatrix (englisch: position weight matrices)) (Messeguer, Escudero et al. 2002).

### **Identifizierung potentieller Promotorbereiche und einer TSS**

Zur Detektion eines potentiellen Promotorbereichs und einer TSS des potentiellen humanen Homologs wurde der 5'-upstream-Bereich um 2000 bp (chr17: 62,841,244 - 62,843,536; alternate assembly CHM1\_1.1) erweitert und eine Suche mit dem BDGP Promoter Prediction Tool durchgeführt, welches auf neuronalen Netzwerken basiert (Reese 2001) (Standardparameter; Organismus: Eukaryoten, ohne Gegenstrang, Promotor Score>0,8 (80%)). Um den genauen Promotorbereich zu identifizieren, wurde zusätzlich ein Sequenzalignment gegen den Chast-Promotor mithilfe von BLAST (bl2seq; <http://blast.ncbi.nlm.nih.gov/Blast.cgi>) durchgeführt.

### **4.3.2 Systembiologische Untersuchungen zu den hypoxie-induzierten lncRNAs LINC00323-003 und MIR503HG**

In den experimentellen Arbeiten wurden menschliche HUVEC-Zellen 24h hypoxischen/non-hypoxischen Bedingungen ausgesetzt, wobei die beiden intergenisch liegenden lncRNAs LINC00323-003 (lnc-DSCAM-1) und MIR503HG (lnc-PLAC1-1) als signifikant hochregulierte lncRNAs (ca. 8-fach hochreguliert) nach Hypoxie identifiziert wurden. Es konnte experimentell gezeigt werden, dass die LINC00323-003 eine zytoplasmatische Expression besitzt, wohingegen die lncRNA MIR503HG im Zellkern und Zytoplasma exprimiert ist. Ebenfalls wurde herausgefunden, dass beide lncRNAs zelltyp-spezifisch in



Endothelzellen und vaskularisierten Organen, z.B. Nieren, nach Hypoxie exprimiert sind, nicht aber in kardiovaskulären Fibroblasten oder glatten Muskelzellen der Aorta. Die experimentellen Arbeiten wurden von Dr. Jan Fiedler am Institut für Molekulare und Translationale Therapiestrategien (Leitung: Prof. Dr. Dr. med. Thomas Thum, Medizinische Hochschule Hannover) durchgeführt, für weitere Details zu den experimentellen Arbeiten siehe Fiedler, Breckwoldt et al. (2015).

Da es in diesem Projekt um eine erste experimentelle Identifizierung potentieller hypoxie-induzierter lncRNAs in Endothelzellen ging, haben sich die durchgeführten systembiologischen Analysen lediglich auf das proteinkodierende Potential und eine Sequenz-Struktur-Konservierung konzentriert. Diese wurden entsprechend der Analysen zu Chast aus Punkt 4.3.1 durchgeführt.

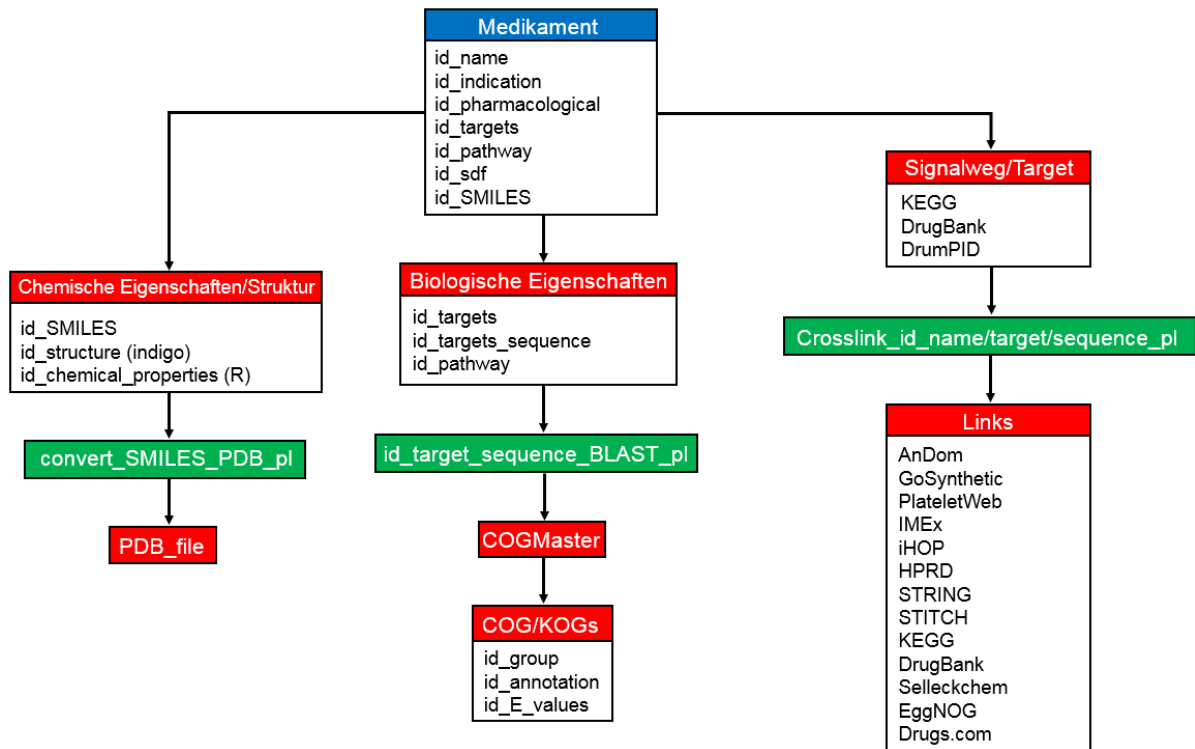
Alle verwendeten Sequenzen zu Punkt 4.3 wurden aus den Datenbanken Ensemble und NCBI entnommen.

#### **4.4 Analysen zur DrumPID-Datenbank**

Abbildung 4.4 gibt einen Überblick über die verwendeten Daten (als blauer Kasten) und eigenen Berechnungen (als roter und grüner Kasten) der DrumPID-Datenbank. Es ist anzumerken, dass Vorarbeiten bereits in meiner Masterarbeit durchgeführt wurden, die Datenbank jedoch im Rahmen der Dissertation grundlegend bearbeitet und erweitert wurde. Zusätzlich wurden die Ergebnisse der endgültigen Berechnungen inkludiert.

Als Datengrundlage (Abb. 4.4, blauer Kasten) wurden die sdf- und SMILES-Annotation (Simplified Molecular Input Line Entry Specification) für FDA-zugelassene Medikamente aus der DrugBank-Datenbank (Knox, Law et al. 2011) verwendet. Entsprechende pharmakologische Informationen und Beschreibungen zur Indikation für jedes Medikament wurden aus den Datenbanken DrugBank (Knox, Law et al. 2011) und Drugs.com (<http://www.drugs.com/>) verwendet, alle bekannten Zielproteine und zugehörigen Signalnetzwerke für jedes Medikament wurden aus den Datenbanken DrugBank (Knox, Law et al. 2011) und KEGG (Kanehisa, Sato et al. 2016) heruntergeladen. In einem ersten Schritt wurden basierend auf der SMILES-Annotation spezifische chemische Deskriptoren und Eigenschaften unter Verwendung des Chemoinformatik-R-Pakets rcdk (molekulare und atomic

descriptors) (R. 2007) berechnet (Abb. 4.4, roter Kasten Links). Zusätzlich wurde für jedes Medikament die entsprechende Struktur (als SVG Format) mittels indigo-depict aus dem Chemoinformatik-Programm indigo (<http://lifescience.opensource.epam.com/indigo/>) berechnet. Mittels eines Anbindungsskripts (Abb. 4.4, grüner Kasten Links) in Perl wurde der SMILES-String in eine PDB-Struktur konvertiert (als Download für Nutzer). In einem weiteren Schritt wurde basierend auf der Proteinsequenz für jedes Zielprotein die orthologe Gruppe (COG/KOGs (Tatusov, Fedorova et al. 2003)) unter Verwendung des COGmasters aus dem JANE-Paket (Liang, Schmid et al. 2009) berechnet (Abb. 4.4, roter und grüner Kasten Mitte). Zusätzlich wurde eine Ähnlichkeitssuche für SMILES-Annotationen mittels der Fingerprint-Funktion aus dem Chemoinformatik-R-Paket rcdk (fingerprint function; 'tanimoto' similarity metric method, Threshold>0,66) (R. 2007) eingebaut. Alle heruntergeladenen und berechneten Daten wurden in einer MySQL-Datenbank (phpMyAdmin ([http://www.phpmyadmin.net/home\\_page/index.php](http://www.phpmyadmin.net/home_page/index.php))) gespeichert und jedem Medikament eine individuelle ID-Nummer zugeordnet. DrumPID bietet zusätzlich zahlreiche Links (Abb. 4.4, roter und grüner Kasten Rechts) zu zahlreichen Datenbanken und Programmen für weiterführende Informationen und Analysen (etwa Datenbanken PlateletWeb (Boyanova, Nilla et al. 2012) und GoSynthetic (Liang, Kruger et al. 2013)). Die entsprechende Gestaltung der Benutzeroberfläche erfolgte mit den Programmiersprachen HTML (<http://www.w3.org/html/>) und Java (<http://www.oracle.com/technetwork/java/index.html>), als Betriebssysteme wurden Windows und Linux verwendet. DrumPID ist unentgeltlich zugänglich unter <http://drumpid.bioapps.biozentrum.uni-wuerzburg.de/compounds/index.php>.



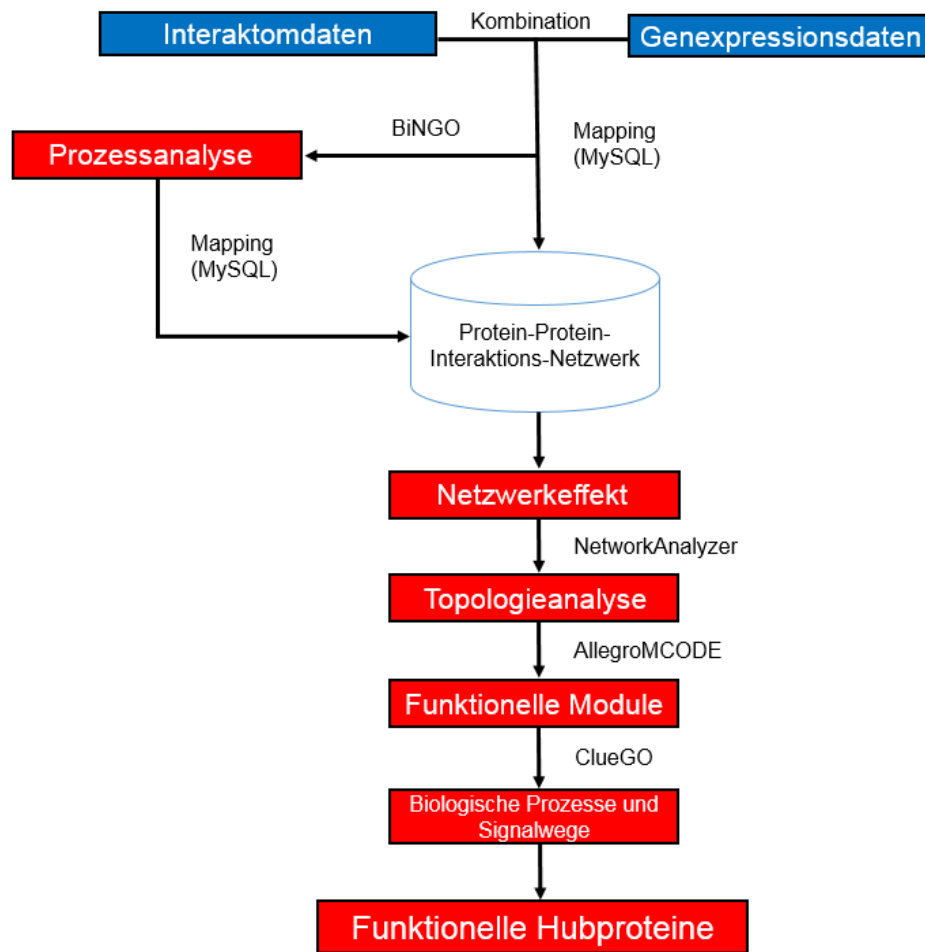
**Abbildung 4.4: Überblick über die DrugPID-Datenbank.** Die Abbildung fasst die verwendeten Daten (als blauer Kasten) und darauf basierenden eigenen Berechnungen (als roter und grüner Kasten) zusammen. (Details siehe Text; Abbildung wurde in ähnlicher Form in Kunz, Liang et al. (2016) veröffentlicht.)

## 5 Ergebnisse

Alle Daten und Materialien zu den in der Thesis dargestellten Ergebnissen sind zusätzlich in den Anhängen (supplementary material) der entsprechenden eigenen Publikationen zu finden. Zusätzlich habe ich eine CD zur Doktorarbeit angefertigt, die auch noch einmal alle Daten der Thesis gesammelt hat.

### 5.1 Integrierte Netzwerkanalysen am Beispiel von *A. thaliana* infiziert mit *Pst* DC3000

Um experimentelle Genexpressionsdatensätze hinsichtlich ihres Interaktionskontextes und ihrer biologischen Funktion zu untersuchen, um darauf basierend entsprechend relevante funktionelle Hubproteine zu detektieren, wurde eine eigenständige integrierte systembiologische Methode entwickelt (Naseem, Kunz et al. 2014) (Kunz, Dandekar et al. 2016). Die entwickelte Methode ist in Abbildung 5.1A vereinfacht dargestellt. Die Methode ist in der Lage, verschiedene Datenebenen (als blaue Kästen), in diesem Falle Genexpression (Transkriptom, z.B. aus Microarray) und Interaktom (z.B. aus AINM), zu kombinieren. Die überlappenden Gene zwischen dem Transkriptom (Genexpression) und Interaktom bilden ein PPI-Netzwerk, welches anschließend netzwerkbiologisch und funktionell (z.B. NetworkAnalyzer und AllegroMCODE) untersucht werden kann, um funktionelle Hubproteine zu identifizieren (Abb. 5.1A, rote Kästen). Die Methode beinhaltet zusätzlich eine weitere funktionelle Prozessanalyse der experimentellen Daten mittels BiNGO (Abb. 5.1A, roter Kasten Links), was eine vorherige Reduktion der Gene/Proteine ermöglicht, anhand derer ein funktionelles PPI-Netzwerk erstellt und anschließend hinsichtlich funktioneller Hubproteine untersucht werden kann. Es sei angemerkt, dass die Methode einen Überblick über mögliche Analyseschritte gibt, welche selbstverständlich je nach der jeweiligen wissenschaftlichen Fragestellung individuell geändert und angewendet werden können.



**Abbildung 5.1A: Überblick über die entwickelte systembiologische Methode.** Die Abbildung fasst die entwickelte Methode zusammen, welche verschiedene Datenebenen (als blauer Kasten) mit netzwerkbiologischen und funktionellen Analysen (als roter Kasten) kombiniert, um funktionelle Hubproteine zu identifizieren. (Details siehe Text; Abbildung wurde in ähnlicher Form in Kunz, Dandekar et al. (2016) veröffentlicht.)

In dieser Arbeit sollten mithilfe der entwickelten Methode eine Infektion von *A. thaliana* mit dem Pathogen *Pst* DC3000 untersucht werden und neue funktionelle Hubproteine und Modulatoren einer CK-vermittelten Immunabwehr gefunden werden, deren Ergebnisse im Folgenden vorgestellt werden.

Als Transkriptomdaten dienten Genexpressionsdaten von *A. thaliana* infiziert mit *Pst* DC3000 vs. Mock (24h), *Pst* COR<sup>-</sup> vs. Mock (24h) und *Pst* COR<sup>-</sup> *hrpS* vs. Mock (24h) (GSE5520). Der Datensatz *Pst* DC3000 vs. Mock besitzt 4551 DEGe (Abb. 5.1B (Oben, Links)), wobei hiervon 1023 eine Überlappung mit dem Interaktom zeigen (Transkriptom-Interaktom-Mapping). Das aus den 1023 DEGen rekonstruierte EBN besteht aus 2554 Knoten und 3956 Interaktionen (Abb. 5.1C (Oben, Links)); DEGe in rot; IP aus dem Interaktom in blau; Interaktionen in grau).

Der Datensatz *Pst* COR<sup>-</sup> vs. Mock besitzt 2328 DEGe (Abb. 5.1B (Oben, Mitte)), wobei 491 eine Überlappung mit dem Interaktom zeigen. Das rekonstruierte EBN basierend auf den 491 DEGen verfügt über 1628 Knoten und 2142 Interaktionen (Abb. 5.1C (Oben, Mitte); DEGe in rot; IP aus dem Interaktom in blau; Interaktionen in grau).

Der Datensatz der Doppelmutante *Pst* COR<sup>-</sup> *hrpS* vs. Mock besitzt 609 DEGe (Abb. 5.1B (Oben, Rechts)), wobei das Transkriptom-Interaktom-Mapping eine Überlappung von 59 Knoten gezeigt hat, was in einem EBN von 374 Knoten und 348 Interaktionen resultiert (Abb. 5.1C (Oben, Rechts); DEGe in rot; IP aus dem Interaktom in blau; Interaktionen in grau).

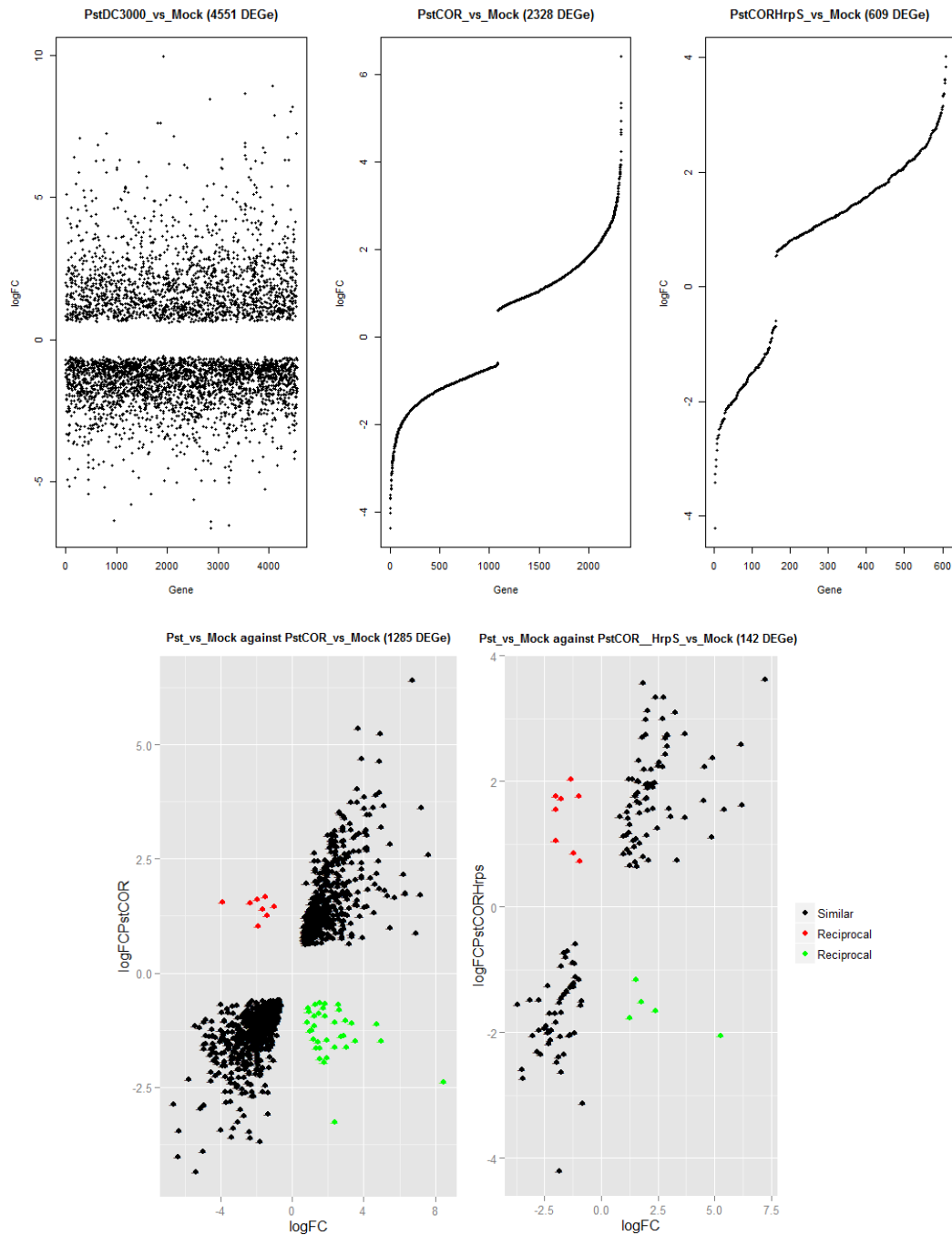
Von den DEGen der einzelnen Mutanten zeigt sich eine Überlappung von 1285 DEGen (43 reziprok reguliert) zwischen *Pst* DC3000 und *Pst* COR<sup>-</sup> und 142 DEGen (13 reziprok reguliert) zwischen *Pst* DC3000 und *Pst* COR<sup>-</sup> *hrpS* (Abb. 5.1B (Unten); Links: DEGe *Pst* DC3000 und *Pst* COR<sup>-</sup>; Rechts: DEGe *Pst* DC3000 und *Pst* COR<sup>-</sup> *hrpS*; DEGe mit derselben Regulierung in schwarz, DEGe mit reziproker Regulierung in rot und grün).

Das Mapping der Knoten der EBNe ist in einem Venn-Diagramm in Abbildung 5.1C (Unten, Links) zusammengefasst (*Pst* DC3000 in blau: 2554 Knoten; *Pst* COR<sup>-</sup> in gelb: 1628 Knoten; *Pst* COR<sup>-</sup> *hrpS* in grün: 374 Knoten).

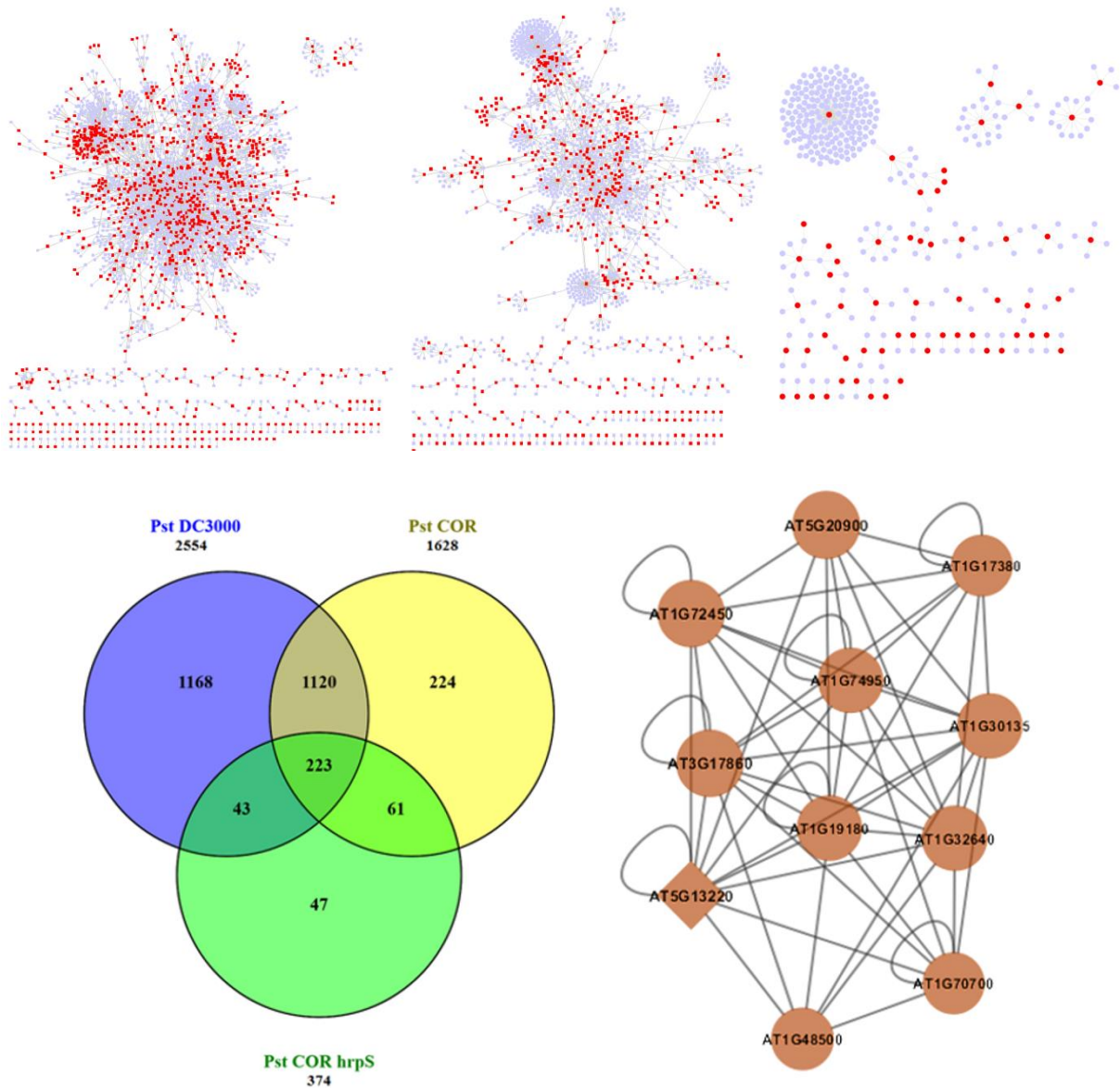
Das Mapping der Knoten der EBNe gegen 137 in *A. thaliana* bekannte Effektorproteine hat eine Überlappung von 115 Knoten für *Pst* DC3000, 99 für *Pst* COR<sup>-</sup> und 21 für *Pst* COR<sup>-</sup> *hrpS* ergeben, wobei sich 17 Effektorproteine in allen drei EBNe finden lassen (hier nicht gezeigt).

Als Nächstes wurden die rekonstruierten EBNe der drei Datensätze auf ihre Netzwerktopologie und funktionellen Cluster untersucht. Die Topologieanalyse ist in Tabelle 5.1A zusammengefasst, wobei sich auf die wichtigen Parameter „durchschnittliche Anzahl an Interaktionsnachbarn“, „Netzwerkzentralität“ und „Netzwerkhomogenität“ konzentriert wurde.

Die Clusteranalyse hat für das *Pst* DC3000-regulierte EBN 27 funktionelle Cluster mit einem Score>1 gefunden. Hierbei hat ein Cluster um JAZ (Score=4,182; elf Proteine; in braun) die stärkste Vernetzung mit einer Immunfunktion gezeigt (Abb. 5.1C Unten, Rechts). Für das *Pst* COR<sup>-</sup>-regulierte EBN zeigen sich zwölf Cluster, wobei ebenfalls ein Cluster um JAZ (Score=2,556; neun Proteine) gefunden wurde (hier nicht dargestellt). Für die Doppelmutante *Pst* COR<sup>-</sup> *hrpS* wurden zwei Cluster mit einem Score>1 gefunden, allerdings hat sich kein Cluster um JAZ gezeigt.



**Abbildung 5.1B: Darstellung der differentiell-exprimierten Gene der verwendeten Datensätze. (Oben)** Die Plots fassen die Genexpression (logFC auf y-Achse) der DEGe (logFC  $\geq +0,5$  /  $\leq -0,5$ ; p-Value  $< 0,05$ ) zusammen (Gene auf x-Achse; einzelne DEGe als schwarze Punkte dargestellt). Der Datensatz *Pst* DC3000 vs. Mock verfügt über 4551 DEGe (Links), *Pst* COR<sup>-</sup> vs. Mock über 2328 DEGe (Mitte) und *Pst* COR<sup>-</sup> *hrpS* vs. Mock über 609 DEGe (Rechts). **(Unten)** Die beiden Plots fassen die überlappenden DEGe zwischen *Pst* DC3000 vs. Mock und den beiden Mutanten zusammen (Links: 1285 DEGe zwischen *Pst* DC3000 und *Pst* COR<sup>-</sup>; Rechts: 142 DEGe zwischen *Pst* DC3000 und *Pst* COR<sup>-</sup> *hrpS*; in schwarz sind DEGe mit derselben Regulierung, in rot und grün sind DEGe mit einer reziproken Regulierung zwischen den Datensätzen dargestellt; x-Achse repräsentiert jeweils die Expression (logFC) von *Pst* DC3000 vs. Mock, y-Achse die dazugehörige Expression (logFC) der Mutante; einzelne DEGe als farbige Punkte dargestellt).



**Abbildung 5.1C: Übersicht der rekonstruierten expressions-basierten Netzwerke. (Oben)** Das EBN für *Pst* DC3000 vs. Mock basierend auf den 1023 überlappenden DEGen verfügt über 2554 Knoten und 3956 Interaktionen (Links). Für *Pst* COR<sup>-</sup> vs. Mock resultieren die 491 überlappenden DEGe in 1628 Knoten und 2142 Interaktionen (Mitte). Für *Pst* COR<sup>-</sup> *hrpS* vs. Mock resultieren die 59 überlappenden DEGe in 374 Knoten und 348 Interaktionen (Rechts). (DEGe in rot; IP aus dem Interaktom in blau; Interaktionen in grau) **(Unten) Links:** Das Venn-Diagramm fasst das Mapping der Knoten der EBNe zusammen (*Pst* DC3000 vs. Mock (blau): 2554 Knoten; *Pst* COR<sup>-</sup> vs. Mock (gelb): 1628 Knoten; *Pst* COR<sup>-</sup> *hrpS* vs. Mock (grün): 374 Knoten). **Rechts:** Die Clusteranalyse zeigt für das von *Pst* DC3000 vs. Mock regulierte EBN ein immunrelevantes Cluster um JAZ (Score=4,182; in braun). (Details siehe Text.)



**Tabelle 5.1A: Überblick der Topologieanalyse für die rekonstruierten Netzwerke.** Die Tabelle fasst wichtige Parameter der rekonstruierten EBNe und Immunnetzwerke für die drei Datensätze aus der Topologieanalyse sowie die Anzahl an Knoten und Interaktionen für jedes Netzwerk zusammen. (IN=Immunnetzwerk; durch. Anzahl=durchschnittliche Anzahl.)

Parameter	<i>Pst</i> DC3000 vs. Mock		<i>Pst</i> COR vs. Mock		<i>Pst</i> COR <i>hrpS</i> vs. Mock	
	EBN	IN	EBN	IN	EBN	IN
<b>Knoten</b>	2554	1216	1628	925	374	30
<b>Interaktionen</b>	3956	1596	2142	1022	348	24
<b>durch. Anzahl an Nachbarn</b>	3,013	2,530	2,566	2,151	1,818	1,6
<b>Netzwerkzentralität</b>	0,051	0,108	0,110	0,194	0,483	0,200
<b>Netzwerkhomogenität</b>	1,991	2,229	2,383	3,424	5,160	0,816

In einem weiteren Schritt sollten für jeden Datensatz funktionelle Hubproteine, die sich als relevant für die Immunabwehr erweisen, identifiziert werden. Hierzu wurden die DEGe der drei Datensätze hinsichtlich biologischer Prozesse (GO) mit dem Cytoscape-Plug-in BiNGO analysiert. Hierbei zeigen sich für die 1023 DEGe aus dem Transkriptom-Interaktom-Mapping für *Pst* DC3000 vs. Mock 402 biologische Prozesse, von denen 110 Funktionen und damit resultierende 310 Gene mit dem Immunsystem assoziiert sind (15 wichtige Funktionen sind in Tabelle 5.1B dargestellt; alle 110 ausgewählten Funktionen siehe Anhang T1; 157 Gene ohne Annotation in BiNGO). Die korrespondierenden 310 Gene resultieren wiederum in einem Immunnetzwerk bestehend aus 1216 Knoten und 1596 Interaktionen (Abb. 5.1D (Links), 310 Immungene in grün, IP aus dem Interaktom in blau, Interaktionen in grau).

**Tabelle 5.1B: Übersicht der immunrelevanten biologischen Prozesse für *Pst* DC3000 vs. Mock.** Die funktionelle Analyse hat 402 biologische Prozesse ergeben, wobei 110 Funktionen mit dem Immunsystem assoziiert sind (157 Gene ohne BiNGO-Annotation). Die Tabelle fasst wichtige immunrelevante Funktionen zusammen (alle ausgewählten Funktionen siehe Anhang T1). (x=Anzahl der korrespondierenden Proteine für den biologischen Prozess; n=Anzahl aller annotierten Proteine in BiNGO; p-Value=Wahrscheinlichkeit, ein identisches Ergebnis zu erzielen.)

<b>p-Value</b>	<b>x</b>	<b>n</b>	<b>Biologischer Prozess (GO)</b>
1,43E-05	9	39	cytokinin mediated signaling pathway
7,09E-08	54	637	defense response
8,24E-09	27	193	defense response to bacterium
2,15E-04	13	105	defense response, incompatible interaction
1,02E-02	19	272	immune response
8,58E-03	2	4	inflammatory response
1,19E-02	18	257	innate immune response
8,75E-04	9	65	regulation of defense response
3,25E-10	33	241	response to bacterium
3,19E-11	55	528	response to other organism
5,34E-55	304	3207	response to stimulus
2,13E-27	171	1853	response to stress
7,92E-19	34	133	response to wounding
6,40E-03	5	31	systemic acquired resistance
8,70E-04	4	12	systemic acquired resistance, salicylic acid mediated signaling pathway

Für die Mutante *Pst* COR<sup>-</sup> vs. Mock ergaben sich für die 491 überlappenden DEGe 207 biologische Prozesse, wobei 82 Funktionen mit insgesamt 148 Genen eine immunrelevante Funktion besitzen (15 wichtige Funktionen sind in Tabelle 5.1C dargestellt; alle 82 ausgewählten Funktionen siehe Anhang T2; 77 Gene ohne BiNGO-Annotation). Das entsprechende Immunnetzwerk verfügt über 925 Knoten und 1022 Interaktionen (Abb. 5.1D (Mitte), 148 Immungene in grün, IP aus dem Interaktom in blau, Interaktionen in grau).

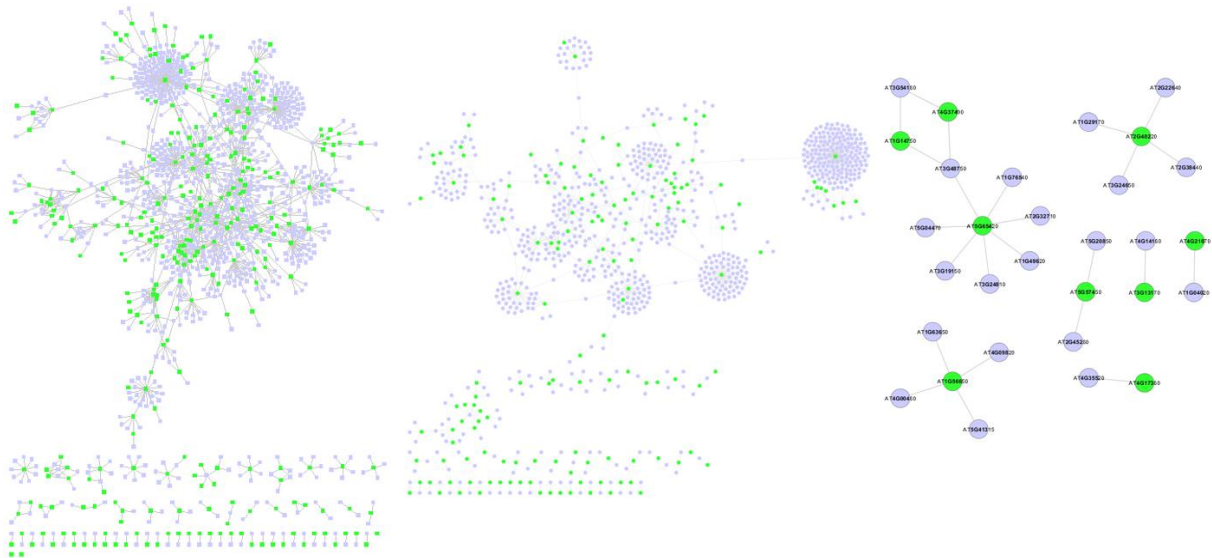
Für *Pst* COR<sup>-</sup> *hrpS* vs. Mock hingegen zeigen neun der 59 überlappenden DEGe aus dem Transkriptom-Interaktom-Mapping eine immunrelevante Funktion (95 biologische Prozesse, davon zwölf assoziiert mit dem Immunsystem; Tab. 5.1D; sechs Gene ohne BiNGO-Annotation). Das Immunnetzwerk zeigt hier 30 Knoten und 24 Interaktionen (Abb. 5.1D (Rechts), neun Immungene in grün, IP aus dem Interaktom in blau, Interaktionen in grau).

**Tabelle 5.1C: Übersicht der immunrelevanten biologischen Prozesse für *Pst* COR<sup>-</sup> vs. Mock.** Die funktionelle Analyse zeigt 207 biologische Prozesse, von denen 82 eine Verbindung mit dem Immunsystem besitzen (77 Gene ohne BiNGO-Annotation). Die Tabelle fasst wichtige immunrelevante Funktionen zusammen (alle ausgewählten Funktionen siehe Anhang T2). (x=Anzahl der korrespondierenden Proteine für den biologischen Prozess; n=Anzahl aller annotierten Proteine in BiNGO; p-Value=Wahrscheinlichkeit, ein identisches Ergebnis zu erzielen.)

p-Value	x	n	Biologischer Prozess (GO)
9,54E-06	29	637	defense response
2,32E-08	18	193	defense response to bacterium
6,04E-04	14	272	immune response
2,86E-04	15	283	immune system process
3,44E-04	14	257	innate immune response
2,87E-08	11	65	regulation of defense response
9,41E-06	7	41	regulation of immune response
4,91E-03	2	6	regulation of plant-type hypersensitive response
2,05E-08	12	78	regulation of response to stress
1,40E-08	32	550	response to biotic stimulus
5,30E-09	32	528	response to other organism
2,26E-26	145	3207	response to stimulus
1,91E-07	14	133	response to wounding
4,65E-07	7	27	salicylic acid mediated signaling pathway
1,23E-03	3	12	systemic acquired resistance, salicylic acid mediated signaling pathway

**Tabelle 5.1D: Übersicht der immunrelevanten biologischen Prozesse für *Pst* COR<sup>-</sup> *hrpS* vs. Mock.** Die funktionelle Analyse hat 95 biologische Prozesse ergeben (sechs Gene ohne Annotation in BiNGO). Die Tabelle fasst die ausgewählten zwölf mit dem Immunsystem verbundenen Funktionen zusammen. (x=Anzahl der korrespondierenden Proteine für den biologischen Prozess; n=Anzahl aller annotierten Proteine in BiNGO; p-Value=Wahrscheinlichkeit, ein identisches Ergebnis zu erzielen.)

p-Value	x	n	Biologischer Prozess (GO)
2,16E-03	2	29	cellular response to carbohydrate stimulus
4,75E-03	1	2	cellular response to disaccharide stimulus
1,61E-03	7	729	cellular response to stimulus
2,18E-04	6	364	cellular response to stress
4,75E-03	1	2	cellular response to sucrose stimulus
8,61E-03	3	177	response to carbohydrate stimulus
8,56E-05	3	36	response to disaccharide stimulus
4,24E-05	5	163	response to DNA damage stimulus
2,46E-04	2	10	response to gamma radiation
6,51E-04	2	16	response to ionizing radiation
7,86E-05	3	35	response to sucrose stimulus
2,38E-03	1	1	response to trehalose stimulus



**Abbildung 5.1D: Überblick der rekonstruierten Immunnetzwerke. (Links)** Für *Pst* DC3000 vs. Mock resultieren die 310 korrespondierenden Gene für 110 immunrelevante Funktionen in einem Immunnetzwerk aus 1216 Knoten und 1596 Interaktionen. **(Mitte)** Bei der Mutante *Pst* COR<sup>-</sup> vs. Mock ergeben die 148 assoziierten Immungene aus 82 immunrelevanten Funktionen ein Immunnetzwerk mit 925 Knoten und 1022 Interaktionen. **(Rechts)** Für die Doppelmutante *Pst* COR<sup>-</sup> *hrpS* vs. Mock zeigt das basierend auf den neun Immungenen aus zwölf immunrelevanten Funktionen rekonstruierte Immunnetzwerk 30 Knoten und 24 Interaktionen. (Immungene in grün, IP aus Interaktom in blau, Interaktionen in grau; Details siehe Text.)

Die Immunnetzwerke wurden anschließend auf ihre Topologie untersucht, was in Tabelle 5.1A zusammengefasst ist. Anhand einer hohen Anzahl an IPn ( $\geq 4$ ) wurden in den Immunnetzwerken funktionelle Hubproteine für die pflanzliche Immunabwehr detektiert. Insgesamt haben sich für *Pst* DC3000 vs. Mock 185, für *Pst* COR<sup>-</sup> vs. Mock 73 und für *Pst* COR<sup>-</sup> *hrpS* vs. Mock drei Hubproteine gezeigt (siehe Anhang T3). Tabelle 5.1E fasst wichtige Hubknoten mit einer immunrelevanten Funktion für *Pst* DC3000 und überlappende Hubproteine mit den beiden Mutanten zusammen. Es sei angemerkt, dass sich die angegebene Genexpression (logFC) auf die Arraydaten beziehen, wobei sich teils keine signifikante Genexpression in den Arraydaten gezeigt hat, was bedeutet, dass der Hubknoten ein entsprechender IP der DEGe ist (die Anzahl der Interaktionen bezieht sich auf das jeweilige Immunnetzwerk).

**Tabelle 5.1E: Überblick der identifizierten funktionellen Hubproteine für die Immunabwehr.** Die Tabelle fasst gefundene immunrelevante Hubproteine für *Pst* DC3000 vs. Mock und überlappende Hubproteine mit den beiden Mutanten mit wichtigen Parametern und der Annotation zusammen. In rot sind Hubknoten markiert, welche in *Pst* DC3000 vs. Mock und *Pst* COR<sup>-</sup> vs. Mock vorhanden sind, in grün hinterlegt repräsentiert Hubproteine zwischen allen drei Datensätzen. (Anz.=Anzahl der IP; p-Value=Wahrscheinlichkeit, ein identisches Ergebnis zu erzielen.)

Hubprotein	Anz.	p-Value	logFC	Name	Beschreibung
<b><i>Pst</i> DC3000</b>					
AT1G80840	7	3,56E-04	3,86	WRKY40	putative WRKY transcription factor 40
AT2G40670	6	6,96E-03	-1,33	ARR16	two-component response regulator ARR16
AT3G63010	6	4,95E-04	2,33	GID1B	gibberellin receptor GID1B
AT5G57050	13	1,70E-02	1,05	ABI2	protein phosphatase 2C 77
AT1G32640	12	3,45E-06	3,95	MYC2	transcription factor MYC2
AT3G57040	7	2,65E-03	-1,27	ARR9	two-component response regulator ARR9
AT1G10470	11	1,18E-02	-1,25	ARR4	two-component response regulator ARR4
AT5G13930	6	1,60E-04	6,30	TT4	chalcone synthase
AT1G64280	15	2,65E-02	1,12	NPR1	Regulatory protein NPR1
AT3G16857	10	2,68E-02	-1,11	ARR1	response regulator 1
AT5G13220	14	1,15E-07	6,93	JAZ10	jasmonate-zim-domain protein 10
AT1G73500	8	9,66E-05	2,08	MKK9	mitogen-activated protein kinase 9
AT4G28720	5	4,68E-03	-1,68	YUC8	flavin-containing monooxygenase YUCCA8
AT4G26080	23	8,89E-05	2,28	ABI1	protein phosphatase 2C 56
AT5G62920	5	1,52E-03	-2,15	ARR6	two-component response regulator ARR6
AT5G13080	5	9,44E-03	2,76	WRKY75	putative WRKY transcription factor 75
AT1G19180	21	2,66E-05	4,59	JAZ1	jasmonate-zim-domain protein 1
AT1G56650	4	6,59E-05	6,21	PAP1	production of anthocyanin pigment 1, MYB75
AT4G23650	7	4,79E-02	0,62	CDPK6	calcium-dependent protein kinase 6
AT2G01760	59	6,69E-03	-1,59	ARR14	two-component response regulator ARR14
AT4G39400	11	1,30E-02	-1,03	BRI1	protein brassinosteroid insensitive 1
AT1G74890	7	2,22E-04	-3,67	ARR15	two-component response regulator ARR15
AT1G59940	10	5,00E-02	-1,47	ARR3	response regulator 3
AT5G35750	5	8,08E-02	-1,68	AHK2	histidine kinase 2
AT3G21510	6	3,52E-01	0,26	AHP1	histidine-containing phosphotransmitter 1
AT3G29350	6	5,98E-01	-0,14	AHP2	histidine-containing phosphotransfer protein 2
AT5G39340	18	3,85E-02	-0,71	AHP3	histidine-containing phosphotransfer protein 3
AT1G03430	9	2,58E-01	-0,39	AHP5	histidine-containing phosphotransfer protein 5
<b><i>Pst</i> COR<sup>-</sup></b>					
AT1G80840	7	2,29E-02	2,41	WRKY40	putative WRKY transcription factor 40
AT3G63010	6	4,37E-02	0,95	GID1B	gibberellin receptor GID1B
AT1G32640	12	4,19E-02	0,78	MYC2	transcription factor MYC2
AT1G64280	15	6,53E-03	1,19	NPR1	Regulatory protein NPR1
AT5G13220	14	1,62E-02	0,86	JAZ10	jasmonate-zim-domain protein 10
AT1G73500	8	5,46E-03	0,99	MKK9	mitogen-activated protein kinase 9
AT4G28720	5	1,90E-02	-1,06	YUC8	flavin-containing monooxygenase YUCCA8
AT4G26080	23	2,36E-02	0,76	ABI1	protein phosphatase 2C 56
AT1G19180	21	2,19E-02	1,32	JAZ1	jasmonate-zim-domain protein 1
AT1G56650	4	1,16E-02	2,17	PAP1	production of anthocyanin pigment 1, MYB75
AT4G23650	7	4,78E-03	1,02	CDPK6	calcium-dependent protein kinase 6
AT2G01760	59	1,20E-02	-1,08	ARR14	two-component response regulator ARR14
AT4G39400	11	4,72E-02	-0,65	BRI1	protein brassinosteroid insensitive 1
<b><i>Pst</i> COR<sup>-</sup> hrpS</b>					
AT1G56650	4	4,25E-02	1,62	PAP1	production of anthocyanin pigment 1, MYB75

## **5.2 Dynamische Modellierung am Beispiel des *in silico/in vitro*-Lungentumormodells**

### **5.2.1 Integrierte systembiologische Analysen der experimentellen Arraydaten**

Um eine Korrelation eines 3D-Tumortestsystems basierend auf der SISmuc mit herkömmlichen 2D-Tumortestsystemen und klinischen Beobachtungen zu untersuchen, wurden drei verschiedene NSCLC-Zelllinien (HCC827, A549 und H441) mit dem HSP90-Inhibitor 17AAG behandelt. Die systembiologischen Analysen haben sich auf zwei Teile konzentriert. In einem ersten Teil wurden die Behandlungsunterschiede zwischen dem 2D- und 3D-System untersucht, wohingegen im zweiten Teil für das 3D-System potentielle Marker für ein Behandlungsansprechen und therapeutische Targets für die beiden KRAS-mutierten Zelllinien A549 und H441 gefunden werden sollten.

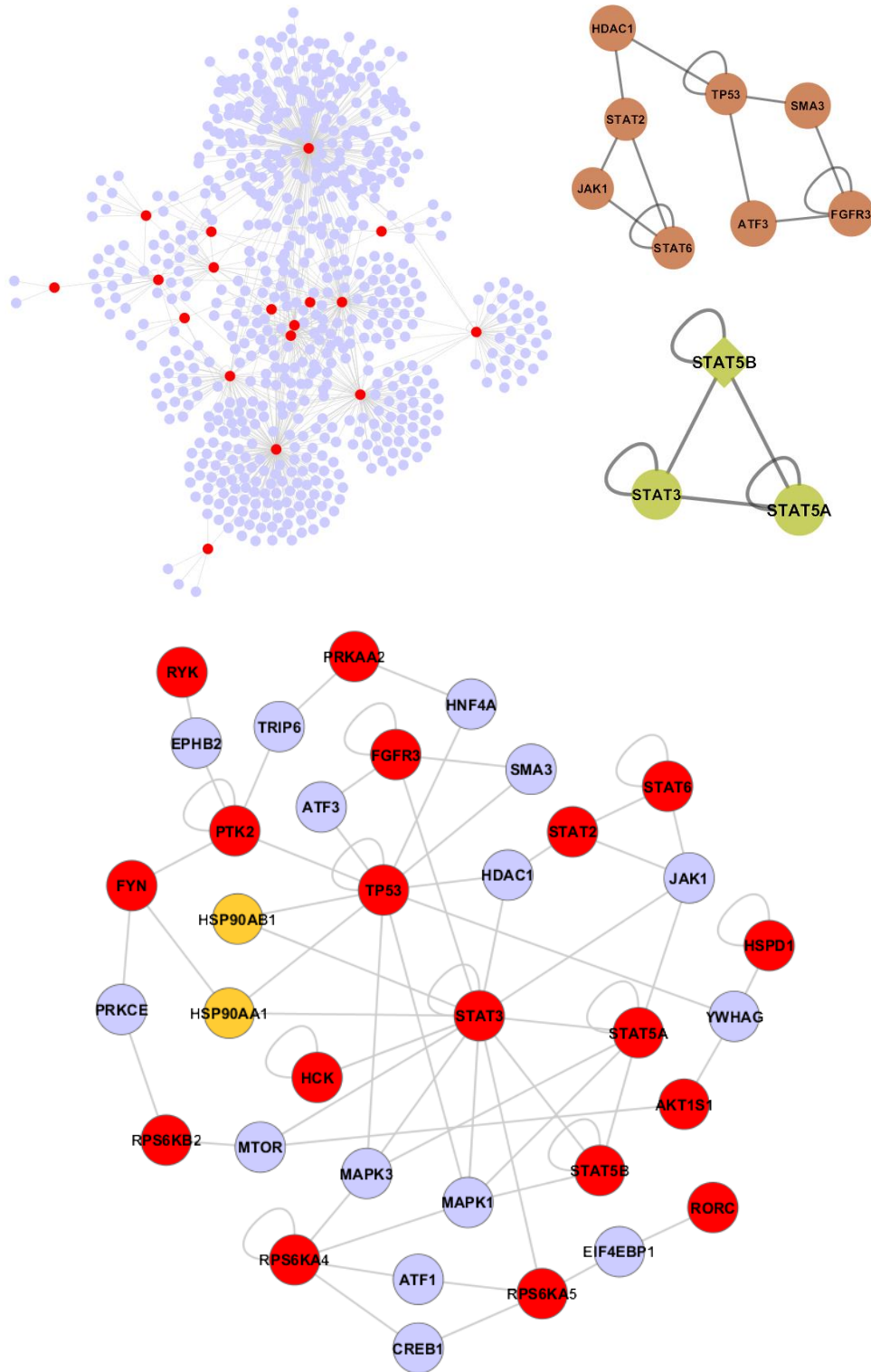
#### **Analyse hinsichtlich der Behandlungsunterschiede zwischen dem 2D- und 3D-System**

Für die Unterschiede im Verhalten zwischen dem 2D- und 3D-System wurden DRPe (Signifikanz  $2D \pm 0,5\%$ ,  $3D \pm 5\%$ ) ausgewählt, die eine reziproke Regulierung zeigen bzw. mindestens in einem der beiden Systeme reguliert sind. Tabelle 5.2.1A fasst die 18 ausgewählten DRPe (STAT5a/b zu STAT5a und STAT5b zugeordnet) zusammen.

**Tabelle 5.2.1A: Übersicht der ausgewählten differentiell-regulierten Proteine zwischen dem 2D- und 3D-System.** Die Tabelle fasst die ausgewählten DRPe (Signifikanz  $2D \pm 0,5\%$ ,  $3D \pm 5\%$ ) zwischen dem 2D- und 3D-Zellkultursystem zusammen (Werte repräsentieren Differenz aus behandelt minus unbehandelt; in grün sind hochregulierte DRPe nach 17AAG-Behandlung, in rot entsprechend herunterregulierte DRPe, nicht farblich hinterlegt stellen nicht signifikant DRPe dar; nr=nicht reguliert vor und nach Behandlung).

	A549		HCC827		H441	
	2D	3D	2D	3D	2D	3D
p53	1,3%	1,5%	-7,6%	-12,6%	-5,3%	11,1%
MSK1/2	-1,4%	-2,9%	-1,0%	2,0%	-1,1%	-4,5%
AMPK $\alpha$ 1	-0,9%	-3,4%	-0,4%	4,5%	-0,5%	-5,4%
TOR	-1,1%	1,4%	-1,0%	0,2%	-0,5%	2,1%
AMPK $\alpha$ 2	-1,0%	-3,0%	-0,7%	7,4%	-0,3%	-0,8%
p70 S6 Kinase	-0,6%	0,0%	-0,2%	0,1%	-0,7%	0,3%
STAT2	-0,7%	-1,8%	-0,5%	9,2%	-0,4%	1,6%
STAT5a	0,2%	-2,7%	-0,3%	5,2%	-0,4%	0,5%
Fyn	-0,9%	0,5%	-0,7%	0,3%	-0,1%	1,8%
STAT6	-0,7%	-2,4%	-0,3%	6,8%	-0,1%	-1,4%
STAT5b	-0,4%	-1,6%	-0,8%	3,5%	-0,4%	-1,1%
Hck	-0,9%	1,3%	-0,4%	2,5%	-0,4%	3,0%
FAK	-1,7%	-2,6%	-1,1%	2,7%	-0,7%	-1,2%
STAT5a/b	-0,6%	-0,8%	-0,7%	5,7%	-0,5%	-1,9%
STAT3	-0,6%	-1,3%	0,1%	-0,3%	-0,9%	0,3%
PRAS40	-0,7%	-10,9%	0,4%	-1,9%	0,4%	-2,4%
HSP60	-4,1%	10,5%	-6,7%	23,5%	5,1%	32,7%
pFGF R3	nr	nr	nr	nr	-3,5%	-5,76%
pRYK	0,5%	0,25%	-5,16%	-0,45%	nr	-1,16%

Das rekonstruierte 2D-3D-Zellkulturnetzwerk basierend auf diesen 18 DRPen (16 PK-Array, 2 RKT-Array; STAT5a/b als STAT5a und STAT5b für Rekonstruktion angesehen) und deren direkten IP aus dem Interaktom besteht aus 679 Knoten und 909 Interaktionen (Abb. 5.2.1A (Oben, Links), 18 DRPe in rot; IP in blau; Interaktionen in grau). Die funktionelle Analyse zeigt zwei Cluster um p53 (acht Proteine; Score=1,125; in braun) und die STAT-Familie (drei Proteine; Score=1; in grün), wobei sich sieben DRPe aus den Arraydaten wiederfinden (Abb. 5.2.1A (Oben, Rechts)). Verbindet man die Proteine der gefundenen Cluster und der 18 DRPe (in rot) mit den beiden HSP90-Isoformen (HSP90AA1 und HSP90AB1 in orange), so zeigt sich, dass die funktionellen Cluster eine starke Vernetzung besitzen und p53 und STAT3 IP von HSP90 sind. Es existiert allerdings auch eine Verbindung zu FYN (Abb. 5.2.1A (Unten)). Für diese gefundenen Proteine zeigt sich vor allem auch eine unterschiedliche Regulierung in den Arraydaten, z.B. Regulation von STATs und FYN in 2D und 3D (Tab. 5.2.1A).



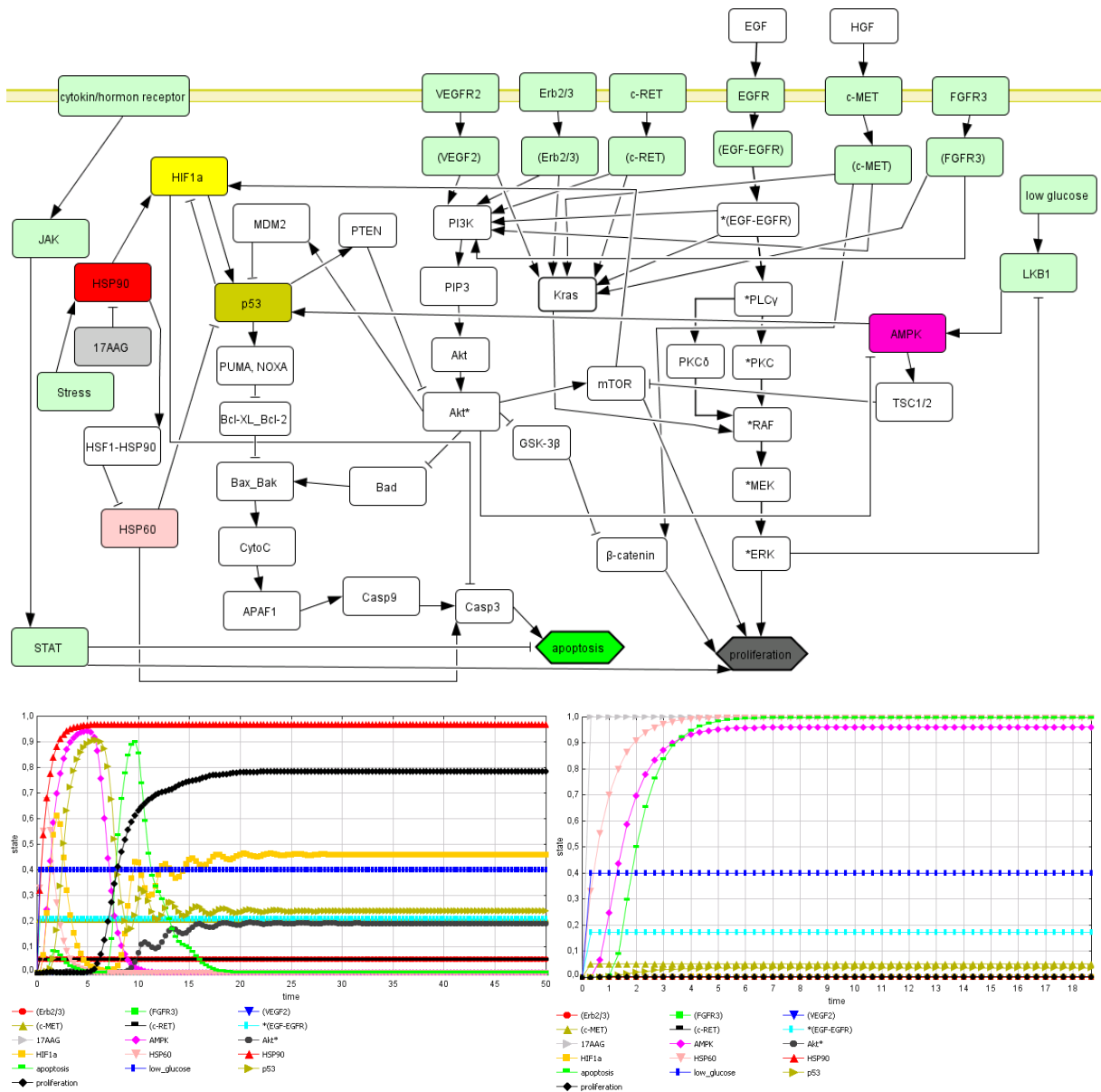
**Abbildung 5.2.1A: Überblick der Analysen zum 2D-3D-Zellkulturnetzwerk. (Oben)** Das rekonstruierte 2D-3D-Zellkulturnetzwerk basierend auf den 18 DRPen verfügt über 679 Knoten und 909 Interaktionen (18 DRPe in rot; IP aus dem Interaktom in blau; Interaktionen in grau), wobei sich zwei funktionelle Cluster um p53 (in braun) und die STAT-Familie (in grün) finden lassen. **(Unten)** Das Netzwerk zeigt die Verbindung der gefundenen Cluster und DRPe (in rot) zu HSP90 (in orange) (Clusterproteine und IP aus dem Interaktom in blau; Interaktionen in grau). (Details siehe Text.)



### ***In silico*-Simulation der HSP90-Inhibition für die HCC827-Zelllinie im 3D-System**

Da sich der stärkste therapeutische Effekt in der HCC827-Zelllinie gezeigt hat, wurden als nächster Schritt die experimentellen Daten in ein *in silico*-Tumormodell integriert und entsprechend simuliert. Die Abbildung 5.2.1B (Oben) zeigt das in CellDesigner 3.5.1 individuell erstellte Netzwerk für die HCC827-Zelllinie. Das Netzwerk wurde anschließend in SQUAD eingelesen, woran sich die *in silico*-Simulationen entsprechend der experimentellen Daten und des spezifischen Mutationshintergrundes angeschlossen haben. Die durchgeführten Simulationen sind in Abbildung 5.2.1B (Unten) dargestellt. Die Grafiken repräsentieren die Aktivierungszustände der Netzwerkknoten (y-Achse) im Zeitverlauf (x-Achse). Der Plot auf der linken Seite zeigt die unbehandelte Zelllinie (17AAG=0), wohingegen die rechte Seite die 17AAG-Behandlung (17AAG=1) verdeutlicht.

Das HCC827-Netzwerk (Abb. 5.2.1B (Oben)) beinhaltet die JAK/STAT- und LKB1/AMPK-Signalnetzwerke aufgrund der Regulierung von STAT und AMPK nach HSP90-Behandlung (Tab. 5.2.1A). Für eine Aktivierung des LKB1/AMPK-Signalwegs wurde eine Glukoseaktivierung eingebaut (siehe Punkt 4.2.2 Material und Methoden). Eine HSP90-Aktivierung wurde mit einer konstanten Aktivierung von 0,7 (über Stress) berücksichtigt, alle anderen eingebauten Rezeptorknoten entsprechend der experimentellen Daten sind im Netzwerk in grün dargestellt (z.B. EGFR und c-MET). Die durchgeführten Simulationen lassen erkennen, dass die mit 17AAG (in grau) behandelten Zellen (Rechts) eine reduzierte Proliferation (in schwarz) und erhöhte Apoptose (in hellgrün) gegenüber den unbehandelten Zellen (Links) aufzeigen, was den experimentellen Ergebnissen entspricht (Tab. 4.2.1). Des Weiteren lassen sich in den Simulationen entsprechend der experimentellen Daten ebenfalls Änderungen der Knoten p53 (in dunkelgrün), HSP60 (in lachsfarben) und AMPK (in violett) nach Behandlung erkennen (AMPK $\alpha$ 2 und HSP60 experimentell 7,4% und 23,5% erhöht, p53 12,6% erniedrigt nach Behandlung; Abb. 5.2.1B (Unten), Tab. 5.2.1A).



**Abbildung 5.2.1B: HCC827-Netzwerk und *in silico*-Simulation der 17AAG-Behandlung im 3D-System. (Oben)** Das Netzwerk veranschaulicht wichtige Signalwege (z.B. EGFR, JAK/STAT und LKB1/AMPK; in grün und violett) entsprechend der Regulierung nach HSP90-Behandlung und deren Verschaltung in der HCC827-Zelllinie. Die Readout-Parameter Proliferation (in schwarz) und Apoptose (in hellgrün) sind als Sechsecke dargestellt, der HSP90-Inhibitor 17AAG ist als graues Rechteck dargestellt; (→) repräsentieren Aktivierungen, (−) Inhibierungen. **(Unten)** Entsprechend des Mutationshintergrundes wurde eine *in silico*-Simulation durchgeführt. Die 17AAG-Behandlung (Rechts) zeigt eine erhöhte Apoptose (in hellgrün) und verminderte Proliferation (in schwarz) verglichen mit der unbehandelten Zelllinie (Links). (Details siehe Text; Input-Parameter siehe Punkt 4.2.2 Material und Methoden.)

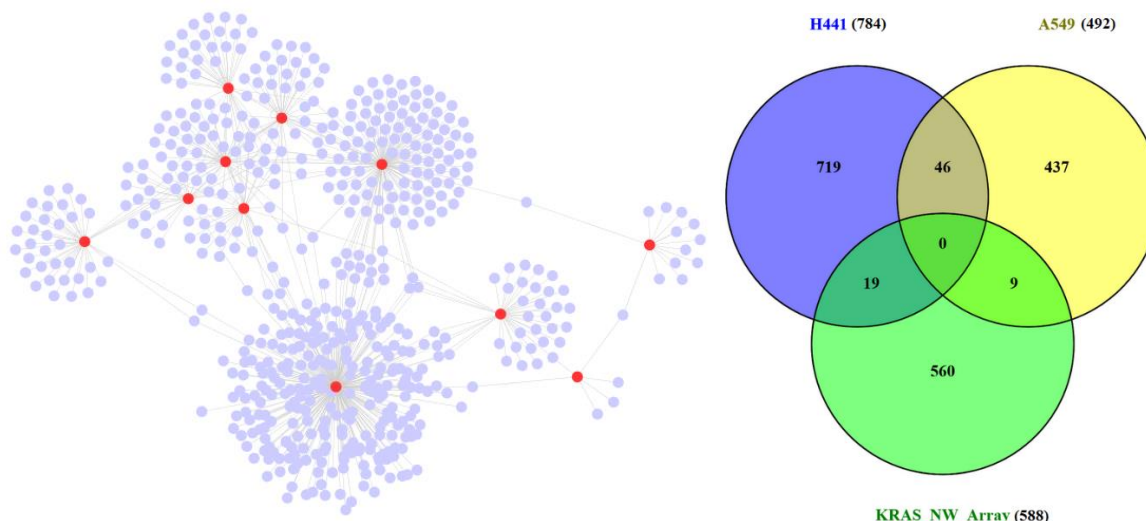
## Identifizierung neuer prognostischer Marker und therapeutischer Targets für die KRAS-mutierten Zelllinien

Die beiden KRAS-mutierten Zelllinien A549 und H441 zeigen in den experimentellen Daten ein unterschiedliches Behandlungsansprechen im 3D-System, z.B. keine Änderung in der Apoptose in H441, wohingegen HSP60 und p53 eine starke Aktivierung besitzen (Tab. 4.2.1, Tab. 5.2.1A). In einem weiteren Teil sollten entsprechende Unterschiede im Signalweg gefunden und darauf aufbauend neue prognostische Marker und therapeutische Drugtargets identifiziert werden. Basierend auf den experimentellen Daten wurden elf DRPe ausgewählt, die eine unterschiedliche Regulierung in den beiden Zelllinien zeigen (Tab. 5.2.1B).

**Tabelle 5.2.1B: Übersicht der ausgewählten differentiell-regulierten Proteine zwischen A549 und H441 im 3D-System.** Die Tabelle fasst die ausgewählten DRPe (Signifikanz $\pm$ 5%) zwischen den beiden KRAS-mutierten Zelllinien im 3D-System zusammen (Werte repräsentieren Differenz aus behandelt minus unbehandelt; in grün sind hochregulierte DRPe, in rot herunterregulierte DRPe nach 17AAG-Behandlung dargestellt; nicht farblich hinterlegt stellen nicht signifikant DRPe dar; nr=nicht reguliert vor und nach Behandlung).

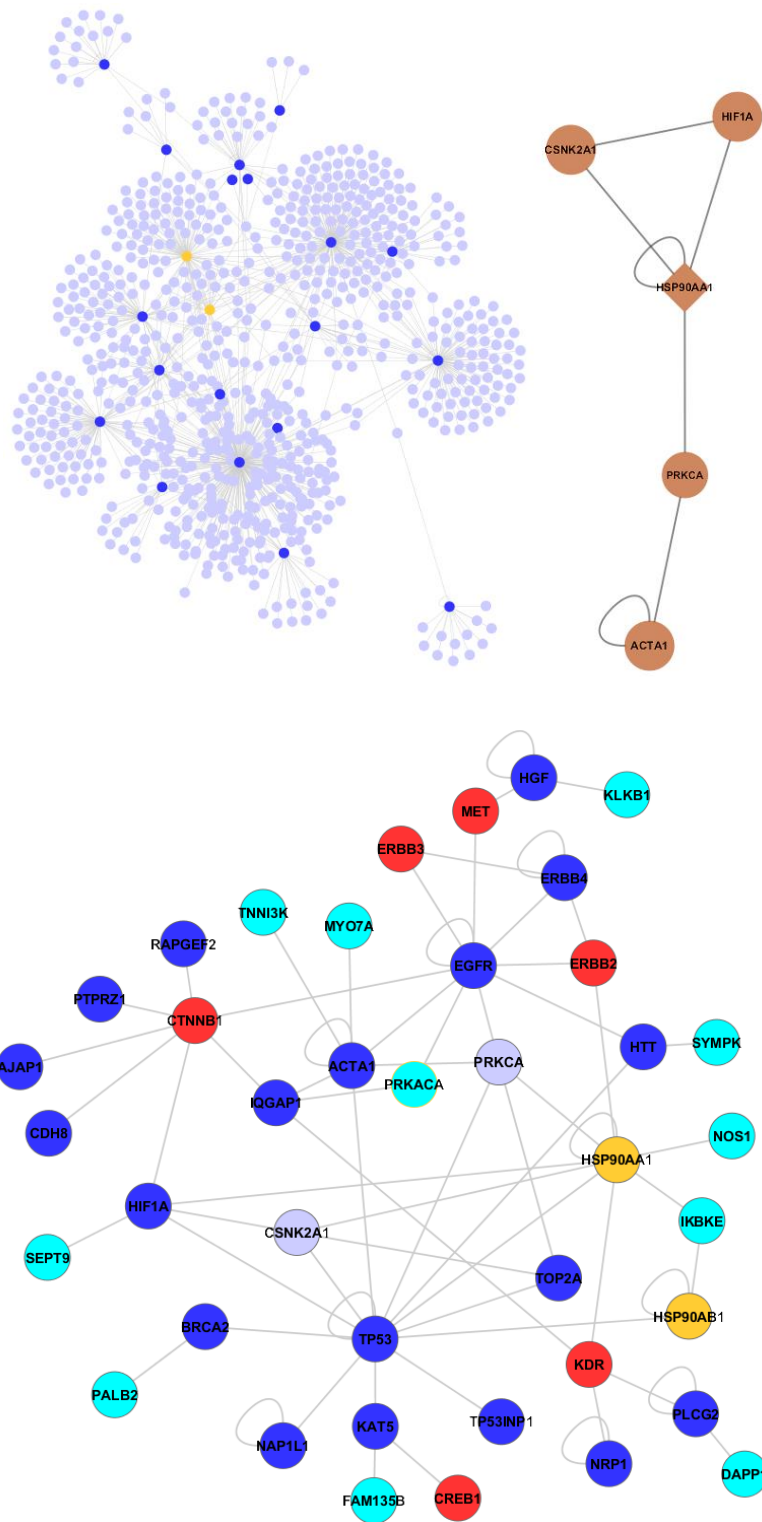
	A549	H441
p53	1,50%	11,10%
AMPK $\alpha$ 1	-3,40%	-5,40%
CREB	-4,80%	-8,90%
$\beta$ -catenin	-2,00%	-8,90%
PRAS40	-10,90%	-2,40%
pErbB2	-0,09%	-19,52%
pErbB3	-0,80%	-25,49%
pFGF R3	nr	-5,76%
pHGFR	-0,15%	-61,05%
pc-RET	nr	-6,39%
pVEGF R2	nr	-16,85%

Das rekonstruierte KRAS-Netzwerk dieser elf DRPe und deren direkten IP aus dem Interaktom besteht aus 588 Knoten und 690 Interaktionen (Abb. 5.2.1C (Links); elf DRPe in rot; IP aus dem Interaktom in blau; Interaktionen in grau). Das Mapping der 588 Knoten auf die zelllinien-spezifischen Mutationen aus der COSMIC-Datenbank hat eine Überlappung von 19 Mutationen für H441 und neun für A549 ergeben (Abb. 5.2.1C (Rechts); insgesamt 784 H441- (blau) und 492 A549- (gelb) Mutationen aus der COSMIC-Datenbank, 588 Knoten aus dem KRAS-Netzwerk in grün).



**Abbildung 5.2.1C: Übersicht des rekonstruierten KRAS-Netzwerks und Mapping der zelllinien-spezifischen Mutationen.** (Links) Das erstellte KRAS-Netzwerk anhand der elf DRPe besteht aus 588 Knoten und 690 Interaktionen (elf DRPe in rot; IP aus dem Interaktom in blau; Interaktionen in grau). (Rechts) Das Venn-Diagramm zeigt das Mapping der 588 Netzwerkknoten auf die zelllinien-spezifischen Mutationen aus der COSMIC-Datenbank (H441-Mutationen (784) in blau, A549-Mutationen (492) in gelb, Netzwerkknoten (588) in grün).

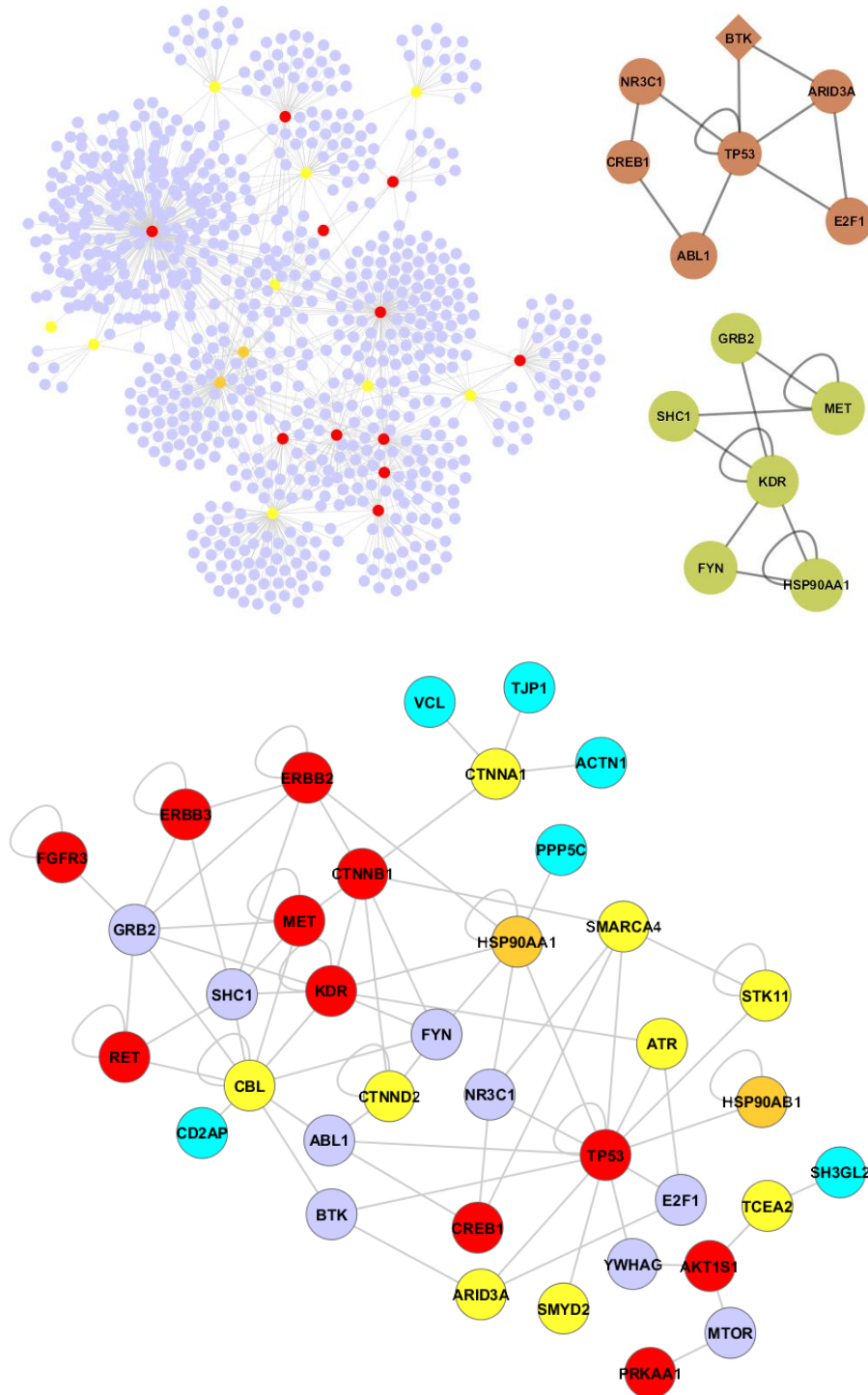
Um eine Verbindung zu HSP90 zu erhalten, wurde ein zelltyp-spezifisches H441-Netzwerk basierend auf den 19 Mutationen (ACTA1, AJAP1, BRCA2, CDH8, EGFR, ERBB4, HGF, HIF1A, HTT, IQGAP1, KAT5, NAP1L1, NRP1, PLCG2, PTPRZ1, RAPGEF2, TOP2A, TP53, TP53INP1) und HSP90 (HSP90AA1 und HSP90AB1) und deren IP aus dem Interaktom erstellt. Das Netzwerk besteht aus 796 Knoten und 982 Interaktionen (Abb. 5.2.1D (Oben, Links)); 19 überlappende Mutationen in dunkelblau, HSP90AA1 und HSP90AB1 in orange, IP aus dem Interaktom in blau; Interaktionen in grau). Die funktionelle Analyse hat hierbei ein Cluster mit fünf Proteinen um HIF1A ergeben (Abb. 5.2.1D (Oben, Rechts); Score=1; in braun). Die Verbindung der 19 H441-Mutationen (in dunkelblau) und der Knoten des Clusters (Clusterknoten CSNK2A1 und PRKCA in blau) mit HSP90 (in orange) sowie der ausgewählten DRPe aus dem Array (in rot; nicht alle vernetzt) ist in Abbildung 5.2.1D (Unten) dargestellt, wobei elf weitere H441-spezifische Mutationen (in cyan), welche nur eine Interaktion von den überlappenden H441-Mutationen entfernt sind oder ebenfalls direkte IP von HSP90 sind, sichtbar werden.



**Abbildung 5.2.1D: Übersicht des rekonstruierten zelltyp-spezifischen H441-Netzwerks.**

**(Oben)** Das erstellte H441-Netzwerk verfügt über 796 Knoten und 982 Interaktionen (Links; 19 Mutationen in dunkelblau, HSP90AA1 und HSP90AB1 in orange, IP aus dem Interaktom in blau; Interaktionen in grau), wobei sich ein Cluster um HIF1A zeigt (Rechts, Score=1, in braun). **(Unten)** Verbindung der H441-Mutationen (in dunkelblau), Clusterknoten, DRPe des Arrays (in rot; nicht alle vernetzt) und weiterer H441-spezifischer Mutationen (in cyan) mit HSP90 (in orange); Clusterknoten CSNK2A1 und PRKCA in blau. (Details siehe Text.)

Das rekonstruierte zelltyp-spezifische A549-Netzwerk der neun Mutationen (ARID3A, ATR, CBL, CTNNA1, CTNND2, SMARCA4, SMYD2, STK11, TCEA2) und HSP90 (HSP90AA1 und HSP90AB1) verfügt über 302 Knoten und 341 Interaktionen (siehe Anhang A1). Da aufgrund der geringen Größe des Netzwerks leider kein Cluster gefunden werden konnte, wurde das Netzwerk für die funktionelle Clusteranalyse mit dem KRAS-Netzwerk (588 Knoten und 690 Interaktionen; Abb. 5.2.1C) verbunden, wobei das kombinierte Netzwerk aus 800 Knoten und 1013 Interaktionen besteht (Abb. 5.2.1E (Oben, Links)); neun A549-Mutationen in gelb, HSP90AA1 und HSP90AB1 in orange, DRPe in rot, IP aus dem Interaktom in blau; Interaktionen in grau). Die funktionelle Analyse hat hierbei zwei Cluster um ARID3A (sieben Proteine; Score=1,286; in braun) und HSP90AA1 (sechs Proteine; Score=1,167; in grün) gefunden (Abb. 5.2.1E (Oben, Rechts)). Die Verbindung der neun A549-Mutationen (in gelb) und der Knoten der beiden gefundenen Cluster (nicht A549-Mutation und nicht DRP in blau) sowie der DRPe aus dem Array (in rot) ist in Abbildung 5.2.1E (Unten) dargestellt, wobei zusätzlich sechs interagierende A549-spezifische Mutationen aus dem Interaktom in cyan markiert sind.



**Abbildung 5.2.1E: Übersicht des rekonstruierten zelltyp-spezifischen A549-Netzwerks.** (Oben) Das erstellte A549-Netzwerk (302 Knoten und 341 Interaktionen; siehe Anhang A1) wurde mit dem KRAS-Netzwerk (Abb. 5.2.1C) verbunden, wobei das erhaltene Netzwerk über 800 Knoten und 1013 Interaktionen verfügt (Links; A549-Mutationen in gelb, HSP90AA1 und HSP90AB1 in orange, DRPe in rot, IP aus dem Interaktom in blau; Interaktionen in grau). Die funktionelle Analyse zeigt zwei Cluster um ARID3A (Score=1,286; in braun) und HSP90AA1 (Score=1,167; in grün). (Unten) Überblick der Verbindung der A549-Mutationen (in gelb), der Clusterknoten (nicht A549-Mutation/nicht DRP in blau), DRPe des Arrays (in rot) und weiterer A549-spezifischer Mutationen (in cyan) mit HSP90 (in orange); IP in blau. (Details siehe Text.)

Unter Verwendung des Mutationshintergrundes, Literaturrecherche, der experimentellen Daten und der Netzwerkverknüpfungen wurde HIF1A für die H441-Zelllinie und STK11 (auch bekannt als LKB1) für die A549-Zelllinie als potentielle therapeutische Drugtargets ausgewählt. Deren therapeutischer Effekt auf Apoptose und Proliferation wurde in einem weiteren Schritt mithilfe einer *in silico*-Simulation untersucht, was im Folgenden gezeigt werden soll. Zusätzlich wurde für beide Zelllinien eine 17AAG-Behandlung simuliert.

### ***In silico*-Simulation einer AMPK-Aktivierung in der A549-Zelllinie**

Das erstellte Netzwerk ist in Abbildung 5.2.1F (Oben) dargestellt. Neben einer *KRAS*-Mutation zeigt die A549-Zelllinie eine Mutation von *LKB1* (COSMIC-Datenbank; Abb. 5.2.1F (Oben), wobei beide Mutationen in der Simulation mit einer konstanten Aktivierung von 0,4 berücksichtigt wurden (low glucose für *LKB1*; siehe Punkt 4.2.2 Material und Methoden). Eine HSP90-Aktivierung wurde mit einer konstanten Aktivierung von 0,7 (über Stress) berücksichtigt, zusätzlich wurde entsprechend der unbehandelten A549-Zellen für die 17AAG-Behandlung EGFR mit einer konstanten Aktivierung von 0,15 und AKT\* mit einer konstanten Aktivierung von 0,2 berücksichtigt (siehe Punkt 4.2.2 Material und Methoden). Es sei angemerkt, dass bei der Simulation der Behandlung lediglich die Parameter AICAR und 17AAG auf 1,0 gesetzt wurden, alle anderen Werte wurden nicht verändert und entsprechen den initialen Input-Parametern für die Simulation der unbehandelten Zellen (*KRAS*=0,4; low glucose (für *LKB1*)=0,4; EGFR=0,15; AKT\*=0,2; Stress=0,7; siehe Punkt 4.2.2 Material und Methoden). Da für eine *in vitro*-Testung kein experimenteller *LKB1*-Aktivator existiert, wurde an dieser Stelle eine AMPK-Aktivierung untersucht, welche downstream von *LKB1* liegt (Abb. 5.2.1F (Oben, in grün)), wofür die Substanz AICAR verfügbar ist (siehe Punkt 5.3.2 zur DrumPID-Datenbank). Die Ergebnisse der durchgeführten Simulationen für eine AMPK-Aktivierung sind in Abbildung 5.2.1F (Mitte) dargestellt. Hierbei zeigt sich, dass die mit AICAR behandelten Zellen (Mitte Rechts) eine gesteigerte Apoptose (in dunkelgrün) und verminderte Proliferation (in violett) besitzen, aber auch ein erhöhtes p53 (in grau) und AMPK (in cyan) und vermindertes mechanistic Target Of Rapamycin (mTOR; in cyan) gegenüber den unbehandelten Zellen erkennen lassen (Mitte Links).

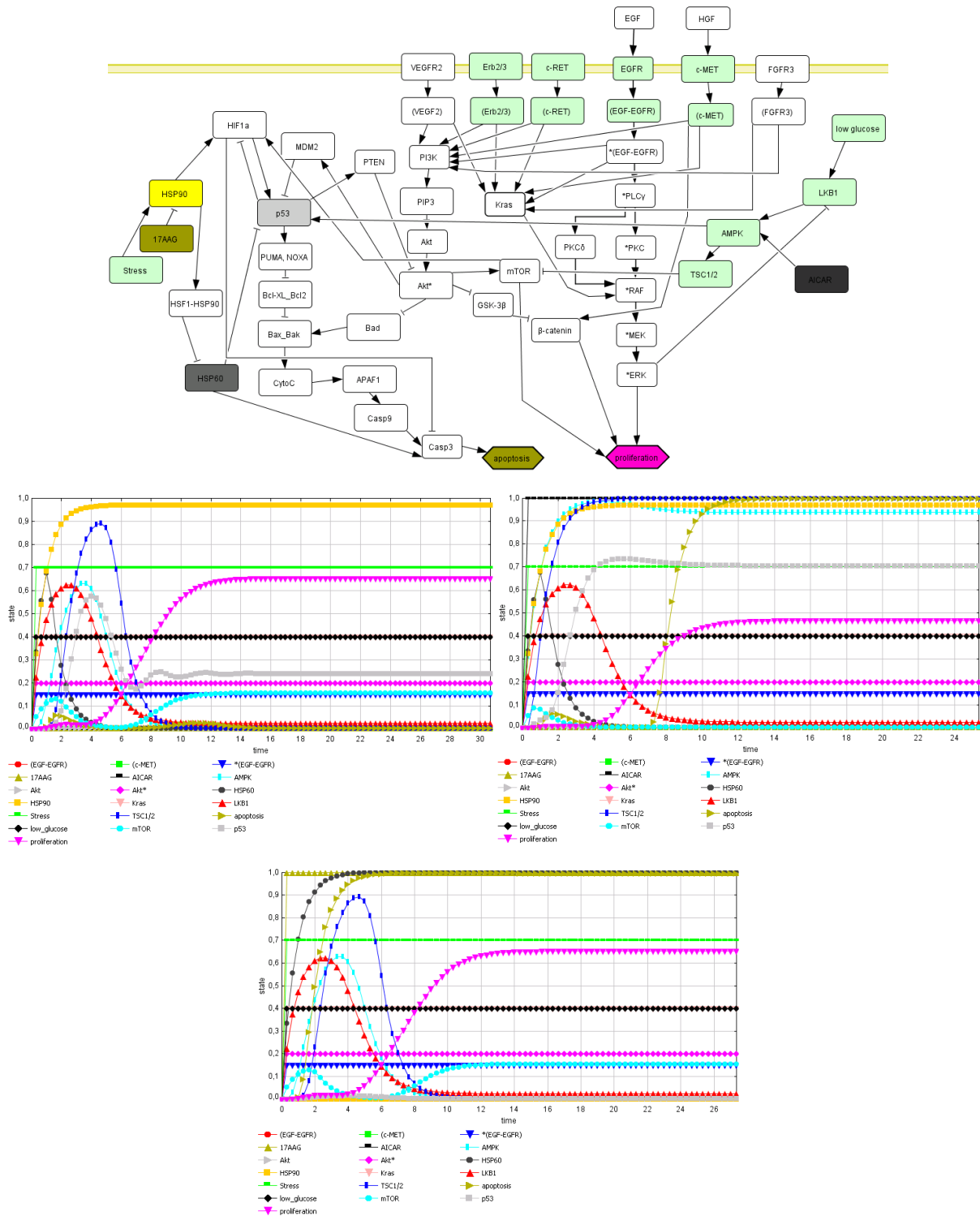
Um den geringen therapeutischen Effekt insbesondere auf die Proliferation einer 17AAG-Behandlung für die A549-Zelllinie zu veranschaulichen, wurde an dieser Stelle zusätzlich eine Simulation einer 17AAG-Aktivierung durchgeführt (Abb.5.2.1F (Unten)). Diese lässt eine erhöhte Apoptose (in dunkelgrün) und HSP60 (in schwarz) sowie keine Änderung der



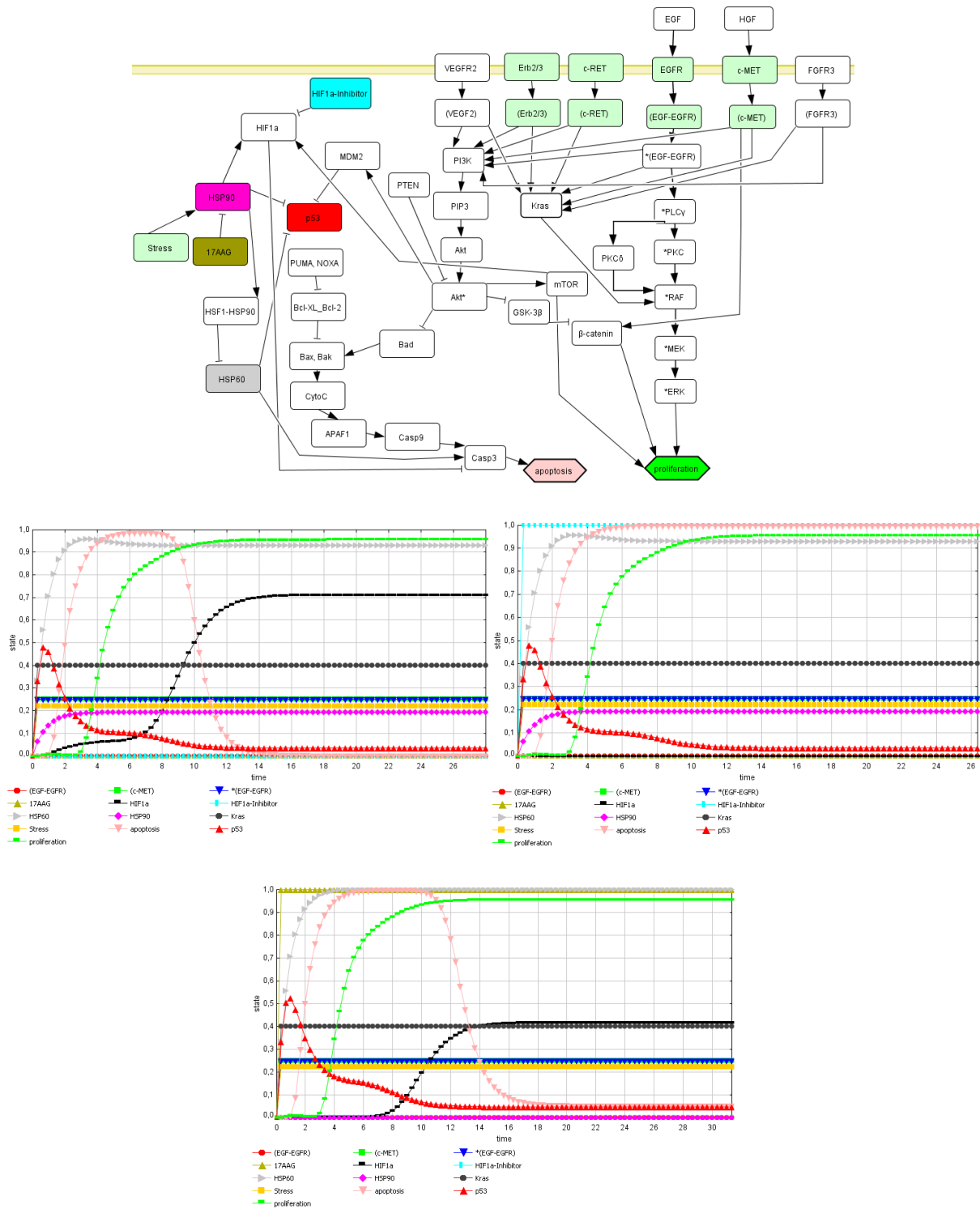
Proliferation (in violett) gegenüber den unbehandelten Zellen (Mitte links) erkennen (Abb. 5.2.1F (Unten)). Zu einem ähnlichen Ergebnis führte auch das vorab durchgeführte Experiment (Tab. 4.2.1, Tab. 5.2.1A). An dieser Stelle sei angemerkt, dass p53 (in grau) entsprechend der zu erwartenden Regulierung aus dem Signalnetzwerk (HSP90 inhibierend auf HSP60, dieses ist wiederum inhibierend auf p53) eine reduzierte Aktivierung nach 17AAG-Behandlung in der Simulation zeigt, wohingegen die experimentellen Daten keine signifikante Änderung aufgewiesen haben (Tab. 5.2.1A).

### ***In silico*-Simulation einer HIF1A-Inhibition in der H441-Zelllinie**

Das erstellte Netzwerk ist in Abbildung 5.2.1G (Oben) abgebildet. Die H441-Zelllinie weist eine *KRAS*-Mutation und eine Loss Of Function (LOF)-Mutation im *p53*-Gen auf (COSMIC-Datenbank) (Abb. 5.2.1G (Oben)), wobei für die *KRAS*-Mutation eine konstante Aktivierung von 0,4 verwendet wurde. Die LOF von p53 wurde im Netzwerk durch eine fehlende Verbindung von p53 zu anderen Knoten berücksichtigt, sodass p53 zwar keine Knoten aktivieren und inhibieren kann, aber dennoch von anderen Knoten reguliert werden kann (Abb. 5.2.1G (Oben); siehe Punkt 4.2.2 Material und Methoden). Entsprechend der experimentellen Daten für die unbehandelten Zellen der 17AAG-Behandlung wurden als weitere initiale Input-Parameter für EGFR (als \*(EGF-EGFR)) und c-MET (als (c-MET)) eine konstante Aktivierung von 0,25 sowie für eine HSP90-Aktivierung (über Stress) 0,22 verwendet (siehe Punkt 4.2.2 Material und Methoden). Bei den Simulationen der Behandlung wurden ebenfalls lediglich die Parameter HIF1a-Inhibitor und 17AAG auf 1,0 gesetzt, alle anderen Werte entsprechen denen der Simulation der unbehandelten Zellen ( $KRAS=0,4$ ; \*(EGF-EGFR)=0,25; (c-MET)=0,25; Stress=0,22; siehe Punkt 4.2.2 Material und Methoden). Die Simulation der HIF1A-Inhibition ist in Abbildung 5.2.1G (Mitte) dargestellt. Es lässt sich erkennen, dass H441-Zellen mit einer HIF1A-Inhibition (Mitte Rechts) eine über der Proliferation (in hellgrün) liegende Apoptose (in lachsfarben) gegenüber den unbehandelten Zellen (Mitte Links) aufzeigen. Zusätzlich wurde eine Simulation für die 17AAG-Behandlung durchgeführt (Abb. 5.2.1G (Unten)). Diese zeigt, verglichen mit den unbehandelten Zellen (Mitte Links), keinen therapeutischen Effekt auf Proliferation (in hellgrün) und Apoptose (in lachsfarben) im zeitlichen Verlauf, wohingegen sich eine Zunahme von HSP60 (in grau) und p53 (in rot, gering) beobachten lässt (17AAG in dunkelgrün). Dies entspricht weitestgehend den experimentellen Daten (Tab. 4.2.1, Tab. 5.2.1A).



**Abbildung 5.2.1F: A549-Netzwerk und *in silico*-Simulation im 3D-System. (Oben)** Das Netzwerk beinhaltet wichtige Signalwege (z.B. EGFR, KRAS und AMPK; in weiß und grün) und deren Verbindung für die A549-Zelllinie. Proliferation in violett, Apoptose und 17AAG in dunkelgrün, AICAR in schwarz, (→) Aktivierung, (−) Inhibierung. **(Mitte)** Eine AMPK-Aktivierung (Rechts) resultiert in einer erhöhten Apoptose und reduzierten Proliferation verglichen mit den unbehandelten Zellen (Links). **(Unten)** Eine 17AAG-Behandlung zeigt eine höhere Apoptose, aber keine Änderung der Proliferation gegenüber den unbehandelten Zellen (Mitte Links). (Details siehe Text; Input-Parameter siehe Punkt 4.2.2 Material und Methoden.)



**Abbildung 5.2.1G: H441-Netzwerk und *in silico*-Simulation im 3D-System. (Oben)** Das Netzwerk zeigt wichtige Signalwege (z.B. EGFR, KRAS und p53; in grün, weiß und rot) und deren Interaktionen für die H441-Zelllinie. Proliferation in hellgrün, Apoptose in lachsfarben, 17AAG-Inhibitor in dunkelgrün, HIF1a-Inhibitor in cyan, (→) Aktivierung, (−|) Inhibierung. **(Mitte)** Eine HIF1A-Inhibierung (Rechts) resultiert in einer über der Proliferation liegenden Apoptose gegenüber unbehandelten Zellen (Links). **(Unten)** Verglichen mit den unbehandelten Zellen (Mitte Links) zeigt die 17AAG-Behandlung eine geringe Änderung der Proliferation und Apoptose. (Details siehe Text; Input-Parameter siehe Punkt 4.2.2 Material und Methoden.)

## 5.2.2 *In silico*-Analysen der Resistenzmechanismen nach Gefitinib-Behandlung

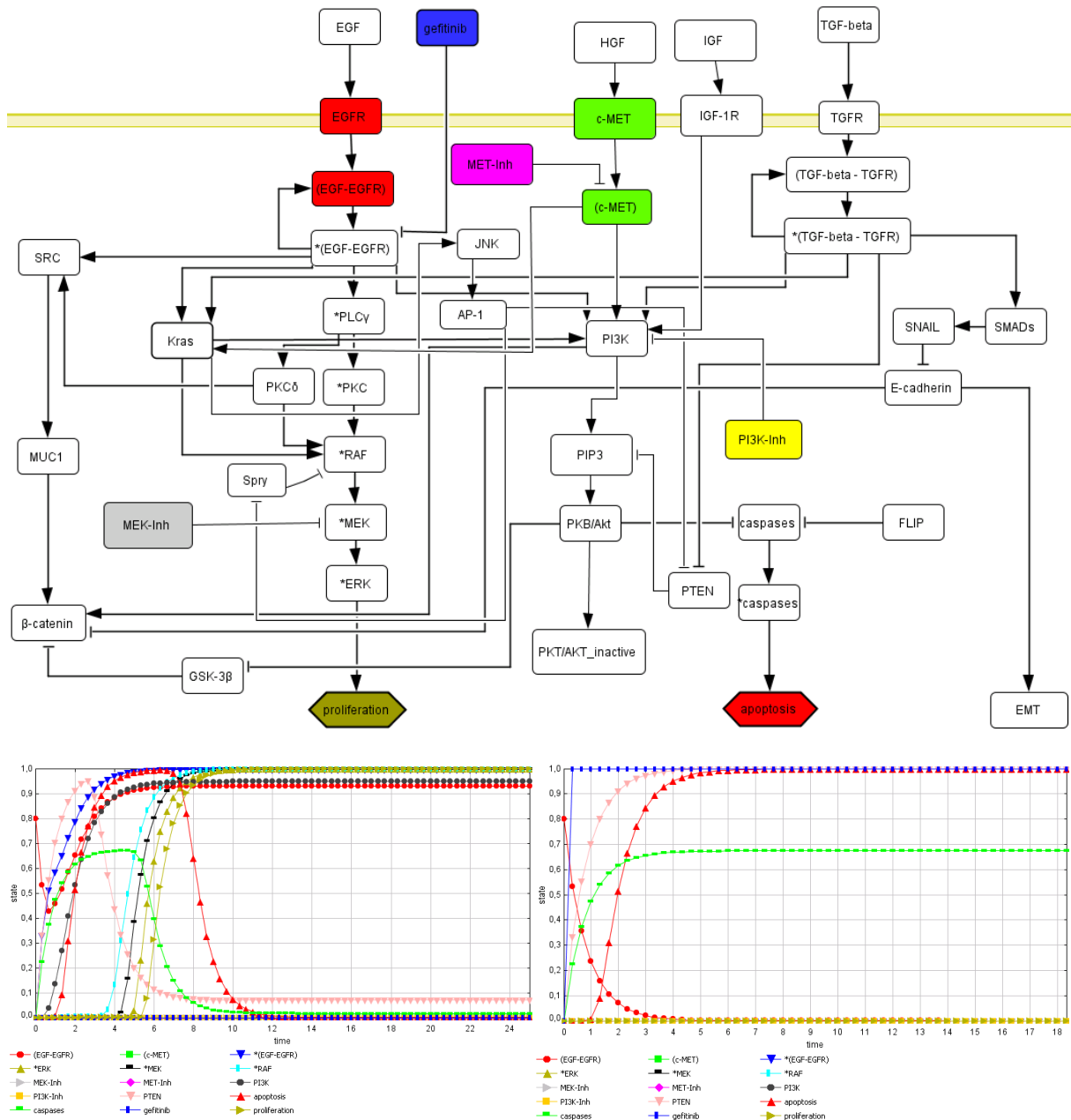
Das *in silico*-Tumormodell sollte im weiteren Verlauf verschiedene Resistenzmechanismen nach Gefitinib-Behandlung mit bekanntem Mutationsstatus für die Zelllinien A549 und HCC827 abbilden und anschließend eine geeignete Behandlungsstrategie simulieren.

### Gefitinib-Resistenz in der HCC827-Zelllinie

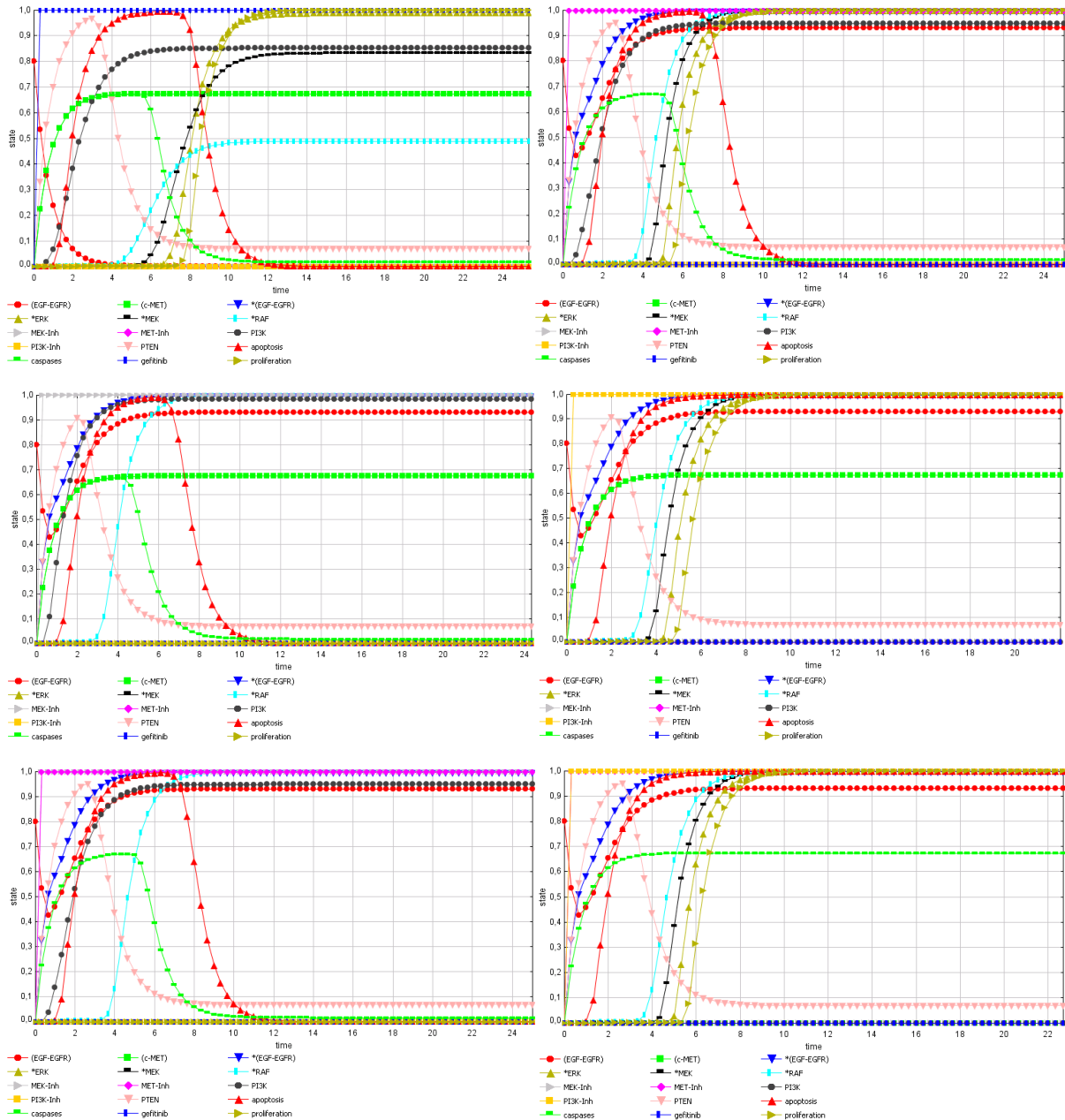
Das erstellte Netzwerk in CellDesigner ist in Abbildung 5.2.2A (Oben) dargestellt. Es integriert die für HCC827 bekannte aktivierende Mutation im EGFR-Rezeptor (in rot), den c-MET-Rezeptor (in grün) und beteiligte Signalwege für Apoptose (PI3K/AKT) und Proliferation (RAS/RAF/ERK) (Proliferation in dunkelgrün, Apoptose in rot; Abb. 5.2.2A (Oben)). Das Netzwerk wurde anschließend in SQUAD zur weiteren Simulation genutzt, deren Ergebnisse in den Abbildungen 5.2.2A (Unten), 5.2.2B und 5.2.2C abgebildet sind. In einem ersten Schritt wurde der unbehandelte Systemzustand, d.h. ein Tumorzellwachstum, simuliert (Abb. 5.2.2A (Unten, Links)). Die unbehandelten Zellen ((EGF-EGFR)-Value=0,8, in rot) zeigen eine erhöhte Proliferation (dunkelgrüne Kurve) und reduzierte Apoptose (rote Kurve), was entsprechende Tumorbedingungen reflektiert (Abb. 5.2.2A (Unten, Links)). Die folgende Simulation der Behandlung mit Gefitinib (in blau) zeigt einen gegensätzlichen Effekt, was auf ein erfolgreiches Therapieansprechen hindeutet (Abb. 5.2.2A (Unten, Rechts)). Um eine entsprechende Resistenz nachzubilden, wurde zusätzlich der c-MET-Rezeptor aktiviert (in grün). Die Simulation der Gefitinib-Behandlung (in blau) zeigt in diesem Falle eine Aktivierung der Proliferation (in dunkelgrün) und Reduktion der Apoptose (rote Kurve), erfolgt verglichen mit den unbehandelten Zellen zu einem späteren Zeitpunkt, im Zeitablauf, was eine entsprechende Gefitinib-Resistenz für die HCC827-Zelllinie reflektiert (Abb. 5.2.2B (Oben, Links)). Anhand des Kurvenverlaufs der Gefitinib-Behandlung wird deutlich, dass neben c-MET (in grün) die beiden Knoten PI3K (in dunkelgrau) und MEK (in schwarz) eine erhöhte Aktivierung aufweisen (Abb. 5.2.2B (Oben, Links)), sodass deren Inhibierung im Folgenden als neue Targets untersucht wurden.

Die alleinige Inhibierung von c-MET (MET-Inh in violett) zeigt, dass die Apoptose (in rot) im Laufe der Zeit abnimmt und die Proliferation (in dunkelgrün) ansteigt (Abb. 5.2.2B (Oben, Rechts)). Die Simulation der Inhibierung von MEK (MEK-Inh in grau) zeigt eine reduzierte Proliferation (in dunkelgrün), führt aber im Zeitverlauf auch zu einer Abnahme der Apoptose (in rot) (Abb. 5.2.2B (Mitte, Links)). Die alleinige Inhibierung von PI3K (PI3K-Inh in gelb)

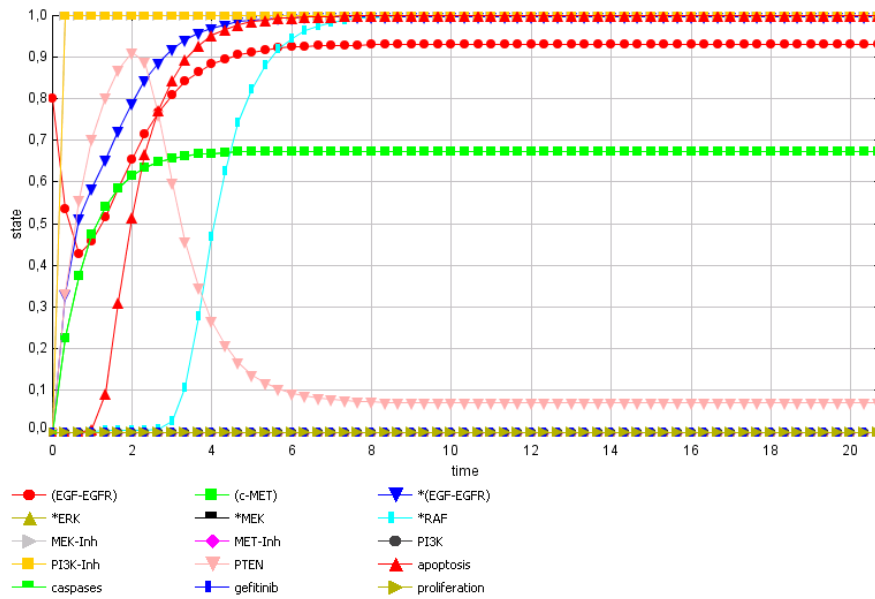
führt zu einer erhöhten Apoptose (in rot), die Proliferationsrate (dunkelgrüne Kurve) zeigt ein gleich hohes Aktivierungslevel (Abb. 5.2.2B (Mitte, Rechts)). Die kombinierte Inhibierung von c-MET (MET-Inh in violett) und MEK (MEK-Inh in grau) lässt eine reduzierte Apoptose (in rot) und Proliferation (in dunkelgrün) im zeitlichen Verlauf erkennen (Abb. 5.2.2B (Unten, Links)). Eine Inhibierung von c-MET (MET-Inh in violett) und PI3K (PI3K-Inh in gelb) zeigt eine erhöhte Proliferation (in dunkelgrün) und Apoptose (in rot) im Zeitverlauf (Abb. 5.2.2B (Unten, Rechts), wohingegen eine kombinierte Inhibierung von MEK (MEK-Inh in grau) und PI3K (PI3K-Inh in gelb) zu einer Aktivierung der Apoptose (rote Kurve) und einer verminderten Proliferation (dunkelgrüne Kurve) im zeitlichen Verlauf führt (Abb. 5.2.2C).



**Abbildung 5.2.2A: Netzwerk und *in silico*-Simulation der Gefitinib-Resistenz in der HCC827-Zelllinie.** (Oben) Das Netzwerk integriert wichtige Signalwege (EGFR in rot, c-MET in grün) und deren IP für die HCC827-Zelllinie. Readout-Parameter für Tumorwachstum (Proliferation in dunkelgrün) und Zelltod (Apoptose in rot) sind als Sechsecke dargestellt, therapeutische Inhibitoren MET-Inh in violett, MEK-Inh in grau, PI3K-Inh in gelb, Gefitinib in blau; Inh=Inhibitor; (→) veranschaulichen Aktivierungen, (−) Inhibierungen. (Unten) Links: Eine EGFR-Überexpression (in rot) resultiert in einer verminderten Apoptose (in rot) und erhöhten Proliferation (in dunkelgrün) im Zeitverlauf, was das Tumorwachstum verdeutlicht. Rechts: Eine Gefitinib-Behandlung (in blau) zeigt einen therapeutischen Effekt (erhöhte Apoptose in rot, reduzierte Proliferation in dunkelgrün).



**Abbildung 5.2.2B: *In silico*-Simulation der Gefitinib-Resistenz in der HCC827-Zelllinie.** (Oben) Links: Eine EGFR- (in rot) und c-MET- (in grün) Komutation resultiert in einer Gefitinib-Resistenz im zeitlichen Verlauf (Proliferation (dunkelgrün) über Apoptose (rot)). Rechts: Die alleinige Inhibierung von c-MET (MET-Inh in violett) zeigt einen geringen therapeutischen Effekt (reduzierte Apoptose (rot) und erhöhte Proliferation (dunkelgrün) im Zeitablauf). (Mitte) Eine Inhibierung von MEK (MEK-Inh in grau) resultiert in einer verminderten Apoptose (in rot) und Proliferation (in dunkelgrün) im zeitlichen Verlauf (Links), wohingegen die alleinige PI3K-Inhibierung (PI3K-Inh in gelb) zu einer erhöhten Proliferation (in dunkelgrün) und Apoptose (in rot) führt (Rechts). (Unten) Die kombinierte Inhibierung von c-MET (MET-Inh in violett) und MEK (MEK-Inh in grau) zeigt eine geringe Apoptose (in rot) und Proliferation (in dunkelgrün) im Zeitverlauf (Links), wohingegen eine Inhibierung von c-MET (MET-Inh in violett) und PI3K (PI3K-Inh in gelb) zu einer erhöhten Proliferation (in dunkelgrün) und Apoptose (in rot) führt (Rechts). (Details siehe Text; Input-Parameter siehe Punkt 4.2.3 Material und Methoden.)



**Abbildung 5.2.2C: *In silico*-Simulation einer kombinierten PI3K- und MEK-Inhibierung in der HCC827-Zelllinie.** Die kombinierte Inhibierung von PI3K (PI3K-Inh in gelb) und MEK (MEK-Inh in grau) resultiert in einer erhöhten Apoptose (in rot) und reduzierten Proliferation (in dunkelgrün) im zeitlichen Verlauf. (Details siehe Text; Input-Parameter siehe Punkt 4.2.3 Material und Methoden.)

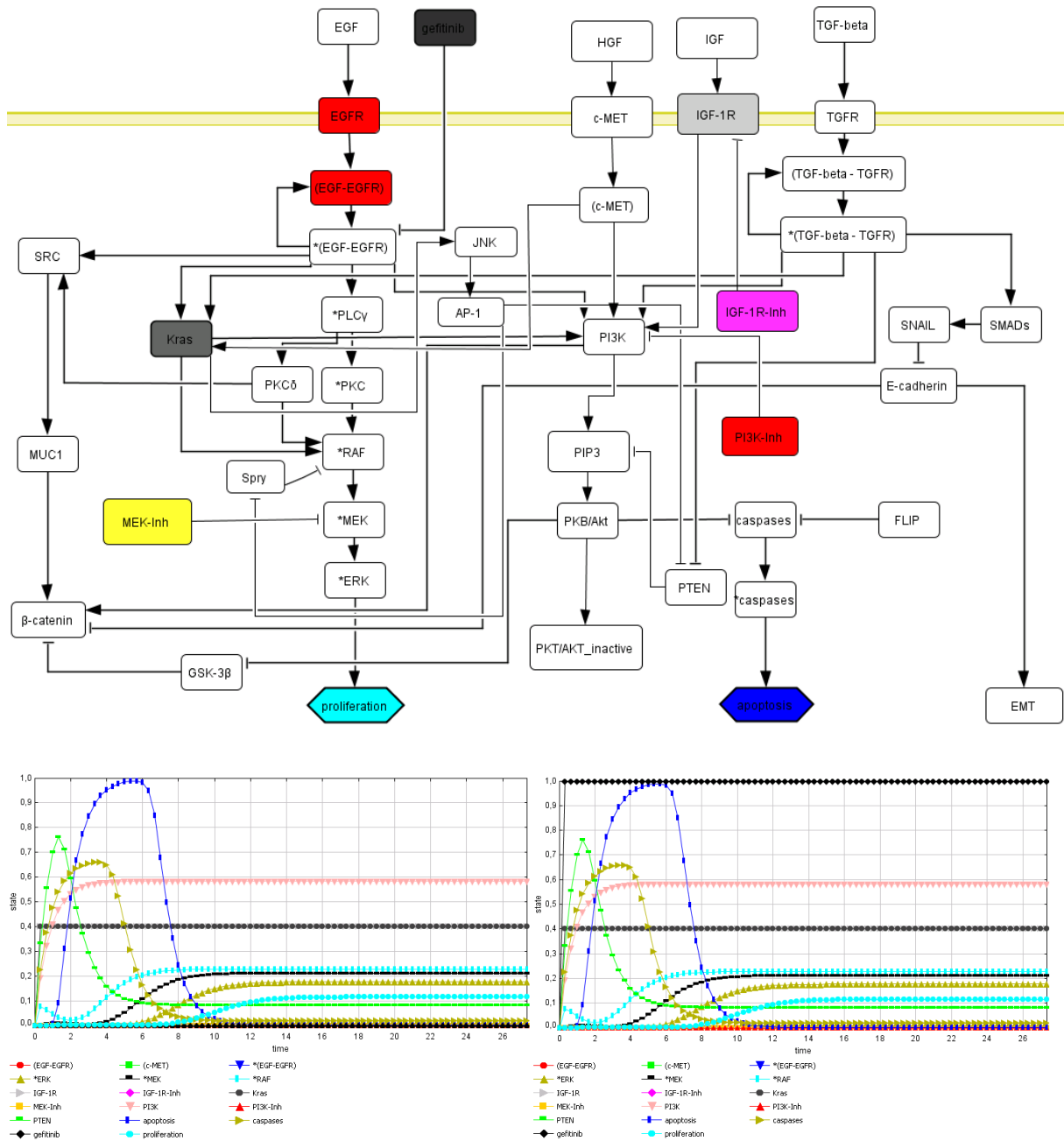
### Gefitinib-Resistenz in der A549-Zelllinie

Das Netzwerk ist in Abbildung 5.2.2D (Oben) dargestellt. Eine Deregulierung von IGF-1R (in grau) und eine KRAS-Mutation (in dunkelgrau) wurden in der Simulation durch eine konstante Aktivierung von 0,4 berücksichtigt (siehe Punkt 4.2.3 Material und Methoden). Die durchgeführten Simulationen sind in den Abbildungen 5.2.2D (Unten) und Abb. 5.2.2E abgebildet. Die unbehandelten Zellen (Komutation von KRAS (in dunkelgrau) und IGF-1R (in grau)) zeigen eine über der Apoptose (in blau) liegende Proliferation (in hellblau) im zeitlichen Verlauf, was ein entsprechendes Tumorwachstum darstellt (5.2.2D (Unten, Links)). Eine Gefitinib-Gabe (in schwarz) zeigt keinen Effekt auf Apoptose und Proliferation im Zeitablauf (5.2.2D (Unten, Rechts)). Als Nächstes wurde der therapeutische Effekt einer IGF-1R-, PI3K- und MEK-Behandlung untersucht (Abb. 5.2.2E). Eine alleinige Inhibierung von IGF-1R (IGF-1R-Inh in violett) zeigt keine Änderung in der Apoptose (in blau) und Proliferation (in hellblau) gegenüber den unbehandelten Zellen im Zeitverlauf (Abb. 5.2.2E (Oben, Links)). Eine Inhibierung von MEK (MEK-Inh in gelb) resultiert in einer reduzierten Proliferation (in hellblau), jedoch lässt sich kein Effekt auf die Apoptose (in blau) im Zeitablauf erkennen (Abb. 5.2.2E (Oben, Rechts)). Eine Inhibierung von PI3K (PI3K-Inh in rot; Abb. 5.2.2E (Mitte,

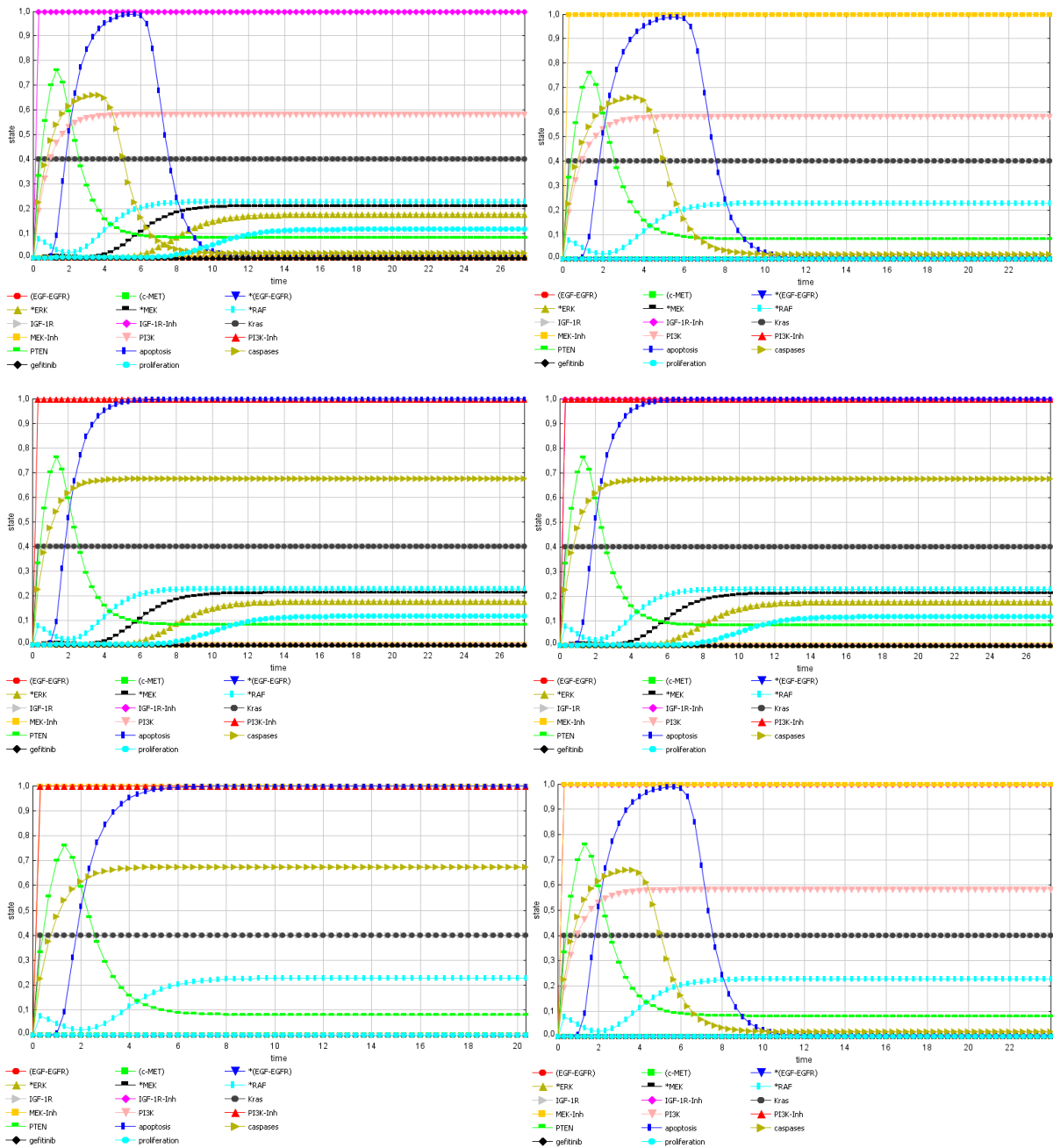


Links)) sowie eine kombinierte Inhibierung von PI3K (PI3K-Inh in rot) und IGF-1R (IGF-1R-Inh in violett) (Abb. 5.2.2E (Mitte, Rechts)) zeigen einen ähnlichen Effekt und resultieren in einer Induktion der Apoptose (in blau), sie haben allerdings keinen Effekt auf die Proliferation (in hellblau) verglichen mit den unbehandelten Zellen. Die kombinierte Inhibierung von PI3K (PI3K-Inh in rot) und MEK (MEK-Inh in gelb) zeigt eine erhöhte Apoptose (in blau) und verminderte Proliferation (in hellblau) im Zeitverlauf (Abb. 5.2.2E (Unten, Links)). Eine Inhibierung von MEK (MEK-Inh in gelb) und IGF-1R (IGF-1R-Inh in violett) hat einen Effekt auf die Proliferation (in hellblau), wobei sich ebenfalls eine Abnahme der Apoptose (in blau) im zeitlichen Verlauf zeigt (Abb. 5.2.2E (Unten, Rechts)).

Alles in allem zeigen die durchgeführten Simulationen in der H441- und A549-Zelllinie, dass eine kombinierte Inhibierung von PI3K und MEK den bestmöglichen therapeutischen Effekt auf Apoptose und Proliferation für beide Zelllinien liefert.



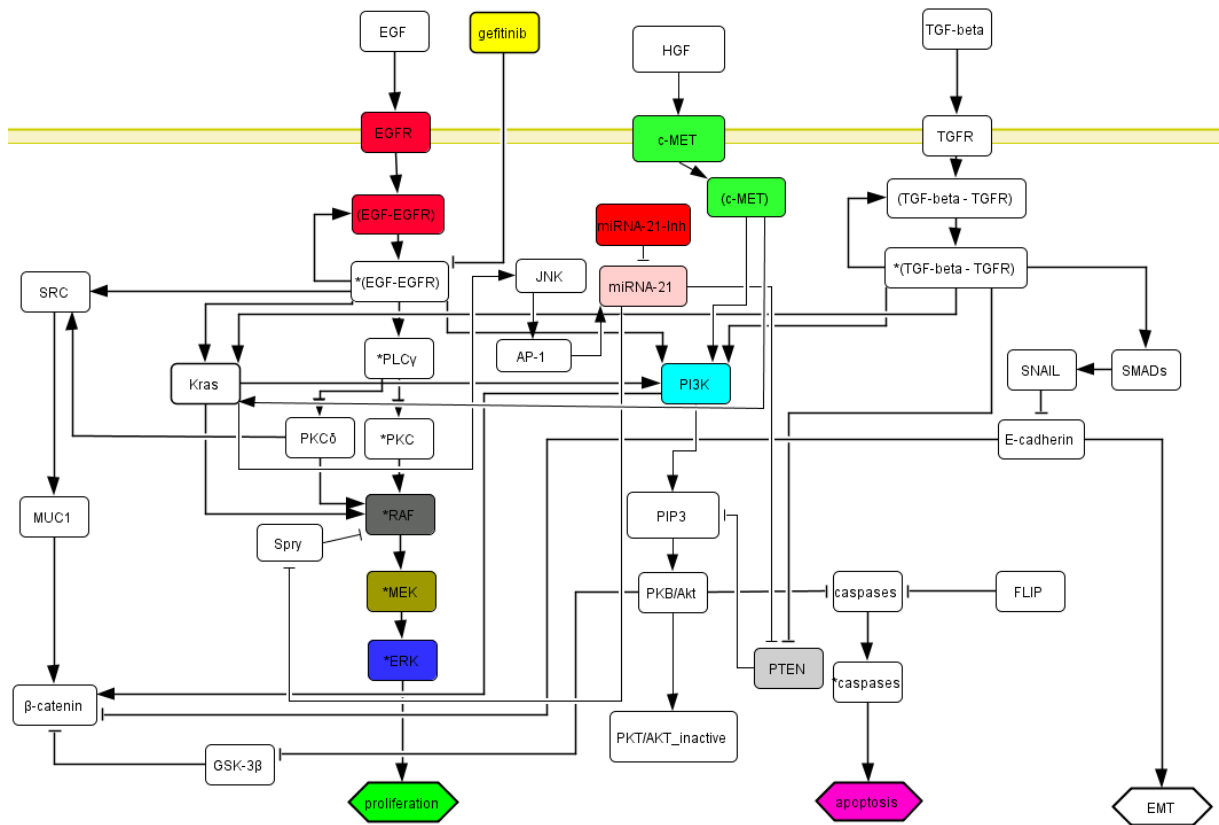
**Abbildung 5.2.2D: Netzwerk und *in silico*-Simulation der Gefitinib-Resistenz in der A549-Zelllinie.** (Oben) Das Netzwerk integriert wichtige Signalwege (EGFR in rot und IGF-1R in grau) und deren resultierende IP für die A549-Zelllinie. Readout-Parameter Proliferation (in hellblau) und Apoptose (in blau) sind als Sechsecke dargestellt, KRAS in dunkelgrau. Therapeutische Inhibitoren PI3K-Inh in rot, MEK-Inh in gelb, IGF-1R-Inh in violett, Gefitinib in schwarz. (→) veranschaulichen Aktivierungen, (-) Inhibierungen. (Unten) Links: Eine KRAS- (in dunkelgrau) und IGF-1R- (in grau) Mutation resultiert in einer verminderten Apoptose (in blau) und erhöhten Proliferation (in hellblau) im zeitlichen Verlauf, was eine entsprechende Tumorbildung veranschaulicht. Rechts: Eine Gefitinib-Behandlung (in schwarz) zeigt keinen therapeutischen Effekt auf Proliferation (in hellblau) und Apoptose (in blau). (Details siehe Text; Input-Parameter siehe Punkt 4.2.3 Material und Methoden.)



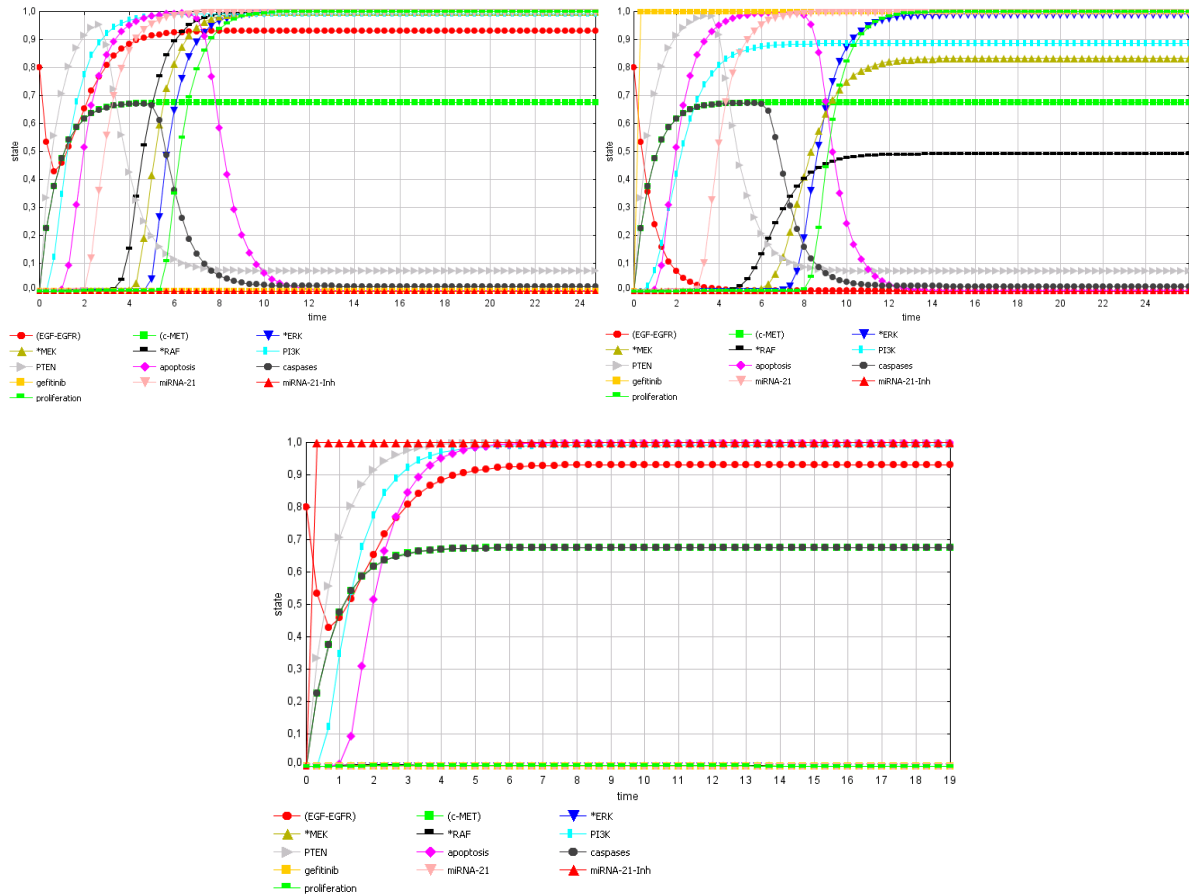
**Abbildung 5.2.2E: *In silico*-Simulation der Gefitinib-Resistenz in der A549-Zelllinie.** (Oben) Links: Eine IGF-1R-Inhibierung (in violett) zeigt keine Änderung in der Apoptose (in blau) und Proliferation (in hellblau) im Zeitverlauf verglichen mit den unbehandelten Zellen. Rechts: Eine MEK-Inhibierung (in gelb) lässt eine reduzierte Proliferation (in hellblau) erkennen, allerdings zeigt sich kein Effekt auf die Apoptose (in blau). (Mitte) Sowohl eine alleinige PI3K-Inhibierung (in rot; Links) als auch die kombinierte Inhibierung (Rechts) von PI3K und IGF-1R (IGF-1R-Inh in violett) zeigen eine erhöhte Apoptose (in blau), aber keinen Effekt auf die Proliferation (in hellblau) im Zeitablauf verglichen mit den unbehandelten Zellen. (Unten) Links: Eine Inhibierung von MEK (MEK-Inh in gelb) und PI3K (PI3K-Inh in rot) resultiert in einer verminderten Proliferation (in hellblau) und erhöhten Apoptose (in blau) im Zeitverlauf. Rechts: Die Inhibierung von MEK (MEK-Inh in gelb) und IGF-1R (IGF-1R-Inh in violett) resultiert in einer verminderten Proliferation (in hellblau), allerdings ergibt sich kein Effekt auf die Apoptose (in blau). (Details siehe Text; Input-Parameter siehe Punkt 4.2.3.)

### **Einfluss der miRNA-21 auf die Gefitinib-Resistenz in der HCC827-Zelllinie**

In einem weiteren Teil sollte die miRNA-21, welche bei NSCLC überexprimiert ist und mit einer Gefitinib-Resistenzentwicklung in Verbindung steht, *in silico* für die HCC827-Zelllinie untersucht werden. Das erstellte Netzwerk ist in Abbildung 5.2.2F abgebildet. Die miRNA-21 (in lachsfarben) ist EGFR (in rot) aktiviert und inhibiert u. a. Spry (in weiß) und PTEN (in grau), womit sie einen Einfluss auf die Apoptose und Proliferation hat (Abb. 5.2.2F). Für die Simulationen wurde eine EGFR-Überexpression (in rot) durch eine Aktivierung von 0,8 sowie eine Überexpression von c-MET (in grün) durch eine konstante Aktivierung von 0,4 berücksichtigt (siehe Punkt 4.2.3 Material und Methoden). Die Ergebnisse der Simulationen sind in Abbildung 5.2.2G dargestellt. Die unbehandelten Zellen spiegeln mit einer erhöhten Proliferation (in hellgrün) und verminderten Apoptose (in violett) im zeitlichen Verlauf ein entsprechendes Tumorwachstum wider (Abb. 5.2.2G (Oben, Links)). Eine Gefitinib-Behandlung (in gelb) zeigt ein ähnliches Verhalten (Proliferation über der Apoptoserate), allerdings erfolgt die Zunahme der Proliferation und Abnahme der Apoptose gegenüber den unbehandelten Zellen (Zeitpunkt ca. 5,8) zu einem späteren Zeitpunkt (ca. 8) (Abb. 5.2.2G (Oben, Rechts)). Betrachtet man den Graphen genauer, so lässt sich ebenfalls eine Zunahme der Expression der miRNA-21 (in lachsfarben) erkennen, aber auch eine Abnahme von PTEN (in grau) und Zunahme von RAF-MEK-ERK (in schwarz-dunkelgrün-blau) beobachten (Abb. 5.2.2G (Oben, Rechts)). Eine Inhibierung der miRNA-21 (miRNA-21-Inh in rot) resultiert in einer erhöhten Apoptose (in violett) und reduzierten Proliferation (in hellgrün) im Zeitverlauf, sodass hier ein potentieller therapeutischer Effekt zu beobachten ist (Abb. 5.2.2G (Unten)). Die Inhibierung der miRNA-21 lässt ebenfalls eine Zunahme von PTEN (in grau) und Abnahme von RAF-MEK-ERK (in schwarz-dunkelgrün-blau) erkennen, wobei dies unabhängig vom PI3K-Level (in hellblau) ist.



**Abbildung 5.2.2F: Netzwerk der miRNA-21 in der HCC827-Zelllinie.** Das Netzwerk veranschaulicht wichtige Signalwege (z.B. EGFR in rot, c-MET in grün, miRNA-21 in lachsfarben) und deren IP für die HCC827-Zelllinie. Readout-Parameter Proliferation (in hellgrün) und Apoptose (in violett) sind als Sechsecke dargestellt, miRNA-21-Inh in rot, Gefitinib in gelb. (→) veranschaulichen Aktivierungen, (-) Inhibierungen.



**Abbildung 5.2.2G: *In silico*-Simulation des Einflusses der miRNA-21 auf die Gefitinib-Resistenz in der HCC827-Zelllinie. (Oben) Links:** Unbehandelte Zellen zeigen eine verminderte Apoptose (in violett) und erhöhte Proliferation (in hellgrün) im zeitlichen Verlauf, was ein entsprechendes Tumorwachstum verdeutlicht. **Rechts:** Eine Gefitinib Behandlung (in gelb) zeigt einen ähnlichen Verlauf, wobei der Anstieg der Proliferation (in hellgrün) und Abnahme der Apoptose (in violett) zu einem späteren Zeitpunkt erfolgt. **(Unten)** Eine Inhibierung der miRNA-21 (miRNA-21-Inh in rot) resultiert in einer erhöhten Apoptose (in violett) und reduzierten Proliferation (in hellgrün) im Zeitablauf. (Details siehe Text; Input-Parameter siehe Punkt 4.2.3 Material und Methoden.)

### 5.3 Identifizierung von RNA- und Protein-Interaktionen

#### 5.3.1 Ergebnisse der Analysen von hypertrophie- und hypoxie-assoziierten lncRNAs

In einer Zusammenarbeit mit dem Institut für Molekulare und Translationale Therapiestrategien (Leitung: Prof. Dr. Dr. med. Thomas Thum, Medizinische Hochschule Hannover) wurde in einem ersten Projekt nach einer neuen therapeutischen lncRNA bei Herzhypertrophie (Punkt 5.3.1.1) gesucht, wohingegen in einem zweiten Projekt neue hypoxie-induzierte lncRNAs in Endothelzellen (Punkt 5.3.1.2) identifiziert werden sollten. Meine Aufgabe war es im Rahmen der Dissertation, die lncRNAs systembiologisch zu charakterisieren, worauf im Folgenden eingegangen werden soll.

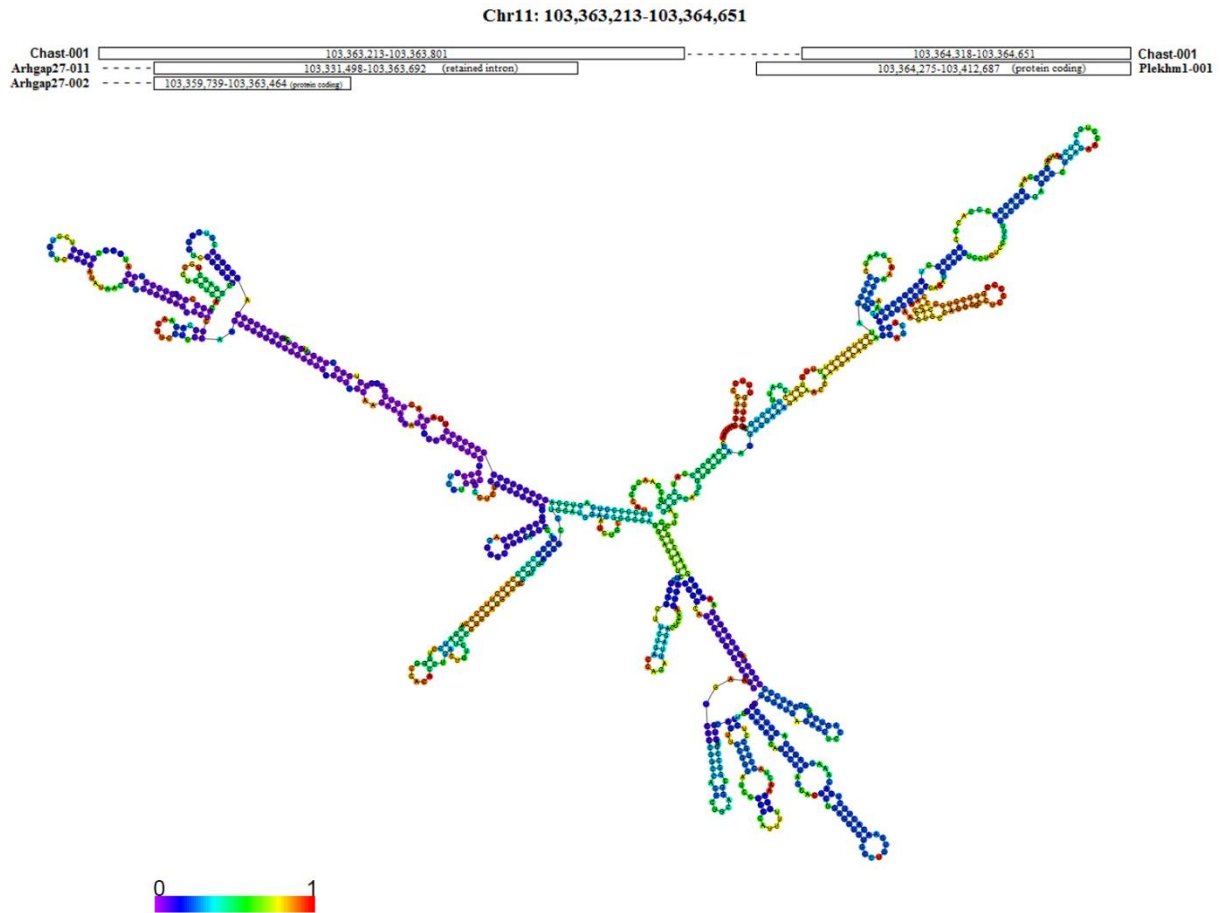
##### 5.3.1.1 Herzhypertrophie-assoziierte lncRNA Chast

Die Chast lncRNA (Gm11641, ENSMUSG00000085733, ENSMUST00000130556) liegt auf Chromosom (chr) 11 an Position 103,363,213 - 103,364,651 (Forward Strang) und besitzt eine Transkriptlänge von 923bp (Abb. 5.3.1.1A). Sie liegt antisense zu den beiden proteinkodierenden Genen Rho GTPase Activating Protein 27 (Arhgap27; ENSMUSG00000034255; Arhgap27\_002 (ENSMUST00000092557, chr11: 103,359,739 - 103,363,464); Arhgap27\_011 (ENSMUST00000116102, chr11: 103,331,498 - 103,363,692) und Pleckstrin homology domain containing, family M (with RUN domain) member 1 (Plekhm1; ENSMUSG00000034247; chr11: 103,364,275 - 103,412,687; ENSMUST00000041272) (Abb. 5.3.1.1A). Hinsichtlich des proteinkodierenden Potentials hat die Analyse einen Score von -0,79724 ergeben.

Die durchgeführte Sequenzanalyse konnte homologe Gene in verschiedenen Säugern finden, welche in Tabelle 5.3.1.1A zusammengefasst sind.

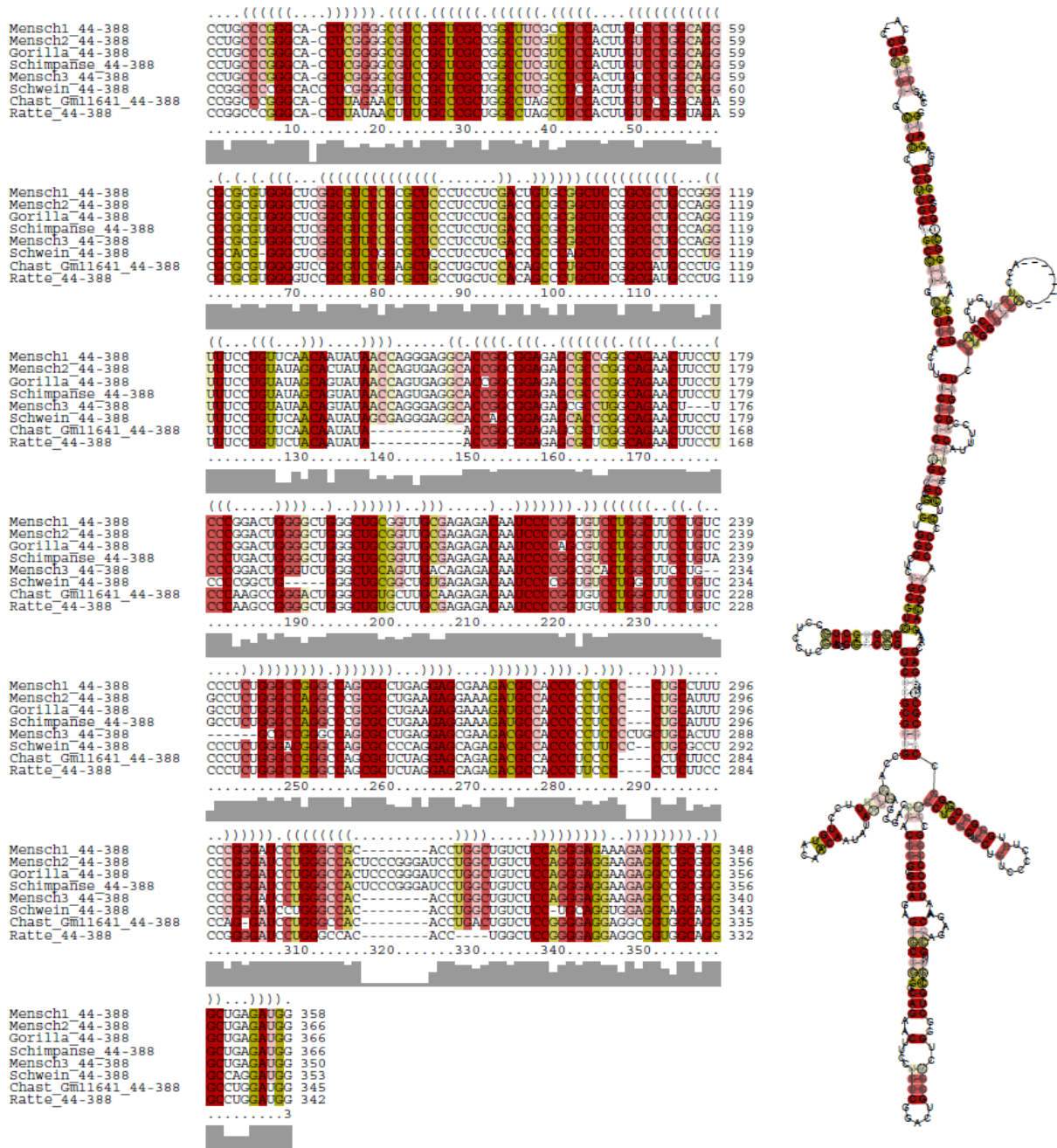
**Tabelle 5.3.1.1A: Sequenzhomologie der hypertrophie-assoziierten lncRNA Chast.** Die Tabelle fasst die Sequenzhomologie der lncRNA Chast in verschiedenen Säugetieren mit der Sequenzübereinstimmung (Identität) und der jeweiligen Position in Chast zusammen.

	Ratte		Mensch		Schwein		Schimpanse		Gorilla	
	Identität	Position	Identität	Position	Identität	Position	Identität	Position	Identität	Position
<b>Gm11641 (Chast)</b>	88 - 94%	1 - 590 588 - 921	66 - 80%	44 - 388 646 - 922	80%	8 - 388	66 - 78%	44 - 380 44 - 388 44 - 441 646 - 922	66 - 77%	44 - 388 44 - 453 85 - 388 646 - 920

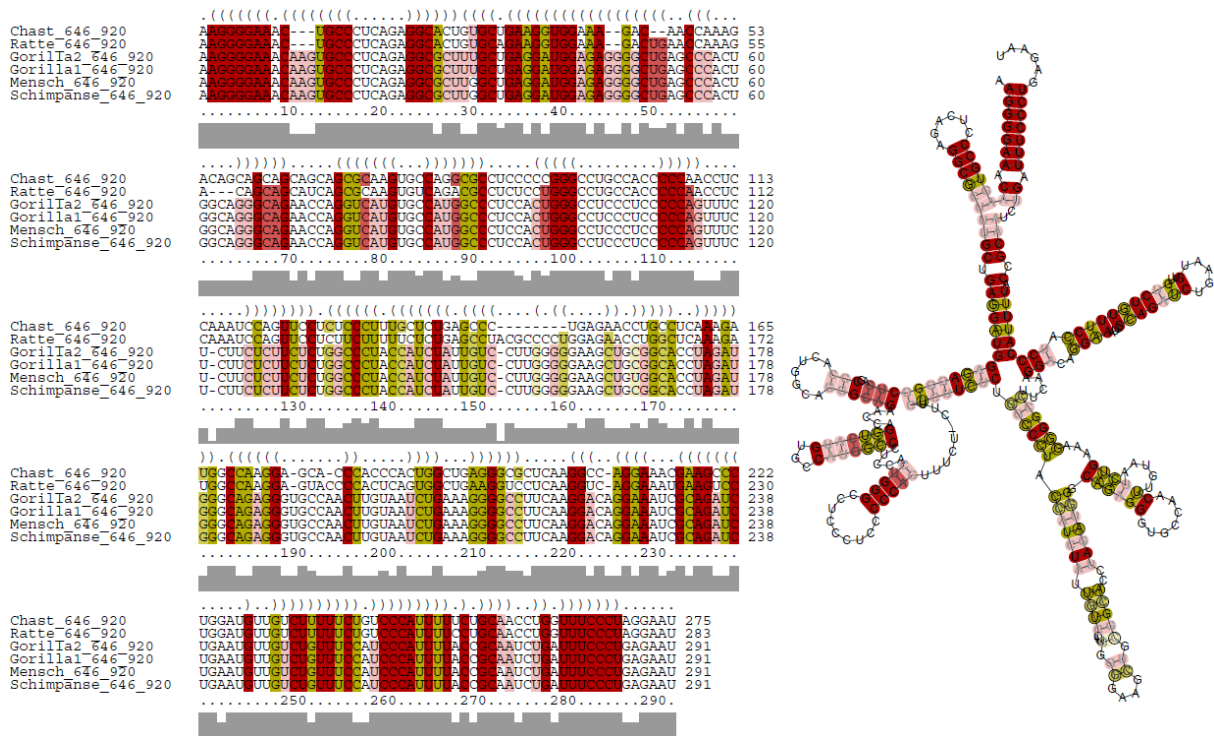


**Abbildung 5.3.1.1A: Überblick über die genomische Position, Sekundärstruktur und mRNA-Interaktion der lncRNA Chast.** (Oben) Chast liegt mit einer Transkriptlänge von 923bp auf chr11: 103,363,213 - 103,364,651 (Forward Strang) in antisense Richtung zu den proteinkodierenden Genen *Arhgap27* und *Plekhl1*. (Mitte) Sekundärstrukturfaltung von Chast, wobei die entsprechende komplementäre Basenpaarung (0=keine Variation; 1=Variation) farbig dargestellt ist (violett und blau zeigen hohe Übereinstimmung der Nukleotide). (Unten) Unter Verwendung des genomischen Kontextes wurde eine potentielle Interaktion von Chast (Position 2045 - 2079) mit der mRNA des antisense liegenden Gens *Plekhl1* (ENSMUST00000041272; Position 818 - 845) gefunden (Energie: -12,39 kcal/mol). (Details siehe Text.)





**Abbildung 5.3.1.1B: Überblick der globalen Sequenz-Struktur-Konservierung für den Sequenzbereich 44 - 388 der lincRNA Chast. (Links)** Globales Sequenz-Struktur-Alignment für den überlappenden Sequenzbereich von 44 - 388 in der Maus. Konservierte Sequenzblöcke sind in verschiedenen Farben entsprechend ihrer Sequenzübereinstimmung und Basenpaarung (z.B. C-G, A-U oder U-G) zwischen den Organismen dargestellt (rot verdeutlicht hohe Konservierung/Basenpaarung). **(Rechts)** Konservierte Sekundärstruktur von Chast basierend auf dem Sequenzalignment (gleicher Farbcode wie das Sequenz-Struktur-Alignment).



**Abbildung 5.3.1.1C: Überblick der globalen Sequenz-Struktur-Konservierung für den Sequenzbereich 646 - 920 der lncRNA Chast. (Links)** Globales Sequenz-Struktur-Alignment für die überlappende Sequenzposition 646 - 920 der lncRNA Chast. Konservierte Sequenzblöcke sind in verschiedenen Farben entsprechend ihrer Sequenzübereinstimmung und Basenpaarung (z.B. C-G, A-U oder U-G) zwischen den Organismen dargestellt (rot verdeutlicht hohe Konservierung/Basenpaarung). **(Rechts)** Konservierte Sekundärstruktur von Chast basierend auf dem Sequenzalignment (gleicher Farbcode wie das Sequenz-Struktur-Alignment).

Die weitere Strukturkonservierung konzentrierte sich auf die entsprechenden homologen Sequenzen für Mensch, Ratte, Schwein, Gorilla und Schimpanse und den überlappenden Sequenzbereich von 44 - 388 in der Maus. Die durchgeführte Analyse hinsichtlich einer globalen Sequenz-Struktur-Konservierung zwischen diesen Säugern ist in Abbildung 5.3.1.1B dargestellt (Sequenzalignment Links, konservierte Struktur Rechts). Da die humane homologe Sequenz an Chast-Position 646 - 922 experimentell validiert werden konnte, wurde zusätzlich eine Sequenz-Struktur-Konservierung für diesen Bereich durchgeführt, auch wenn sich für die Position keine Sequenzkonservierung im Schwein gezeigt hat (Tab. 5.3.1.1A). Das Ergebnis der globalen Sequenz-Struktur-Konservierung ist in Abbildung 5.3.1.1C dargestellt, wobei der überlappende Sequenzbereich von 646 - 920 in der Maus für die homologen Sequenzen für Mensch, Ratte, Gorilla und Schimpanse verwendet wurde (Sequenzalignment Links,

konservierte Struktur Rechts). Die gesamte Sekundärstrukturfaltung von Chast mithilfe von RNAfold ist in Abbildung 5.3.1.1A (Mitte) dargestellt.

Die Analyse nach potentiellen IPn hat folgendes Ergebnis ergeben:

Hinsichtlich der beiden antisense liegenden Gene wurde nach einer potentiellen mRNA-Interaktion für die beiden proteinkodierenden Transkripte Arhgap27\_002 (ENSMUST00000092557) und Plekhm1\_001 (ENSMUST00000041272) gesucht. Hier konnte für Arhgap27 an Position 55 - 480 eine mögliche Interaktion mit Chast (Position 1 - 427) gefunden werden (Energie: -604,99kcal/mol; hier nicht gezeigt), wohingegen sich für Plekhm1 eine potentielle mRNA-Interaktion an Position 818 - 845 mit Chast (Position 2045 - 2079) gezeigt hat (Energie: -12,39 kcal/mol; Abb. 5.3.1.1A (Unten)).

Die Analyse nach PIPn hat 827 potentielle Interaktionen im gesamten murinen Proteom ergeben, wobei wichtige PIn in Tabelle 5.3.1.1B zusammengefasst sind (Discriminative Power>0,75; Interaction Strength>0,5; insgesamt 1920 PIn (hier nicht gezeigt)). Eine funktionelle Analyse dieser Proteine hat 191 verschiedene biologische Prozesse und Signalwege in 19 verschiedenen Funktionsgruppen ergeben, welche in Abbildung 5.3.1.1D dargestellt sind.

Des Weiteren konnten neun regulatorische Elemente in der cDNA von Chast gefunden werden, welche in Tabelle 5.3.1.1C mit Position und bekannten Binfaktoren zusammengefasst sind.

Die Analyse nach miRNA-Interaktionen hat ergeben, dass Chast 546 potentielle miRNA-IP besitzt (hier nicht gezeigt). Die anschließende Filterung unter Verwendung eines selbst erstellten Katalogs von 75 miRNAs (hier nicht gezeigt), welche eine bekannte Rolle bei Herzhypertrophie und/oder eine generelle Funktion im Herzen besitzen, hat eine Überlappung für 15 miRNAs ergeben (bei mehreren Bindestellen nur Führungsstrang (guide strand), Interaktionen mit Passagierstrang (passenger strand) wurden nicht berücksichtigt). Tabelle 5.3.1.1D fasst die gefundenen 15 miRNAs mit deren Bindestelle in Chast zusammen, wobei bei mehreren Bindestellen nur diejenige mit der minimalen Energie berücksichtigt wurde.

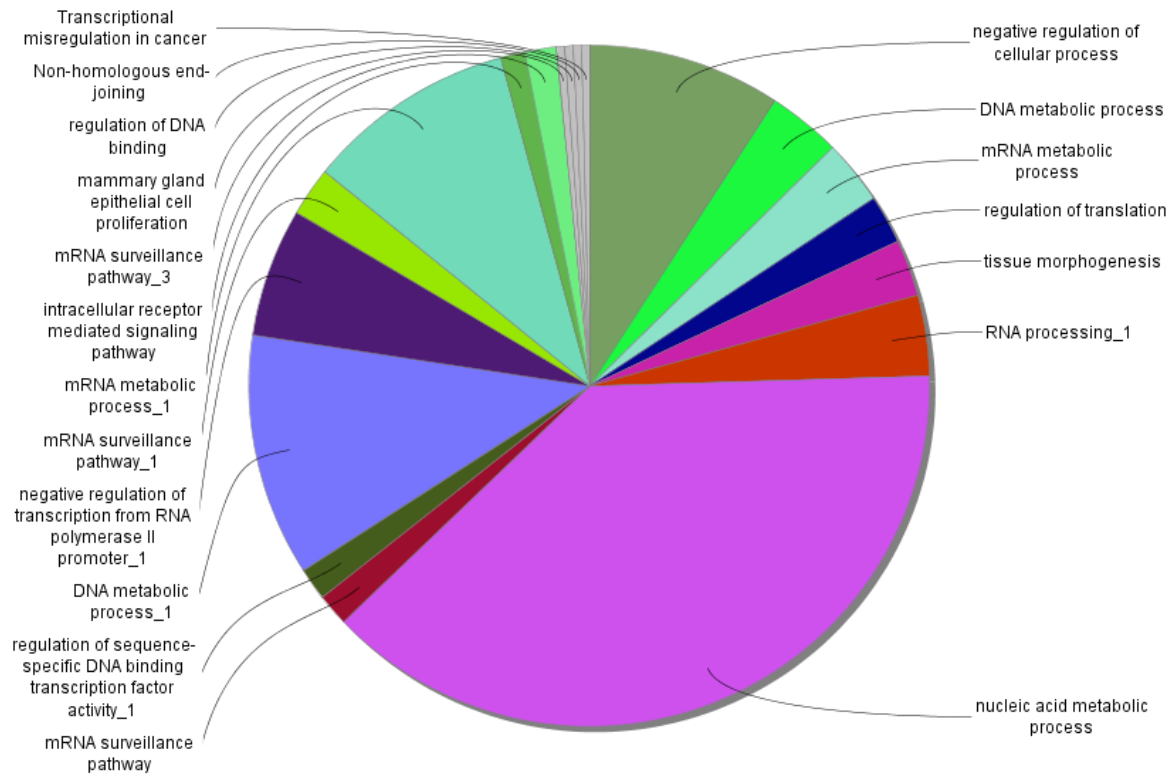
**Tabelle 5.3.1.1B: Überblick der gefundenen Protein-Interaktions-Partner für die lncRNA Chast.** Die Tabelle fasst potentielle murine PIP von Chast mit wichtigen Parametern zusammen, welche ebenfalls eine Verbindung zu den gefundenen COG/KOGs, regulatorischen Elementen und TFBS zeigen (Protein-ID entsprechend UniprotKB-Datenbank (Version 2012\_11); Discriminative Power>0,75 und Interaction Strength>0,5 weisen auf ein hohes Potential für eine PI hin). (Details siehe Text.)

Protein-ID	Z-Score	Discriminative Power	Interaction Strength
Q61188 (EZH2)	2,01	0,99	0,81
Q01147 (CREB1)	0,66	0,94	0,85
P22361 (HNF1A)	1,58	0,99	0,85
Q9QWV9 (CCNT1)	2,37	0,99	0,85
P42225 (STAT1)	1,87	0,99	0,87
P05627 (JUN)	0,55	0,91	0,88
Q60765 (ATF3)	0,37	0,84	0,89
Q99J64 (RBM43)	0,88	0,96	0,91
Q9WVH4 (FOXO3)	2,18	0,99	0,91
P51448 (RORA)	1,93	0,99	0,92
P42228 (STAT4)	2,48	0,99	0,93
Q6SJ95 (TBPL2)	1,28	0,98	0,97

**Tabelle 5.3.1.1C: Überblick der regulatorischen Elemente für die lncRNA Chast.** Die Tabelle fasst die identifizierten regulatorischen Elemente in der cDNA von Chast (Motiv) mit Position und bekannten Binfaktoren zusammen. (Na=nicht angegeben.)

Regulatorisches Element	Binfaktor	Motiv
STAT BS2	Putativer TF	392' ATcTCCATCC '383
Pu.1 (SPI-1)	Pu.1 (SPI-1)	774' AGAGGAACT '766
CyA	C/EBP alpha	516' tTTTgGAAACT '526
IE-A4 B	IRF-7	677' GAAgGTGgAAAAG '688
IE-A4 D	IRF-3	644' GAAAGGgGAAACT '656
NRAP	GABP	215' AAGCCGGgACT '225
C/EBP	C/EBP beta and delta	535' CTTGGGAAAG '544
Site I	C/EBP	887' ACAGaAAAaGACA '875
ISRE	Na	423' CAGAAgCTGAgAGA '410

**Abbildung 5.3.1.1D: Überblick der funktionellen Analyse der potentiellen Protein-Interaktions-Partner von Chast.** Die Interaktionsanalyse hat 827 potentielle PIP im gesamten murinen Proteom ergeben (Discriminative Power>0,75; Interaction Strength>0,5; insgesamt 1920 PIen), welche in 19 Funktionsgruppen mit insgesamt 191 biologischen Prozessen und Signalwegen gruppiert sind (hier: Funktionsgruppen dargestellt).



**Tabelle 5.3.1.1D: Überblick der herz- und hypertrophie-assoziierten miRNA-Interaktionspartner der lncRNA Chast.** Die Tabelle fasst die gefundenen IP mit der Bindeenergie, Bindestelle und deren Sequenzübereinstimmung in der lncRNA zusammen (es wurde jeweils nur die Bindestelle mit dem minimalen Energiebetrag berücksichtigt; bei mehreren Bindestellen nur guide strand; mmu=Mus musculus).

<b>Gm11641 (Chast)</b>			
	<b>Bindestelle</b>	<b>Energie (kCal/Mol)</b>	<b>Übereinstimmung (%)</b>
mmu-miRNA-133b	831 - 851	-21,49	73,91%
mmu-miRNA-133a	830 - 851	-21,49	75,00%
mmu-let-7b	332 - 360	-21,71	75,00%
mmu-miRNA-185	120 - 141	-28,47	78,26%
mmu-miRNA-125b-5p	103 - 126	-21,83	68,00%
mmu-miRNA-125b-1-3p	719 - 739	-24,75	76,19%
mmu-miRNA-125b-2-3p	330 - 355	-24,83	80,77%
mmu-miRNA-103	135 - 165	-24,52	73,33%
mmu-miRNA-208a-5p	275 - 295	-27,79	80,95%
mmu-miRNA-214	125 - 149	-23,29	68,00%
mmu-miRNA-18b	709 - 733	-21,74	72,00%
mmu-miRNA-24	267 - 289	-26,64	82,61%
mmu-miRNA-210	99 - 125	-23,79	67,86%
mmu-miRNA-145	330 - 352	-22,63	72,00%
mmu-miRNA-150	830 - 858	-21,09	65,52%

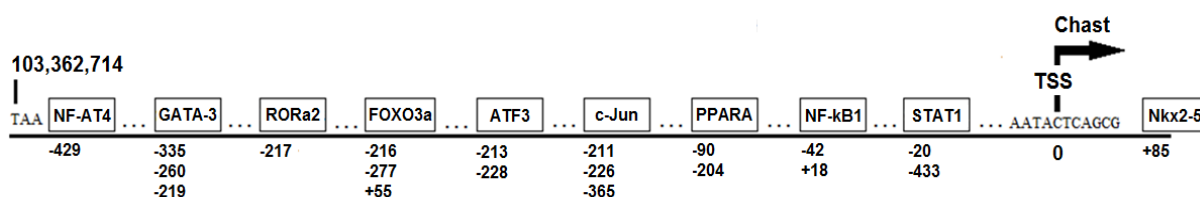
Die COG/KOGs-Suche hat für die cDNA von Chast den KOG1778 (CREB binding protein/P300 and related TAZ Zn-finger proteins) identifiziert (Tab. 5.3.1.1E). Für die Maus sind die beiden Proteine Crebbp und Ep300 für diese orthologe Gruppe annotiert (orthologe Gene aus EggNOG-Datenbank). Die beiden Proteine Crebbp und Ep300 sind ebenfalls der orthologen Gruppe COG5076 (Transcription factor involved in chromatin remodeling, contains bromodomain) zugehörig (orthologe Gruppe aus STRING-Datenbank). Für den COG5076 konnten zehn bekannte funktionelle Interaktionsgruppen gefunden werden (Funktionsvorhersage über STRING-Datenbank), welche in Tabelle 5.3.1.1E mit bekannten orthologen Proteinen (aus STRING- und/oder EggNOG-Datenbank) der Maus zusammengefasst sind.

**Tabelle 5.3.1.1E: Überblick der COG/KOGs-Suche für die lncRNA Chast.** Die lncRNA Chast zeigt eine hohe funktionelle Ähnlichkeit zum KOG1778. Die Tabelle fasst die identifizierten COG/KOGs und orthologe Interaktionsgruppen mit wichtigen Vertretern der Maus zusammen. (Details siehe Text.)

Gruppe	Annotation	Score	Maus-Protein
KOG1778	CREB binding protein/P300 and related TAZ Zn-finger proteins	1	Crebbp, Ep300
COG5076	Transcription factor involved in chromatin remodeling, contains bromodomain	1	Crebbp, Ep300
COG0553	Superfamily II DNA/RNA helicases, SNF2 family	0,990	Smarca1
COG2940	Proteins containing SET domain	0,988	Ezh2
COG5141	PHD zinc finger-containing protein	0,986	Kdm4a
COG5179	Transcription initiation factor TFIID, subunit TAF1	0,984	Taf1
COG2319	FOG-WD40 repeat	0,980	Wdfy1
COG5032	Phosphatidylinositol kinase and protein kinases of the PI-3 kinase family	0,975	Pik3c3
NOG12952	non supervised orthologous group	0,974	Ncor1
COG2101	TATA-box binding protein (TBP), component of TFIID and TFIIB	0,969	Tbpl2
COG5333	Cdk activating kinase (CAK)/RNA polymerase II transcription initiation/nucleotide excision repair factor TFIIH/TFIK, cyclin H subunit	0,969	Ccnt1
NOG85021	non supervised orthologous group	0,965	Rbm43

Für den Promotorbereich von Chast wurde der Bereich von -499 bis +100 bp (chr11: 103,362,714 - 103,363,313) vor dem Transkriptstart gewählt. Die Promotoranalyse hat insgesamt 1226 potentielle TFBS für 311 TF-Motive gefunden, wobei wichtige herz- und stress-assoziierte TFBS exemplarisch in Abbildung 5.3.1.1E dargestellt sind.

**Abbildung 5.3.1.1E: Promotoranalyse der lncRNA Chast.** Die Promotoranalyse (Chast-Position chr11: 103,362,714 - 103,363,313) hat insgesamt 1226 TFBS für 311 TF-Motive ergeben (TFBS-Ähnlichkeitsscore>85%), wovon wichtige herz- und stressassoziierte TFBS mit ihrer relativen Position zur TSS (0; chr11: 103,363,213) zusammengefasst sind.



Chast besitzt vier potentielle humane homologe Sequenzen, wobei drei an Position 44 - 388 und eines an Position 646 - 922 eine Homologie zu Chast zeigen (Tab. 5.3.1.1A). Von diesen konnte für die humane Sequenz chr17: 62,843,244 - 62,843,536 (Chast-Position 646 - 922; NC\_018928.2, 66% Identität) eine entsprechende experimentelle Expression gefunden werden. Für dieses humane Homolog wurde eine Analyse nach einem potentiellen Promotor und einer TSS durchgeführt, wobei für den Bereich chr17: 62,841,244 - 62,843,536 drei potentielle Promotorbereiche gefunden wurden (Tab. 5.3.1.1F; TSS jeweils in Klammern).

**Tabelle 5.3.1.1F: Überblick über die potentiellen Promotorbereiche des humanen Homologs der lncRNA Chast.** Die Tabelle fasst die gefundenen Promotorbereiche des humanen Homologs zu Chast zusammen (als Promotorbereich wurde der 5'-upstream-Bereich der homologen Sequenz um 2000 bp (chr17: 62,841,244 - 62,843,536) verlängert; Position bezieht sich auf den entsprechenden Promotorbereich). (Details siehe Text.)

Position	Score	Promotorsequenz (TSS in Klammern)
45 - 95	0,84	GGCGCGCGGAGCTCCCGGGTGCGGACCCGCCACCTGCCGC(A)CCCTTCCCT
1655 -1705	0,92	GGGGTTTCTCTATATTGGCCAGGCTGGTCTTGAACCTG(A)CCTCAAATG
1971 - 2021	0,80	GTGACTCCGTCTCAAAAAACATGGGTGAGAAGGGGAAAC(A)AGTGCCCTC

Um den genauen Promotorbereich zu finden, wurde im Anschluss ein Sequenzalignment gegen den Chast-Promotor der Maus (chr11: 103,362,714 - 103,363,313) durchgeführt, wobei hier eine Sequenzähnlichkeit von 68% mit dem Bereich um den berechneten Promotor an Position 45 - 95 (Score=0,84) gefunden wurde (Abb. 5.3.1.1F; Query=Chast, Sbjct (Subject)=Mensch). Ebenfalls zeigt sich eine hohe Überlappung der vorhergesagten Promotorsequenz (Position 45 - 95) mit dem Bereich um die TSS von Chast (Abb. 5.3.1.1F; vorhergesagte Promotorposition 45 - 95 (Mensch) als blauer Kasten, Bereich um TSS (Position 439 - 500) von Chast als rote Sterne).



**Abbildung 5.3.1.1F: Darstellung der Überlappung zwischen dem Promotor von Chast und dem des Menschen.** Das Sequenzalignment zeigt eine Übereinstimmung des Promotorbereichs zwischen Mensch (chr17: 62,841,244 - 62,843,536) und Chast (chr11: 103,362,714 - 103,363,313) von 68%, wobei sich eine hohe Überlappung der vorhergesagten Promotorsequenz im Menschen (Position 45 - 95) mit dem Bereich um die TSS von Chast (-61 Nte relativ zur TSS) erkennen lässt (vorhergesagter Promotor und TSS im Menschen als blauer Kasten, Promotorsequenz relativ zur TSS (Position 439 - 500) von Chast als rote Sterne; Query=Chast, Sbjct (Subject)=Mensch). (Details siehe Text.)

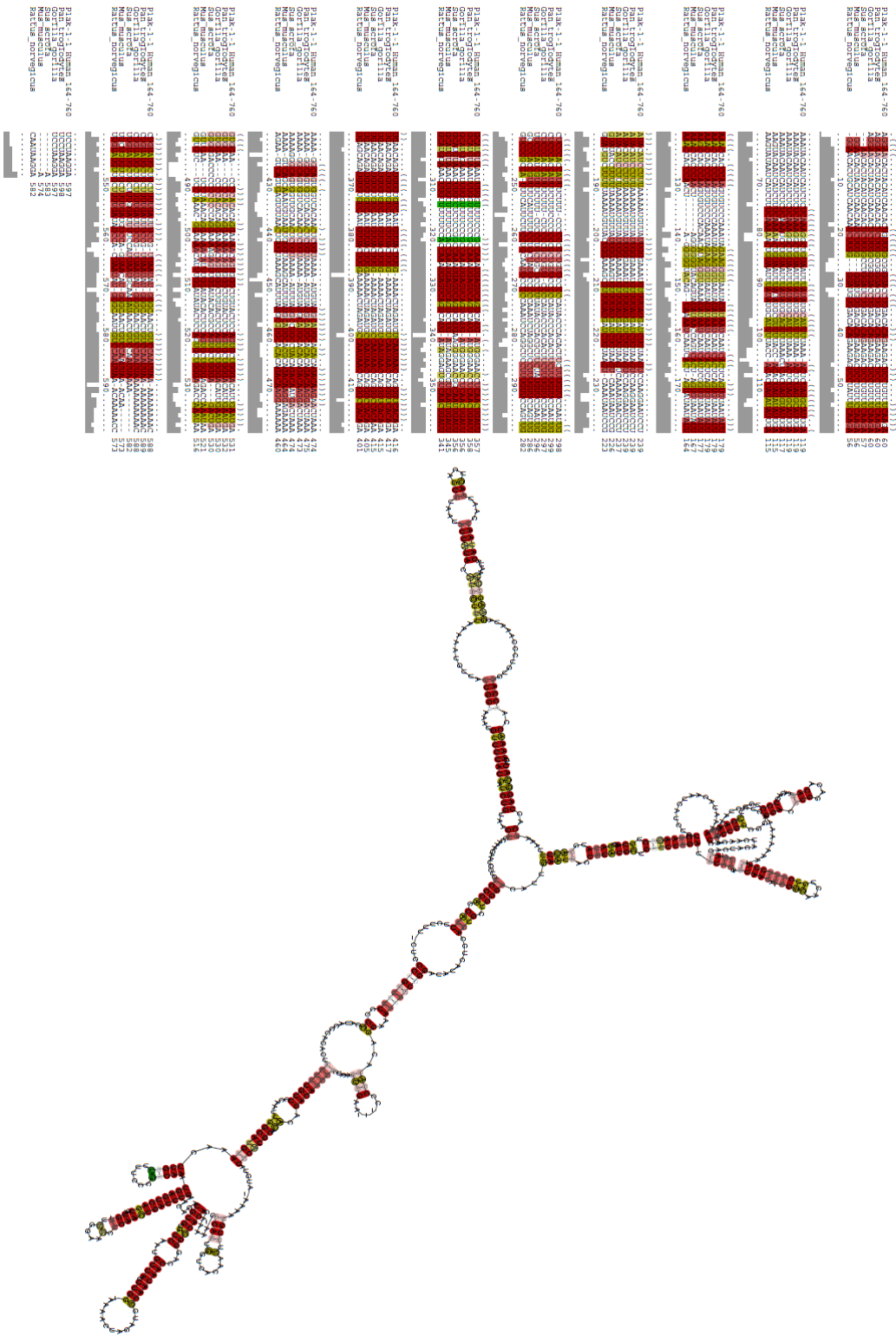
Score	Expect	Identities	Gaps	Strand
62.6 bits(68)	2e-13	145/214(68%)	26/214(12%)	Plus/Plus
Query 407	GCTCTAGCCCCAGTGCAGCGAGAAGGGCAGGC	GGTGC	CGCGGGGCTCTAGGGTGGGGACG-	465
Sbjct 13	GCTCCGCCCCAGCACGGAGACAAGGGGAGCA	GGCG	CGGGAGCTCCC	72
Query 466	GGCACCCTGGGGGCACCTTTCCCGCCCTCCCAATACTCAGCGCCAT---CCTCCCCAACGC			522
Sbjct 73	GCCACCTGCCG-CACCCTTCCCTCCC---CAACCCT--GCCCTTTCCCCACCCCCACCC			126
Query 523	CAT-TCCCCCTCAG-TCCAGCCCGGC-----CGGGCACCTTAGAACTTT			566
Sbjct 127	CACCTCCACCTCAACTCCCGCTCGGGCGGCCACTGCCTGCCCGGCACCTCGGGGCGTC			186
Query 567	CGCCCGCTGGCCTAGCTTCCACTTGTCCCGGCAG		600	
Sbjct 187	CGCTCGCCGGCCTCGTCTCCACTTGTCCCGGCAG		220	

### 5.3.1.2 Hypoxie-induzierte lncRNAs LINC00323-003 und MIR503HG

In einem zweiten Projekt wurden hypoxie-induzierte lncRNAs in Endothelzellen systembiologischen untersucht. Die Analysen haben für die lncRNA LINC00323-003 hinsichtlich des proteinkodierenden Potentials einen Score von -0,743374 ergeben, wohingegen sich für die lncRNA MIR503HG ein Score von -0,173539 ergeben hat. Bei der Sequenzanalyse wurde sich auf die Säugetiere Maus, Ratte, Schwein, Gorilla und Schimpanse konzentriert. Die Sequenzanalyse ist in Tabelle 5.3.1.2 zusammengefasst. Die lncRNA MIR503HG besitzt in all diesen Organismen homologe Gene, dagegen wurden für die LINC00323-003 lediglich im Gorilla und Schimpansen homologe Sequenzen gefunden. Für die weitere Strukturkonservierung wurde dementsprechend lediglich die lncRNA MIR503HG berücksichtigt. Des Weiteren wurde für den Gorilla und Schimpansen nur der Sequenzbereich 164 - 760 und für die MIR503HG nur der überlappende Sequenzbereich von 164 - 760 verwendet. Die Ergebnisse zeigen neben einer Sequenzkonservierung auch eine globale Sequenz-Struktur-Konservierung in Säugetieren. Das entsprechende Sequenzalignment und die konservierte Sekundärstruktur ist in Abbildung 5.3.1.2 dargestellt.

**Tabelle 5.3.1.2: Sequenzhomologie der beiden hypoxie-induzierten lncRNAs LINC00323-003 und MIR503HG.** Die Tabelle fasst die Sequenzhomologie der beiden lncRNAs in verschiedenen Säugetieren mit der Sequenzähnlichkeit (Identität) und der jeweiligen Position in der lncRNA zusammen.

	Maus		Ratte		Schwein		Schimpanse		Gorilla	
	Identität	Position	Identität	Position	Identität	Position	Identität	Position	Identität	Position
<b>LINC00323-003</b>	---	---	---	---	---	---	98 - 100%	84 - 346 342 - 580 1 - 86	98 - 99%	84 - 346 342 - 580 1 - 86
<b>MIR503HG-002</b>	80%	165 - 747	80%	165 - 760	86%	167 - 745	99 - 100%	164 - 760 75 - 166 1 - 81	99 - 100%	164 - 760 75 - 166 1 - 81



**Abbildung 5.3.1.2: Überblick der Sequenz-Struktur-Konservierung der lncRNA MIR503HG.** **(Oben)** Globales Sequenz-Struktur-Alignment. Konservierte Sequenzblöcke sind in unterschiedlichen Farben entsprechend ihrer Sequenzübereinstimmung und Basenpaarung (z.B. C-G, A-U oder U-G) zwischen den Organismen dargestellt (rot verdeutlicht hohe Konservierung/Basenpaarung). **(Unten)** Darstellung der konservierten Sekundärstruktur der lncRNA MIR503HG basierend auf dem Sequenzalignment (gleicher Farbcode wie das Sequenz-Struktur-Alignment).

### 5.3.2 Die DrumPID-Datenbank

Die DrumPID-Datenbank beinhaltet über 1383 FDA-zugelassene Medikamente (und über 5000 nicht-FDA-zugelassene Medikamente) einschließlich deren chemischer Eigenschaften, Targets, Signalwege und assoziierten orthologen Gruppen (Tab. 5.3.2A). Sie verfügt über fünf Suchkategorien inklusive Data-Mining-Anwendungen, Suchabfragen mit mehreren Schlüsselwörtern und Kategorien sowie heuristischer Autovervollständigungsverfahren und Demonstrationssuchen (Abb. 5.3.2A, Tab. 5.3.2A). Zusätzlich bietet DrumPID dem Nutzer zahlreiche Links für weitere individuelle Analysen, etwa zu COG/KOGs (z.B. Datenbanken STRING (Szklarczyk, Franceschini et al. 2015) und EggNOG (Huerta-Cepas, Szklarczyk et al. 2016)), zu organismen- und gewebespezifischen Interaktionen (z.B. Datenbanken HPRD (Keshava Prasad, Goel et al. 2009), IMEx (Orchard, Kerrien et al. 2012) und PlateletWeb (Boyanova, Nilla et al. 2012)), für detaillierte Struktur- und Funktionsvorhersagen (AnDom-Software (Schmidt, Bork et al. 2002) und GoSynthetic-Datenbank (Liang, Kruger et al. 2013)) oder den Screening-Katalog Selleckchem (<http://www.selleckchem.com/>) zur Identifizierung verfügbarer chemischer Substanzen für ein Target.

Die folgenden Anwendungsbeispiele sollen die DrumPID-Datenbank veranschaulichen (Abb. 5.3.2B). Die Beispiele wurden speziell ausgewählt, um das Potenzial und die Stärken der Datenbank anhand typischer wissenschaftlicher Fragestellungen und Herausforderungen zu verdeutlichen.

#### **Beispiel 1: Analyse des Interaktionskontextes für ein Medikament**

Die DrumPID-Datenbank kann für Standardsuchen für Informationen zu einem Medikament verwendet werden oder um ein Medikament für ein bestimmtes Zielprotein zu finden. In diesem Fallbeispiel soll ein geeignetes Medikament gefunden werden, welches die AMP-Kinase (AMPK) in Tumorzellen aktiviert (siehe A549-Zelllinie Punkt 5.2.1). Hierfür findet die DrumPID-Datenbank vier mögliche Medikamente ('AMP-activated protein kinase' in Kategorie „Affected Protein“), z.B. Metformin (Abb. 5.3.2B (A)). Mithilfe von DrumPID kann man nun den therapeutischen Effekt (z.B. Aktivator oder Inhibitor) und den Interaktionskontext genauer untersuchen. Hinsichtlich Metformin wird deutlich, dass es primär zur Behandlung von Diabetes verwendet wird. Betrachtet man die P1en und beteiligten Netzwerke genauer, so lässt sich erkennen, dass Metformin nicht nur eine Verbindung zum mTOR-Signalweg hat, sondern auch zum AMPK- und Langlebigkeits-Signalnetzwerk (Abb. 5.3.2B (A); AMPK-Interaktion

wurde aus KEGG und DrugBank gefunden; das gesamte AMPK-Interaktionsnetzwerk kann durch Bewegen des Mauszeigers über das Signalnetzwerk oder durch Folgen der implementierten Links, z.B. zu KEGG oder IMEx, genauer betrachtet werden, hier nicht gezeigt). Diese Verbindungen zeigen nicht nur einen Einfluss von Metformin auf eine mögliche längere Lebenszeit bei Diabetespatienten (Bannister, Holden et al. 2014), sondern auch einen Effekt auf AMPK und mTOR. Es ist bekannt, dass AMPK in Tumorzellen herunterreguliert ist und sowohl p53 aktiviert (assoziiert mit Apoptose) als auch mTOR inhibiert (assoziiert mit Metabolismus und Proliferation; siehe Punkt 5.2.1). Somit könnte Metformin ein potentieller Aktivator für den AMPK-Signalweg in Tumorzellen sein (siehe auch Link zu Selleckchem, hier nicht gezeigt), um auf diesem Weg die Apoptose und den mTOR-Signalweg zu beeinflussen (Abb. 5.3.2B (A); mTOR-Signalnetzwerk ko04150). Demnach könnte Metformin nicht nur zur Behandlung von Diabetes verwendet werden, sondern auch als Off-Label-Medikament zur Behandlung von AMPK-mutierten Tumorerkrankungen von Interesse sein. Hinsichtlich einer spezifischen AMPK-Aktivierung für die A549-Zelllinien (siehe Punkt 5.2.1) ist Metformin aufgrund seiner verschiedenen Netzwerkregulation und seines breiten Wirkmechanismus eher ungeeignet. Hierfür kann dem bereits angesprochenen Link zu Selleckchem gefolgt werden, um einen selektiven AMPK-Aktivator zu finden, im konkreten Fall AICAR (hier nicht gezeigt).

### **Beispiel 2: Analyse des gewebespezifischen Interaktionskontextes für ein Protein**

In diesem Beispiel soll gezeigt werden, wie mit der DrumPID-Datenbank ein Medikament unter Berücksichtigung seines gewebespezifischen Interaktionskontextes gefunden werden kann. Für das Target Platelet-Derived Growth Factor Receptor Beta (PDGFRB; assoziiert mit Wundheilung und Zellregenerierung/-wachstum nach einem Infarkt) findet DrumPID vier Medikamente, z.B. Dasatinib ('platelet-derived growth factor receptor' in Kategorie „Affected Protein“; Abb. 5.3.2B (B)). Durch einfaches Scrollen auf der Ergebnisseite der Datenbank wird ersichtlich, dass all diese Medikamente nicht nur PDGFRB beeinflussen, sondern auch viele weitere Targets. Insbesondere für Medikamente mit einem breiten Wirkmechanismus ist es häufig von großem Interesse, seine IP und Signalwegregulation in einem gewebespezifischen Kontext zu untersuchen, um so das bestmögliche Medikament zu finden. Solch eine spezifische Analyse kann die DrumPID-Datenbank leisten. Um den Effekt von PDGFRB in Blutplättchen zu untersuchen, kann dem PDGFRB-Targetlink zur PlateletWeb-Datenbank gefolgt werden, wo für PDGFRB insgesamt 46 IP im Blutplättchen und 20 IP in anderen menschlichen Geweben gefunden werden (weitere Möglichkeit z.B. über Link zur STRING-Datenbank, hier

nicht gezeigt). Das generierte Netzwerk für die IP von PDGFRB im Blutplättchen ist in Abbildung 5.3.2B (B) dargestellt (46 experimentell validierte IP von PDGFRB aus PlateletWeb heruntergeladen und mit Cytoscape 2.8.3 (Shannon, Markiel et al. 2003) visualisiert; PDGFRB-PDGFRB Dimer in blau; SRC, YES1 und FYN in grün; Dasatinib in rot; Interaktionen in grau). Betrachtet man nun den Interaktionskontext im Blutplättchen mit den gefundenen Medikamenten in DrumPID für PDGFRB und deren weitere Targets, so wird eine Verbindung zwischen PDGFRB, SRC, YES1 und FYN deutlich, die auf Dasatinib (beeinflusst alle vier) als geeigneten Kandidaten für das PDGFRB-Netzwerk im Blutplättchen hinweist. Dies verdeutlicht, dass die DrumPID-Datenbank in der Lage ist ein Medikament zu finden, welches den bestmöglichen gewebespezifischen Interaktionskontext abdeckt.

### **Beispiel 3: Analyse des organismenspezifischen Interaktionskontextes mittels COG/KOGs**

Neben einem gewebespezifischen Kontext erlaubt die DrumPID-Datenbank auch organismenspezifische Vergleiche. Hierdurch können z.B. Unterschiede für ein bestimmtes Target zwischen Organismen gefunden werden, was insbesondere für die Entwicklung neuer Medikamente bzw. Substanzen von großem Interesse ist, etwa um spezifisch das bakterielle Protein zu inhibieren, ohne einen schädlichen Einfluss auf das menschliche Protein zu haben. Als Fallbeispiel sollen hier die beiden Enzyme Ergosterol und Riboflavin für den Schimmelpilz *Aspergillus fumigatus* (*A. fumigatus*) gezeigt werden. *A. fumigatus* verursacht die sogenannte Aspergillose, eine Infektionserkrankung bei (meist) immungeschwächten Patienten. Eine Behandlung ist dabei schwierig, da es aufgrund der hohen Ähnlichkeiten zwischen *A. fumigatus* und den menschlichen Proteinen häufig zu toxischen Nebenwirkungen bei den Patienten kommt. Die DrumPID-Datenbank findet für das in *A. fumigatus* bereits bekannte Enzym Ergosterol sechs Medikamente, welche sich allesamt bereits in der klinischen Anwendung befinden, etwa Amphotericin B (hier nicht gezeigt). In den meisten Fällen ist aber ein Medikament bzw. eine Substanz für ein bestimmtes, neu identifiziertes Target noch nicht bekannt. Hierfür kann die in DrumPID implementierte COG/KOGs-Suche helfen, um orthologe Gene in verschiedenen Organismen (hier Mensch und *Aspergillus*) zu finden. Für die Riboflavinkinase findet DrumPID das Vitamin B2 Riboflavin ('Riboflavin kinase' in Kategorie „Affected Protein“) sowie den KOG3110 (Riboflavin kinase) (Abb. 5.3.2B (C)). Mithilfe des implementierten Links zur EggNOG-Datenbank können die entsprechenden orthologen Gene im Menschen (Riboflavin-Kinase-Protein RFK) und in verschiedenen *Aspergillus*-Stämmen

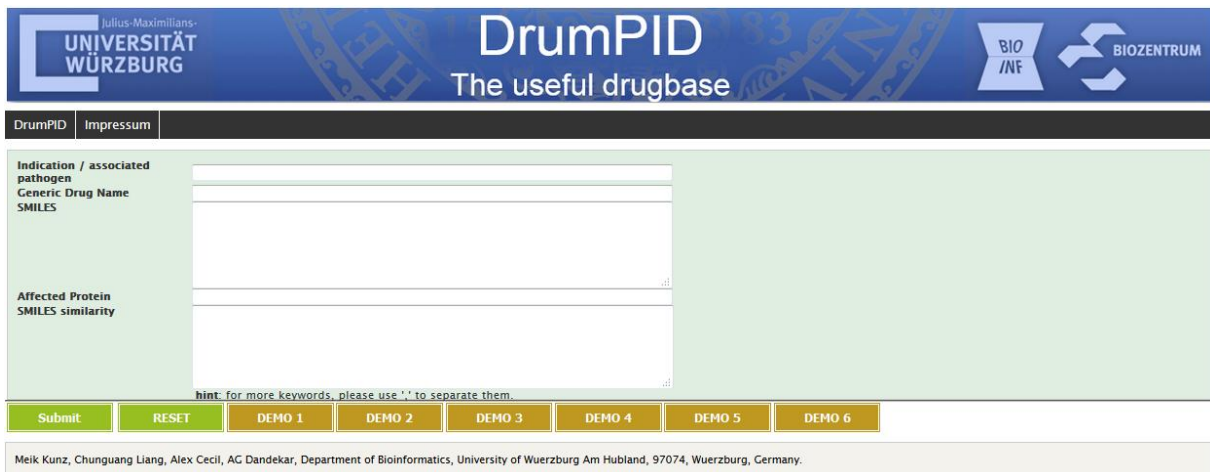
gefunden werden, z.B. *Aspergillus terreus* ATEG\_06773 und *Aspergillus niger* FMN1 (hier nicht gezeigt). In einem nächsten Schritt kann für die gefundenen Proteinsequenzen anhand eines Sequenzvergleichs mittels des BLAST-Algorithmus die entsprechende Riboflavinkinase in *A. fumigatus* (Af293) identifiziert werden (hier nicht gezeigt). Ein hieran anschließendes Sequenzalignment mittels BLAST zeigt lediglich eine Identität (Sequenzübereinstimmung) von 38 - 41% zwischen den drei *Aspergillus*-Stämmen und der humanen Riboflavinkinase (Tab. 5.3.2B, Links), wohingegen die Ähnlichkeit zwischen *A. fumigatus* und den beiden *Aspergillus*-Stämmen *Aspergillus terreus* und *Aspergillus niger* über 70% beträgt (Tab. 5.3.2B, Rechts). Dies weist für potentielle Medikamente auf flexible Regionen zwischen Mensch und *Aspergillus* hin. Hinsichtlich der Medikamentenentwicklung sind die funktionellen Bindestellen in einem Protein von Bedeutung. Entsprechend der UniProt-Datenbank besitzt die Riboflavinkinase im Menschen und den drei *Aspergillus*-Stämmen eine ähnlich aktive Seite (siehe Stern unter Nukleotid E) und Metallbindestelle (siehe Stern über Nukleotid T) (Abb. 5.3.2C, Links). Betrachtet man die flankierenden Regionen dieser Bindestellen im Sequenzalignment genauer, so wird deutlich, dass sich diese unterscheiden (Abb. 5.3.2C, Links; flankierende Region um die aktive Seite in blau; flankierende Region um die Metallbindestelle in violett; Suche (Query)=Mensch, Sbjct (Subjekt)=*Aspergillus*). Im Sequenzalignment wird ebenfalls deutlich, dass zwei Gapregionen (in rot und grün) zwischen *Aspergillus* und Mensch bestehen (Abb. 5.3.2C, Links). Dies deutet darauf hin, dass neben unterschiedlichen flankierenden Regionen zusätzlich auch flexible Bereiche zwischen *Aspergillus* und Mensch existieren, die dazu dienen können, spezifisch die Kinase in *Aspergillus* zu blocken, ohne einen Einfluss auf die menschliche Riboflavinkinase zu haben. Zusätzlich ist es oft notwendig, stammspezifische Unterschiede zu finden, um z.B. geeignete Medikamente gegen jeden einzelnen Pilzstamm zu entwickeln, die noch effizienter bei einer entsprechenden Infektion wirken. Das Sequenzalignment zwischen den *Aspergillus*-Stämmen zeigt, dass die grüne Gapregion zur humanen Riboflavinkinase, trotz erkennbarer Sequenzunterschiede, in allen drei Pilzstämmen vorhanden ist, sich jedoch die rote Gapregion zwischen den Pilzstämmen unterscheidet (Abb. 5.3.2C, Rechts; Suche (Query)= *Aspergillus fumigatus*, Sbjct (Subjekt)=beide anderen *Aspergillus*-Stämme). Dies lässt vermuten, dass die grüne Gapregion für ein Medikament mit breiter antibiotischer Wirkung dienen kann, wohingegen die rote Gapregion für unterschiedliche Substanzen zwischen den verschiedenen Pilzstämmen geeignet ist. Die gefundenen und weiteren Informationen aus der DrugPID-Datenbank, etwa 3D-Struktur von Riboflavin als PDB-Download und Homologiemodelle der Riboflavinkinase von

*A. fumigatus* (drei Strukturen verfügbar über AnDom und PDB), ermöglichen im weiteren Verlauf die Medikamenten-Target-Interaktion genauer zu untersuchen und das Riboflavin geeignet zu modifizieren, etwa durch Einfügen oder Entfernen von Seitenketten, um spezifisch die Riboflavinkinase in *A. fumigatus* zu blockieren, ohne einen toxischen Effekt im Menschen zu erzeugen (hier nicht gezeigt).

#### Beispiel 4: Identifizierung ähnlicher Zielstrukturen (Nebenwirkungen) für ein Medikament/eine Substanz

DrumPID ermöglicht zusätzlich ähnliche Substrukturen für ein Medikament/eine Substanz zu finden. Die Ähnlichkeitssuche für den SMILES-String (Dexamethason):

[H][C@@@]12C[C@@H](C)[C@](O)(C(=O)CO)[C@@]1(C)C[C@H](O)[C@@]1(F)[C@@]2([H])CCC2=CC(=O)C=C[C@]12C (Kategorie „SMILES similarity“) berechnet 61 Substrukturen (Tanimoto-Ähnlichkeitsscore>0,66), darunter Dexamethason und Betamethason mit einem Score von 1 (Abb. 5.3.2B (D); hier Treffer>0,96 gezeigt). Eine anschließende Suche mit den gefundenen Substrukturen in der Kategorie „SMILES“ zeigt, dass die gefundenen Treffer ähnliche Substrukturen u. a. mit gleichen Targets, z.B. Glukokortikoid-Rezeptor, aber auch weitere Nebeneffekte aufweisen (hier nicht gezeigt). Diese eingebaute Suchfunktion in der DrumPID-Datenbank ist somit hilfreich, um konservierte Strukturbereiche und mögliche Targets zu identifizieren, etwa für neu entwickelte Substanzen mit noch unbekannter Funktion.



**Abbildung 5.3.2A: Benutzeroberfläche und Suchkategorien der DrumPID-Datenbank.** DrumPID ermöglicht dem Nutzer Suchen in den Kategorien: „Indication/associated pathogen“, „Generic Drug Name“, „SMILES“, „Affected Protein“ und „SMILES similarity“. Hierbei sind Suchanfragen mit mehreren Wörtern in einer sowie mehreren Kategorien möglich, zusätzlich stehen eine Autovervollständigung, Data-Mining sowie Demoversionen zur Verfügung. (Abbildung wurde in ähnlicher Form in Kunz, Liang et al. (2016) veröffentlicht.)



**Tabelle 5.3.2A: Überblick über die einzelnen Suchmöglichkeiten und der Ergebnisseite der DrumPID-Datenbank.** (Tabelle wurde in ähnlicher Form in Kunz, Liang et al. (2016) veröffentlicht.)

Kategorie	Beschreibung
<b><u>Suchmöglichkeiten</u></b>	
Indication	Diese Kategorie sucht nach der Indikation für ein Medikament (Data-Mining-Suche).
Associated pathogen	Diese Kategorie sucht nach krankheitsauslösenden Pathogenen, wogegen ein Medikament wirkt (Data-Mining-Suche).
Generic Drug Name	In dieser Suchkategorie kann nach einem Medikament gesucht werden.
SMILES	Diese Kategorie erlaubt Suchen mit dem SMILES-String für ein Medikament.
Affected Protein	Diese Suchkategorie findet alle Medikamente für ein entsprechendes Target.
SMILES similarity	Die Kategorie berechnet für einen SMILES-String die vorhandenen ähnlichen Substrings in der Datenbank basierend auf dem Tanimoto-Ähnlichkeitsscore (Fingerprint-Funktion aus dem R-Paket rcdk; Ergebnisse mit einem Threshold>0,66 werden auf einer neuen Seite angezeigt).
<b><u>Ergebnisseite</u></b>	
Drug ID	Die DrumPID-ID für ein Medikament ist angegeben.
Generic Name	Der Name des Medikaments wird angezeigt.
External Links	Links zu externen Datenbanken (z.B. DrugBank, Drugs.com und Selleckchem) werden angezeigt.
Pharmacological properties	Die pharmakologische Beschreibung für ein Medikament entsprechend der DrugBank ist gegeben.
Indication properties	Informationen zur Indikation für ein Medikament entsprechend der DrugBank werden angezeigt.
Structure	Die Struktur für ein Medikament wird gezeigt.
SMILES and PDB structure	Der SMILES-String für ein Medikament und der PDB-Strukturfile werden angezeigt inklusive Download-Option.
Chemical formula	Die chemische Formel für ein Medikament wurde berechnet.
Atom count	Die Anzahl der Atome für ein Medikament wurde berechnet.
Mass	Die molare Masse für ein Medikament wurde berechnet (Lipinski's rule of five).
H-bond donor count	Die Wasserstoffbrücken-Donore für ein Medikament wurden berechnet (Lipinski's rule of five).
H-bond acceptor count	Die Wasserstoffbrücken-Akzeptoren für ein Medikament wurden berechnet (Lipinski's rule of five).
logP	Der logP-Wert für ein Medikament wurde berechnet (Lipinski's rule of five).
Ring count	Die Anzahl der Atomringe für ein Medikament wurde berechnet.
Polar surface area	Die Polar surface area für ein Medikament wurde berechnet.
van-der-Waals surface area	Die van-der-Waals surface area für ein Medikament wurde berechnet.
Target pathways	Der Signalweg für ein Target (aus DrugBank und KEGG inkl. Herkunftsschema) ist angegeben.
Protein binding	Der Prozentsatz der Proteinbindung für ein Medikament ist angezeigt.
Drug target	Das Drugtarget (aus DrugBank und KEGG inkl. Herkunftsschema) ist angegeben mit zusätzlichen Links zu den Datenbanken PlateletWeb, AnDom, GoSynthetic, HPRD, iHOP, STRING, KEGG und IMEX.
Ortholog Group of target protein	Für jedes Target wurde die orthologe Gruppe (COG/KOGs) berechnet und wird mit Annotation und E-Value angegeben inkl. zusätzlichen Links zu den Datenbanken GoSynthetic, STRING und EggNOG.

**Tabelle 5.3.2B: Sequenzalignment der Riboflavinkinase zwischen dem Menschen und *Aspergillus*.** Das Sequenzalignment der Riboflavinkinase zeigt eine Sequenzübereinstimmung (Identität) von 38 - 41% zwischen dem Menschen und den drei *Aspergillus*-Stämmen (siehe Spalte Mensch), wohingegen die beiden *Aspergillus*-Stämme *Aspergillus terreus* und *Aspergillus niger* eine Sequenzähnlichkeit von über 70% zu *A. fumigatus* aufweisen (siehe Spalte *A. fumigatus*).

	<b>Mensch</b>		<b><i>A. fumigatus</i></b>	
	<b>Identität</b>	<b>Coverage</b>	<b>Identität</b>	<b>Coverage</b>
<b><i>A. fumigatus</i></b>	38%	84%	-	-
<b><i>Aspergillus terreus</i></b>	39%	82%	74%	100%
<b><i>Aspergillus niger</i></b>	41%	84%	73%	92%

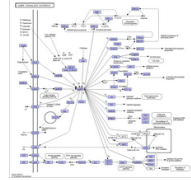
**A)**

Drug ID: 233    Generic Name: Metformin

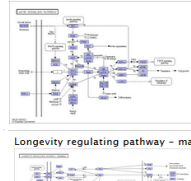
Resource - Target pathways

Drugbank: -

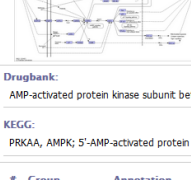
KEGG: AMPK signaling pathway [PATH:ko04152]



mTOR signaling pathway [PATH:ko04150]



Longevity regulating pathway - mammal [PATH:ko04211]



Resource Drug target

Drugbank: AMP-activated protein kinase subunit beta-1

KEGG: PRKAA, AMPK; 5'-AMP-activated protein kinase, catalytic alpha subunit [EC:2.7.11.11]

Orthologous Group of target protein

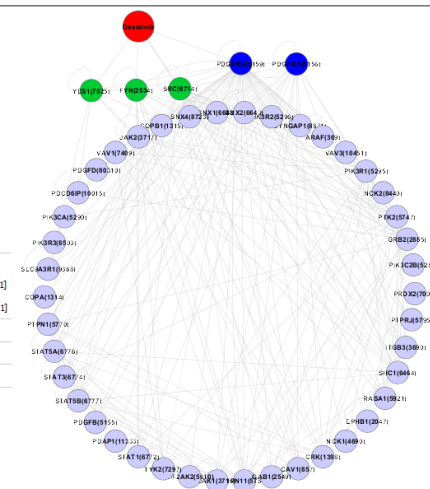
#	Group	Annotation	E-Value
1	KOG1616	Protein involved in Snf1 protein kinase complex assembly	0.0

**B)**

Drugbank: Proto-oncogene tyrosine-protein kinase ABL1, Tyrosine-protein kinase ABL2, Proto-oncogene tyrosine-protein kinase Src, Proto-oncogene tyrosine-protein kinase Fyn, Proto-oncogene tyrosine-protein kinase LCK, Proto-oncogene tyrosine-protein kinase Yes, Mast/stem cell growth factor receptor, Beta platelet-derived growth factor receptor, Ephrin type-A receptor 2, Signal transducer and activator of transcription 5B

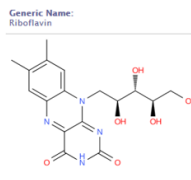
Resource Drug target

KEGG: PDGFRB; platelet-derived growth factor receptor beta [EC:2.7.10.1], KIT, SCFR; proto-oncogene tyrosine-protein kinase Kit [EC:2.7.10.1], EphA2, EOK; Eph receptor A2 [EC:2.7.10.1], SRC; tyrosine-protein kinase Src [EC:2.7.10.2], ABL1; abelson tyrosine-protein kinase 1 [EC:2.7.10.2]



**C)**

Generic Name: Riboflavin



Drugbank: Flavin reductase, Riboflavin kinase, Riboflavin synthase alpha chain

Resource Drug target

KEGG: -

Orthologous Group of target protein

#	Group	Annotation	E-Value
1	KOG3110	Riboflavin kinase	1,00E-121
2	COG0196	For more details please follow the cross link	
3	COG0702	EggNOG CoSynthetic KOG3110 Riboflavin kinase	
4	COG9451	Nucleoside-diphosphate-sugar epimerases	7,00E-05
5	COG0307	Riboflavin synthase alpha chain	1,00E-159
6	KOG3310	Riboflavin synthase alpha chain	3,00E-37

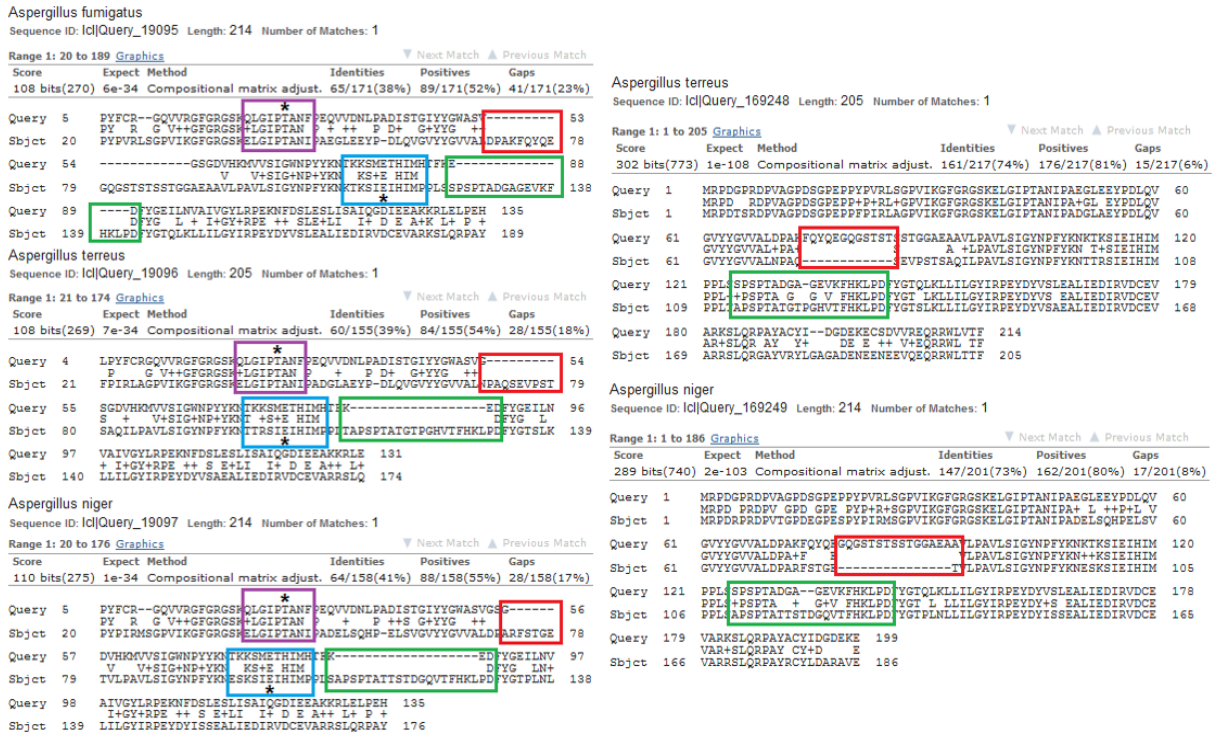
**D)**

SMILES-String: [H][C@H]1[C@@H]([C@@H]([C@H]([C@H]1O)O)O)O

SCORE...hit#

[H][C@H]1[C@@H]([C@@H]([C@H]([C@H]1O)O)O)O	1.0000000
[H][C@H]1[C@@H]([C@@H]([C@H]([C@H]1O)O)O)O	1.0000000
[H][C@H]1[C@@H]([C@@H]([C@H]([C@H]1O)O)O)O	0.9800000
[H][C@H]1[C@@H]([C@@H]([C@H]([C@H]1O)O)O)O	0.9607843
[H][C@H]1[C@@H]([C@@H]([C@H]([C@H]1O)O)O)O	0.9600000
[H][C@H]1[C@@H]([C@@H]([C@H]([C@H]1O)O)O)O	0.9600000

**Abbildung 5.3.2B: Suchbeispiele mit der DrugPID-Datenbank. (A)** Suche um ein Medikament für ein bestimmtes Target zu finden (hier ‘AMP-activated protein kinase’ in Kategorie „Affected Protein” liefert Metformin; mit Informationen zu Signalwegen, Interaktionen und COG/KOGs gezeigt). **(B)** Suche nach einem Medikament unter Berücksichtigung seines gewebespezifischen Interaktionskontextes (hier ‘platelet-derived growth factor receptor’ in Kategorie „Affected Protein” liefert Dasatinib; mit Informationen zu Signalwegen und Interaktionen gezeigt; ebenfalls ist das PDGFRB-Interaktionsnetzwerk im Blutplättchen dargestellt (PDGFRB-PDGFRB Dimer in blau; SRC, YES1 und FYN in grün; Dasatinib in rot; Interaktionen in grau). **(C)** Suche nach COG/KOGs um organismenspezifische Unterschiede für ein Target zu finden (hier ‘Riboflavin kinase’ in Kategorie „Affected Protein” liefert Riboflavin und z.B. den KOG3110). **(D)** Suche nach ähnlichen Substrukturen für einen SMILES-String (hier Treffer>0,96 für Suche mit dem SMILES-String für Dexamethason in Kategorie „SMILES similarity” gezeigt). (Details siehe Text; Abbildung verändert nach Kunz, Liang et al. (2016).)



**Abbildung 5.3.2C: Darstellung der funktionellen Bindestellen der Riboflavin kinase. (Links)** Das Sequenzalignment zeigt, dass der Mensch und die drei *Aspergillus*-Stämme eine ähnliche aktive Seite (siehe Stern unter Nukleotid E) und Metallbindestelle (siehe Stern über Nukleotid T) besitzen, sich allerdings die flankierenden Bereiche unterscheiden (flankierende Region um die aktive Seite in blau; flankierende Region um die Metallbindestelle in violett). Zusätzlich zeigen sich zwei Gapregionen (in rot und grün) zwischen *Aspergillus* und dem Menschen (Suche (Query)=Mensch, Sbjct (Subjekt)=*Aspergillus*). **(Rechts)** Das Sequenzalignment zwischen den *Aspergillus*-Stämmen zeigt, dass die grüne Gapregion, trotz sichtbarer Sequenzunterschiede, in allen drei Pilzstämmen vorhanden ist, sich aber die rote Gapregion unterscheidet (Suche (Query)= *Aspergillus fumigatus*, Sbjct (Subjekt)=beide anderen *Aspergillus*-Stämme). (Details siehe Text; Abbildung wurde in ähnlicher Form in Kunz, Liang et al. (2016) veröffentlicht.)

## 6 Diskussion

### 6.1 Eine grundlegende systembiologische Methode für verschiedene wissenschaftliche Fragestellungen und biologische Aspekte

Der zunehmende Bedarf an biomedizinischer (z.B. Krebserkrankungen) und biotechnologischer (z.B. Nahrungsknappheit) Forschung produziert eine Vielzahl an biologischen Daten. Hierfür sind Analyseverfahren notwendig, welche insbesondere ermöglichen, die generierten experimentellen Datensätze individuell zu analysieren. PPI-Netzwerke sind dynamisch in funktionellen Modulen organisiert, welche durch Hubproteine kontrolliert werden (Tully, Hill et al. 2014) (Dong, Jiang et al. 2015). Änderungen der transkriptionellen Genexpression beeinflussen zwar nicht zwangsläufig die PPIen, jedoch haben Untersuchungen gezeigt, dass eine Überlappung zwischen beiden Datenebenen eine entsprechende Regulierung und Anpassung des zellulären Systems erfordert (Tully, Hill et al. 2014) (Tsuda and Somssich 2015). Ein Problem ist häufig, dass wichtige zentrale Netzwerkknoten oftmals nicht direkt in den experimentellen Datensätzen sichtbar werden und sich indirekte IP stark deregulierter Knoten als geeignete Drugtargets anbieten, welche mittels netzwerkbiologischer Untersuchungen gefunden werden können. Es ist leicht nachvollziehbar, dass der Fortschritt in der biowissenschaftlichen Forschung auch mit einer enormen Entwicklung an systembiologischen Analyseprogrammen und -software einhergeht, welche sich meist auf ein Problem konzentrieren, allerdings oftmals leider kein direkter methodischer Ansatz für eine integrierte Analyse existiert, der effizient von Jedermann angewendet werden kann.

Um diesbezüglich eine Analysemöglichkeit zur Verfügung zu stellen, wurde in dieser Arbeit eine neue systembiologische Methode entwickelt (Naseem, Kunz et al. 2014) (Kunz, Dandekar et al. 2016). Diese erlaubt verschiedene genomische Datenebenen, etwa Transkriptom (z.B. Genexpressionsarrays) und Interaktom (z.B. PPIen von Pflanzen), miteinander zu kombinieren und in einem integrierten Analyseansatz mit dem Ziel zu untersuchen, neue funktionelle Hubproteine zu identifizieren. Die gemeinsamen Gene aus dem Transkriptom und Interaktom rekonstruieren ein PPI-Netzwerk, welches netzwerkbiologisch und funktionell (z.B. Module) analysiert werden kann, um funktionelle Hubproteine zu detektieren. Die entwickelte Methode gibt einen grundlegenden Überblick über einzelne Analyseschritte und ist auf verschiedene biologische Aspekte (z.B. Immunabwehr und Zellwachstum) und Organismen (z.B. Pflanzen und Mensch) übertragbar. So wurden etwa mit der Methode HIF1A und AMPK *in silico* als

neue Drugtargets bei Lungenkrebs sowie AHK2 und ARR14 als neue Modulatoren einer CK-vermittelten Immunabwehr gefunden. Die entwickelte systembiologische Methode liefert demnach das Handwerkszeug für individuelle wissenschaftliche Fragestellungen und verschiedene biologische Aspekte.

## **6.2 Einfluss verschiedener Virulenzfaktoren auf die Pathogenabwehr von *A. thaliana***

Die pflanzliche Immunabwehr wird durch Hormone und deren Interaktionen reguliert, wobei eine antagonistische Interaktion zwischen SA und JA/ET die Grundlage der Immunabwehr bildet (Verhage, van Wees et al. 2010) (Robert-Seilaniantz, Grant et al. 2011) (Grant and Jones 2009) (Caarls, Pieterse et al. 2015). Eine erst seit kurzem bekannte Rolle in der pflanzlichen Immunabwehr wird dem Hormon CK zugesprochen (Naseem, Kunz et al. 2014) (Kunz M 2013). Die Signalübertragung von CK erfolgt über ein Zweikomponentensystem, wobei AHKs CKe erkennen und AHPs das CK-Signal an ARRn übertragen (Naseem and Dandekar 2012) (Naseem, Kunz et al. 2013) (Kunz M 2013). CKe sind in der Lage, die Immunabwehr über SA zu beeinflussen und können zu einer gesteigerten Resistenz beitragen, die zu Grunde liegenden Mechanismen sind jedoch bislang noch nicht vollständig aufgeklärt (Naseem, Kunz et al. 2014) (Kunz M 2013). Es ist bekannt, dass pathogene Abwehrstoffe mit Hubproteinen und regulatorischen Proteinen interagieren, wichtige pflanzliche Signalwege und u. a. den Stoffwechsel beeinflussen, wobei sie auch die Genexpression des Wirts regulieren und somit einen direkten Einfluss auf das zelluläre Interaktom haben (Tully, Hill et al. 2014) (Mukhtar, Carvunis et al. 2011) (Sarris, Duxbury et al. 2015) (Kazan and Lyons 2014) (Attaran, Major et al. 2014). Dies verdeutlicht ein Wechselspiel zwischen Wirt und Pathogen in der pflanzlichen Immunabwehr, was ein komplexes Zusammenspiel einzelner pflanzlicher Komponenten und Signalwege erfordert. Unsere Arbeitsgruppe konnte als erste Gruppe ein semiquantitatives Modell zur Wirkung der CKe in der Immunabwehr aufzeigen (Naseem, Philippi et al. 2012) (Naseem, Kunz et al. 2013). Eine konzeptionelle Darstellung des gegenseitigen Wechselspiels liefert das ZigZag-Modell, welches grundlegend die pflanzlichen Abwehrmechanismen (PTI und ETI) beschreibt, allerdings keine vollständigen Auskünfte über die dahinterstehenden komplexen Interaktionen und beteiligten Signalwege liefert (Cui, Tsuda et al. 2015) (Dong, Jiang et al. 2015).

In diesem Zusammenhang wurden in dieser Arbeit drei verschiedene Infektionsszenarien hinsichtlich des Einflusses auf die Funktion und Topologie des pflanzlichen Interaktoms untersucht. Die untersuchten Datensätze beinhalten wildtypische *Pst* DC3000 (verfügt über Effektorproteine, COR-Phytotoxine und PAMPs) sowie die beiden Mutanten *Pst* COR<sup>-</sup> (verfügt über Effektorproteine und PAMPs) und *Pst* COR<sup>-</sup> *hrpS* (verfügt über PAMPs). Allgemein betrachtet haben die durchgeführten Analysen gezeigt, dass *Pst* DC3000 wichtige pflanzliche Hormonsignalwege für die Immunabwehr beeinflusst und sich eine gemeinsame Überlagerung in der Immunantwort gegenüber pathogener Virulenzfaktoren zeigt. Insbesondere zeigen Hubproteine um JAZ und CK eine gegensätzliche Regulierung, was eine Verbindung zwischen dem JA- und CK-Signalweg in der pflanzlichen Pathogenabwehr aufzeigt.

### **6.2.1 *Pst* DC3000 beeinflusst das zelluläre Wirtsznetzwerk und zeigt, wie sich die zwei Phasen der Immunantwort (PTI und ETI) in *A. thaliana* überlappen**

In einem ersten Schritt wurden die drei Datensätze mit der entwickelten Methode hinsichtlich ihres Einflusses auf das pflanzliche Interaktom untersucht. Der Transkriptomdatensatz von *A. thaliana* infiziert mit *Pst* DC3000 vs. Mock (24h) verfügt über 4551 DEGe (19,95% aller 22810 Gene des Datensatzes), wobei 1023 (22,48% der DEGe) eine Überlappung mit dem Interaktom zeigen. Das erstellte EBN basierend auf den 1023 DEGen verfügt über 2554 Knoten und 3956 Interaktionen. Die Mutante *Pst* COR<sup>-</sup> vs. Mock (24h) verfügt über 2328 DEGe (10,21% aller 22810 Gene des Datensatzes), wobei 491 (21,09% der DEGe) eine Überlappung mit dem Interaktom zeigen. Das rekonstruierte EBN besitzt 1628 Knoten und 2142 Interaktionen. Der Transkriptomdatensatz der Doppelmutante *Pst* COR<sup>-</sup> *hrpS* vs. Mock (24h) verfügt über 609 DEGe (2,67% aller 22810 Gene des Datensatzes), wobei sich eine Überlappung von 59 DEGen (9,69% der DEGe) mit dem Interaktom gezeigt hat. Das auf den 59 DEGen basierende EBN besteht aus 374 Knoten und 348 Interaktionen.

Anhand der Ergebnisse lässt sich erkennen, dass die Anzahl der regulierten DEGe und das resultierende EBN in *Pst* DC3000 gegenüber den beiden Mutanten deutlich höher ist. Deutlich wird dies vor allem in der Doppelmutante *Pst* COR<sup>-</sup> *hrpS*, bei der nach einer 24-stündigen Infektion weniger als 3% an DEGen des gesamten Datensatzes zu verzeichnen sind. Dies veranschaulicht, dass die von *Pst* DC3000 sekretierten Abwehrstoffe die Genexpression und das zelluläre Interaktom von *A. thaliana* beeinflussen, zeigt zudem aber auch die Bedeutung

der COR-Phytotoxine und Effektorproteine in der *Pst* DC3000-vermittelten Anfälligkeit von *A. thaliana*. Deutlich wird dies auch an den überlappenden DEGen zwischen *Pst* DC3000 und den beiden Mutanten. Hier zeigen sich 1285 DEGe zwischen *Pst* DC3000 und *Pst* COR<sup>-</sup> und 142 DEGe zwischen *Pst* DC3000 und *Pst* COR<sup>-</sup> *hrpS*, von denen lediglich 54 DEGe (hier nicht gezeigt) in allen drei Datensätzen vorhanden sind (152 DEGe zwischen *Pst* COR<sup>-</sup> und *Pst* COR<sup>-</sup> *hrpS*; hier nicht gezeigt). Es lässt sich zudem eine höhere Anzahl an reziprok regulierter DEGe zwischen *Pst* DC3000 und *Pst* COR<sup>-</sup> erkennen (43 zwischen *Pst* DC3000 und *Pst* COR<sup>-</sup> gegenüber 13 zwischen *Pst* DC3000 und *Pst* COR<sup>-</sup> *hrpS*). Vergleicht man die EBNe untereinander, so zeigen sich 1343 überlappende Knoten zwischen *Pst* DC3000 und *Pst* COR<sup>-</sup>, 266 Knoten zwischen *Pst* DC3000 und *Pst* COR<sup>-</sup> *hrpS* und 284 Knoten zwischen *Pst* COR<sup>-</sup> und *Pst* COR<sup>-</sup> *hrpS*, wobei 223 Knoten gemeinsam in allen drei EBNe vorhanden sind. Die höhere Anzahl an DEGen und resultierender überlappende DEGe und Knoten aus dem Interaktom für *Pst* DC3000 und *Pst* COR<sup>-</sup> verglichen mit *Pst* COR<sup>-</sup> *hrpS* legt außerdem die Vermutung nahe, dass der durch Effektoren ausgelöste Einfluss auf das pflanzliche Immunsystem stärker als der von PAMPs und COR-Phytotoxinen ist. Deutlich wird dies vor allem auch an der Überlappung der EBNe mit IPn bekannter Effektorproteinen. Mukhtar et al. (2011) haben 137 Proteine im Interaktom von *A. thaliana* beschrieben, welche eine Interaktion mit Effektorproteinen u. a. von *P. syringae* besitzen (Mukhtar, Carvunis et al. 2011). Vergleicht man die 137 bekannten Proteine mit den Knoten der erstellten EBNe der vorliegenden Arbeit, so finden sich im EBN von *Pst* DC3000 insgesamt 115 und 99 für *Pst* COR<sup>-</sup> wieder, allerdings sind in *Pst* COR<sup>-</sup> *hrpS* mit 21 Knoten deutlich weniger vorhanden. Zwischen allen drei EBNe konnten jedoch 17 Effektorproteine aufgezeigt werden. Da sich einige Effektorproteine auch in dem EBN der Doppelmutante ohne funktionierendes Effektorsystem wiederfinden, verdeutlicht dies zudem eine gegenseitige Überlappung der PTI und ETI und eine Regulierung gemeinsamer Signalwege in *A. thaliana*. Dies veranschaulicht quantitativ die grundlegende Darstellung des ZigZag-Modells im Hinblick auf ein gegenseitiges Wechselspiel zwischen Wirt und Pathogen für die pflanzliche Immunabwehr. Es zeigt in diesem Zusammenhang aber auch, dass je besser ein Pathogen mit spezifischen Virulenzfaktoren ausgestattet ist, desto höher ist der Grad der Modulation der zellulären Signalwege der infizierten Wirtspflanze.

Die generierten EBNe wurden zusätzlich hinsichtlich ihres netzwerkbiologischen Einflusses und ihrer Funktion analysiert. Untersuchungen in Bezug auf die netzwerktopologischen Eigenschaften eignen sich besonders gut, um Einblicke in die hierarchische Organisation



zellulärer Netzwerke zu erhalten (Tully, Hill et al. 2014) (Mukhtar, Carvunis et al. 2011). Die Ergebnisse lassen umfassend erkennen, dass die von *Pst* DC3000 sekretierten Abwehrstoffe wichtige pflanzliche Hormonsignalwege für die Immunabwehr in *A. thaliana* beeinflussen. Die Analysen heben zudem hervor, dass sich der Einfluss auf das Netzwerkverhalten der Effektorproteine und COR-Phytotoxine von dem der PAMPs unterscheidet, sich jedoch auch eine gemeinsame Überlappung in der Immunantwort finden lässt. Die komplexe Immunantwort auf eine Infektion spiegelt sich zudem in der höheren Anzahl an funktionellen Modulen und Hubproteinen in *Pst* DC3000 gegenüber der Doppelmutante *Pst* COR<sup>-</sup> *hrpS* wider, welche lediglich über PAMPs verfügt. Die Ergebnisse lassen ebenfalls erkennen, dass Pathogene die Robustheit der Wirtspflanze durch Beeinflussung des dezentralen Netzwerkverhaltens regulieren, um besser das zelluläre System und die Anfälligkeit des Wirts für eine Infektion zu manipulieren.

Es hat sich hierbei gezeigt, dass das von *Pst* DC3000 regulierte EBN mit einer durchschnittlichen Anzahl von 3,013 Interaktionsnachbarn den höchsten Wert aufweist, gefolgt von *Pst* COR<sup>-</sup> (2,566) und *Pst* COR<sup>-</sup> *hrpS* (1,818). Da dieser Parameter Auskunft über die durchschnittliche Vernetzung der Knoten in einem Netzwerk gibt (Assenov, Ramirez et al. 2008) (Dong and Horvath 2007), lässt sich hieraus schlussfolgern, dass das von *Pst* DC3000 regulierte Netzwerk im Vergleich zu den beiden Mutanten besser vernetzt ist. Ein weiterer wichtiger Parameter ist die Netzwerkzentralität. Zentrale Netzwerke sind anfälliger gegenüber Infektionen, wohingegen dezentrale Netzwerke ein robustes Verhalten aufweisen (Karl and Dandekar 2015). Es zeigt sich, dass das von *Pst* DC3000 regulierte EBN (Wert von 0,051) gegenüber dem EBN von *Pst* COR<sup>-</sup> (Wert von 0,110) eher dezentral ist, wohingegen eine Infektion mit *Pst* COR<sup>-</sup> *hrpS* in einem stark zentralen Netzwerk (Wert von 0,483) resultiert. Des Weiteren wurden die Netzwerke hinsichtlich ihrer Netzwerkheterogenität untersucht, was einen Überblick über die Tendenz des Netzwerks, Hubproteine zu enthalten, gibt (Dong and Horvath 2007). Biologische Signalnetzwerke sind meist heterogen organisiert, wobei die meisten Knoten über wenige IP verfügen und nur einzelne Hubproteine im Netzwerk zu finden sind (Dong and Horvath 2007). Die durchgeführte Analyse lässt erkennen, dass die von *Pst* DC3000 (Wert von 1,991) und *Pst* COR<sup>-</sup> (Wert von 2,383) regulierten EBNe im Vergleich zu einer Infektion mit *Pst* COR<sup>-</sup> *hrpS* (Wert von 5,160) eher homogen sind und über eine gute Vernetzung mit nur wenigen Knoten mit nur einzelnen IPn verfügen.

Die rekonstruierten EBNe wurden darüber hinaus hinsichtlich funktioneller Module (Cluster) untersucht. Funktionelle Module sind stark vernetzte Subgraphen in einem Netzwerk mit einem großen Einfluss auf das gesamte Netzwerkverhalten (Bader and Hogue 2003) (Saito, Smoot et al. 2012). Die Analyse hat für das eher dezentrale EBN von *Pst* DC3000 mit 27 Subgraphen gegenüber den von *Pst* COR<sup>-</sup> (zwölf) und *Pst* COR<sup>-</sup> *hrpS* (zwei) regulierten EBNe mehr funktionelle Cluster ergeben. Hierbei zeigt insbesondere das immunrelevante Cluster um JAZ (Sheard, Tan et al. 2010) die stärkste Vernetzung im *Pst* DC3000-regulierten EBN. Interessanterweise findet sich ein Cluster um JAZ auch im EBN von *Pst* COR<sup>-</sup> wieder, nicht aber in dem von *Pst* COR<sup>-</sup> *hrpS* regulierten EBN. Dies hebt die Rolle des JA-Signalwegs in der pflanzlichen Immunabwehr gegenüber *Pst* DC3000 hervor (Pieterse, Van der Does et al. 2012) (Naseem, Kunz et al. 2013) (Robert-Seilaniantz, Grant et al. 2011).

### **6.2.2 *Pst* DC3000 reguliert wichtige Hubproteine für die Immunabwehr in *A. thaliana***

Anhand der entwickelten Methode wurden funktionelle Hubproteine und neue Modulatoren einer CK-vermittelten Immunabwehr identifiziert. Hierzu wurden die DEGe aus dem Transkriptom-Interaktom-Mapping der drei Datensätze auf ihre biologischen Funktionen (GO) untersucht und die darauf basierenden Immunnetzwerke anschließend auf ihre Topologie analysiert.

Die 1023 überlappenden DEGe für *Pst* DC3000 vs. Mock zeigen eine Beteiligung an 402 biologischen Prozessen. Hiervon zeigen 27,36% (110 der Prozesse) eine immunrelevante Funktion mit insgesamt 310 (30,30% der DEGe) korrespondierenden Genen. Das daraus erstellte Immunnetzwerk verfügt über 1216 Knoten und 1596 Interaktionen, wobei die Knoten mit durchschnittlich 2,530 Nachbarn verknüpft sind. Die mit dem Transkriptom überlappenden 491 DEGe für *Pst* COR<sup>-</sup> vs. Mock resultieren in 207 biologischen Prozessen, wovon 39,61% (82 der Prozesse) mit insgesamt 148 (30,14% der DEGe) Genen eine immunrelevante Funktion aufzeigen. Das resultierende Immunnetzwerk besteht aus 925 Knoten und 1022 Interaktionen mit einer durchschnittlichen Anzahl von 2,151 Interaktionsnachbarn. Für *Pst* COR<sup>-</sup> *hrpS* vs. Mock zeigen die 59 DEGe eine Beteiligung an 95 biologischen Prozessen, wobei 12,63% (12 der Prozesse) mit neun (15,25% der DEGe) Genen eine immunrelevante Funktion zeigen. Das erstellte Immunnetzwerk verfügt über 30 Knoten und 24 Interaktionen mit durchschnittlich 1,6 Nachbarn.

Ähnlich wie die Ergebnisse der vorherigen Analysen zu den EBNeN (Punkt 6.2.1) lässt sich auch hier erkennen, dass pathogene Abwehrstoffe einen Einfluss auf das zelluläre Interaktom der Wirtspflanze haben und zahlreiche biologische immunrelevante Funktionen beeinflussen. Es wird wiederum deutlich, dass das von *Pst* DC3000 regulierte Netzwerk gegenüber den beiden Mutanten einen stärkeren Vernetzungsgrad aufweist (durchschnittliche Interaktionsnachbarn *Pst* DC3000: 2,530; *Pst* COR<sup>-</sup>: 2,151; *Pst* COR<sup>-</sup> *hrpS*: 1,6) und einen stärker dezentralen Netzwerkcharakter besitzt (*Pst* DC3000: 0,108; *Pst* COR<sup>-</sup>: 0,194; *Pst* COR<sup>-</sup> *hrpS*: 0,200). Das von *Pst* DC3000 regulierte Immunnetzwerk (Wert von 2,229) zeigt zudem ein eher homogenes Netzwerkverhalten verglichen mit *Pst* COR<sup>-</sup> (Wert von 3,424). Interessanterweise zeigt das Immunnetzwerk von *Pst* COR<sup>-</sup> *hrpS* (Wert von 0,816) einen stark homogenen Charakter verglichen mit den beiden anderen Immunnetzwerken, was allerdings mit der geringen Anzahl an Knoten für die Rekonstruktion des Netzwerkes zu erklären ist.

Unter Verwendung einer hohen Vernetzung wurden in den Immunnetzwerken funktionelle Hubproteine identifiziert. Hubproteine sind stark vernetzte Knoten in einem Netzwerk, die verschiedene Prozesse und Signalwege verbinden und eine Schlüsselrolle in der zellulären Netzwerkfunktion spielen (Naseem, Kunz et al. 2013). Die Analyse hat gezeigt, dass sich in dem *Pst* DC3000-regulierten Netzwerk eine höhere Anzahl an Hubproteinen mit einer breiteren Funktion in der Immunabwehr finden lässt. Als wichtige Vertreter sind AHK2, AHP3, AAR14, BRI1, MYC2, NPR1, ABI1 und JAZ1 zu nennen, die eine starke Vernetzung und immunrelevante Funktion aufweisen. Die *Arabidopsis* histidine kinase 2 (AHK2) ist Teil des Zweikomponentensystems und fungiert als CK-Rezeptor in der Immunabwehr (Argueso, Ferreira et al. 2012) (Choi, Choi et al. 2011). Hierbei lässt sich erkennen, dass AHK2 keine signifikante Genexpression in den Arraydaten zeigt, sondern ein IP der DEGe ist. Ähnliches zeigt sich für die Phosphotransferasen AHP1, AHP2 und AHP5, welche das CK-Signal auf ARR1s übertragen (Argueso, Ferreira et al. 2012). Dies verdeutlicht die Stärke und den Vorteil der entwickelten Methode, da sich vor allem auch zentrale Netzwerkspieler für die Immunabwehr finden lassen, die nicht direkt in den experimentellen Daten signifikant reguliert sind, was hierdurch weitere Einblicke in die Regulierung erlaubt. Denn bestenfalls eignen sich solche Knoten insbesondere als potentielle biotechnologische und therapeutische Targets. Weitere gefundene Hubknoten des Zweikomponentensystems sind ARR9, ARR4, ARR1, ARR6, ARR14, ARR16, ARR15, ARR3 und AHP3 (Argueso, Ferreira et al. 2012), welche allesamt nach einer Infektion mit *Pst* DC3000 herunterreguliert sind. Ein weiterer

herunterregulierter Hubknoten ist BRI1, welcher Teil des Brassinosteroid-Signalwegs ist, der eine Rolle für Wachstum und Immunabwehr spielt (Belkadir, Yang et al. 2014).

Des Weiteren zeigen die gefundenen Hubproteine eine Assoziation mit Signalwegen für SA, JA, ABA, AUX und GA. Als Beispiel seien hier NPR1 (SA-Signalweg), JAZ1, JAZ10 und MYC2 (JA-Signalweg), ABI1 und ABI2 (ABA-Signalweg), GID1B (GA-Rezeptor) und YUC8 (AUX-Biosynthese) genannt, welche allesamt, mit Ausnahme von YUC8, eine Hochregulierung nach einer *Pst* DC3000-Infektion zeigen. Ein weiterer stark hochregulierter Hubknoten ist TT4, welcher für ein Enzym (Chalkon-Synthase) des Flavonoid-Biosynthesewegs codiert, der ein negativer Regulator von AUX ist und mit der pflanzlichen Pathogenabwehr in Verbindung steht (Dao, Linthorst et al. 2011) (Brown, Rashotte et al. 2001). Die Hubproteine in dem *Pst* DC3000-Immunnetzwerk zeigen ebenfalls eine Verbindung zu zentralen Kinasen, wie etwa MAPK (z.B. MKK9), und WRKY-TFen (z.B. WRKY40) (alle hochreguliert). All diese gefundenen Hubproteine spielen eine Rolle in den Hormonsignalwegen, welche entweder direkt oder indirekt die Immunabwehr in *A. thaliana* gegenüber *Pst* DC3000 beeinflussen (Robert-Seilaniantz, Grant et al. 2011) (Naseem, Kunz et al. 2014) (Naseem, Kunz et al. 2013) (Pieterse, Van der Does et al. 2012) (Pieterse, Leon-Reyes et al. 2009) (Naseem, Philippi et al. 2012) (Kazan and Manners 2013) (Choi, Huh et al. 2010) (Lim and Lee 2015) (Yang, Yao et al. 2012) (Xu, Li et al. 2008) (Pandey, Roccaro et al. 2010) (Hentrich, Böttcher et al. 2013). Interessanterweise finden sich ebenfalls einige der gefundenen Hubproteine (GID1B, MYC2, JAZ1, JAZ10, WRKY40, ARR14, MKK9, CDPK6, YUC8, ABI1, BRI1 und NPR1; alle signifikant reguliert) für *Pst* DC3000 in dem von *Pst* COR regulierten Immunnetzwerk wieder. Dies lässt, auch wenn sich die Expressionswerte teils stark unterscheiden, ein ähnliches Muster in der Genexpression zwischen beiden Infektionen erkennen, was eine grundlegende Rolle dieser Hubproteine in der Immunabwehr und Stressantwort nach einer Infektion mit *Pst* DC3000 verdeutlicht. Des Weiteren zeigt sich in allen drei untersuchten Infektionen eine Überlappung für den Hubknoten PAP1, welcher trotz diverser Unterschiede eine ähnliche Tendenz in der Regulierung der Datensätze zeigt. PAP1 ist mit dem Anthocyan/Flavonoid-Biosyntheseweg verbunden und besitzt eine Verbindung zu den Hormonsignalwegen GA, JA und ABA, wobei eine Überexpression von PAP1 zu einer erhöhten Resistenz beitragen kann (Loreti, Povero et al. 2008) (Mitsunami, Nishihara et al. 2014) (Dao, Linthorst et al. 2011). Dies untermauert an dieser Stelle ebenfalls die gegenseitige Überlappung zwischen den Phasen der Immunantwort (PTI und ETI) in der gemeinsamen Regulierung der Pathogenabwehr in *A. thaliana*.

Zusammengefasst betrachtet trägt die durchgeführte integrierte systembiologische Analyse zu einem besseren Verständnis der Infektion mit dem Pathogen *Pst* DC3000 bei und konnte Unterschiede in den pflanzlichen zellulären Netzwerken aufzeigen sowie funktionelle Hubproteine einer CK-vermittelten Immunabwehr identifizieren. Die gefundenen Hubproteine lassen vor allem eine gegensätzliche Regulierung zwischen JAZ und dem Zweikomponentensystem von CK erkennen, was eine funktionelle Verbindung zwischen dem JA- und CK-Signalweg in der pflanzlichen Immunabwehr hervorhebt. In diesem Zusammenhang sollten sich künftige Arbeiten insbesondere den beiden gefundenen Modulatoren AHK2 und ARR14 (B-type ARRs) hinsichtlich einer generellen Funktion in der Immunabwehr (Effektor und COR) und der Verbindung zu dem gefundenen Cluster um JAZ widmen, aber auch bezüglich deren biotechnologischen Potentials. Hierbei sollte vor allem der CK-Signalübertragung und der funktionellen Rolle in der Regulierung der Immunabwehr nachgegangen werden, wozu auch die weiteren gefundenen Hubproteine für das Zweikomponentensystem mit einbezogen werden sollten. Hierzu würde es sich anbieten, Transkriptomdaten von Einzel- und Doppelmantanten zu verschiedenen Infektionszeitpunkten zu generieren, um Änderungen in der Genexpression und der Aktivität im zeitlichen Verlauf zu untersuchen, aber auch um deren spezifische Funktion in der Immunabwehr und Einfluss auf das Interaktom zu verstehen. Es wäre zudem hilfreich, entsprechende Metabolomdaten zu generieren, um den Einfluss beider Hubproteine auf den Stoffwechsel zu untersuchen. Außerdem wäre die Untersuchung mit weiteren Pathogenen vorteilhaft, um Auskunft über die generelle Funktion beider Hubproteine (und der weiteren Hubknoten) in der CK-vermittelten Immunabwehr zu erhalten.

### **6.3 Systembiologische Analysen am Beispiel eines *in silico/in vitro*-Lungentumormodells**

Lungenkrebs hat die höchste Sterblichkeitsrate aller Tumorerkrankungen aufgrund einer häufig zu späten Diagnose und auftretenden Behandlungsresistenzen (Ferlay J Available from: <http://globocan.iarc.fr>, accessed on 13/08/2015.) (Stahlhut and Slack 2015) (Herbst, Heymach et al. 2008). Neue entwickelte diagnostische und therapeutische Behandlungsansätze haben das Problem, dass vorklinische Testsysteme meist nur ungenau die tatsächliche Situation und Wirksamkeit im Patienten widerspiegeln, was eine folgende klinische Anwendung untersagt (Bhattacharjee 2012) (Kola and Landis 2004) (Morgan, Van Der Graaf et al. 2012) (Arrowsmith and Miller 2013) (Gottlich, Muller et al. 2016) (Stratmann, Fecher et al. 2014) (Arrowsmith 2011). Ein enormes Potential wird neuartigen 3D-Tumortestsystemen, etwa basierend auf einer SISmuc-Matrix, zugeschrieben, welche z.B. eine geringere Proliferationsrate gegenüber herkömmlichen 2D-Zellkultursystemen zeigen und eine Medikamententestung ähnlich den *in vivo*-Tumorbedingungen ermöglichen (Linke, Schanz et al. 2007) (Stratmann, Fecher et al. 2014) (Gottlich, Muller et al. 2016). In diesem Zusammenhang wurde in dieser Arbeit systematisch eine HSP90-Behandlung in einem kombinierten *in silico/in vitro*-Lungentumormodell untersucht (siehe Punkt 6.3.1) sowie individuelle Therapievorhersagen für eine Gefitinib-Resistenz entwickelt (siehe Punkt 6.3.2).

#### **6.3.1 Systematische Analyse einer HSP90-Inhibierung**

Um eine Korrelation zwischen einem 3D-Tumortestsystem, basierend auf einer SISmuc-Matrix, und 2D-Tumortestsystem und klinischen Beobachtungen zu untersuchen, wurden drei verschiedene NSCLC-Zelllinien (HCC827, A549 und H441) experimentell mit dem HSP90-Inhibitor 17AAG behandelt und anschließend systembiologisch untersucht. Das Hitze-Schock-Protein 90 (englisch: Heat Shock Protein 90, HSP90) ist ein stressinduzierter molekularer Chaperon, der die Faltung und Stabilisierung von Signalproteinen, welche mit wichtigen Tumorsignalwegen assoziiert sind, unterstützt (Socinski, Goldman et al. 2013) (Azoitei, Hoffmann et al. 2012). HSP90 hat sich aufgrund seiner breiten Beeinflussung und Wirkung auf verschiedene Tumorsignalwege als ein interessantes therapeutisches Target vor allem für die schwierig zu behandelnden (z.B. keine Wirkung gegenüber EGFR-TKI) und mit einer schlechten Überlebensrate verbundenen KRAS-mutierten Lungentumore entwickelt (Sos,

Michel et al. 2009) (Socinski, Goldman et al. 2013) (Azoitei, Hoffmann et al. 2012). Die experimentellen Ergebnisse konnten im 2D-Modell insbesondere in der A549, aber auch HCC827-Zelllinie, einen therapeutischen Effekt nach 17AAG-Behandlung zeigen, wohingegen sich dies im 3D-Modell lediglich für HCC827 beobachten lässt (die experimentellen Arbeiten wurden von Claudia Göttlich durchgeführt, Lehrstuhl für Tissue Engineering und Regenerative Medizin (Leitung: Prof. Dr. Heike Walles)). Die experimentellen Daten für das 3D-System entsprechen weitestgehend den in der Literatur hinsichtlich einer HSP90-Inhibierung beschriebenen Beobachtungen in der Klinik (Sos, Michel et al. 2009) (Socinski, Goldman et al. 2013). Zahlreiche HSP90-Inhibitoren haben bereits eine erste vorklinische Wirksamkeit bei NSCLC gezeigt, wobei KRAS vor allem als Prädiktor einer erfolgreichen HSP90-Inhibierung berichtet wird. So zeigen etwa die HSP90-Inhibitoren Ganetespib und 17AAG experimentell *in vitro* und *in vivo* mit und ohne Mutationen in EGFR, KRAS und ALK eine therapeutische Wirkung, allerdings lässt sich dieser Effekt insbesondere bei mutierten EGFR und KRAS nicht in der Klinik beobachten (Sos, Michel et al. 2009) (Socinski, Goldman et al. 2013) (Azoitei, Hoffmann et al. 2012). Dies zeigt die Notwendigkeit nach neuen prognostischen Markern und alternativen Therapiestrategien, vor allem für KRAS-mutierte Lungentumore (Socinski, Goldman et al. 2013) (Azoitei, Hoffmann et al. 2012).

Die in dieser Arbeit durchgeführten systembiologischen Untersuchungen konnten in diesem Zusammenhang wichtige netzwerkbiologische Clusterproteine für das unterschiedliche Behandlungsansprechen zwischen dem 2D- und 3D-Tumortestsystem aufzeigen und neue prognostische Marker und therapeutische Drugtargets für die KRAS-mutierte Zelllinien A549 und H441 finden, was im Folgenden diskutiert wird.

### **6.3.1.1 P53 und STAT zeigen eine starke Vernetzung und haben einen Einfluss auf das unterschiedliche Behandlungsansprechen im 2D- und 3D-System**

Anhand der durchgeführten Analyse zeigen sich Unterschiede im Verhalten zwischen dem 2D- und 3D-System in 18 verschiedenen DRPEn. Diese besitzen eine signifikante reziproke Regulierung bzw. eine signifikante Regulierung in mindestens einem der beiden Systeme. Die anschließenden Analysen beruhen wiederum auf der in dieser Arbeit entwickelten Methode (Naseem, Kunz et al. 2014) (Kunz, Dandekar et al. 2016). Das auf den 18 DRPEn basierende 2D-3D-Zellkulturnetzwerk verfügt über 679 Knoten und 909 Interaktionen. Das Mapping auf

das Interaktom ist eine effiziente Methode, um einen Überblick über den gesamten Interaktionskontext zu erhalten und potentielle genregulatorische Effekte in den deregulierten Proteinen zu detektieren. Das rekonstruierte Netzwerk wurde anschließend auf funktionelle Module untersucht. Es wurden zwei Cluster um p53 (acht Proteine) und die STAT-Familie (drei Proteine) gefunden, wobei sich sieben DRPe aus den Arraydaten wiederfinden. Hierbei lässt sich erkennen, dass die funktionellen Cluster eine starke Vernetzung besitzen und p53 und STAT3 IP von HSP90 sind, zudem wird aber auch eine Verbindung zu FYN, ein DRP aus dem Array, ersichtlich. Es lässt sich demnach vermuten, dass die identifizierten Module um p53 und die STAT-Familie einen starken Einfluss auf die gesamte Netzwerktopologie haben und die beobachteten Behandlungsunterschiede nach einer HSP90-Inhibierung zwischen den beiden Systemen erklären können. Dies wird ebenfalls an den experimentellen Arraydaten ersichtlich, die eine unterschiedliche Regulierung z.B. von der STAT-Familie und FYN in den Zelllinien A549 und HCC827, aber auch p53 in A549 und H441 zeigen.

Die STAT-Proteinfamilie wird durch extrazelluläre Bindung von Zytokinen und Wachstumsfaktoren aktiviert und reguliert wichtige zelluläre Prozesse, etwa für Wachstum, Apoptose und Überleben (Akira 1999). In diesem Zusammenhang könnte die komplexe Umgebung des 3D-Zellkultursystems nicht nur die Signalaktivierung und das Behandlungsansprechen, sondern auch eine bessere extrazelluläre Bindung und Differenzierung abbilden (Stratmann, Fecher et al. 2014). Künftige Untersuchungen sollten den unterschiedlichen Auswirkungen zwischen diesen funktionellen Modulen weiter nachgehen. Im weitesten Sinne könnten diese funktionellen Module als potentielle Indikatoren für ein adäquat funktionierendes 3D-Testsystem, z.B. für eine HSP90-Behandlung, dienen.

Es ist bekannt, dass p53 und STAT bereits vielfach diskutierte und wichtige zentrale Signalwege bei Tumorentstehung und -wachstum sind. Neu ist allerdings, dass dies in diesem Zusammenhang erstmalig mittels einer systembiologischen Analyse in einer neuen 3D-Zellkultur klar zu sehen ist. Darauf aufbauend ist es mithilfe des systembiologischen Modells möglich, individuelle Therapievorhersagen durch entsprechendes Einstellen der Signalnetzwerke zu entwickeln, was unter Berücksichtigung des jeweiligen Mutationshintergrunds der Zelllinien erfolgen kann.



### **6.3.1.2 Die dynamische Modellierung reflektiert die experimentellen Daten der HSP90-Inhibierung im 3D-System**

In einem weiteren Schritt wurden wichtige experimentelle Readout-Parameter der 17AAG-Behandlung für die drei Zelllinien des 3D-Testsystems in ein *in silico*-Tumormodell integriert (Stratmann, Fecher et al. 2014) (Gottlich, Muller et al. 2016). Unter Einbeziehung des Mutationshintergrunds konnte die 17AAG-Behandlung für alle drei Zelllinien computergestützt abgebildet werden. Hierbei lassen sich insbesondere analog den experimentellen Daten ein ähnlicher Verlauf in den Readout-Parametern Proliferation und Apoptose, aber auch Änderungen von p53, HSP60 und AMPK nach einer 17AAG-Behandlung erkennen. Interessanterweise zeigt sich für p53 in der A549-Zelllinie entsprechend der zu erwartenden Regulierung aus dem Signalnetzwerk eine reduzierte Aktivierung in der *in silico*-Simulation, wohingegen die experimentellen Daten keine signifikante Änderung gezeigt haben. Da HSP90 inhibierend auf HSP60 wirkt, welches wiederum p53 inaktiviert (Ghosh, Dohi et al. 2008), sollte eine Inhibierung von HSP90 zu einem höheren HSP60- und reduzierten p53-Level führen. In den experimentellen Daten lässt sich für die A549-Zelllinie ein erhöhtes HSP60 nach 17AAG-Behandlung erkennen, was darauf hindeutet, dass eine reduzierte Aktivierung von p53 möglicherweise nicht in den Arraydaten erfasst wurde, dem zusätzliche Westernblot-Analysen genauer nachgehen sollten.

### **6.3.1.3 Die systembiologischen Analysen zeigen eine individuelle Signatur möglicher prognostischer Marker und Drugtargets für die KRAS-mutierten Zelllinien A549 und H441**

Die experimentellen Daten zeigen Unterschiede zwischen den beiden KRAS-mutierten Zelllinien im 3D-Testsystem nach einer 17AAG-Behandlung, z.B. keine Zunahme der Apoptoserate und Aktivierung von HSP60 und p53 in den H441-Zellen, was zelllinien-spezifische Markersignaturen und Therapieansätze notwendig macht. Zudem gibt es wie bereits erwähnt unterschiedliche Berichte hinsichtlich möglicher Prädiktoren einer erfolgreichen HSP90-Behandlung. So berichten Sos et al. (2009) und Azoitei et al. (2012) in durchgeführten *in vitro*- und *in vivo*-Untersuchungen u. a. von KRAS als Prädiktor einer gelungenen HSP90-Behandlung (Sos, Michel et al. 2009) (Azoitei, Hoffmann et al. 2012), wohingegen Socinski et al. (2013) in ihrer klinischen Studie von KRAS als ungeeigneten Prädiktor einer HSP90-

Therapie berichten (Socinski, Goldman et al. 2013). Aus diesem Grund wurde im weiteren Verlauf der Arbeit systematisch nach neuen prognostischen Markern einer HSP90-Inhibierung und neuen therapeutischen Drugtargets für die KRAS-mutierten Zelllinien mithilfe der entwickelten Methode gesucht (Naseem, Kunz et al. 2014) (Kunz, Dandekar et al. 2016).

Es zeigen sich elf unterschiedliche DRPe zwischen beiden Zelllinien nach einer 17AAG-Behandlung, welche in einem KRAS-Netzwerk aus 588 Knoten und 690 Interaktionen resultieren. Zusätzlich wurde der zelllinien-spezifische Mutationshintergrund berücksichtigt. Hierzu wurden die 588 Knoten des erstellten KRAS-Netzwerks auf die 784 H441- und 492 A549-spezifischen Mutationen aus der COSMIC-Datenbank gemappt. Hierbei wurden lediglich die jeweiligen Überlappungen zwischen den einzelnen Zelllinien und dem KRAS-Netzwerk berücksichtigt, wobei sich 19 Mutationen für die H441- und neun Mutationen für die A549-Zelllinie gezeigt haben, anhand derer weitere systematische zelllinien-spezifische Untersuchungen durchgeführt wurden.

Tabelle 6.3.1.3 fasst basierend auf den Analysen ein vereinfachtes Signaturschema ausgewählter potentieller prognostischer Marker für ein Behandlungsansprechen in den KRAS-mutierten Zelllinien und alternativer therapeutischer Drugtargets zusammen. Da bei einer Lungentumorerkrankung mehrere Signalwege und Proteine dereguliert sind, eignet sich in diesem Zusammenhang eine aus mehreren Proteinen bestehende Markersignatur. Es sollten sich unbedingt weitere Untersuchungen anschließen, die das Vorhandensein der Proteine, etwa mittels Westernblot oder immunhistochemischer Antikörperfärbung, messen. So könnte die kombinierte Detektion dieser Signatur von potentieller klinischer Bedeutung nach einer HSP90-Behandlung sein und Aufschluss über den Behandlungserfolg geben sowie gegebenenfalls weitere Ko-Indikationen aufzeigen. Möglicherweise können anhand dieser Signatur die am geeignetsten Drugtargets und alternative marker-basierte personalisierte Therapieansätze abgeleitet werden, dem weiter nachgegangen werden sollte (siehe Punkt 6.3.1.4).

**Tabelle 6.3.1.3: Überblick über die möglichen zelllinien-spezifischen Marker und potentiellen therapeutischen Drugtargets.** In dieser Arbeit wurden zahlreiche zelllinien-spezifische Mutationen für die A549- und H441-Zelllinien identifiziert. Die Tabelle fasst ein vereinfachtes Signaturschema ausgewählter zelllinien-spezifischer Mutationen (in blau hinterlegt jeweils in der entsprechenden Zelllinie mutiert; in orange für beide unbekannter Mutationsstatus) zusammen, welche in funktionellen Clustern liegen und/oder mit HSP90 vernetzt sind und als mögliche prognostische Marker für eine 17AAG-Behandlung fungieren können, gleichzeitig aber auch neue therapeutische Drugtargets darstellen. So zeigt etwa eine Aktivierung von AMPK (STK11) in der A549- und Inhibierung von HIF1A in der H441-Zelllinie einen potentiellen therapeutischen Effekt (nr=nicht reguliert, wt=wildtypisch). (Details siehe Text.)

	Mutationshintergrund (neue prognostische Marker und therapeutische Targets)							
	HIF1A	IKBKE	NOS1	STK11	BRG1	CTNNA1	CSNK2A1	NR3C1
<b>A549</b>	wt	wt	wt	✓	✓	✓	✓	✓
<b>H441</b>	✓	✓	✓	wt	wt	wt	✓	✓

### Prognostische Marker in der H441-Zelllinie

Um netzwerkbiologisch eine Verbindung zu HSP90 zu schaffen, wurde ein H441-Netzwerk basierend auf den 19 zelltyp-spezifischen Mutationen und HSP90 (HSP90AA1 und HSP90AB1) erstellt. Das Netzwerk verfügt über 796 Knoten und 982 Interaktionen, wobei die funktionelle Analyse hierfür ein stark vernetztes Cluster um HIF1A (und HSP90) gefunden hat. Eine genauere Betrachtung zeigt neben einer Verbindung zu p53 zusätzlich auch elf weitere H441-spezifische Mutationen auf, die eine direkte Verbindung zu HSP90 zeigen, z.B. IKBKE (Guo, Kim et al. 2013), oder über einen weiteren IP, sogenannte Bystander, verbunden sind. Da all diese gefundenen Knoten ein gut vernetztes funktionelles Modul um HIF1A bilden, z.B. p53 und HIF1A direkt vernetzt mit HSP90, könnte dies einen großen Einfluss auf das H441-Signalnetzwerk haben und das geringe Behandlungsansprechen erklären. Sodann könnten all diese Knoten als potentielle prognostische Marker für die H441-Zelllinie fungieren. So ist die Serine-Threonin-Kinase (STK) IKBKE ein erst seit kurzem bekanntes Onkogen, welches durch STAT3 und Tabakrauch aktiviert wird, wofür es sich insbesondere als möglicher Marker für raucher-assoziierte NSCLC eignet (Guo, Kim et al. 2013). Ein weiterer interessanter Knoten ist die Stickstoffmonoxid-Synthase 1 (NOS1, englisch: Nitric Oxide Synthase 1), welche mit HSP90 vernetzt ist. Stickstoffmonoxid-Synthasen sind mit der Tumorpathogenese assoziiert und zeigen eine Deregulierung in NSCLC, wobei Untersuchungen zufolge ein erhöhtes NOS1 als günstiger prognostischer Marker fungieren kann, andererseits NOS2 KRAS-induzierte

Lungentumore verstärken kann (Puhakka, Kinnula et al. 2003) (Puhakka, Harju et al. 2006) (Okayama, Saito et al. 2013).

### **Prognostische Marker in der A549-Zelllinie**

Das erstellte Netzwerk für die A549-Zelllinie basierend auf den neun Mutationen und HSP90 (HSP90AA1 und HSP90AB1) besteht aus 302 Knoten und 341 Interaktionen. Da sich aufgrund der geringen Größe des Netzwerks kein funktionelles Cluster gezeigt hat, wurde das Netzwerk mit dem KRAS-Netzwerk (588 Knoten und 690 Interaktionen) verbunden, wobei das resultierende Netzwerk über 800 Knoten und 1013 Interaktionen verfügt. Die funktionelle Analyse konnte für das kombinierte Netzwerk zwei Cluster um ARID3A (sieben Proteine) und HSP90AA1 (sechs Proteine) finden, wobei sich eine starke Vernetzung zu p53, HSP90 und weiteren fünf interagierenden A549-spezifischen Mutationen, etwa Tight Junction Protein 1 (TJP1), zeigt. Auch hier können die gefundenen Knoten als mögliche prognostische Marker für ein Behandlungsansprechen der A549-Zelllinie dienen. Der Tumorsuppressor BRG1 (SMARCA4) fungiert als wichtiger Faktor in der Tumorentstehung und ist mit einer schlechten Überlebensprognose in NSCLC verbunden, wobei ebenfalls eine Komutation u. a. von KRAS und LKB1 (STK11) beobachtet wurde (Medina, Romero et al. 2008) (Reisman, Sciarrotta et al. 2003) (Bell, Chakraborty et al. 2015). Untersuchungen berichten zudem von SMARCA4 als einen neuartigen prognostischen Marker mit einer erhöhten Sensitivität gegenüber einer platinbasierten Chemotherapie in NSCLC (Bell, Chakraborty et al. 2015). Das Zell-Zell-Adhäsionsmolekül Alpha-Catenin (Cadherin-Associated Protein; CTNNA1) ist mit Tumordinvasion und -metastasierung verbunden, wobei Studien berichten, dass es mit einer schlechten Überlebensprognose assoziiert ist und als prognostischer Indikator zur Klassifizierung für Stadium 1 NSCLC-Patienten dienen kann (Lee, Wu et al. 2002) (Kimura, Endo et al. 2000).

Zwei weitere spannende Knoten sind CSNK2A1 und NR3C1 (in orange), welche Teil der gefundenen Cluster sind und als Bystander aus den Analysen sichtbar werden, über deren Mutationsstatus in beiden Zelllinien in der COSMIC-Datenbank nichts bekannt ist. Die Kinase CSNK2A1 (Casein Kinase 2, Alpha 1 Polypeptide) ist eine bei verschiedenen Tumoren überexprimierte Casein-Kinase, welche mit Apoptose, Überleben und dem PI3K/AKT-Signalweg assoziiert ist. Die CSNK2A1 ist ein seit kurzem bekanntes therapeutisches Tumortarget, wobei erste experimentelle Untersuchungen eines Casein-Kinase-2-Inhibitors

bereits therapeutische Erfolge in NSCLC gezeigt haben (Zhang, Long et al. 2013) (So, Rho et al. 2015). Der Glucocorticoid-Rezeptor NR3C1 (Nuclear Receptor Subfamily 3, Group C, Member 1) zeigt eine Verbindung u. a. zu p53 und HSP90. Aktuell berichtet nur eine Studie von einer gefundenen Korrelation von NR3C1 in NSCLC, sodass in der Literatur über dessen Rolle bei Lungenkrebs bislang wenig bekannt ist (Zhao, Liu et al. 2015). Aufgrund der Verbindung zu HSP90 und der jeweiligen gefundenen Cluster könnten die beiden Knoten CSNK2A1 und NR3C1 in diesem Zusammenhang unabhängig vom Mutationsstatus als mögliche neue prognostische und therapeutische Kandidaten bei KRAS-mutierten NSCLC fungieren, dem unbedingt weiter nachgegangen werden muss.

Auf die beiden weiteren gefundenen potentiellen prognostischen Marker und Drugtargets HIF1A (H441-Zelllinie) und STK11 (AMPK; A549-Zelllinie) soll im Folgenden eingegangen werden.

#### **6.3.1.4 HIF1A und AMPK als neue therapeutische Targets für die KRAS-mutierten Zelllinien A549 und H441**

HIF1A könnte als möglicher zelllinien-spezifischer prognostischer Marker für die H441-Zelllinie fungieren. Da sich zudem ein funktionelles Modul um HIF1A gezeigt hat, wurde es als potentielles Drugtarget in der H441-Zelllinie ausgewählt. In diesem Zusammenhang konnte die durchgeführte *in silico*-Simulation einer Inhibition von HIF1A einen potentiellen therapeutischen Effekt (über der Proliferation liegende Apoptose) aufzeigen.

H441-Zelllinien weisen eine LOF-Mutation im *p53*-Gen auf, wobei eine LOF in einem erhöhten HIF1A-Level in Tumorzellen resultiert (Ravi, Mookerjee et al. 2000). Zudem führt eine erhöhte HIF1A-Expression zu einem erhöhten p53-Level (Greijer and van der Wall 2004) (An, Kanekal et al. 1998), was aber aufgrund der LOF von p53 lediglich einen geringen Einfluss auf die Apoptose hat. HIF-1 spielt eine entscheidende Rolle bei der Anpassung von Tumorzellen an Hypoxie, wobei HIF1A und HIF1B einen Heterodimer bilden und die Genexpression wichtiger Gene regulieren, etwa für Angiogenese, Proliferation und Metabolismus (Ravi, Mookerjee et al. 2000) (Greijer and van der Wall 2004) (Volm and Koomagi 2000) (Giatromanolaki, Koukourakis et al. 2001) (Semenza 2003). Eine erhöhte HIF1A-Expression korreliert zudem mit einer erhöhten Mortalität und verringerten Überlebensrate in Patienten in verschiedenen Tumorerkrankungen inklusive NSCLC (Volm

and Koomagi 2000) (Giatromanolaki, Koukourakis et al. 2001) (Semenza 2003). Ebenfalls ist bekannt, dass eine gemeinsame Überexpression von HIF1A und p53 mit einer schlechten Prognose und verringerten Überlebensrate nach einer Chemotherapie in Ovarialkarzinomen assoziiert ist (Birner, Schindl et al. 2001). Eine Inhibierung von HIF1A hat erste therapeutische Erfolge in humanen Xenograft-Modellen gezeigt (Welsh, Williams et al. 2004), wobei sich bereits PX-478 2HCl als ein selektiver HIF1A-Inhibitor bei Prostatakrebs in Phase 1 der klinischen Studie befindet (Palayoor, Mitchell et al. 2008). Demzufolge könnte eine therapeutische Beeinflussung von HIF1A eine vielversprechende Behandlungsstrategie bei Lungenkrebs darstellen, dem weitere experimentelle Untersuchungen nachgehen sollten.

Für die A549-Zelllinie wurde STK11 als möglicher prognostischer Marker gefunden und eine Beeinflussung des STK11-Signalwegs für eine potentielle therapeutische Anwendung als am geeignetsten angesehen. Die Ergebnisse der durchgeführten *in silico*-Simulation einer AMPK-Aktivierung lassen einen möglichen therapeutischen Effekt in Form einer erhöhten Apoptose und reduzierten Proliferation erkennen.

AMPK ist downstream aktiviert von LKB1 (STK11) und beeinflusst Apoptose, Proliferation und Metabolismus durch eine Inhibierung von mTOR und Aktivierung von p53, wobei eine LOF-Mutation von *LKB1* zu einer Deregulierung beider Tumorsuppressoren und unkontrolliertem Tumorzellwachstum führt (Kuhajda 2008) (Luo, Zang et al. 2010) (Alexander and Walker 2011) (Kottakis and Bardeesy 2012). Eine therapeutische Aktivierung von AMPK befindet sich aktuell unter experimenteller Testung für verschiedene Tumorerkrankungen. So zeigt der AMPK-Aktivator AICAR vielversprechende Ergebnisse und konnte bereits die klinische Phase 1/2 zur Behandlung von chronisch lymphatischer Leukämie absolvieren (Van Den Neste, Cazin et al. 2013). Des Weiteren berichten zahlreiche andere Studien, dass AICAR in der Lage ist, *in vitro* und *in vivo* das Tumorwachstum zu inhibieren, auch in A549-Zellen (Rattan, Giri et al. 2005) (Fay, Steele et al. 2009) (Tang, Williams et al. 2011). In diesem Zusammenhang könnte eine gezielte Aktivierung von AMPK, etwa mit AICAR, eine aussichtsreiche Therapieoption darstellen, was dabei auch unabhängig von *LKB1* erfolgen kann (Rattan, Giri et al. 2005).

An dieser Stelle sei angemerkt, dass die Effekte auf Apoptose und Proliferation, z.B. in den H441-Zellen, für beide Zelllinien unter Umständen noch stärker sein können, was daran liegt, dass die hier durchgeführten Simulationen lediglich auf den Aktivierungen der unbehandelten Zellen aus der 17AAG-Behandlung beruhen, sich aber nach einer entsprechend erfolgreichen

Behandlung auch die Signalaktivierungen der Arraydaten, etwa EGFR, ändern sollten. Trotz der verwendeten Aktivierungen der unbehandelten Zellen zeigen die HIF1A-Inhibierung und AMPK-Aktivierung eine über der Proliferation liegende Apoptose, was in diesem Zusammenhang auf einen möglichen therapeutischen Effekt rückschließen lässt. Künftige experimentelle Untersuchungen sollten dieser Fragestellung unbedingt weiter nachgehen, um das Potential einer therapeutischen Beeinflussung von HIF1A und AMPK bei Lungenkrebs zu untersuchen.

### **6.3.2 Systematische Analysen von Resistenzmechanismen gegenüber einer Gefitinib-Behandlung**

#### **6.3.2.1 Eine kombinierte Inhibierung von PI3K und MEK als möglicher Therapieansatz**

Gefitinib ist ein EGFR-TKI, der bei Patienten mit einer Überexpression von EGFR in der Klinik erfolgreich eingesetzt wird, jedoch entwickeln sich nach einem anfänglichen Behandlungsansprechen häufig Therapieresistenzen im zeitlichen Verlauf (Yu, Arcila et al. 2013) (Shen, Zhu et al. 2014) (Gomez, Wykosky et al. 2013). Da sich gezeigt hat, dass das *in silico*-Tumormodell eine Behandlung entsprechend den experimentellen Daten nachbilden kann (siehe Punkt 6.3.1.2) und neue Therapieoptionen (siehe Punkt 6.3.1.3 und 6.3.1.4) entwickelt werden können, wurden in einem weiteren Schritt Resistenzmechanismen nach einer Gefitinib-Behandlung für die HCC827- und A549-Zelllinien untersucht. Aus den Ergebnissen der Untersuchungen wurden neue Therapieansätze abgeleitet, die von potentieller klinischer Bedeutung sein können.

#### **Gefitinib-Resistenz in der HCC827-Zelllinie**

Lungentumorpatienten mit einer Überexpression von EGFR entwickeln nach einem anfänglichen Behandlungsansprechen häufig Therapieresistenzen gegenüber einer Gefitinib-Behandlung im Zeitablauf, was an einer Komutation des c-MET-Rezeptors liegt. c-MET aktiviert KRAS und PI3K und reguliert Apoptose und Proliferation, was Tumorwachstum und Metastasierung verstärkt (Okamoto, Okamoto et al. 2010) (Shen, Zhu et al. 2014) (Garofalo, Romano et al. 2012) (Yu, Arcila et al. 2013) (Goyal, Muzumdar et al. 2013) (Bhardwaj, Cascone et al. 2013). Die durchgeführten Simulationen konnten aufzeigen, dass eine c-MET-

Koexpression zu einer geringeren Ansprechrate und Resistenz gegenüber einer Gefitinib-Behandlung in der HCC827-Zelllinie führt. Es wurde ersichtlich, dass neben c-MET die beiden Knoten PI3K und MEK eine erhöhte Aktivierung aufzeigen, sodass sich eine therapeutische Inhibierung dieser Knoten anbieten könnte. Die anschließenden Simulationen haben hierbei ergeben, dass die Inhibierung von c-MET zu einer erhöhten Apoptose führt, diese aber im zeitlichen Verlauf wiederum abnimmt und die Proliferation ansteigt. Die Simulation der Inhibierung von MEK führt hingegen zu einer reduzierten Proliferation im Zeitablauf, hat aber keinen Einfluss auf die Apoptose. Auf der anderen Seite zeigt die alleinige Inhibierung von PI3K zwar eine erhöhte Apoptose, die Proliferationsrate bleibt jedoch auf einem gleich hohen Aktivierungslevel wie bei den unbehandelten Zellen. Dies verdeutlicht, dass die Inhibierung von PI3K oder c-MET keinen starken Effekt auf die Proliferation hat und eine Inhibierung von MEK keinen Einfluss auf die Apoptose besitzt sowie eine alleinige Inhibierung nicht die beste Behandlungsstrategie darstellt. Demnach wurde eine kombinierte Inhibierung beider Signalwege untersucht. Die Simulationen konnten zeigen, dass die Inhibierung von c-MET und MEK zwar zu einer erhöhten Apoptose und reduzierten Proliferation führt, sich aber im Zeitablauf wiederum eine Abnahme der Apoptose beobachten lässt. Auf der anderen Seite führt eine Inhibierung von c-MET und PI3K zu einer gleich hohen Aktivierung der Proliferation und Apoptose im zeitlichen Verlauf. Erst eine kombinierte Inhibierung von MEK und PI3K resultiert in einer Aktivierung der Apoptose und einer verminderten Proliferation im Zeitverlauf, was eine Gefitinib-Resistenz in der HCC827-Zelllinie umkehren kann (Gottlich, Muller et al. 2016).

### **Gefitinib-Resistenz in der A549-Zelllinie**

Lungentumore zeigen neben einer KRAS-Mutation oftmals eine Deregulierung im IGF-1R, welche mit einer geringen Ansprechrate bzw. Gefitinib-Resistenz in Verbindung stehen (Ge, Chen et al. 2014) (Dong, Qu et al. 2015) (Hurbin, Wislez et al. 2011)). Die durchgeführten Simulationen konnten zeigen, dass eine Komutation von KRAS und IGF-1R und ein dadurch erhöhtes PI3K-Level zu einer geringen Wirkung der Gefitinib-Behandlung beiträgt. Als Therapievoranschlag wurde eine alleinige und kombinierte Inhibierung von IGF-1R, PI3K und MEK untersucht, die zu einer erhöhten Apoptose und reduzierten Proliferation führen sollte. Die diesbezüglich durchgeführten Simulationen lassen erkennen, dass eine alleinige Inhibierung von IGF-1R keine Änderung in der Apoptose und Proliferation gegenüber einer Gefitinib-Gabe im zeitlichen Verlauf zeigt. Eine Inhibierung von MEK resultiert in einer



reduzierten Proliferation, jedoch lässt sich kein Effekt auf die Apoptose erkennen. Eine Inhibierung von PI3K sowie eine kombinierte Inhibierung von PI3K und IGF-1R resultieren in einer erhöhten Apoptose, allerdings lässt sich keine Änderung der Proliferation im Vergleich zu den unbehandelten Zellen beobachten. Eine kombinierte Inhibierung von MEK und IGF-1R hat zwar einen Effekt auf die Proliferation, allerdings führt dies auch zu einer Abnahme der Apoptose im Zeitablauf. Eine kombinierte Inhibierung von MEK und PI3K lässt hingegen eine erhöhte Apoptose und eine reduzierte Proliferation im zeitlichen Verlauf erkennen.

Die Analysen lassen demnach für beide Zelllinien erkennen, dass eine direkte Inhibierung der an der Resistenzentwicklung beteiligten Rezeptoren (c-MET und IGF-1R) in beiden Fällen nicht die bestmögliche Therapiestrategie darstellt. In beiden Zelllinien konnte gezeigt werden, dass eine kombinierte Inhibierung von PI3K und MEK den bestmöglichen therapeutischen Effekt liefert, nicht aber eine alleinige Therapie. Ein ähnlicher Ansatz wurde von Qu, Wu et al. (2014) berichtet, wobei hier eine selektive Inhibierung von MEK und PI3K/mTOR erfolgreich das Tumorwachstum in Gefitinib-resistenten NSCLC-Xenograft-Modellen mit einer EGFR-, c-MET- und KRAS/PIK3CA-Mutation unterdrücken konnte. Demnach könnte eine kombinierte MEK- und PI3K-Inhibierung einen vielversprechenden Therapieansatz bei Gefitinib-resistenten Lungentumorpatienten darstellen, dass künftige Untersuchungen weiter überprüfen sollten.

### **6.3.2.2 Die Analysen zeigen eine Inhibierung der miRNA-21 und resultierende Aktivierung von PTEN als möglichen Therapieansatz und prognostischen Marker**

In einem weiteren Schritt wurde das therapeutische Potential der miRNA-21 in der HCC827-Zelllinie im *in silico*-Tumormodell untersucht. Die miRNA-21 ist EGFR aktiviert und beeinflusst Apoptose und Proliferation u. a. über Spry und PTEN, wobei Gefitinib-resistente Lungentumore ein hohes miRNA-21-Level infolge einer c-MET-Überexpression zeigen (Gomez, Wykosky et al. 2013) (Li, Ren et al. 2014) (Garofalo, Romano et al. 2012) (Zhang, Wang et al. 2010) (Shen, Zhu et al. 2014). Wie bereits in der Einleitung erwähnt, steht die miRNA-21 mit einer Gefitinib-Resistenz in Verbindung, deren Rolle hinsichtlich einer Gefitinib-Resistenz in NSCLC jedoch nicht vollständig verstanden ist (Zhang, Wang et al. 2010) (Shen, Zhu et al. 2014) (Gomez, Wykosky et al. 2013) (Li, Ren et al. 2014).

Die durchgeführten Simulationen lassen erkennen, dass eine Überexpression der miRNA-21 zu

einer Resistenz nach Gefitinib-Behandlung in der HCC827-Zelllinie beitragen kann, wobei sich eine entsprechende Zunahme der miRNA-21 nach Gefitinib-Gabe im zeitlichen Verlauf zeigt. Hieraus lässt sich schlussfolgern, dass eine reduzierte miRNA-21-Expression als möglicher Marker für eine erfolgreiche Gefitinib-Behandlung fungieren kann, anderenfalls weist eine Zunahme der miRNA-21 auf ein geringes Behandlungsansprechen hin. Die Ergebnisse lassen zudem erkennen, dass eine Inhibierung der miRNA-21 eine therapeutische Wirkung erzielen kann. In diesem Zusammenhang lässt sich beobachten, dass dies über eine Induktion von PTEN und eine Inhibierung von RAF-MEK-ERK erfolgt, wobei dies unabhängig vom PI3K-Level und demnach von einer PI3K-aktivierenden Komutation erfolgt. Interessanterweise lässt sich bei den unter Punkt 6.3.2.1 durchgeführten Simulationen einer MEK- und PI3K-Inhibierung als Therapieversuch bei einer Gefitinib-Resistenz in den HCC827- und A549-Zelllinien bei einer erhöhten Apoptose und reduzierten Proliferation ebenfalls eine Abnahme von MEK-ERK im zeitlichen Verlauf erkennen, allerdings keine entsprechende Zunahme von PTEN. Dies deutet darauf hin, dass die therapeutische Wirkung einer Inhibierung der miRNA-21 u. a. über eine Induktion von PTEN erfolgt und somit eine PTEN-Aktivierung als möglicher Marker für eine erfolgreiche Inhibierung der miRNA-21 fungieren könnte, beide aber auch ein hilfreicher prognostischer Marker für eine zielgerichtete EGFR-Therapie sein könnten. Diesem Effekt sollten zusätzliche Untersuchungen unbedingt weiter nachgehen.

Verschiedene Studien haben hinsichtlich einer therapeutischen Nutzung von miRNAs bereits erste Erfolge in Zellkultur und Tiermodellen gezeigt. So berichten Shen et al. (2014) und Li et al. (2014), dass die miRNA-21 eine Rolle in der Resistenzentwicklung nach Gefitinib-Behandlung spielt, wobei hier eine experimentelle Inhibierung der miRNA-21 und resultierender PTEN-Induktion einen therapeutischen Effekt aufzeigt (Shen, Zhu et al. 2014) (Li, Ren et al. 2014). In diesem Zusammenhang könnte eine gezielte Beeinflussung von miRNAs, wie etwa die Herunterregulierung der miRNA-21, einen potentiellen therapeutischen Nutzen bringen und eine Resistenz gegenüber Gefitinib in NSCLC umzukehren. Da miRNAs zudem im Patientenblut nachweisbar sind, könnten sie als vielversprechende nicht-invasive diagnostische und prognostische Marker für NSCLC dienen, worauf sich wiederum geeignete Behandlungsalgorithmen ableiten lassen. So berichten verschiedene Arbeiten, dass etwa die miRNA-21, miRNA-210 und miRNA-221/222 in NSCLC-Patienten ein verändertes Expressionslevel im Tumorgewebe und/oder Blut aufzeigen, was auf einen möglichen nicht-invasiven Blutmarker hinweist (Li, Ren et al. 2014) (Garofalo, Romano et al. 2012) (Shen, Todd et al. 2011) (Boeri, Verri et al. 2011). Dem sollten künftige Untersuchungen in jedem

Falle weiter nachgehen, um das diagnostische, prognostische und therapeutische Potential von miRNAs, aber auch deren zugrunde liegenden Mechanismen auf eine Gefitinib-Resistenz in NSCLC besser zu verstehen, was in diesem Zusammenhang von immenser klinischer Bedeutung ist.

### **6.3.3 Schlussbetrachtung zu den Analysen des Lungentumormodells**

Die durchgeführten Analysen verdeutlichen die Stärke der semiquantitativen *in silico*-Simulation, da sich basierend auf einem individuellen Signalnetzwerkmodell der mögliche Therapieausgang einer Behandlung abschätzen lässt und/oder auch die bestmögliche Therapiestrategie entwickelt werden kann. Dies kann hierbei auch ohne exakte Kenntnisse aller kinetischer Parameter erfolgen, wobei diese anhand mathematischer Differentialgleichungen, z.B. Exponentialfunktionen, berechnet werden können (Gottlich, Muller et al. 2016) (Schlatter, Philippi et al. 2012). Es ermöglicht in diesem Zusammenhang anhand der Netzwerktopologie und weiterer Bedingungen, etwa Behandlung und Mutationshintergrund, Vorhersagen über den Netzwerkverlauf, z.B. von Proliferation und Apoptose, der beteiligten Knoten. Darauf basierend können zahlreiche Simulationen durchgeführt werden und die bestmögliche Therapie für eine weitere experimentelle Untersuchung entwickelt werden. Die experimentellen Daten können wiederum verwendet werden, um das *in silico*-Tumormodell zu optimieren und hierdurch noch exaktere Vorhersagen zu erhalten (Gottlich, Muller et al. 2016). Ein weiterer Vorteil besteht darin, dass sich ein Modell bereits anhand weniger experimenteller Daten erstellen lässt, welches zudem zur Simulation anderer Szenarien verwendet werden kann, etwa Resistenzentwicklung und Wirkung von miRNAs.

Trotz alledem besitzen semiquantitative Simulationen auch einige Nachteile und Limitierungen (Schlatter, Philippi et al. 2012) (Gottlich, Muller et al. 2016). So basieren sie meist nur auf einer reduzierten Netzwerkverschaltung und verwenden nur Schlüsselproteine und -interaktionen. Demnach spiegeln sie nicht immer eine exakte Darstellung aller vorkommenden Proteine und Interaktionen einer lebenden Tumorzelle wider, sondern geben lediglich einen fokussierten Überblick. Aber genau dieser fokussierte Blick auf eine Tumorzelle ermöglicht in diesem Zusammenhang eine spezifische und individualisierte Simulation, z.B. individueller Mutationshintergrund, Resistenzentwicklung oder den bestmöglichen Effekt einer Medikamentengabe (Gottlich, Muller et al. 2016).

Zusammengefasst betrachtet veranschaulichen die durchgeführten Analysen, wie das Zusammenspiel von experimentellen und systembiologischen Untersuchungen in einem kombinierten *in silico/in vitro*-Lungentumormodell zu neuen individuellen Therapiestrategien führen kann. Da Gefitinib standardmäßig in der Klinik angewendet wird und einige HSP90-Inhibitoren sich bereits in klinischen Studien befinden, haben die gefundenen Signaturen und personalisierten Therapieansätze in diesem Zusammenhang eine enorme klinische Relevanz und können direkt in die Klinik transportiert und angewendet werden.

#### **6.4 Ein systembiologisches Analyseprotokoll zur systematischen Charakterisierung von lncRNAs**

In einem weiteren Teil der Arbeit wurde in einer Kollaboration mit dem Institut für Molekulare und Translationale Therapiestrategien (Leitung: Prof. Dr. Dr. med. Thomas Thum; Medizinischen Hochschule Hannover) begleitend zu den experimentellen Untersuchungen eine systembiologische Charakterisierung von lncRNAs durchgeführt. lncRNAs sind eine neue Klasse nicht-kodierender RNAs, die sich aufgrund ihrer Wirkmechanismen als neue diagnostische und therapeutische Werkzeuge eignen. Sie besitzen je nach Lokalisation unterschiedliche genregulatorische Funktionen, etwa Bindung an TFen und Chromatin-modifizierende Proteine oder aber mRNAs und miRNAs (Lee 2012) (Ulitsky and Bartel 2013) (Bell, Long et al. 2014) (Boon, Jaé et al. 2016) (Paralkar and Weiss 2013). Da lncRNAs ein relativ neues Forschungsfeld darstellen, existieren hierzu allerdings relativ wenige systembiologische Algorithmen und Analysemethoden, was die funktionelle Analyse neu identifizierter lncRNAs schwierig macht. Eine weitere Herausforderung ist zudem der bereits erwähnte vielfältige genregulatorische Wirkmechanismus, der eine umfassende und systematische Analyse erfordert. In diesem Zusammenhang wurde in dieser Arbeit ein neuartiges und eigenständiges systembiologisches Analyseprotokoll für eine integrierte Analyse von lncRNAs entwickelt. Es hat sich gezeigt, dass die Analysen hilfreich für die experimentelle Charakterisierung von lncRNAs sind, sodass das entwickelte Analyseprotokoll auch künftig für weitere Analysen entsprechend genutzt werden kann und somit die funktionellen Analysen an neu identifizierten Kandidaten mit potentiell klinischen Nutzen unterstützend vorantreiben kann.

In einem ersten Projekt haben die durchgeführten systembiologischen Analysen hinsichtlich des proteinkodierenden Potentials gezeigt, dass es sich bei der in Folge einer Herzhypertrophie hochregulierten lncRNA Chast um eine nicht-kodierende RNA handelt (Viereck, Kumarswamy et al. 2016). Chast besitzt eine Transkriptlänge von 923bp und liegt antisense zu den beiden proteinkodierenden Genen *Arhgap27* und *Plekhm1*. Die Analysen haben für beide antisense liegenden Gene eine potentielle mRNA-Interaktion gefunden, wobei eine Interaktion mit *Plekhm1* auch experimentell validiert werden konnte (Viereck, Kumarswamy et al. 2016). Obwohl lncRNAs meist eine geringe Konservierung aufweisen, konnten die durchgeführten Analysen eine Sequenz-Struktur-Konservierung in Säugetieren aufzeigen. Hierbei zeigt sich vor allem eine Konservierung in Mensch, Ratte und Schwein, was eine potentielle therapeutische Nutzung, aber auch experimentelle Untersuchungen in den beiden letzten Modellorganismen, erlaubt. Die Sequenzanalyse hat zudem gezeigt, dass Chast vier potentielle homologe Sequenzen im Menschen besitzt. Diesbezüglich konnte eine erste experimentelle Validierung eine signifikante Hochregulierung der humanen Sequenz chr17: 62,843,244 - 62,843,536 (Chast-Position 646 - 922; NC\_018928.2) in hypertrophen Herzgeweben von Patienten mit einer Aortenstenose und nach einem Hypertrophie-Stimulus in aus embryonalen Stammzellen generierten Herzmuskelzellen zeigen (Viereck, Kumarswamy et al. 2016). Hieran anschließende systembiologische Untersuchungen haben für dieses humane Homolog einen potentiellen Promotor und eine TSS an Position chr17: 62,841,289 - 62,841,339 mit einer hohen Sequenzähnlichkeit zum Chast-Promotor der Maus gefunden. Dies deutet darauf hin, dass ein humanes Homolog mit einem ähnlichen Promotorbereich zu Chast existiert, was das Potential von Chast für eine mögliche therapeutische Anwendung hervorhebt.

Die durchgeführten funktionellen Analysen haben zahlreiche potentielle Pfen im gesamten murinen Proteom gefunden, wobei sich insbesondere eine funktionelle Assoziation mit den Prozessen Transkription, Translation, Metabolismus, Histonmodifizierungen und Signalübertragung, aber auch mit dem kardiovaskulärem System, der Herzentwicklung und Hämatopoese zeigt. Ebenfalls wurde eine funktionelle COG/KOGs-Suche durchgeführt. Die COG/KOGs-Annotation gruppiert orthologe Gene entsprechend ihrer biologischen Funktionen (Jensen, Julien et al. 2008) (Tatusov, Fedorova et al. 2003). Die COG/KOGs-Analyse gibt einen ersten Überblick über die potentielle Funktion der lncRNA Chast, zeigt aber auch weitere funktionelle IP auf und ermöglicht so Aussagen über potentielle regulatorische Effekte. Die Analyse hat eine hohe funktionelle Ähnlichkeit von Chast mit dem KOG1778 (CREB binding protein/P300 and related TAZ Zn-finger proteins) gefunden. Vertreter für den KOG1778 in der

Maus sind hierbei die beiden Proteine Crebbp und Ep300, wobei diese beiden ebenfalls der orthologen Gruppe COG5076 (Transcription factor involved in chromatin remodeling, contains bromodomain) zugeordnet sind. CREBBP interagiert u. a. mit CREB (Kwok, Lundblad et al. 1994) (Chrivia, Kwok et al. 1993), HNF1A (Soutoglou, Papafotiou et al. 2000) und STAT1 (Zhang, Vinkemeier et al. 1996), wobei sich für diese in der Maus auch ein hohes Potential für eine PI mit Chast gezeigt hat. Für den COG5076 konnten zehn funktionelle Interaktionsgruppen gefunden werden. Diese weisen vor allem auf eine funktionelle Interaktion von Chast mit Transkription, Chromatin-/Histonmodifizierungen und Signalübertragung hin, z.B. COG5179, COG2101 und COG5333 (Transkriptionsinitiation), COG5032 (PI3K-Signalweg) und COG2940 (Histonmethyltransferasen). Von diesen lässt sich z.B. für Tbp12 (COG2101), Ezh2 (COG2940), Ccnt1 (COG5333) und Rbm43 (NOG85021) auch eine potentielle PI mit Chast aus der genomweiten Interaktionsvorhersage erkennen. Eine potentielle PI von Chast mit STAT wird zudem durch die Analyse nach regulatorischen Elementen untermauert. Regulatorische Elemente sind spezifische Regionen in der Sequenz, die an Elemente binden bzw. von Elementen gebunden werden, z.B. Proteine und TFen, und diese regulieren können bzw. von diesen reguliert werden (Solovyev, Shahmuradov et al. 2010). Für Chast lassen sich neun regulatorische Elemente finden, die potentiell z.B. von SPI.1, STAT und IRF-3 gebunden werden können. Überdies haben die durchgeführten Analysen eine Interaktion mit 546 miRNAs gezeigt, wobei 15 miRNAs eine bekannte Funktion im Herzen bzw. bei Herzhypertrophie aufweisen. Ebenfalls wurde der Promotor von Chast nach TFBS untersucht. Hierbei wurden insgesamt 1226 putative TFBS für 311 TF-Motive gefunden, wobei sich auch Bindestellen für herz- und stress-assoziierte TFen zeigen, z.B. NF-AT4, Nkx2-5, NF-kB und ATF3, was auf eine Verbindung zu stressinduzierter Herzhypertrophie hinweist. Eine experimentelle Validierung konnte hierbei eine Induktion von Chast aufgrund einer Bindung von Nfatc3 (Nfat4) im Promotor nachweisen (Viereck, Kumarswamy et al. 2016). Interessanterweise zeigen sich für einige der gefundenen TFen auch eine potentielle PI mit Chast, z.B. c-Jun, Stat1, Atf3, Rora und Foxo3.

Die durchgeführten funktionellen Analysen lassen erkennen, dass Chast Proteine und TFen binden kann, welche die Signalübertragung und Genexpression regulieren, aber auch eine Verbindung zum kardiovaskulären System besitzen. Es lässt sich demnach vermuten, dass Chast direkt und indirekt (i) Proteine binden und die Translation beeinflussen kann, aber auch eine Chromatin-modifizierende Funktion besitzt und so die Transkription z.B. für herz- und

stressassoziierte Gene reguliert, und/oder (ii) in einem negativen Feedbackloop seine eigene Transkription reguliert. Alles in allem deutet dies auf eine komplexe multifunktionale genregulatorische Funktion von Chast hin, was entsprechend der Literatur für lncRNAs bekannt ist (Lee 2012) (Ulitsky and Bartel 2013) (Bell, Long et al. 2014) (Boon, Jaé et al. 2016) (Paralkar and Weiss 2013). Ferner ist dies im Einklang mit den durchgeführten experimentellen Untersuchungen, die gezeigt haben, dass Chast sowohl im Zellkern, Zytoplasma als auch Chromatin exprimiert ist (Viereck, Kumarswamy et al. 2016). Künftige Arbeiten sollten den gefundenen überlappenden Pfen nachgehen, um den Wirkmechanismus von Chast besser zu verstehen und langfristig einen therapeutischen Effekt in Patienten mit Herzhypertrophie zu erzielen. Dies ist vor allem auch dahingehend notwendig, da bioinformatische Vorhersagealgorithmen eine natürliche Limitierung aufgrund der hohen Falsch-Positiv-Rate aufweisen, sodass die gefundenen Interaktionen in jedem Falle durch experimentelle Untersuchungen validiert werden sollten.

In einem zweiten Projekt ging es um eine erste experimentelle Identifizierung potentieller hypoxie-induzierter lncRNAs in Endothelzellen, sodass sich die systembiologischen Arbeiten vorerst lediglich auf das proteinkodierende Potential und eine Sequenz-Struktur-Konservierung konzentriert haben. Die durchgeführten Analysen haben für die infolge einer Hypoxie induzierten lncRNAs gezeigt, dass es sich bei den beiden lncRNAs LINC00323-003 und MIR503HG aufgrund ihres geringen proteinkodierenden Potentials um nicht-kodierende RNAs handelt. Die Analysen zeigen zudem eine hohe Sequenz-Struktur-Konservierung für die lncRNA MIR503HG in Säugetieren, wohingegen die LINC00323-003 eine geringe Konservierung aufweist (z.B. homologe Sequenzen in Gorilla und Schimpanse, nicht aber in Maus und Ratte). Dies untermauert die Tatsache, dass lncRNAs häufig eine geringe Konservierung aufweisen, was experimentelle Untersuchungen in Modellorganismen hinsichtlich einer therapeutischen Nutzung schwierig machen (Fiedler, Breckwoltd et al. 2015). In diesem Zusammenhang sollten sich weitere Untersuchungen dem funktionellen Effekt dieser beiden und weiterer lncRNAs widmen, um deren Funktionen im Hinblick auf Angiogenese und Vaskularisierung besser zu verstehen, was eine potentielle therapeutische Rolle, etwa bei Ischämie- und Tumorerkrankungen, spielen kann (Fiedler, Breckwoltd et al. 2015). Hierzu eignen sich insbesondere die durchgeführten systembiologischen Analysemethoden für die lncRNA Chast.

## **6.5 Die DrumPID-Datenbank als nützliche Datenbank für den täglichen wissenschaftlichen Gebrauch auf dem Gebiet der Analyse von Protein-Pharmaka-Interaktionen**

Als ein generelles Analysewerkzeug wurde zum Abschluss der Arbeit eine Datenbank entwickelt, welche systematische Untersuchungen von Drugtargets, Protein-Pharmaka-Interaktionen und Signalwegen ermöglicht. Die DrumPID-Datenbank verfügt über 1383 FDA-zugelassene Medikamente (und mehr als 5000 nicht-FDA-zugelassene Medikamente) und stellt Informationen über deren chemische Eigenschaften und Interaktionskontext bereit. DrumPID integriert verschiedene Datenbanken und stellt die Interaktion zwischen einem Medikament und seinem Target in den Fokus, bietet jedoch zusätzlich verschiedene Links zu anderen Datenbanken an. Somit wird eine Datenüberflutung vermieden, ermöglicht aber dennoch individuelle weiterführende Informationen und Analysen. Die Benutzeroberfläche beschränkt sich auf das Wesentliche und ist daher einfach und überschaubar gehalten, leicht zu bedienen und frei zugänglich. DrumPID ist demnach für den täglichen wissenschaftlichen Gebrauch auf dem Gebiet der Analyse von PIen geeignet. Dass ein hohes Interesse auf dem Forschungsgebiet der Protein-Pharmaka-Interaktionen besteht und die DrumPID-Datenbank sehr willkommen ist, lässt auch ein Blick auf die Statistik der Datenbank erkennen. So haben innerhalb von 20 Tagen nach Publikation bereits 513 Personen aus 52 Ländern, insbesondere Deutschland, USA und China, DrumPID genutzt (ermittelt in Linux).

Eine Stärke von DrumPID ist die implementierte COG/KOGs-Suche. Diese gibt einen Überblick über die potentielle Funktion und beteiligte Gene in anderen Organismen (z.B. Riboflavinkinase in *Aspergillus* und Mensch). Der Benutzer ist somit in der Lage, entsprechende orthologe Gene in anderen Organismen zu finden und miteinander zu vergleichen, etwa auf Sequenzunterschiede. Als weiteren Vorteil ermöglicht DrumPID zusätzlich den gewebespezifischen Wirkmechanismus (z.B. über PlateletWeb-Datenbank) zu untersuchen. Dies erlaubt ein Medikament zu finden, welches bestmöglich den gewebespezifischen Interaktionskontext abdeckt. Des Weiteren beinhaltet DrumPID eine SMILES-Ähnlichkeitssuche. Diese ermöglicht ähnliche Substrukturen für ein Medikament/eine Substanz zu finden und gibt Aufschluss über mögliche Targets, Nebenwirkungen und potentielle Funktionen. Anhand dieser eingebauten Funktion in der DrumPID-Datenbank ist es möglich, Auskunft über konservierte Strukturbereiche und potentielle IP zu erhalten, was hierbei vor allem für neu entwickelte Substanzen mit noch



unbekannter Wirkmechanismen hilfreich ist. Da Substrukturen zusätzlich andere Targets beeinflussen können, erlaubt dies hierdurch auch Informationen über mögliche Nebeneffekte für ein Medikament bzw. eine Substanz zu finden, was für weitere Substanzmodifizierungen, etwa für Off-Label-Einsatz, von Bedeutung sein kann (Kunz, Liang et al. 2016).

Alternative Datenbanken konzentrieren sich auf verschiedene anders gelagerte Schwerpunkte, z.B. ChEMBL (Willighagen, Waagmeester et al. 2013) zu Informationen über Medikamente, ChemSpider (Williams, Tkachenko et al. 2010) zur chemischen Struktur und KEGG und DrugBank (Kanehisa, Sato et al. 2016) (Knox, Law et al. 2011) zu einem Medikament und Target, wohingegen DrumPID den Schwerpunkt auf Protein-Pharmaka-Interaktionen inklusive eigener Berechnungen legt, allerdings diese (z.B. SMILES-String und PDB-Struktur) und weitere Informationen über Links für individuelle Analysen bereitstellt. DrumPID schließt hierbei eine Nische zwischen aktuellen Datenbanken und ist vor allem nützlich, ein Medikament, dessen Zielprotein und Interaktionskontext im Detail zu untersuchen (Kunz, Liang et al. 2016). Mit DrumPID wird der Nutzer in die Lage versetzt, u. a. einzelne Interaktionen oder auch ganze Signalnetzwerke zu analysieren, hat zudem aber auch die Möglichkeit, potentielle Targets in anderen Organismen zu untersuchen sowie Funktions- und Strukturvorhersagen durchzuführen. Hierfür werden Daten aus jeweils mindestens zwei verschiedenen Datenbanken integriert bzw. bereitgestellt. So werden Interaktions- und Signalnetzwerkdaten (inklusive Herkunftsschema) aus DrugBank (Knox, Law et al. 2011) und KEGG (Kanehisa, Sato et al. 2016) und Informationen zu Medikamenten aus DrugBank (Knox, Law et al. 2011), KEGG (Kanehisa, Sato et al. 2016) und Drugs.com (<http://www.drugs.com/>) integriert. Entsprechende Struktur- und Funktionsvorhersagen werden mit AnDom (Schmidt, Bork et al. 2002) und GoSynthetic (Liang, Kruger et al. 2013) sowie eigene Berechnungen mit dem COGmaster (Liang, Schmid et al. 2009) bereitgestellt. Darüber hinaus liefert DrumPID zusätzliche Informationen zu Krankheiten und gewebespezifische, experimentelle und vorhergesagte PIen durch Links z.B. zu den Datenbanken DrugBank, PlateletWeb, IMEx und STRING (Knox, Law et al. 2011) (Szkarczyk, Franceschini et al. 2015) (Orchard, Kerrien et al. 2012). Anhand der integrierten und bereitgestellten Informationen (z.B. PDB-Struktur als Download von DrumPID; Homologiemodelle über Link zu AnDom) ist es mit der DrumPID-Datenbank möglich, ein geeignetes Medikament für ein Zielprotein zu finden und dessen Wirkmechanismus und Effekt zu untersuchen, was zu einem besseren experimentellen Verständnis beitragen kann. Der Anwender ist zudem in der Lage, eine potentielle chemische

Leitstruktur für ein Zielprotein zu entwickeln, etwa durch Einfügen oder Entfernen von Seitenketten, was z.B. spezifisch ein parasitisches Protein inhibiert, ohne einen toxischen Effekt im Menschen zu haben (Kunz, Liang et al. 2016).

Eine generelle Limitierung von DrumPID liegt auf dem Fokus der Datenbank und der daraus resultierenden individuellen Datenmenge sowie der Gestaltung der möglichen Suchkategorien (Kunz, Liang et al. 2016). Eine Suche funktioniert meist über Schlüsselwörter und Data-Mining, etwa in der Kategorie „Indication/associated pathogen“, sodass entsprechende Ergebnisse lediglich anhand der Suchwörter in der Datenbank ausgegeben werden (Kunz, Liang et al. 2016). Dies liefert leider in seltenen Fällen nicht das bestmögliche Ergebnis und oftmals auch zu viele Treffer. Als Beispiel sei an dieser Stelle Methotrexat genannt, ein Medikament zur Behandlung u. a. von Chorionadenomen/-karzinomen (Plazentatumoren) und rheumatoider Arthritis (‘Methotrexate’ in Kategorie „Generic Drug Name“). Hierfür würde eine Suche nach ‘cancer’ in der Kategorie „Indication/associated pathogen“ leider nicht das gewünschte Medikament anzeigen, wohingegen eine Suche mit ‘chorioadenoma’ als Ergebnis Methotrexat liefert (alternativ findet z.B. ‘carcinoma’ 37 Ergebnisse inklusive Methotrexat). Bei einer Suche sollte der Anwender deshalb immer zuerst prüfen, ob die Autovervollständigung für das gewünschte Suchwort einen Treffer anzeigt bzw. welche Schlüsselwörter hierzu vorgeschlagen werden, denn so ist sichergestellt, den korrekten Eintrag zu erhalten. Zudem ist es hilfreich, wenn mehrere Suchwörter und Kategorien miteinander kombiniert werden. Dies spezifiziert die Suche und reduziert die Anzahl der Treffer. Werden diese Anwenderhinweise beachtet, liefert DrumPID zu einer Suchanfrage auch immer einen bestmöglichen Überblick über Medikamente, Pfen und Signalnetzwerke (Kunz, Liang et al. 2016).

Künftige Arbeiten an der Datenbank sollten sich insbesondere auf die Optimierung der Schlüsselwörter und Data-Mining der Suchkategorien konzentrieren, z.B. eine Synonym-Tabelle, um noch exaktere Suchergebnisse und das bestmögliche Ergebnis für den Nutzer zu ermöglichen. Auch sollten regelmäßige Updates (z.B. quartalsweise) der Datenbank durchgeführt werden, um DrumPID auf dem aktuellsten Stand zu halten. Eine Verknüpfung der Datenbank mit dem lncRNA-Analyseprotokoll wäre ebenfalls vorteilhaft, denn so kann direkt für eine gefundene lncRNA-PI (Drugtarget) ein potentielles Medikament und der Interaktionskontext, insbesondere gewebespezifisch, untersucht werden.

## 7 Schlussfolgerung und Ausblick

Der Einsatz von computergestützten Analysen ist ein fester Bestandteil in der heutigen biowissenschaftlichen Forschung. In dieser Arbeit wurden in einem systematischen Ansatz verschiedene biologische Themengebiete untersucht. Hierfür wurden neuartige und eigenständige systembiologische Analysemethoden und -protokolle entwickelt, welche integrierte netzwerkbiologische und funktionelle Analysen sowie die Identifizierung funktioneller Hubproteine erlauben. Verschiedene Arbeitsgruppen beschäftigen sich mit Interaktomanalysen, z.B. Tully, Hill et al. (2014), Dong, Jiang et al. (2015) und Mukhtar, Carvunis et al. (2011). Die Stärke des eigenen Ansatzes liegt in der schnellen und effizienten Analyse-Pipeline, aber auch der Verknüpfung mit semiquantitativen Analysen (SQUAD) und einer Reihe von Sequenz-Struktur-Analysen und Funktionsanalysen mit Netzwerkanalysen (z.B. ClueGO). Die entwickelten und gezeigten Analysemethoden können zudem auf unterschiedliche biologische Fragestellungen und Organismen übertragen werden, was einen innovativen und interdisziplinären methodischen Ansatz darstellt. So wurden systematisch Analysen beginnend mit (i) pflanzlichen Wirts-Pathogen-Netzwerken über (ii) Netzwerkanalysen und dynamischen Modellierungen in einem Lungentumormodell bis hin zu (iii) Analysen von RNA- und Protein-Interaktionen durchgeführt. Als Ergebnis wurden (i) u. a. eine Überlagerung der Pathogenabwehr (PTI und ETI) und neue Modulatoren einer CK-vermittelten Immunabwehr (z.B. AHK2 und ARR14) gefunden. Zudem wurden (ii) *in silico* neue prognostische Marker und Drugtargets (z.B. AMPK und HIF1A) gefunden und individuelle Therapieansätze bei Lungenkrebs entwickelt sowie (iii) eine funktionelle Charakterisierung neuer hypertrophie- und hypoxie-assoziiierter lncRNAs durchgeführt und die DrumPID-Datenbank entwickelt. Die durchgeführten Analysen tragen zu einem besseren Verständnis der experimentellen Daten bei und erlauben ein systematisches Verständnis der zugrunde liegenden Signalwege und Netzwerkeffekte in verschiedenen Organismen, gleichwohl aber auch die Entwicklung neuer therapeutischer Strategien für weitere experimentelle Untersuchungen. Die Analysen veranschaulichen exemplarisch verschiedene systembiologische Analyseansätze und deren Möglichkeit und hohes Potential, auf beliebige Aspekte biologischer Fragestellungen angewendet zu werden. So wurden die Untersuchungen in dieser Arbeit zwar für Pflanzenpathogene, Herz- und Tumorerkrankungen durchgeführt, sie sind aber jederzeit auch auf andere Themengebiete mit hohem Forschungsbedarf übertragbar, etwa Schlaganfall-, Alzheimer- und Infektionsforschung. In diesem Zusammenhang ist es

möglich, verschiedene biologische Datensätze und Datenebenen miteinander zu kombinieren und beginnend vom Genom und Interaktionskontext über semiquantitative Simulationen hin zu neuen Interventionen und Experimenten zu gelangen.

An diese Arbeit anschließende Untersuchungen sollten sich unbedingt genauer den gefundenen Hubproteinen und Therapieansätzen widmen, um langfristig gesehen zu neuen therapeutischen und biotechnologischen Anwendungen zu gelangen. So besitzen die gefundene Markersignatur und die vorhergesagten Therapiestrategien einen direkten klinischen Bezug und können zu neuen vielversprechenden therapeutischen Interventionen bei Lungentumorpatienten führen. Andererseits können die Ergebnisse auch ein wichtiger Schritt zu einem besseren Verständnis der Pathogenabwehr sein, woraus sich neue biotechnologische Ansätze ableiten lassen, die den Pflanzenschutz und die Nahrungssicherung nachhaltig voranbringen. Weitere Bestrebungen sollten auch eine detailliertere Untersuchung der gefundenen überlappenden und flexiblen Regionen in der Riboflavinkinase im Menschen und *A. fumigatus* umfassen, um hieraus bestenfalls neue geeignete antimykotische Substanzen ableiten zu können. Sogleich kann auch das entwickelte lncRNA-Analyseprotokoll ein nützliches Werkzeug für künftige wissenschaftliche Arbeiten sein und die Analysen für neu detektierte lncRNAs mit potentiell klinischen Nutzen vorantreiben.

Systembiologische Analysen haben aber auch natürliche Probleme und Limitierungen, etwa eine hohe Falsch-Positiv-Rate der vorhergesagten IP und TFBS oder verfügen über keine exakten kinetischen Parameter. Hierbei ist vor allem das Wissen über mögliche Fehlerquellen derartiger Analysen hilfreich und insbesondere die richtige Anwendung von Programmen wichtig. Es sollten in diesem Zusammenhang immer verschiedene Analyseschritte und Programme miteinander kombiniert werden, um eine integrierte Blickweise auf die erhaltenen Ergebnisse zu bekommen. Dies ermöglicht fundierte Aussagen über die potentielle Funktion und IP zu bekommen und die bestmöglichen experimentellen Kandidaten zu finden, was die umfangreichen und zeitaufwendigen experimentellen Untersuchungen unterstützen kann. Eine experimentelle Untersuchung ist aber im Umkehrschluss unbedingt erforderlich, um die vorhergesagten Treffer zu validieren. So kann der kombinierte Einsatz systembiologischer Analysen und experimenteller Arbeiten unnötige Experimente und Kosten vermeiden und zu neuen biotechnologischen und therapeutischen Ansätzen führen. In diesem Zusammenhang bleibt es für die Zukunft wünschenswert, dass eine intensive Vernetzung von experimentellen und systembiologischen Arbeiten zu immer besser werdenden systembiologischen Modellen

führt, die zu einem umfassenden Verständnis grundlegender biologischer Fragestellungen beitragen und die biowissenschaftliche Forschung nachhaltig voranbringen.

## 8 Literaturverzeichnis

- Adjei, A. A. (2001). "Blocking Oncogenic Ras Signaling for Cancer Therapy." Journal of the National Cancer Institute **93**(14): 1062-1074.
- Agostini, F., A. Zanzoni, P. Klus, D. Marchese, D. Cirillo and G. G. Tartaglia (2013). "catRAPID omics: a web server for large-scale prediction of protein-RNA interactions." Bioinformatics **29**(22): 2928-2930.
- Akira, S. (1999). "Functional roles of STAT family proteins: lessons from knockout mice." Stem Cells **17**(3): 138-146.
- Alexander, A. and C. L. Walker (2011). "The role of LKB1 and AMPK in cellular responses to stress and damage." FEBS Lett **585**(7): 952-957.
- An, W. G., M. Kanekal, M. C. Simon, E. Maltepe, M. V. Blagosklonny and L. M. Neckers (1998). "Stabilization of wild-type p53 by hypoxia-inducible factor 1alpha." Nature **392**(6674): 405-408.
- Anderson, J. P., E. Badruzsaufari, P. M. Schenk, J. M. Manners, O. J. Desmond, C. Ehlert, D. J. Maclean, P. R. Ebert and K. Kazan (2004). "Antagonistic interaction between abscisic acid and jasmonate-ethylene signaling pathways modulates defense gene expression and disease resistance in Arabidopsis." Plant Cell **16**(12): 3460-3479.
- Argueso, C. T., F. J. Ferreira, P. Epple, J. P. To, C. E. Hutchison, G. E. Schaller, J. L. Dangel and J. J. Kieber (2012). "Two-component elements mediate interactions between cytokinin and salicylic acid in plant immunity." PLoS Genet **8**(1): e1002448.
- Arrowsmith, J. (2011). "Trial watch: Phase II failures: 2008–2010." Nat Rev Drug Discov **10**(5): 328-329.
- Arrowsmith, J. and P. Miller (2013). "Trial Watch: Phase II and Phase III attrition rates 2011-2012." Nat Rev Drug Discov **12**(8): 569-569.
- Asai, T., G. Tena, J. Plotnikova, M. R. Willmann, W. L. Chiu, L. Gomez-Gomez, T. Boller, F. M. Ausubel and J. Sheen (2002). "MAP kinase signalling cascade in Arabidopsis innate immunity." Nature **415**(6875): 977-983.
- Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin and G. Sherlock (2000). "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium." Nat Genet **25**(1): 25-29.
- Assenov, Y., F. Ramirez, S. E. Schelhorn, T. Lengauer and M. Albrecht (2008). "Computing topological parameters of biological networks." Bioinformatics **24**(2): 282-284.
- Attaran, E., I. T. Major, J. A. Cruz, B. A. Rosa, A. J. Koo, J. Chen, D. M. Kramer, S. Y. He and G. A. Howe (2014). "Temporal Dynamics of Growth and Photosynthesis Suppression in Response to Jasmonate Signaling." Plant Physiol **165**(3): 1302-1314.
- Azoitei, N., C. M. Hoffmann, J. M. Ellegast, C. R. Ball, K. Obermayer, U. Gößele, B. Koch, K. Faber, F. Genze, M. Schrader, H. A. Kestler, H. Döhner, G. Chiosis, H. Glimm, S. Fröhling and C. Scholl (2012). "Targeting of KRAS mutant tumors by HSP90 inhibitors involves degradation of STK33." The Journal of Experimental Medicine **209**(4): 697-711.
- Bader, G. D. and C. W. Hogue (2003). "An automated method for finding molecular complexes in large protein interaction networks." BMC Bioinformatics **4**: 2.
- Baines, A. T., D. Xu and C. J. Der (2011). "Inhibition of Ras for cancer treatment: the search continues." Future Med Chem **3**(14): 1787-1808.
- Bandi, N. and E. Vassella (2011). "miR-34a and miR-15a/16 are co-regulated in non-small cell lung cancer and control cell cycle progression in a synergistic and Rb-dependent manner." Mol Cancer **10**: 55.

- Bannister, C. A., S. E. Holden, S. Jenkins-Jones, C. L. Morgan, J. P. Halcox, G. Schernthaner, J. Mukherjee and C. J. Currie (2014). "Can people with type 2 diabetes live longer than those without? A comparison of mortality in people initiated with metformin or sulphonylurea monotherapy and matched, non-diabetic controls." *Diabetes Obes Metab* **16**(11): 1165-1173.
- Belkhadir, Y., L. Yang, J. Hetzel, J. L. Dangl and J. Chory (2014). "The growth-defense pivot: crisis management in plants mediated by LRR-RK surface receptors." *Trends Biochem Sci* **39**(10): 447-456.
- Bell, E. H., A. R. Chakraborty, X. Mo, Z. Liu, K. Shilo, S. Kirste, P. Stegmaier, M. McNulty, N. Karachaliou, R. Rosell, G. Bepler, D. P. Carbone and A. Chakravarti (2015). "SMARCA4/BRG1 Is a Novel Prognostic Biomarker Predictive of Cisplatin-Based Chemotherapy Outcomes in Resected Non-Small Cell Lung Cancer." *Clin Cancer Res*.
- Bell, R. D., X. Long, M. Lin, J. H. Bergmann, V. Nanda, S. L. Cowan, Q. Zhou, Y. Han, D. L. Spector, D. Zheng and J. M. Miano (2014). "Identification and Initial Functional Characterization of a Human Vascular Cell Enriched Long Non-Coding RNA." *Arteriosclerosis, thrombosis, and vascular biology* **34**(6): 1249-1259.
- Bernhart, S., I. Hofacker, S. Will, A. Gruber and P. Stadler (2008). "RNAalifold: improved consensus structure prediction for RNA alignments." *BMC Bioinformatics* **9**(1): 474.
- Bhardwaj, V., T. Cascone, M. A. Cortez, A. Amini, J. Evans, R. U. Komaki, J. V. Heymach and J. W. Welsh (2013). "Modulation of c-Met signaling and cellular sensitivity to radiation: potential implications for therapy." *Cancer* **119**(10): 1768-1775.
- Bhattacharjee, Y. (2012). "Biomedicine. Pharma firms push for sharing of cancer trial data." *Science* **338**(6103): 29.
- Bianchi, F., F. Nicassio, M. Marzi, E. Belloni, V. Dall'olio, L. Bernard, G. Pelosi, P. Maisonneuve, G. Veronesi and P. P. Di Fiore (2011). "A serum circulating miRNA diagnostic test to identify asymptomatic high-risk individuals with early stage lung cancer." *EMBO Mol Med* **3**(8): 495-503.
- Bindea, G., B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W. H. Fridman, F. Pages, Z. Trajanoski and J. Galon (2009). "ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks." *Bioinformatics* **25**(8): 1091-1093.
- Birner, P., M. Schindl, A. Obermair, G. Breitenecker and G. Oberhuber (2001). "Expression of hypoxia-inducible factor 1alpha in epithelial ovarian tumors: its impact on prognosis and on response to chemotherapy." *Clin Cancer Res* **7**(6): 1661-1668.
- Bjaanæs, M. M., A. R. Halvorsen, S. Solberg, L. Jørgensen, T. A. Dragani, A. Galvan, F. Colombo, M. Anderlini, U. Pastorino, E. Kure, A.-L. Børresen-Dale, O. T. Brustugun and Å. Helland (2014). "Unique microRNA-profiles in EGFR-mutated lung adenocarcinomas." *International Journal of Cancer. Journal International du Cancer* **135**(8): 1812-1821.
- Boeri, M., U. Pastorino and G. Sozzi (2012). "Role of microRNAs in lung cancer: microRNA signatures in cancer prognosis." *Cancer J* **18**(3): 268-274.
- Boeri, M., C. Verri, D. Conte, L. Roz, P. Modena, F. Facchinetti, E. Calabrò, C. M. Croce, U. Pastorino and G. Sozzi (2011). "MicroRNA signatures in tissues and plasma predict development and prognosis of computed tomography detected lung cancer." *Proceedings of the National Academy of Sciences* **108**(9): 3713-3718.
- Bommer, G. T., I. Gerin, Y. Feng, A. J. Kaczorowski, R. Kuick, R. E. Love, Y. Zhai, T. J. Giordano, Z. S. Qin, B. B. Moore, O. A. MacDougald, K. R. Cho and E. R. Fearon (2007). "p53-mediated activation of miRNA34 candidate tumor-suppressor genes." *Curr Biol* **17**(15): 1298-1307.

- Boon, R. A., N. Jaé, L. Holdt and S. Dimmeler (2016). "Long Noncoding RNAs From Clinical Genetics to Therapeutic Targets?" Journal of the American College of Cardiology **67**(10): 1214-1226.
- Boyanova, D., S. Nilla, I. Birschmann, T. Dandekar and M. Dittrich (2012). "PlateletWeb: a systems biologic analysis of signaling networks in human platelets." Blood **119**(3): e22-34.
- Brighenti, M. (2015). "MicroRNA and MET in lung cancer." Annals of Translational Medicine **3**(5): 68.
- Brown, D. E., A. M. Rashotte, A. S. Murphy, J. Normanly, B. W. Tague, W. A. Peer, L. Taiz and G. K. Muday (2001). "Flavonoids act as negative regulators of auxin transport in vivo in arabidopsis." Plant Physiol **126**(2): 524-535.
- Caarls, L., C. M. Pieterse and S. C. Van Wees (2015). "How salicylic acid takes transcriptional control over jasmonic acid signaling." Front Plant Sci **6**: 170.
- Cabili, M. N., C. Trapnell, L. Goff, M. Koziol, B. Tazon-Vega, A. Regev and J. L. Rinn (2011). "Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses." Genes Dev **25**(18): 1915-1927.
- Chang, T. C., E. A. Wentzel, O. A. Kent, K. Ramachandran, M. Mullendore, K. H. Lee, G. Feldmann, M. Yamakuchi, M. Ferlito, C. J. Lowenstein, D. E. Arking, M. A. Beer, A. Maitra and J. T. Mendell (2007). "Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis." Mol Cell **26**(5): 745-752.
- Charkiewicz, R., L. Pilz, A. Sulewska, M. Kozlowski, W. Niklinska, M. Moniuszko, J. Reszec, C. Manegold and J. Niklinski (2015). "Validation for histology-driven diagnosis in non-small cell lung cancer using hsa-miR-205 and hsa-miR-21 expression by two different normalization strategies." Int J Cancer.
- Chen, Z., H. Zeng, Y. Guo, P. Liu, H. Pan, A. Deng and J. Hu (2010). "miRNA-145 inhibits non-small cell lung cancer cell proliferation by targeting c-Myc." J Exp Clin Cancer Res **29**: 151.
- Chinchilla, D., C. Zipfel, S. Robatzek, B. Kemmerling, T. Nurnberger, J. D. Jones, G. Felix and T. Boller (2007). "A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence." Nature **448**(7152): 497-500.
- Cho, W. C., A. S. Chow and J. S. Au (2009). "Restoration of tumour suppressor hsa-miR-145 inhibits cancer cell growth in lung adenocarcinoma patients with epidermal growth factor receptor mutation." Eur J Cancer **45**(12): 2197-2206.
- Cho, W. C., A. S. Chow and J. S. Au (2011). "MiR-145 inhibits cell proliferation of human lung adenocarcinoma by targeting EGFR and NUDT1." RNA Biol **8**(1): 125-131.
- Choi, J., D. Choi, S. Lee, C. M. Ryu and I. Hwang (2011). "Cytokinins and plant immunity: old foes or new friends?" Trends Plant Sci **16**(7): 388-394.
- Choi, J., S. U. Huh, M. Kojima, H. Sakakibara, K. H. Paek and I. Hwang (2010). "The cytokinin-activated transcription factor ARR2 promotes plant immunity via TGA3/NPR1-dependent salicylic acid signaling in Arabidopsis." Dev Cell **19**(2): 284-295.
- Chrivia, J. C., R. P. Kwok, N. Lamb, M. Hagiwara, M. R. Montminy and R. H. Goodman (1993). "Phosphorylated CREB binds specifically to the nuclear protein CBP." Nature **365**(6449): 855-859.
- Coate, L. E., T. John, M. S. Tsao and F. A. Shepherd (2009). "Molecular predictive and prognostic markers in non-small-cell lung cancer." Lancet Oncol **10**(10): 1001-1010.
- Corney, D. C., A. Flesken-Nikitin, A. K. Godwin, W. Wang and A. Y. Nikitin (2007). "MicroRNA-34b and MicroRNA-34c Are Targets of p53 and Cooperate in Control of Cell Proliferation and Adhesion-Independent Growth." Cancer Research **67**(18): 8433-8438.
- Cui, H., K. Tsuda and J. E. Parker (2015). "Effector-triggered immunity: from pathogen perception to robust defense." Annu Rev Plant Biol **66**: 487-511.



- Dalli, D., A. Wilm, I. Mainz and G. Steger (2006). "STRAL: progressive alignment of non-coding RNA using base pairing probability vectors in quadratic time." *Bioinformatics* **22**(13): 1593-1599.
- Dao, T. T. H., H. J. M. Linthorst and R. Verpoorte (2011). "Chalcone synthase and its functions in plant resistance." *Phytochemistry Reviews* **10**(3): 397-412.
- deAlmeida, A. C., R. J. van Oort and X. H. Wehrens (2010). "Transverse aortic constriction in mice." *J Vis Exp*(38).
- DeRoy, S., R. Thilmony, Y. B. Kwack, K. Nomura and S. Y. He (2004). "A family of conserved bacterial effectors inhibits salicylic acid-mediated basal immunity and promotes disease necrosis in plants." *Proc Natl Acad Sci U S A* **101**(26): 9927-9932.
- Di Cara, A., A. Garg, G. De Micheli, I. Xenarios and L. Mendoza (2007). "Dynamic simulation of regulatory networks using SQUAD." *BMC Bioinformatics* **8**(1): 1-10.
- Dong, J. and S. Horvath (2007). "Understanding network concepts in modules." *BMC Systems Biology* **1**: 24-24.
- Dong, S., X. Qu, W. Li, X. Zhong, P. Li, S. Yang, X. Chen, M. Shao and L. Zhang (2015). "The long non-coding RNA, GAS5, enhances gefitinib-induced cell death in innate EGFR tyrosine kinase inhibitor-resistant lung adenocarcinoma cells with wide-type EGFR via downregulation of the IGF-1R expression." *J Hematol Oncol* **8**: 43.
- Dong, X., Z. Jiang, Y. L. Peng and Z. Zhang (2015). "Revealing shared and distinct gene network organization in Arabidopsis immune responses by integrative analysis." *Plant Physiol* **167**(3): 1186-1203.
- Dong, Z., L. Ren, L. Lin, J. Li, Y. Huang and J. Li (2015). "Effect of microRNA-21 on multidrug resistance reversal in A549/DDP human lung cancer cells." *Mol Med Rep* **11**(1): 682-690.
- Eberhard, D. A., B. E. Johnson, L. C. Amler, A. D. Goddard, S. L. Heldens, R. S. Herbst, W. L. Ince, P. A. Janne, T. Januario, D. H. Johnson, P. Klein, V. A. Miller, M. A. Ostland, D. A. Ramies, D. Sebisanoovic, J. A. Stinson, Y. R. Zhang, S. Seshagiri and K. J. Hillan (2005). "Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib." *J Clin Oncol* **23**(25): 5900-5909.
- Engelman, J. A., K. Zejnullahu, T. Mitsudomi, Y. Song, C. Hyland, J. O. Park, N. Lindeman, C. M. Gale, X. Zhao, J. Christensen, T. Kosaka, A. J. Holmes, A. M. Rogers, F. Cappuzzo, T. Mok, C. Lee, B. E. Johnson, L. C. Cantley and P. A. Janne (2007). "MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling." *Science* **316**(5827): 1039-1043.
- Enright, A. J., B. John, U. Gaul, T. Tuschl, C. Sander and D. S. Marks (2003). "MicroRNA targets in Drosophila." *Genome Biol* **5**(1): R1.
- Erson, A. E. and E. M. Petty (2009). "miRNAs and cancer: New research developments and potential clinical applications." *Cancer Biology & Therapy* **8**(24): 2317-2322.
- Esquela-Kerscher, A. and F. J. Slack (2006). "Oncomirs - microRNAs with a role in cancer." *Nat Rev Cancer* **6**(4): 259-269.
- Fang, R., T. Xiao, Z. Fang, Y. Sun, F. Li, Y. Gao, Y. Feng, L. Li, Y. Wang, X. Liu, H. Chen, X. Y. Liu and H. Ji (2012). "MicroRNA-143 (miR-143) regulates cancer glycolysis via targeting hexokinase 2 gene." *J Biol Chem* **287**(27): 23227-23235.
- Fay, J. R., V. Steele and J. A. Crowell (2009). "Energy homeostasis and cancer prevention: the AMP-activated protein kinase." *Cancer Prev Res (Phila)* **2**(4): 301-309.
- Ferlay J, S. I., Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. (Available from: <http://globocan.iarc.fr>, accessed on 13/08/2015.). "GLOBOCAN 2012 v1.0,

- Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013.
- Ferlay, J., E. Steliarova-Foucher, J. Lortet-Tieulent, S. Rosso, J. W. Coebergh, H. Comber, D. Forman and F. Bray (2013). "Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012." Eur J Cancer **49**(6): 1374-1403.
- Fiedler, J., K. Breckwoldt, C. W. Remmele, D. Hartmann, M. Dittrich, A. Pfanne, A. Just, K. Xiao, M. Kunz, T. Muller, A. Hansen, R. Geffers, T. Dandekar, T. Eschenhagen and T. Thum (2015). "Development of Long Noncoding RNA-Based Strategies to Modulate Tissue Vascularization." J Am Coll Cardiol **66**(18): 2005-2015.
- Fong, K. M., Y. Sekido, A. F. Gazdar and J. D. Minna (2003). "Lung cancer • 9: Molecular biology of lung cancer: clinical implications." Thorax **58**(10): 892-900.
- Frankel, L. B., N. R. Christoffersen, A. Jacobsen, M. Lindow, A. Krogh and A. H. Lund (2008). "Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells." J Biol Chem **283**(2): 1026-1033.
- Frezzetti, D., M. D. Menna, P. Zoppoli, C. Guerra, A. Ferraro, A. M. Bello, P. D. Luca, C. Calabrese, A. Fusco, M. Ceccarelli, M. Zollo, M. Barbacid, R. D. Lauro and G. D. Vita (2011). "Upregulation of miR-21 by Ras in vivo and its role in tumor growth." Oncogene **30**(3): 275-286.
- Friedman, R. C., K. K. Farh, C. B. Burge and D. P. Bartel (2009). "Most mammalian mRNAs are conserved targets of microRNAs." Genome Res **19**(1): 92-105.
- Fukuoka, M., S. Yano, G. Giaccone, T. Tamura, K. Nakagawa, J.-Y. Douillard, Y. Nishiwaki, J. Vansteenkiste, S. Kudoh, D. Rischin, R. Eek, T. Horai, K. Noda, I. Takata, E. Smit, S. Averbuch, A. Macleod, A. Feyereislova, R.-P. Dong and J. Baselga (2003). "Multi-Institutional Randomized Phase II Trial of Gefitinib for Previously Treated Patients With Advanced Non-Small-Cell Lung Cancer." Journal of Clinical Oncology **21**(12): 2237-2246.
- Funahashi, A., M. Morohashi, H. Kitano and N. Tanimura (2003). "CellDesigner: a process diagram editor for gene-regulatory and biochemical networks." BIOSILICO **1**(5): 159-162.
- Gao, W., Y. Yu, H. Cao, H. Shen, X. Li, S. Pan and Y. Shu (2010). "Deregulated expression of miR-21, miR-143 and miR-181a in non small cell lung cancer is related to clinicopathologic characteristics or patient prognosis." Biomed Pharmacother **64**(6): 399-408.
- Garg, A., A. Di Cara, I. Xenarios, L. Mendoza and G. De Micheli (2008). "Synchronous versus asynchronous modeling of gene regulatory networks." Bioinformatics **24**(17): 1917-1925.
- Garofalo, M. and C. M. Croce (2011). "microRNAs: Master regulators as potential therapeutics in cancer." Annu Rev Pharmacol Toxicol **51**: 25-43.
- Garofalo, M., G. Romano, G. Di Leva, G. Nuovo, Y. J. Jeon, A. Ngankeu, J. Sun, F. Lovat, H. Alder, G. Condorelli, J. A. Engelman, M. Ono, J. K. Rho, L. Cascione, S. Volinia, K. P. Nephew and C. M. Croce (2012). "EGFR and MET receptor tyrosine kinase-altered microRNA expression induces tumorigenesis and gefitinib resistance in lung cancers." Nat Med **18**(1): 74-82.
- Ge, X., Q. Chen, Y. P. Wu, Y. Zhang, H. Xia, D. Yuan, Q. Chen, W. Leng, L. Chen, Q. Tang, X. Pang and F. Bi (2014). "Induced IGF-1R activation contributes to gefitinib resistance following combined treatment with paclitaxel, cisplatin and gefitinib in A549 lung cancer cells." Oncol Rep **32**(4): 1401-1408.
- Geng, X., J. Cheng, A. Gangadharan and D. Mackey (2012). "The coronatine toxin of *Pseudomonas syringae* is a multifunctional suppressor of Arabidopsis defense." Plant Cell **24**(11): 4763-4774.
- Ghosh, J. C., T. Dohi, B. H. Kang and D. C. Altieri (2008). "Hsp60 regulation of tumor cell apoptosis." J Biol Chem **283**(8): 5188-5194.

- Giatromanolaki, A., M. I. Koukourakis, E. Sivridis, H. Turley, K. Talks, F. Pezzella, K. C. Gatter and A. L. Harris (2001). "Relation of hypoxia inducible factor 1 alpha and 2 alpha in operable non-small cell lung cancer to angiogenic/molecular profile of tumours and survival." *Br J Cancer* **85**(6): 881-890.
- Glazebrook, J. (2005). "Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens." *Annu Rev Phytopathol* **43**: 205-227.
- Gohre, V. and S. Robatzek (2008). "Breaking the barriers: microbial effector molecules subvert plant immunity." *Annu Rev Phytopathol* **46**: 189-215.
- Gomes, C. P., J. H. Cho, L. Hood, O. L. Franco, R. W. Pereira and K. Wang (2013). "A Review of Computational Tools in microRNA Discovery." *Front Genet* **4**: 81.
- Gomez, G. G., J. Wykosky, C. Zanca, F. B. Furnari and W. K. Cavenee (2013). "Therapeutic resistance in cancer: microRNA regulation of EGFR signaling networks." *Cancer biology & medicine* **10**(4): 192-205.
- Gomez, G. G., J. Wykosky, C. Zanca, F. B. Furnari and W. K. Cavenee (2013). "Therapeutic resistance in cancer: microRNA regulation of EGFR signaling networks." *Cancer Biol Med* **10**(4): 192-205.
- Gottlich, C., L. C. Muller, M. Kunz, F. Schmitt, H. Walles, T. Walles, T. Dandekar, G. Dandekar and S. L. Nietzer (2016). "A Combined 3D Tissue Engineered In Vitro/In Silico Lung Tumor Model for Predicting Drug Effectiveness in Specific Mutational Backgrounds." *J Vis Exp*(110).
- Goyal, L., M. D. Muzumdar and A. X. Zhu (2013). "Targeting the HGF/c-MET Pathway in Hepatocellular Carcinoma." *Clinical Cancer Research* **19**(9): 2310-2318.
- Grant, M. R. and J. D. Jones (2009). "Hormone (dis)harmony moulds plant health and disease." *Science* **324**(5928): 750-752.
- Greijer, A. E. and E. van der Wall (2004). "The role of hypoxia inducible factor 1 (HIF-1) in hypoxia induced apoptosis." *J Clin Pathol* **57**(10): 1009-1014.
- Guo, J., D. Kim, J. Gao, C. Kurtyka, H. Chen, C. Yu, D. Wu, A. Mittal, A. A. Beg, S. P. Chellappan, E. B. Haura and J. Q. Cheng (2013). "IKBKE is induced by STAT3 and tobacco carcinogen and determines chemosensitivity in non-small cell lung cancer." *Oncogene* **32**(2): 151-159.
- Han, P., W. Li, C. H. Lin, J. Yang, C. Shang, S. T. Nurnberg, K. K. Jin, W. Xu, C. Y. Lin, C. J. Lin, Y. Xiong, H. C. Chien, B. Zhou, E. Ashley, D. Bernstein, P. S. Chen, H. S. Chen, T. Quertermous and C. P. Chang (2014). "A long noncoding RNA protects the heart from pathological hypertrophy." *Nature* **514**(7520): 102-106.
- Hartung, T. (2009). "Toxicology for the twenty-first century." *Nature* **460**(7252): 208-212.
- Hatley, M. E., D. M. Patrick, M. R. Garcia, J. A. Richardson, R. Bassel-Duby, E. van Rooij and E. N. Olson (2010). "Modulation of K-Ras-dependent lung tumorigenesis by MicroRNA-21." *Cancer Cell* **18**(3): 282-293.
- Havgaard, J. H., R. B. Lyngsø and J. Gorodkin (2005). "The foldalign web server for pairwise structural RNA alignment and mutual motif search." *Nucleic Acids Research* **33**(suppl 2): W650-W653.
- He, L., X. He, L. P. Lim, E. de Stanchina, Z. Xuan, Y. Liang, W. Xue, L. Zender, J. Magnus, D. Ridzon, A. L. Jackson, P. S. Linsley, C. Chen, S. W. Lowe, M. A. Cleary and G. J. Hannon (2007). "A microRNA component of the p53 tumour suppressor network." *Nature* **447**(7148): 1130-1134.
- Heegaard, N. H. H., A. J. Schetter, J. A. Welsh, M. Yoneda, E. D. Bowman and C. C. Harris (2012). "Circulating microRNA Expression Profiles in Early Stage Non-Small Cell Lung Cancer." *International Journal of Cancer. Journal International du Cancer* **130**(6): 1378-1386.

- Hentrich, M., C. Böttcher, P. Düchting, Y. Cheng, Y. Zhao, O. Berkowitz, J. Masle, J. Medina and S. Pollmann (2013). "The jasmonic acid signaling pathway is linked to auxin homeostasis through the modulation of YUCCA8 and YUCCA9 gene expression." The Plant journal : for cell and molecular biology **74**(4): 626-637.
- Herbst, R. S., J. V. Heymach and S. M. Lippman (2008). "Lung Cancer." New England Journal of Medicine **359**(13): 1367-1380.
- Herbst, R. S. and A. Sandler (2008). "Bevacizumab and Erlotinib: A Promising New Approach to the Treatment of Advanced NSCLC." The Oncologist **13**(11): 1166-1176.
- Hofacker, I. L. (2003). "Vienna RNA secondary structure server." Nucleic Acids Res **31**(13): 3429-3431.
- Hofacker, I. L., M. Fekete and P. F. Stadler (2002). "Secondary structure prediction for aligned RNA sequences." J Mol Biol **319**(5): 1059-1066.
- Huerta-Cepas, J., D. Szklarczyk, K. Forslund, H. Cook, D. Heller, M. C. Walter, T. Rattei, D. R. Mende, S. Sunagawa, M. Kuhn, L. J. Jensen, C. von Mering and P. Bork (2016). "eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences." Nucleic Acids Res **44**(D1): D286-293.
- Hurbin, A., M. Wislez, B. Busser, M. Antoine, C. Tenaud, N. Rabbe, S. Dufort, F. de Fraipont, D. Moro-Sibilot, J. Cadranet, J. L. Coll and E. Brambilla (2011). "Insulin-like growth factor-1 receptor inhibition overcomes gefitinib resistance in mucinous lung adenocarcinoma." J Pathol **225**(1): 83-95.
- Hwang, I., J. Sheen and B. Muller (2012). "Cytokinin signaling networks." Annu Rev Plant Biol **63**: 353-380.
- Izzotti, A., G. A. Calin, P. Arrigo, V. E. Steele, C. M. Croce and S. De Flora (2009). "Downregulation of microRNA expression in the lungs of rats exposed to cigarette smoke." Faseb j **23**(3): 806-812.
- Jansson, M. D. and A. H. Lund (2012). "MicroRNA and cancer." Mol Oncol **6**(6): 590-610.
- Jensen, L. J., P. Julien, M. Kuhn, C. von Mering, J. Muller, T. Doerks and P. Bork (2008). "eggNOG: automated construction and annotation of orthologous groups of genes." Nucleic Acids Research **36**(Database issue): D250-D254.
- Johnson, S. M., H. Grosshans, J. Shingara, M. Byrom, R. Jarvis, A. Cheng, E. Labourier, K. L. Reinert, D. Brown and F. J. Slack (2005). "RAS is regulated by the let-7 microRNA family." Cell **120**(5): 635-647.
- Jones, J. D. and J. L. Dangl (2006). "The plant immune system." Nature **444**(7117): 323-329.
- Kanehisa, M., Y. Sato, M. Kawashima, M. Furumichi and M. Tanabe (2016). "KEGG as a reference resource for gene and protein annotation." Nucleic Acids Res **44**(D1): D457-462.
- Karl, S. and T. Dandekar (2015). "Convergence behaviour and Control in Non-Linear Biological Networks." Sci Rep **5**: 9746.
- Kasinski, A. L., K. Kelnar, C. Stahlhut, E. Orellana, J. Zhao, E. Shimer, S. Dysart, X. Chen, A. G. Bader and F. J. Slack (2015). "A combinatorial microRNA therapeutics approach to suppressing non-small cell lung cancer." Oncogene **34**(27): 3547-3555.
- Kasinski, A. L. and F. J. Slack (2012). "miRNA-34 Prevents Cancer Initiation and Progression in a Therapeutically Resistant K-ras and p53-Induced Mouse Model of Lung Adenocarcinoma." Cancer Research **72**(21): 5576-5587.
- Kayani, M., M. A. Kayani, F. A. Malik and R. Faryal (2011). "Role of miRNAs in breast cancer." Asian Pac J Cancer Prev **12**(12): 3175-3180.
- Kazan, K. and R. Lyons (2014). "Intervention of Phytohormone Pathways by Pathogen Effectors." Plant Cell **26**(6): 2285-2309.
- Kazan, K. and J. M. Manners (2009). "Linking development to defense: auxin in plant-pathogen interactions." Trends Plant Sci **14**(7): 373-382.

- Kazan, K. and J. M. Manners (2013). "MYC2: the master in action." *Mol Plant* **6**(3): 686-703.
- Keshava Prasad, T. S., R. Goel, K. Kandasamy, S. Keerthikumar, S. Kumar, S. Mathivanan, D. Telikicherla, R. Raju, B. Shafreen, A. Venugopal, L. Balakrishnan, A. Marimuthu, S. Banerjee, D. S. Somanathan, A. Sebastian, S. Rani, S. Ray, C. J. Harrys Kishore, S. Kanth, M. Ahmed, M. K. Kashyap, R. Mohmood, Y. L. Ramachandra, V. Krishna, B. A. Rahiman, S. Mohan, P. Ranganathan, S. Ramabadran, R. Chaerkady and A. Pandey (2009). "Human Protein Reference Database--2009 update." *Nucleic Acids Res* **37**(Database issue): D767-772.
- Kim, N. H., H. S. Kim, X.-Y. Li, I. Lee, H.-S. Choi, S. E. Kang, S. Y. Cha, J. K. Ryu, D. Yoon, E. R. Fearon, R. G. Rowe, S. Lee, C. A. Maher, S. J. Weiss and J. I. Yook (2011). "A p53/miRNA-34 axis regulates Snail1-dependent cancer cell epithelial–mesenchymal transition." *The Journal of Cell Biology* **195**(3): 417-433.
- Kimura, K., Y. Endo, Y. Yonemura, C. W. Heizmann, B. W. Schafer, Y. Watanabe and T. Sasaki (2000). "Clinical significance of S100A4 and E-cadherin-related adhesion molecules in non-small cell lung cancer." *Int J Oncol* **16**(6): 1125-1131.
- Knox, C., V. Law, T. Jewison, P. Liu, S. Ly, A. Frolkis, A. Pon, K. Banco, C. Mak, V. Neveu, Y. Djoumbou, R. Eisner, A. C. Guo and D. S. Wishart (2011). "DrugBank 3.0: a comprehensive resource for 'omics' research on drugs." *Nucleic Acids Res* **39**(Database issue): D1035-1041.
- Kola, I. and J. Landis (2004). "Can the pharmaceutical industry reduce attrition rates?" *Nat Rev Drug Discov* **3**(8): 711-716.
- Kong, L., Y. Zhang, Z. Q. Ye, X. Q. Liu, S. Q. Zhao, L. Wei and G. Gao (2007). "CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine." *Nucleic Acids Res* **35**(Web Server issue): W345-349.
- Kottakis, F. and N. Bardeesy (2012). "LKB1-AMPK axis revisited." *Cell Res* **22**(12): 1617-1620.
- Kris, M. G., R. B. Natale, R. S. Herbst, T. J. Lynch, Jr., D. Prager, C. P. Belani, J. H. Schiller, K. Kelly, H. Spiridonidis, A. Sandler, K. S. Albain, D. Cella, M. K. Wolf, S. D. Averbuch, J. J. Ochs and A. C. Kay (2003). "Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: a randomized trial." *Jama* **290**(16): 2149-2158.
- Kuhajda, F. P. (2008). "AMP-activated protein kinase and human cancer: cancer metabolism revisited." *Int J Obes (Lond)* **32 Suppl 4**: S36-41.
- Kumar, M. S., S. J. Erkeland, R. E. Pester, C. Y. Chen, M. S. Ebert, P. A. Sharp and T. Jacks (2008). "Suppression of non-small cell lung tumor development by the let-7 microRNA family." *Proc Natl Acad Sci U S A* **105**(10): 3903-3908.
- Kumarswamy, R., C. Bauters, I. Volkmann, F. Maury, J. Fetisch, A. Holzmann, G. Lemesle, P. de Groote, F. Pinet and T. Thum (2014). "Circulating long noncoding RNA, LIPCAR, predicts survival in patients with heart failure." *Circ Res* **114**(10): 1569-1575.
- Kunz M, A. N., Dandekar T, and Naseem M. (2013). "Hormone Signaling Networks Open Multiple Routes for Immunity and Disease in Plants." *Biohelikon: Immunity and Diseases* **1:2**.
- Kunz, M., C. Liang, S. Nilla, A. Cecil and T. Dandekar (2016). "The drug-minded protein interaction database (DrumPID) for efficient target analysis and drug development." *Database (Oxford)* **2016**.
- Kunz, M., K. Xiao, C. Liang, J. Viereck, C. Pachel, S. Frantz, T. Thum and T. Dandekar (2014). "Bioinformatics of cardiovascular miRNA biology." *J Mol Cell Cardiol*.
- Kunz M., Dandekar T. and Naseem M. (2016). "A systems biology methodology combining transcriptome and interactome datasets assess the implications of cytokinin signaling for plant immune networks." *Methods Mol Biol*, accepted manuscript.

- Kwok, R. P., J. R. Lundblad, J. C. Chrivia, J. P. Richards, H. P. Bachinger, R. G. Brennan, S. G. Roberts, M. R. Green and R. H. Goodman (1994). "Nuclear protein CBP is a coactivator for the transcription factor CREB." *Nature* **370**(6486): 223-226.
- Landi, M. T., Y. Zhao, M. Rotunno, J. Koshiol, H. Liu, A. W. Bergen, M. Rubagotti, A. M. Goldstein, I. Linnoila, F. M. Marincola, M. A. Tucker, P. A. Bertazzi, A. C. Pesatori, N. E. Caporaso, L. M. McShane and E. Wang (2010). "MicroRNA expression differentiates histology and predicts survival of lung cancer." *Clinical cancer research : an official journal of the American Association for Cancer Research* **16**(2): 430-441.
- Le Quesne, J. and C. Caldas (2010). "Micro-RNAs and breast cancer." *Mol Oncol* **4**(3): 230-241.
- Lee, J. T. (2012). "Epigenetic regulation by long noncoding RNAs." *Science* **338**(6113): 1435-1439.
- Lee, Y. C., C. T. Wu, C. S. Chen, H. H. Hsu and Y. L. Chang (2002). "The significance of E-cadherin and alpha-, beta-, and gamma-catenin expression in surgically treated non-small cell lung cancers of 3 cm or less in size." *J Thorac Cardiovasc Surg* **123**(3): 502-507.
- Lee, Y. S. and A. Dutta (2007). "The tumor suppressor microRNA let-7 represses the HMGA2 oncogene." *Genes Dev* **21**(9): 1025-1030.
- Leever, S. J., B. Vanhaesebroeck and M. D. Waterfield (1999). "Signalling through phosphoinositide 3-kinases: the lipids take centre stage." *Curr Opin Cell Biol* **11**(2): 219-225.
- Lewis, B. P., I. H. Shih, M. W. Jones-Rhoades, D. P. Bartel and C. B. Burge (2003). "Prediction of mammalian microRNA targets." *Cell* **115**(7): 787-798.
- Li, B., S. Ren, X. Li, Y. Wang, D. Garfield, S. Zhou, X. Chen, C. Su, M. Chen, P. Kuang, G. Gao, Y. He, L. Fan, K. Fei, C. Zhou and G. Schmit-Bindert (2014). "MiR-21 overexpression is associated with acquired resistance of EGFR-TKI in non-small cell lung cancer." *Lung Cancer* **83**(2): 146-153.
- Liang, C., B. Kruger and T. Dandekar (2013). "GoSynthetic database tool to analyse natural and engineered molecular processes." *Database (Oxford)* **2013**: bat043.
- Liang, C., A. Schmid, M. J. Lopez-Sanchez, A. Moya, R. Gross, J. Bernhardt and T. Dandekar (2009). "JANE: efficient mapping of prokaryotic ESTs and variable length sequence reads on related template genomes." *BMC Bioinformatics* **10**: 391.
- Lim, C. W. and S. C. Lee (2015). "Arabidopsis abscisic acid receptors play an important role in disease resistance." *Plant Mol Biol* **88**(3): 313-324.
- Linke, K., J. Schanz, J. Hansmann, T. Walles, H. Brunner and H. Mertsching (2007). "Engineered liver-like tissue on a capillarized matrix for applied research." *Tissue engineering* **13**(11): 2699-2707.
- Loreti, E., G. Povero, G. Novi, C. Solfanelli, A. Alpi and P. Perata (2008). "Gibberellins, jasmonate and abscisic acid modulate the sucrose-induced expression of anthocyanin biosynthetic genes in Arabidopsis." *New Phytol* **179**(4): 1004-1016.
- Luo, Z., M. Zang and W. Guo (2010). "AMPK as a metabolic tumor suppressor: control of metabolism and cell growth." *Future Oncol* **6**(3): 457-470.
- Ma, Q., Q. Jiang, Q. Pu, X. Zhang, W. Yang, Y. Wang, S. Ye, S. Wu, G. Zhong, J. Ren, Y. Zhang, L. Liu and W. Zhu (2013). "MicroRNA-143 inhibits migration and invasion of human non-small-cell lung cancer and its relative mechanism." *Int J Biol Sci* **9**(7): 680-692.
- Maere, S., K. Heymans and M. Kuiper (2005). "BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks." *Bioinformatics* **21**(16): 3448-3449.
- Maiwald, T. and J. Timmer (2008). "Dynamical modeling and multi-experiment fitting with PottersWheel." *Bioinformatics* **24**(18): 2037-2043.

- Marchetti, A., C. Martella, L. Felicioni, F. Barassi, S. Salvatore, A. Chella, P. P. Campese, T. Iarussi, F. Mucilli, A. Mezzetti, F. Cuccurullo, R. Sacco and F. Buttitta (2005). "EGFR mutations in non-small-cell lung cancer: analysis of a large series of cases and development of a rapid and sensitive method for diagnostic screening with potential implications on pharmacologic treatment." J Clin Oncol **23**(4): 857-865.
- Medina, P. P., O. A. Romero, T. Kohno, L. M. Montuenga, R. Pio, J. Yokota and M. Sanchez-Cespedes (2008). "Frequent BRG1/SMARCA4-inactivating mutations in human lung cancer cell lines." Hum Mutat **29**(5): 617-622.
- Messeguer, X., R. Escudero, D. Farre, O. Nunez, J. Martinez and M. M. Alba (2002). "PROMO: detection of known transcription regulatory elements using species-tailored searches." Bioinformatics **18**(2): 333-334.
- Minna, J. D., J. A. Roth and A. F. Gazdar (2002). "Focus on lung cancer." Cancer Cell **1**(1): 49-52.
- Mitsunami, T., M. Nishihara, I. Galis, K. M. Alamgir, Y. Hojo, K. Fujita, N. Sasaki, K. Nemoto, T. Sawasaki and G.-i. Arimura (2014). "Overexpression of the PAP1 Transcription Factor Reveals a Complex Regulation of Flavonoid and Phenylpropanoid Metabolism in Nicotiana tabacum Plants Attacked by Spodoptera litura." PLoS ONE **9**(9): e108849.
- Mohr, P. G. and D. M. Cahill (2007). "Suppression by ABA of salicylic acid and lignin accumulation and the expression of multiple genes, in Arabidopsis infected with Pseudomonas syringae pv. tomato." Funct Integr Genomics **7**(3): 181-191.
- Moll, C., J. Reboredo, T. Schwarz, A. Appelt, S. Schurlein, H. Walles and S. Nietzer (2013). "Tissue engineering of a human 3D in vitro tumor test system." J Vis Exp(78).
- Monaghan, J. and C. Zipfel (2012). "Plant pattern recognition receptor complexes at the plasma membrane." Curr Opin Plant Biol **15**(4): 349-357.
- Montani, F., M. J. Marzi, F. Dezi, E. Dama, R. M. Carletti, G. Bonizzi, R. Bertolotti, M. Bellomi, C. Rampinelli, P. Maisonneuve, L. Spaggiari, G. Veronesi, F. Nicassio, P. P. Di Fiore and F. Bianchi (2015). "miR-Test: a blood test for lung cancer early detection." J Natl Cancer Inst **107**(6): djv063.
- Morgan, P., P. H. Van Der Graaf, J. Arrowsmith, D. E. Feltner, K. S. Drummond, C. D. Wegner and S. D. Street (2012). "Can the flow of medicines be improved? Fundamental pharmacokinetic and pharmacological principles toward improving Phase II survival." Drug Discov Today **17**(9-10): 419-424.
- Morgillo, F., W. Y. Kim, E. S. Kim, F. Ciardiello, W. K. Hong and H. Y. Lee (2007). "Implication of the insulin-like growth factor-IR pathway in the resistance of non-small cell lung cancer cells to treatment with gefitinib." Clin Cancer Res **13**(9): 2795-2803.
- Mortazavi, F., J. Lu, R. Phan, M. Lewis, K. Trinidad, A. Aljilani, G. Pezeshkpour and F. Tamanoi (2015). "Significance of KRAS/PAK1/Crk pathway in non-small cell lung cancer oncogenesis." BMC Cancer **15**: 381.
- Mukhtar, M. S., A. R. Carvunis, M. Dreze, P. Epple, J. Steinbrenner, J. Moore, M. Tasan, M. Galli, T. Hao, M. T. Nishimura, S. J. Pevzner, S. E. Donovan, L. Ghamsari, B. Santhanam, V. Romero, M. M. Poulin, F. Gebreab, B. J. Gutierrez, S. Tam, D. Monachello, M. Boxem, C. J. Harbort, N. McDonald, L. Gai, H. Chen, Y. He, J. Vandenhoute, F. P. Roth, D. E. Hill, J. R. Ecker, M. Vidal, J. Beynon, P. Braun and J. L. Dangl (2011). "Independently evolved virulence effectors converge onto hubs in a plant immune system network." Science **333**(6042): 596-601.
- Naseem, M. and T. Dandekar (2012). "The role of auxin-cytokinin antagonism in plant-pathogen interactions." PLoS Pathog **8**(11): e1003026.
- Naseem, M., M. Kunz, N. Ahmed and T. Dandekar (2013). "Integration of boolean models on hormonal interactions and prospects of cytokinin-auxin crosstalk in plant immunity." Plant Signal Behav **8**(4): e23890.

- Naseem, M., M. Kunz and T. Dandekar (2014). "Probing the unknowns in cytokinin-mediated immune defense in Arabidopsis with systems biology approaches." Bioinform Biol Insights **8**: 35-44.
- Naseem, M., N. Philippi, A. Hussain, G. Wangorsch, N. Ahmed and T. Dandekar (2012). "Integrated systems view on networking by hormones in Arabidopsis immunity reveals multiple crosstalk for cytokinin." Plant Cell **24**(5): 1793-1814.
- Nicholson, R. I., J. M. W. Gee and M. E. Harper (2001). "EGFR and cancer prognosis." European Journal of Cancer **37, Supplement 4**: 9-15.
- Nietzer, S. L., F. Baur, S. Sieber, J. Hansmann, T. Schwarz, C. Stoffer, H. Haefner, M. Gasser, A. M. Waaga-Gasser, H. Walles and G. Dandekar (2016). "Mimicking metastases including tumor stroma: a new technique to generate a 3D colorectal cancer model based on a biological decellularized intestinal scaffold." Tissue Engineering Part C: Methods.
- Okamoto, W., I. Okamoto, K. Tanaka, E. Hatashita, Y. Yamada, K. Kuwata, H. Yamaguchi, T. Arao, K. Nishio, M. Fukuoka, P. A. Jänne and K. Nakagawa (2010). "TAK-701, a Humanized Monoclonal Antibody to Hepatocyte Growth Factor, Reverses Gefitinib Resistance Induced by Tumor-Derived HGF in Non-Small Cell Lung Cancer with an EGFR Mutation." Molecular Cancer Therapeutics **9**(10): 2785-2792.
- Okayama, H., M. Saito, N. Oue, J. M. Weiss, J. Stauffer, S. Takenoshita, R. H. Wiltrout, S. P. Hussain and C. C. Harris (2013). "NOS2 enhances KRAS-induced lung carcinogenesis, inflammation and microRNA-21 expression." International journal of cancer. Journal international du cancer **132**(1): 9-18.
- Okayama, H., M. Saito, N. Oue, J. M. Weiss, J. Stauffer, S. Takenoshita, R. H. Wiltrout, S. P. Hussain and C. C. Harris (2013). "NOS2 enhances KRAS-induced lung carcinogenesis, inflammation and microRNA-21 expression." Int J Cancer **132**(1): 9-18.
- Oliveros, J. C. ((2007-2015)). "Venny. An interactive tool for comparing lists with Venn's diagrams. <http://bioinfogp.cnb.csic.es/tools/venny/index.html>."
- Orchard, S., S. Kerrien, S. Abbani, B. Aranda, J. Bhate, S. Bidwell, A. Bridge, L. Briganti, F. S. Brinkman, G. Cesareni, A. Chatr-aryamontri, E. Chautard, C. Chen, M. Dumousseau, J. Goll, R. E. Hancock, L. I. Hannick, I. Jurisica, J. Khadake, D. J. Lynn, U. Mahadevan, L. Perfetto, A. Raghunath, S. Ricard-Blum, B. Roechert, L. Salwinski, V. Stumpflen, M. Tyers, P. Uetz, I. Xenarios and H. Hermjakob (2012). "Protein interaction data curation: the International Molecular Exchange (IMEx) consortium." Nat Methods **9**(4): 345-350.
- Palayoor, S. T., J. B. Mitchell, D. Cerna, W. Degraff, M. John-Aryankalayil and C. N. Coleman (2008). "PX-478, an inhibitor of hypoxia-inducible factor-1alpha, enhances radiosensitivity of prostate carcinoma cells." Int J Cancer **123**(10): 2430-2437.
- Pampaloni, F., E. G. Reynaud and E. H. Stelzer (2007). "The third dimension bridges the gap between cell culture and live tissue." Nat Rev Mol Cell Biol **8**(10): 839-845.
- Pandey, S. P., M. Roccaro, M. Schon, E. Logemann and I. E. Somssich (2010). "Transcriptional reprogramming regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of Arabidopsis." Plant J **64**(6): 912-923.
- Pao, W., T. Y. Wang, G. J. Riely, V. A. Miller, Q. Pan, M. Ladanyi, M. F. Zakowski, R. T. Heelan, M. G. Kris and H. E. Varmus (2005). "KRAS Mutations and Primary Resistance of Lung Adenocarcinomas to Gefitinib or Erlotinib." PLoS Med **2**(1): e17.
- Paralkar, V. R. and M. J. Weiss (2013). "Long noncoding RNAs in biology and hematopoiesis." Blood **121**(24): 4842-4846.
- Philippi, N., D. Walter, R. Schlatter, K. Ferreira, M. Ederer, O. Sawodny, J. Timmer, C. Borner and T. Dandekar (2009). "Modeling system states in liver cells: Survival, apoptosis and their modifications in response to viral infection." BMC Systems Biology **3**(1): 1-11.



- Pieterse, C. M., A. Leon-Reyes, S. Van der Ent and S. C. Van Wees (2009). "Networking by small-molecule hormones in plant immunity." *Nat Chem Biol* **5**(5): 308-316.
- Pieterse, C. M., D. Van der Does, C. Zamioudis, A. Leon-Reyes and S. C. Van Wees (2012). "Hormonal modulation of plant immunity." *Annu Rev Cell Dev Biol* **28**: 489-521.
- Puhakka, A., V. Kinnula, U. Napankangas, M. Saily, P. Koistinen, P. Paakko and Y. Soini (2003). "High expression of nitric oxide synthases is a favorable prognostic sign in non-small cell lung carcinoma." *Apmis* **111**(12): 1137-1146.
- Puhakka, A. R., T. H. Harju, P. K. Paakko, Y. M. Soini and V. L. Kinnula (2006). "Nitric oxide synthases are associated with bronchial dysplasia." *Lung Cancer* **51**(3): 275-282.
- Qi, J. and D. Mu (2012). "MicroRNAs and lung cancers: from pathogenesis to clinical implications." *Frontiers of medicine* **6**(2): 134-155.
- Qu, Y., X. Wu, Y. Yin, Y. Yang, D. Ma and H. Li (2014). "Antitumor activity of selective MEK1/2 inhibitor AZD6244 in combination with PI3K/mTOR inhibitor BEZ235 in gefitinib-resistant NSCLC xenograft models." *J Exp Clin Cancer Res* **33**: 52.
- R., G. (2007). "Chemical Informatics Functionality in R." *J. Stat. Software* **18**, 1-16.
- Ramalingam, S. S., T. K. Owonikoko and F. R. Khuri (2011). "Lung cancer: New biological insights and recent therapeutic advances." *CA Cancer J Clin* **61**(2): 91-112.
- Ramaswamy, S., N. Nakamura, F. Vazquez, D. B. Batt, S. Perera, T. M. Roberts and W. R. Sellers (1999). "Regulation of G1 progression by the PTEN tumor suppressor protein is linked to inhibition of the phosphatidylinositol 3-kinase/Akt pathway." *Proc Natl Acad Sci U S A* **96**(5): 2110-2115.
- Rattan, R., S. Giri, A. K. Singh and I. Singh (2005). "5-Aminoimidazole-4-carboxamide-1-beta-D-ribofuranoside inhibits cancer cell proliferation in vitro and in vivo via AMP-activated protein kinase." *J Biol Chem* **280**(47): 39582-39593.
- Ravi, R., B. Mookerjee, Z. M. Bhujwalla, C. H. Sutter, D. Artemov, Q. Zeng, L. E. Dillehay, A. Madan, G. L. Semenza and A. Bedi (2000). "Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1alpha." *Genes Dev* **14**(1): 34-44.
- Reeder, J. and R. Giegerich (2005). "Consensus shapes: an alternative to the Sankoff algorithm for RNA consensus structure prediction." *Bioinformatics* **21**(17): 3516-3523.
- Reese, M. G. (2001). "Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome." *Comput Chem* **26**(1): 51-56.
- Reisman, D. N., J. Sciarrotta, W. Wang, W. K. Funkhouser and B. E. Weissman (2003). "Loss of BRG1/BRM in human lung cancer cell lines and primary lung cancers: correlation with poor prognosis." *Cancer Res* **63**(3): 560-566.
- Riely, G. J., M. G. Kris, D. Rosenbaum, J. Marks, A. Li, D. A. Chitale, K. Nafa, E. R. Riedel, M. Hsu, W. Pao, V. A. Miller and M. Ladanyi (2008). "Frequency and Distinctive Spectrum of KRAS Mutations in Never Smokers with Lung Adenocarcinoma." *Clinical Cancer Research* **14**(18): 5731-5734.
- Robert-Seilantantz, A., M. Grant and J. D. Jones (2011). "Hormone crosstalk in plant disease and defense: more than just jasmonate-salicylate antagonism." *Annu Rev Phytopathol* **49**: 317-343.
- Roine, E., W. Wei, J. Yuan, E. L. Nurmiäho-Lassila, N. Kalkkinen, M. Romantschuk and S. Y. He (1997). "Hrp pilus: an hrp-dependent bacterial surface appendage produced by *Pseudomonas syringae* pv. tomato DC3000." *Proc Natl Acad Sci U S A* **94**(7): 3459-3464.
- Sachdeva, M. and Y.-Y. Mo (2010). "MicroRNA-145 suppresses cell invasion and metastasis by directly targeting mucin 1." *Cancer research* **70**(1): 378-387.
- Saito, R., M. E. Smoot, K. Ono, J. Ruschinski, P. L. Wang, S. Lotia, A. R. Pico, G. D. Bader and T. Ideker (2012). "A travel guide to Cytoscape plugins." *Nat Methods* **9**(11): 1069-1076.

- Sampson, V. B., N. H. Rong, J. Han, Q. Yang, V. Aris, P. Soteropoulos, N. J. Petrelli, S. P. Dunn and L. J. Krueger (2007). "MicroRNA Let-7a Down-regulates MYC and Reverts MYC-Induced Growth in Burkitt Lymphoma Cells." *Cancer Research* **67**(20): 9762-9770.
- Sankoff, D. (1985). "Simultaneous Solution of the RNA Folding, Alignment and Protosequence Problems." *SIAM Journal on Applied Mathematics* **45**(5): 810-825.
- Sarris, P. F., Z. Duxbury, S. U. Huh, Y. Ma, C. Segonzac, J. Sklenar, P. Derbyshire, V. Cevik, G. Rallapalli, S. B. Saucet, L. Wirthmueller, F. L. Menke, K. H. Sohn and J. D. Jones (2015). "A Plant Immune Receptor Detects Pathogen Effectors that Target WRKY Transcription Factors." *Cell* **161**(5): 1089-1100.
- Schanz, J., J. Pusch, J. Hansmann and H. Walles (2010). "Vascularised human tissue models: a new approach for the refinement of biomedical research." *J Biotechnol* **148**(1): 56-63.
- Schlatter, R., N. Philippi, G. Wangorsch, R. Pick, O. Sawodny, C. Borner, J. Timmer, M. Ederer and T. Dandekar (2012). "Integration of Boolean models exemplified on hepatocyte signal transduction." *Briefings in Bioinformatics* **13**(3): 365-376.
- Schmidt, S., P. Bork and T. Dandekar (2002). "A versatile structural domain analysis server using profile weight matrices." *J Chem Inf Comput Sci* **42**(2): 405-407.
- Sekido, Y., K. M. Fong and J. D. Minna (2003). "Molecular Genetics of Lung Cancer." *Annual Review of Medicine* **54**(1): 73-87.
- Semenza, G. L. (2003). "Targeting HIF-1 for cancer therapy." *Nat Rev Cancer* **3**(10): 721-732.
- Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker (2003). "Cytoscape: a software environment for integrated models of biomolecular interaction networks." *Genome Res* **13**(11): 2498-2504.
- Sheard, L. B., X. Tan, H. Mao, J. Withers, G. Ben-Nissan, T. R. Hinds, Y. Kobayashi, F. F. Hsu, M. Sharon, J. Browse, S. Y. He, J. Rizo, G. A. Howe and N. Zheng (2010). "Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor." *Nature* **468**(7322): 400-405.
- Shen, H., F. Zhu, J. Liu, T. Xu, D. Pei, R. Wang, Y. Qian, Q. Li, L. Wang, Z. Shi, J. Zheng, Q. Chen, B. Jiang and Y. Shu (2014). "Alteration in Mir-21/PTEN expression modulates gefitinib resistance in non-small cell lung cancer." *PLoS One* **9**(7): e103305.
- Shen, J., N. W. Todd, H. Zhang, L. Yu, X. Lingxiao, Y. Mei, M. Guarnera, J. Liao, A. Chou, C. L. Lu, Z. Jiang, H. Fang, R. L. Katz and F. Jiang (2011). "Plasma microRNAs as potential biomarkers for non-small-cell lung cancer." *Lab Invest* **91**(4): 579-587.
- Shigematsu, H., L. Lin, T. Takahashi, M. Nomura, M. Suzuki, I. I. Wistuba, K. M. Fong, H. Lee, S. Toyooka, N. Shimizu, T. Fujisawa, Z. Feng, J. A. Roth, J. Herz, J. D. Minna and A. F. Gazdar (2005). "Clinical and Biological Features Associated With Epidermal Growth Factor Receptor Gene Mutations in Lung Cancers." *Journal of the National Cancer Institute* **97**(5): 339-346.
- Siemens, H., R. Jackstadt, S. Hunten, M. Kaller, A. Menssen, U. Gotz and H. Hermeking (2011). "miR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions." *Cell Cycle* **10**(24): 4256-4271.
- Smith, C., S. Heyne, A. S. Richter, S. Will and R. Backofen (2010). "Freiburg RNA Tools: a web server integrating INTARNA, EXPARNA and LOCARNA." *Nucleic Acids Res* **38**(Web Server issue): W373-377.
- So, K. S., J. K. Rho, Y. J. Choi, S. Y. Kim, C. M. Choi, Y. J. Chun and J. C. Lee (2015). "AKT/mTOR down-regulation by CX-4945, a CK2 inhibitor, promotes apoptosis in chemorefractory non-small cell lung cancer cells." *Anticancer Res* **35**(3): 1537-1542.
- Socinski, M. A., J. Goldman, I. El-Hariry, M. Koczywas, V. Vukovic, L. Horn, E. Paschold, R. Salgia, H. West, L. V. Sequist, P. Bonomi, J. Brahmer, L. C. Chen, A. Sandler, C. P. Belani, T. Webb, H. Harper, M. Huberman, S. Ramalingam, K. K. Wong, F. Teofilovici, W. Guo and G.

- I. Shapiro (2013). "A multicenter phase II study of ganetespib monotherapy in patients with genotypically defined advanced non-small cell lung cancer." *Clin Cancer Res* **19**(11): 3068-3077.
- Solovyev, V. V., I. A. Shahmuradov and A. A. Salamov (2010). "Identification of promoter regions and regulatory sites." *Methods Mol Biol* **674**: 57-83.
- Sos, M. L., K. Michel, T. Zander, J. Weiss, P. Frommolt, M. Peifer, D. Li, R. Ullrich, M. Koker, F. Fischer, T. Shimamura, D. Rauh, C. Mermel, S. Fischer, I. Stuckrath, S. Heynck, R. Beroukhim, W. Lin, W. Winckler, K. Shah, T. LaFramboise, W. F. Moriarty, M. Hanna, L. Tolosi, J. Rahnenfuhrer, R. Verhaak, D. Chiang, G. Getz, M. Hellmich, J. Wolf, L. Girard, M. Peyton, B. A. Weir, T. H. Chen, H. Greulich, J. Barretina, G. I. Shapiro, L. A. Garraway, A. F. Gazdar, J. D. Minna, M. Meyerson, K. K. Wong and R. K. Thomas (2009). "Predicting drug susceptibility of non-small cell lung cancers based on genetic lesions." *J Clin Invest* **119**(6): 1727-1740.
- Soutoglou, E., G. Papafotiou, N. Katrakili and I. Talianidis (2000). "Transcriptional activation by hepatocyte nuclear factor-1 requires synergism between multiple coactivator proteins." *J Biol Chem* **275**(17): 12515-12520.
- Sozzi, G., M. Boeri, M. Rossi, C. Verri, P. Suatoni, F. Bravi, L. Roz, D. Conte, M. Grassi, N. Sverzellati, A. Marchiano, E. Negri, C. La Vecchia and U. Pastorino (2014). "Clinical utility of a plasma-based miRNA signature classifier within computed tomography lung cancer screening: a correlative MILD trial study." *J Clin Oncol* **32**(8): 768-773.
- Stahlhut, C. and F. J. Slack (2015). "Combinatorial Action of MicroRNAs let-7 and miR-34 Effectively Synergizes with Erlotinib to Suppress Non-small Cell Lung Cancer Cell Proliferation." *Cell Cycle* **14**(13): 2171-2180.
- Steffen, P., B. Voss, M. Rehmsmeier, J. Reeder and R. Giegerich (2006). "RNAshapes: an integrated RNA analysis package based on abstract shapes." *Bioinformatics* **22**(4): 500-503.
- Stratmann, A. T., D. Fecher, G. Wangorsch, C. Gottlich, T. Walles, H. Walles, T. Dandekar, G. Dandekar and S. L. Nietzer (2014). "Establishment of a human 3D lung cancer model based on a biological tissue matrix combined with a Boolean in silico model." *Mol Oncol* **8**(2): 351-365.
- Sun, S., J. H. Schiller and A. F. Gazdar (2007). "Lung cancer in never smokers [mdash] a different disease." *Nat Rev Cancer* **7**(10): 778-790.
- Suzuki, T., T. Nakagawa, H. Endo, T. Mitsudomi, A. Masuda, Y. Yatabe, T. Sugiura, T. Takahashi and T. Hida (2003). "The sensitivity of lung cancer cell lines to the EGFR-selective tyrosine kinase inhibitor ZD1839 ('Iressa') is not related to the expression of EGFR or HER-2 or to K-ras gene status." *Lung Cancer* **42**(1): 35-41.
- Szklarczyk, D., A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. P. Tsafou, M. Kuhn, P. Bork, L. J. Jensen and C. von Mering (2015). "STRING v10: protein-protein interaction networks, integrated over the tree of life." *Nucleic Acids Res* **43**(Database issue): D447-452.
- Takamizawa, J., H. Konishi, K. Yanagisawa, S. Tomida, H. Osada, H. Endoh, T. Harano, Y. Yatabe, M. Nagino, Y. Nimura, T. Mitsudomi and T. Takahashi (2004). "Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival." *Cancer Res* **64**(11): 3753-3756.
- Tang, D., Y. Shen, M. Wang, R. Yang, Z. Wang, A. Sui, W. Jiao and Y. Wang (2013). "Identification of plasma microRNAs as novel noninvasive biomarkers for early detection of lung cancer." *Eur J Cancer Prev* **22**(6): 540-548.
- Tang, X., H. Shigematsu, B. N. Bekele, J. A. Roth, J. D. Minna, W. K. Hong, A. F. Gazdar and I. I. Wistuba (2005). "EGFR Tyrosine Kinase Domain Mutations Are Detected in Histologically Normal Respiratory Epithelium in Lung Cancer Patients." *Cancer Research* **65**(17): 7568-7572.

- Tang, Y. C., B. R. Williams, J. J. Siegel and A. Amon (2011). "Identification of aneuploidy-selective antiproliferation compounds." *Cell* **144**(4): 499-512.
- Tatusov, R. L., N. D. Fedorova, J. D. Jackson, A. R. Jacobs, B. Kiryutin, E. V. Koonin, D. M. Krylov, R. Mazumder, S. L. Mekhedov, A. N. Nikolskaya, B. S. Rao, S. Smirnov, A. V. Sverdlov, S. Vasudevan, Y. I. Wolf, J. J. Yin and D. A. Natale (2003). "The COG database: an updated version includes eukaryotes." *BMC Bioinformatics* **4**: 41.
- Tazawa, H., N. Tsuchiya, M. Izumiya and H. Nakagama (2007). "Tumor-suppressive miR-34a induces senescence-like growth arrest through modulation of the E2F pathway in human colon cancer cells." *Proc Natl Acad Sci U S A* **104**(39): 15472-15477.
- Templeton, A. K., S. Miyamoto, A. Babu, A. Munshi and R. Ramesh (2014). "Cancer stem cells: progress and challenges in lung cancer." *Stem Cell Investigation* **1**(4).
- Thakar, J. and R. Albert (2010). "Boolean models of within-host immune interactions." *Current Opinion in Microbiology* **13**(3): 377-381.
- The R Development Core Team. Vienna, Austria, 2010. R Foundation for Statistical Computing, ISBN 3-900051-07-0). R: A Language and Environment for Statistical Computing.
- Tokarz, P. and J. Blasiak (2012). "The role of microRNA in metastatic colorectal cancer and its significance in cancer prognosis and treatment." *Acta Biochim Pol* **59**(4): 467-474.
- Tsuda, K. and I. E. Somssich (2015). "Transcriptional networks in plant immunity." *New Phytol* **206**(3): 932-947.
- Tully, J. P., A. E. Hill, H. M. Ahmed, R. Whitley, A. Skjellum and M. S. Mukhtar (2014). "Expression-based network biology identifies immune-related functional modules involved in plant defense." *BMC Genomics* **15**: 421.
- Tursz, T., F. Andre, V. Lazar, L. Lacroix and J. C. Soria (2011). "Implications of personalized medicine--perspective from a cancer center." *Nat Rev Clin Oncol* **8**(3): 177-183.
- Ulitsky, I. and D. P. Bartel (2013). "lincRNAs: genomics, evolution, and mechanisms." *Cell* **154**(1): 26-46.
- Van Den Neste, E., B. Cazin, A. Janssens, E. González-Barca, M. J. Terol, V. Levy, J. Pérez de Oteyza, P. Zachee, A. Saunders, M. de Frias and C. Campàs (2013). "Acadesine for patients with relapsed/refractory chronic lymphocytic leukemia (CLL): a multicenter phase I/II study." *Cancer Chemotherapy and Pharmacology* **71**(3): 581-591.
- van Rens, M. T., A. B. de la Riviere, H. R. Elbers and J. M. van Den Bosch (2000). "Prognostic assessment of 2,361 patients who underwent pulmonary resection for non-small cell lung cancer, stage I, II, and IIIA." *Chest* **117**(2): 374-379.
- Verhage, A., S. C. van Wees and C. M. Pieterse (2010). "Plant immunity: it's the hormones talking, but what do they say?" *Plant Physiol* **154**(2): 536-540.
- Vickers, M. M., J. Bar, I. Gorn-Hondermann, N. Yarom, M. Daneshmand, J. E. Hanson, C. L. Addison, T. R. Asmis, D. J. Jonker, J. Maroun, I. A. Lorimer, G. D. Goss and J. Dimitroulakos (2012). "Stage-dependent differential expression of microRNAs in colorectal cancer: potential role as markers of metastatic disease." *Clin Exp Metastasis* **29**(2): 123-132.
- Viereck, J., C. Bang, A. Foinquinos and T. Thum (2014). "Regulatory RNAs and paracrine networks in the heart." *Cardiovascular Research* **102**(2): 290-301.
- Viereck, J., R. Kumarswamy, A. Foinquinos, K. Xiao, P. Avramopoulos, M. Kunz, M. Dittrich, T. Maetzig, K. Zimmer, J. Remke, A. Just, J. Fendrich, K. Scherf, E. Bolesani, A. Schambach, F. Weidemann, R. Zweigerdt, L. J. de Windt, S. Engelhardt, T. Dandekar, S. Batkai and T. Thum (2016). "Long noncoding RNA Chast promotes cardiac remodeling." *Sci Transl Med* **8**(326): 326ra322.
- Volm, M. and R. Koomagi (2000). "Hypoxia-inducible factor (HIF-1) and its relationship to apoptosis and proliferation in lung cancer." *Anticancer Res* **20**(3a): 1527-1533.

- Wakeling, A. E., S. P. Guy, J. R. Woodburn, S. E. Ashton, B. J. Curry, A. J. Barker and K. H. Gibson (2002). "ZD1839 (Iressa): An Orally Active Inhibitor of Epidermal Growth Factor Signaling with Potential for Cancer Therapy." Cancer Research **62**(20): 5749-5754.
- Wangorsch, G., E. Butt, R. Mark, K. Hubertus, J. Geiger, T. Dandekar and M. Dittrich (2011). "Time-resolved in silico modeling of fine-tuned cAMP signaling in platelets: feedback loops, titrated phosphorylations and pharmacological modulation." BMC Systems Biology **5**: 178-178.
- WD. Travis, E. B., HK Müller-Hermelink, CC. Harris (2004). "World Health Organization Classification of Tumors. Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart." IARC Press; Lyon, France.
- Welsh, S., R. Williams, L. Kirkpatrick, G. Paine-Murrieta and G. Powis (2004). "Antitumor activity and pharmacodynamic properties of PX-478, an inhibitor of hypoxia-inducible factor-1alpha." Mol Cancer Ther **3**(3): 233-244.
- Williams, A. J., V. Tkachenko, S. Golotvin, R. Kidd and G. McCann (2010). "ChemSpider - building a foundation for the semantic web by hosting a crowd sourced databasing platform for chemistry." Journal of Cheminformatics **2**(Suppl 1): O16-O16.
- Willighagen, E. L., A. Waagmeester, O. Spjuth, P. Ansell, A. J. Williams, V. Tkachenko, J. Hastings, B. Chen and D. J. Wild (2013). "The ChEMBL database as linked open data." J Cheminform **5**(1): 23.
- Woodburn, J. R. (1999). "The Epidermal Growth Factor Receptor and Its Inhibition in Cancer Therapy." Pharmacology & Therapeutics **82**(2-3): 241-250.
- Xia, H. and K. M. Hui (2012). "MicroRNAs involved in regulating epithelial-mesenchymal transition and cancer stem cells as molecular targets for cancer therapeutics." Cancer Gene Ther **19**(11): 723-730.
- Xin, M., E. M. Small, L. B. Sutherland, X. Qi, J. McAnally, C. F. Plato, J. A. Richardson, R. Bassel-Duby and E. N. Olson (2009). "MicroRNAs miR-143 and miR-145 modulate cytoskeletal dynamics and responsiveness of smooth muscle cells to injury." Genes & Development **23**(18): 2166-2178.
- Xin, X. F. and S. Y. He (2013). "Pseudomonas syringae pv. tomato DC3000: a model pathogen for probing disease susceptibility and hormone signaling in plants." Annu Rev Phytopathol **51**: 473-498.
- Xu, J., Y. Li, Y. Wang, H. Liu, L. Lei, H. Yang, G. Liu and D. Ren (2008). "Activation of MAPK kinase 9 induces ethylene and camalexin biosynthesis and enhances sensitivity to salt stress in Arabidopsis." J Biol Chem **283**(40): 26996-27006.
- Yan, B., J. Yao, J. Y. Liu, X. M. Li, X. Q. Wang, Y. J. Li, Z. F. Tao, Y. C. Song, Q. Chen and Q. Jiang (2015). "lncRNA-MIAT regulates microvascular dysfunction by functioning as a competing endogenous RNA." Circ Res **116**(7): 1143-1156.
- Yanaihara, N., N. Caplen, E. Bowman, M. Seike, K. Kumamoto, M. Yi, R. M. Stephens, A. Okamoto, J. Yokota, T. Tanaka, G. A. Calin, C. G. Liu, C. M. Croce and C. C. Harris (2006). "Unique microRNA molecular profiles in lung cancer diagnosis and prognosis." Cancer Cell **9**(3): 189-198.
- Yang, D. L., J. Yao, C. S. Mei, X. H. Tong, L. J. Zeng, Q. Li, L. T. Xiao, T. P. Sun, J. Li, X. W. Deng, C. M. Lee, M. F. Thomashow, Y. Yang, Z. He and S. Y. He (2012). "Plant hormone jasmonate prioritizes defense over growth by interfering with gibberellin signaling cascade." Proc Natl Acad Sci U S A **109**(19): E1192-1200.
- Yang, M., H. Shen, C. Qiu, Y. Ni, L. Wang, W. Dong, Y. Liao and J. Du (2013). "High expression of miR-21 and miR-155 predicts recurrence and unfavourable survival in non-small cell lung cancer." European Journal of Cancer **49**(3): 604-615.

- Yin, R., S. Zhang, Y. Wu, X. Fan, F. Jiang, Z. Zhang, D. Feng, X. Guo and L. Xu (2011). "microRNA-145 suppresses lung adenocarcinoma-initiating cell proliferation by targeting OCT4." *Oncol Rep* **25**(6): 1747-1754.
- Yu, H. A., M. E. Arcila, N. Rekhtman, C. S. Sima, M. F. Zakowski, W. Pao, M. G. Kris, V. A. Miller, M. Ladanyi and G. J. Riely (2013). "Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers." *Clin Cancer Res* **19**(8): 2240-2247.
- Yu, S. L., H. Y. Chen, G. C. Chang, C. Y. Chen, H. W. Chen, S. Singh, C. L. Cheng, C. J. Yu, Y. C. Lee, H. S. Chen, T. J. Su, C. C. Chiang, H. N. Li, Q. S. Hong, H. Y. Su, C. C. Chen, W. J. Chen, C. C. Liu, W. K. Chan, W. J. Chen, K. C. Li, J. J. Chen and P. C. Yang (2008). "MicroRNA signature predicts survival and relapse in lung cancer." *Cancer Cell* **13**(1): 48-57.
- Zhang, J. G., J. J. Wang, F. Zhao, Q. Liu, K. Jiang and G. H. Yang (2010). "MicroRNA-21 (miR-21) represses tumor suppressor PTEN and promotes growth and invasion in non-small cell lung cancer (NSCLC)." *Clin Chim Acta* **411**(11-12): 846-852.
- Zhang, J. J., U. Vinkemeier, W. Gu, D. Chakravarti, C. M. Horvath and J. E. Darnell, Jr. (1996). "Two contact regions between Stat1 and CBP/p300 in interferon gamma signaling." *Proc Natl Acad Sci U S A* **93**(26): 15092-15096.
- Zhang, S., H. Long, Y. L. Yang, Y. Wang, D. Hsieh, W. Li, A. Au, H. J. Stoppler, Z. Xu, D. M. Jablons and L. You (2013). "Inhibition of CK2alpha down-regulates Notch1 signalling in lung cancer cells." *J Cell Mol Med* **17**(7): 854-862.
- Zhao, N., Y. Liu, Z. Chang, K. Li, R. Zhang, Y. Zhou, F. Qiu, X. Han and Y. Xu (2015). "Identification of Biomarker and Co-Regulatory Motifs in Lung Adenocarcinoma Based on Differential Interactions." *PLoS One* **10**(9): e0139165.
- Zhou, J. Y., X. Chen, J. Zhao, Z. Bao, X. Chen, P. Zhang, Z. F. Liu and J. Y. Zhou (2014). "MicroRNA-34a overcomes HGF-mediated gefitinib resistance in EGFR mutant lung cancer cells partly by targeting MET." *Cancer Lett* **351**(2): 265-271.
- Zipfel, C., G. Kunze, D. Chinchilla, A. Caniard, J. D. Jones, T. Boller and G. Felix (2006). "Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation." *Cell* **125**(4): 749-760.
- Zuker, M. (2003). "Mfold web server for nucleic acid folding and hybridization prediction." *Nucleic Acids Res* **31**(13): 3406-3415.
- Zwiesler-Vollick, J., A. E. Plovianich-Jones, K. Nomura, S. Bandyopadhyay, V. Joardar, B. N. Kunkel and S. Y. He (2002). "Identification of novel hrp-regulated genes through functional genomic analysis of the *Pseudomonas syringae* pv. tomato DC3000 genome." *Mol Microbiol* **45**(5): 1207-1218.

## 9. Anhang

### 9.1 Abkürzungsverzeichnis

ABA	Abscisinsäure (englisch: Abscisic acid)
Abb.	Abbildung
<i>A. fumigatus</i>	<i>Aspergillus fumigatus</i>
AGO2	Argonaute 2
AHKs	Arabidopsis Histidine Kinases (Arabidopsis Histidinkinasen)
AHPs	Arabidopsis Histidine Phosphotransfer Proteins
AICAR	5-Aminoimidazole-4-carboxamide ribonucleotide
AKT	V-Akt Murine Thymoma Viral Oncogene Homolog
AMPK	AMP-aktivierte Proteinkinase
Arhgap27	Rho GTPase Activating Protein 27
ARRs	Arabidopsis Response Regulators
<i>A. thaliana</i>	<i>Arabidopsis thaliana</i>
AUX	Auxin
bzw.	beziehungsweise
chr	Chromosom
CK	Cytokinin
c-MET	MET Proto-Oncogene
COG/KOGs	Clusters of Orthologous Groups/Eukaryotic Orthologous Groups
COR	Coronatin
CSNK2A1	Casein Kinase 2, Alpha 1 Polypeptide
CTNNA1	Alpha-Catenin (Cadherin-Associated Protein)
DEGe	differenziell-exprimierte Gene
DGCR8	DiGeorge Syndrome Critical Region 8
DRP	differenziell-reguliertes Protein
EBN	expressions-basiertes Netzwerk
EGFR	Epidermal Growth Factor Receptor

ET	Ethylen
ETI	Effector-triggered immunity
ETS	Effector-triggered susceptibility
GA	Gibberellinsäure (englisch: Gibberellic acid)
GEO	Gene Expression Omnibus
GO	Gene-Ontology
HIF1A	Hypoxia-inducible factor 1-alpha
Hrp	hypersensitive response and pathogenicity hrp
HSP90	Hitze-Schock-Protein 90 (englisch: Heat Shock Protein 90)
IGF-1R	Insulin-Like Growth Factor 1 Receptor
IP	Interaktionspartner
JA	Jasmonsäure (englisch: Jasmonic acid)
KRAS	Kirsten Rat Sarcoma Viral Oncogene Homolog
let-7	lethal-7
lncRNAs	long noncoding RNAs (lange nicht-kodierende Ribonukleinsäuren)
LOF	Loss Of Function
MAMPs	Microbial-associated molecular patterns
MAPK	Mitogen-aktivierte Proteinkinasen
MIAT	Myocardial infarction-associated transcript
miRNAs	microRNAs
mRNA	messenger-RNA
mTOR	mechanistic Target Of Rapamycin
Myheart	Myosin heavy-chain-associated RNA transcript
ncRNAs	noncoding Ribonucleic Acids (nicht-kodierende Ribonukleinsäuren)
NOS1	Nitric Oxide Synthase 1
NR3C1	Nuclear Receptor Subfamily 3, Group C, Member 1
NSCLC	Non-Small Cell Lung Cancer
Nte	Nukleotide
PAMPs	Pathogen-associated molecular patterns



PDCD4	Programmed cell death protein 4
PDGFRB	Platelet-Derived Growth Factor Receptor Beta
PI	Protein-Interaktion
PIP	Protein-Interaktions-Partner
PI3K	Phosphoinositide-3-Kinase
PK	Phosphokinase
Plekhm1	Pleckstrin homology domain containing, family M (with RUN domain) member 1
PPI	Protein-Protein-Interaktion
pre-miRNA	precursor-miRNA
pri-miRNA	primary-miRNA
PRRs	Pattern recognition receptors
PR1	Pathogenesis-related protein 1
<i>Pst</i> DC3000	<i>Pseudomonas syringae</i> pv. <i>tomato</i> DC3000
PTEN	Phosphatase and Tensin homolog
PTI	PAMP-triggered immunity
p53	Tumor-Protein P53 (auch TP53)
RISC	RNA-inducible silencing complex
RNAs	Ribonucleic Acids
RNase	Ribonuklease
RTK	Rezeptor-Tyrosinkinase (englisch: Receptor Tyrosine Kinase)
SA	Salicylsäure (englisch: Salicylic acid)
SENCR	Smooth muscle and Endothelial cell-enriched migration/differentiation-associated long Non-Coding RNA
SISmuc	Small Intestinal Submucosa
SMARCA4	SWI/SNF-Related Matrix-Associated Actin-Dependent Regulator Of Chromatin Subfamily A Member 4 (BRG1)
SMILES	Simplified Molecular Input Line Entry Specification
Spry1	Sprouty RTK Signaling Antagonist 1
STK	Serine-Threonin-Kinase

Tab.	Tabelle
TF	Transkriptionsfaktor
TFBS	Transkriptionsfaktorbindestellen
TKI	Tyrosin-Kinase-Inhibitor
TRBP	Trans-Activation-Responsive RNA-Binding Protein
TSS	Transkriptionsstartstelle
u. a.	unter anderem
UTR	untranslatierte Region (englisch: untranslated region)
VEGF	Vascular Endothelial Growth Factor
Y2H	Yeast Two-Hybrid
z.B.	zum Beispiel
17AAG	17-allylamino-geldanamycin

## 9.2 Verzeichnis der Abbildungen und Tabellen

### 9.2.1 Abbildungsverzeichnis

3.2	Überblick über die pflanzliche Immunabwehr .....	21
3.4	Biogenese von miRNAs .....	27
4.4	Überblick über die DrumPID-Datenbank .....	51
5.1A	Überblick über die entwickelte systembiologische Methode .....	53
5.1B	Darstellung der differentiell-exprimierten Gene der verwendeten Datensätze .....	55
5.1C	Übersicht der rekonstruierten expressions-basierten Netzwerke .....	56
5.1D	Überblick der rekonstruierten Immunnetzwerke .....	60
5.2.1A	Überblick der Analysen zum 2D-3D-Zellkulturnetzwerk .....	64
5.2.1B	HCC827-Netzwerk und <i>in silico</i> -Simulation der 17AAG-Behandlung im 3D-System .....	66
5.2.1C	Übersicht des rekonstruierten KRAS-Netzwerks und Mapping der zelllinien-spezifischen Mutationen .....	68
5.2.1D	Übersicht des rekonstruierten zelltyp-spezifischen H441-Netzwerks .....	69
5.2.1E	Übersicht des rekonstruierten zelltyp-spezifischen A549-Netzwerks .....	71
5.2.1F	A549-Netzwerk und <i>in silico</i> -Simulation im 3D-System .....	74
5.2.1G	H441-Netzwerk und <i>in silico</i> -Simulation im 3D-System .....	75
5.2.2A	Netzwerk und <i>in silico</i> -Simulation der Gefitinib-Resistenz in der HCC827-Zelllinie .....	78
5.2.2B	<i>In silico</i> -Simulation der Gefitinib-Resistenz in der HCC827-Zelllinie .....	79
5.2.2C	<i>In silico</i> -Simulation einer kombinierten PI3K- und MEK-Inhibierung in der HCC827-Zelllinie .....	80
5.2.2D	Netzwerk und <i>in silico</i> -Simulation der Gefitinib-Resistenz in der A549-Zelllinie .....	82
5.2.2E	<i>In silico</i> -Simulation der Gefitinib-Resistenz in der A549-Zelllinie .....	83
5.2.2F	Netzwerk der miRNA-21 in der HCC827-Zelllinie .....	85
5.2.2G	<i>In silico</i> -Simulation des Einflusses der miRNA-21 auf die Gefitinib-Resistenz in der HCC827-Zelllinie .....	86
5.3.1.1A	Überblick über die genomische Position, Sekundärstruktur und mRNA-Interaktion der lncRNA Chast .....	88
5.3.1.1B	Überblick der globalen Sequenz-Struktur-Konservierung für den Sequenzbereich 44 - 388 der lncRNA Chast .....	89
5.3.1.1C	Überblick der globalen Sequenz-Struktur-Konservierung für den Sequenzbereich 646 - 920 der lncRNA Chast .....	90
5.3.1.1D	Überblick der funktionellen Analyse der potentiellen Protein-Interaktions-Partner von Chast .....	93

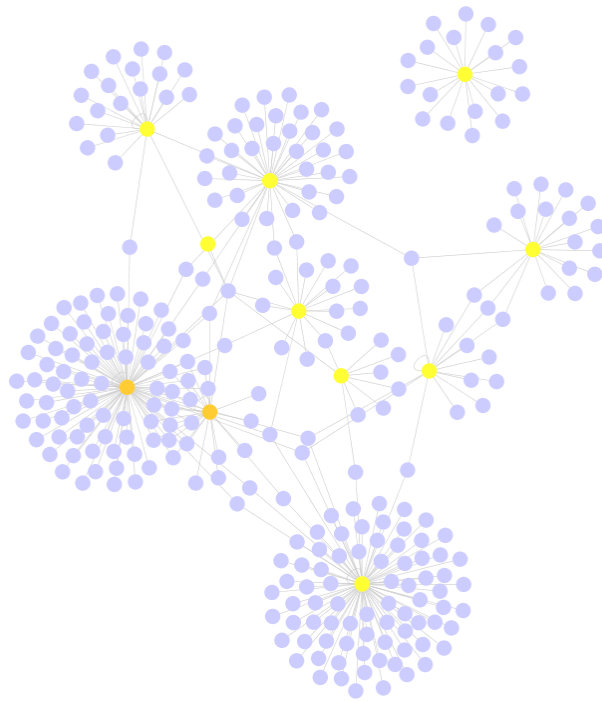
5.3.1.1E	Promotoranalyse der lncRNA Chast .....	96
5.3.1.1F	Darstellung der Überlappung zwischen dem Promotor von Chast und dem des Menschen .....	97
5.3.1.2	Überblick der Sequenz-Struktur-Konservierung der lncRNA MIR503HG ...	99
5.3.2A	Benutzeroberfläche und Suchkategorien der DrumPID-Datenbank .....	104
5.3.2B	Suchbeispiele mit der DrumPID-Datenbank .....	107
5.3.2C	Darstellung der funktionellen Bindestellen der Riboflavinkinase .....	108
Anhang A1	Übersicht über das zelltyp-spezifische A549-Netzwerk .....	165

## 9.2.2 Tabellenverzeichnis

4.2.1	Übersicht der Änderungen auf Apoptose und Proliferation nach 17AAG-Behandlung .....	40
5.1A	Überblick der Topologieanalyse für die rekonstruierten Netzwerke .....	57
5.1B	Übersicht der immunrelevanten biologischen Prozesse für <i>Pst</i> DC3000 vs. Mock .....	58
5.1C	Übersicht der immunrelevanten biologischen Prozesse für <i>Pst</i> COR <sup>-</sup> vs. Mock .....	59
5.1D	Übersicht der immunrelevanten biologischen Prozesse für <i>Pst</i> COR <sup>-</sup> <i>hrpS</i> vs. Mock .....	59
5.1E	Überblick der identifizierten funktionellen Hubproteine für die Immunabwehr .....	61
5.2.1A	Übersicht der ausgewählten differentiell-regulierten Proteine zwischen dem 2D- und 3D-System .....	63
5.2.1B	Übersicht der ausgewählten differentiell-regulierten Proteine zwischen A549 und H441 im 3D-System .....	67
5.3.1.1A	Sequenzhomologie der hypertrophie-assoziierten lncRNA Chast .....	87
5.3.1.1B	Überblick der gefundenen Protein-Interaktions-Partner für die lncRNA Chast .....	92
5.3.1.1C	Überblick der regulatorischen Elemente für die lncRNA Chast .....	92
5.3.1.1D	Überblick der herz- und hypertrophie-assoziierten miRNA-Interaktionspartner der lncRNA Chast .....	94
5.3.1.1E	Überblick der COG/KOGs-Suche für die lncRNA Chast .....	95
5.3.1.1F	Überblick über die potentiellen Promotorbereiche des humanen Homologs der lncRNA Chast .....	96
5.3.1.2	Sequenzhomologie der beiden hypoxie-induzierten lncRNAs LINC00323-003 und MIR503HG .....	98
5.3.2A	Überblick über die einzelnen Suchmöglichkeiten und der Ergebnisseite der DrumPID-Datenbank .....	105

5.3.2B	Sequenzalignment der Riboflavinkinase zwischen dem Menschen und <i>Aspergillus</i> .....	106
6.3.1.3	Überblick über die möglichen zelllinien-spezifischen Marker und potentiellen therapeutischen Drugtargets .....	123
Anhang T1	Übersicht der ausgewählten immunrelevanten biologischen Prozesse für <i>Pst</i> DC3000 vs. Mock .....	166
Anhang T2	Übersicht der ausgewählten immunrelevanten biologischen Prozesse für <i>Pst</i> COR <sup>-</sup> vs. Mock .....	168
Anhang T3	Liste der funktionellen Hubproteine in den Immunnetzwerken .....	171

### 9.3 Anhang der Abbildungen und Tabellen



**Abbildung Anhang A1: Übersicht über das zelltyp-spezifische A549-Netzwerk.** Das rekonstruierte zelltyp-spezifische A549-Netzwerk (neun A549-Mutationen ARID3A, ATR, CBL, CTNNA1, CTNND2, SMARCA4, SMYD2, STK11, TCEA2 in gelb; HSP90AA1 und HSP90AB1 in orange) verfügt über 302 Knoten und 341 Interaktionen (Interaktionspartner aus dem Interaktom in blau, Interaktionen in grau).

**Tabelle Anhang T1: Übersicht der ausgewählten immunrelevanten biologischen Prozesse für *Pst* DC3000 vs. Mock.** Die Tabelle fasst die ausgewählten 110 immunrelevanten Funktionen zusammen. (x=Anzahl der korrespondierenden Proteine für den biologischen Prozess; n=Anzahl aller annotierten Proteine in BiNGO; p-Value=Wahrscheinlichkeit, ein identisches Ergebnis zu erzielen.)

p-Value	x	n	Biologischer Prozess (GO)
9,41E-04	8	53	abscisic acid mediated signaling pathway
1,06E-03	5	21	auxin biosynthetic process
1,51E-04	7	31	auxin mediated signaling pathway
4,92E-09	15	59	cellular response to abiotic stimulus
9,41E-04	8	53	cellular response to abscisic acid stimulus
2,29E-04	7	33	cellular response to auxin stimulus
2,51E-08	11	33	cellular response to brassinosteroid stimulus
1,79E-25	65	361	cellular response to chemical stimulus
1,43E-05	9	39	cellular response to cytokinin stimulus
2,97E-23	52	254	cellular response to endogenous stimulus
1,12E-02	11	128	cellular response to external stimulus
1,12E-02	11	128	cellular response to extracellular stimulus
6,09E-04	6	28	cellular response to gibberellin stimulus
1,06E-19	45	227	cellular response to hormone stimulus
1,43E-05	9	39	cellular response to jasmonic acid stimulus
1,62E-07	12	47	cellular response to light stimulus
4,12E-25	61	323	cellular response to organic substance
1,62E-07	12	47	cellular response to radiation
2,82E-03	3	8	cellular response to red light
3,08E-07	11	41	cellular response to red or far red light
4,95E-04	6	27	cellular response to salicylic acid stimulus
2,51E-08	11	33	cellular response to steroid hormone stimulus
1,31E-22	90	729	cellular response to stimulus
1,43E-05	9	39	cytokinin mediated signaling pathway
7,09E-08	54	637	defense response
8,24E-09	27	193	defense response to bacterium
3,46E-03	5	27	defense response to bacterium, incompatible interaction
1,80E-04	14	117	defense response to fungus
4,11E-03	3	9	defense response to oomycetes
2,15E-04	13	105	defense response, incompatible interaction
1,32E-03	5	22	ethylene biosynthetic process
1,32E-03	5	22	ethylene metabolic process
6,09E-04	6	28	gibberellic acid mediated signaling pathway
1,78E-05	8	31	gibberellin mediated signaling pathway
6,09E-20	45	224	hormone-mediated signaling pathway
1,02E-02	19	272	immune response
8,58E-03	2	4	inflammatory response
1,19E-02	18	257	innate immune response
8,31E-04	5	20	jasmonic acid biosynthetic process

1,43E-05	9	39	jasmonic acid mediated signaling pathway
1,84E-05	7	23	jasmonic acid metabolic process
4,32E-05	11	66	negative regulation of response to stimulus
8,58E-03	2	4	positive regulation of gibberellic acid mediated signaling pathway
8,75E-04	9	65	regulation of defense response
4,40E-03	2	3	regulation of defense response to virus
1,51E-03	2	2	regulation of defense response to virus by host
1,22E-03	4	13	regulation of gibberellic acid mediated signaling pathway
4,40E-03	2	3	regulation of immune effector process
7,62E-03	3	11	regulation of response to biotic stimulus
9,51E-09	24	157	regulation of response to stimulus
8,53E-04	10	78	regulation of response to stress
1,51E-03	2	2	regulation of transcription in response to stress
9,25E-37	145	1168	response to abiotic stimulus
1,35E-10	36	272	response to abscisic acid stimulus
5,13E-09	34	282	response to auxin stimulus
3,25E-10	33	241	response to bacterium
5,11E-11	56	550	response to biotic stimulus
3,39E-07	12	50	response to blue light
8,31E-08	13	53	response to brassinosteroid stimulus
3,68E-05	26	279	response to cadmium ion
5,55E-04	17	177	response to carbohydrate stimulus
2,12E-49	203	1710	response to chemical stimulus
2,49E-07	28	241	response to cold
1,51E-03	2	2	response to continuous far red light stimulus by the high-irradiance response system
9,84E-05	11	72	response to cytokinin stimulus
4,47E-03	4	18	response to desiccation
4,04E-04	7	36	response to disaccharide stimulus
2,52E-46	134	835	response to endogenous stimulus
2,77E-05	16	125	response to ethylene stimulus
3,03E-07	27	229	response to external stimulus
8,48E-03	12	140	response to extracellular stimulus
2,77E-06	10	41	response to far red light
3,86E-08	23	156	response to fungus
1,51E-07	18	107	response to gibberellin stimulus
5,52E-06	10	44	response to gravity
5,80E-04	14	131	response to heat
1,06E-30	106	767	response to hormone stimulus
1,62E-07	41	434	response to inorganic substance
3,58E-03	4	17	response to insect
6,70E-23	40	148	response to jasmonic acid stimulus
5,59E-04	10	74	response to light intensity
1,00E-16	60	455	response to light stimulus
1,87E-05	31	350	response to metal ion
9,87E-03	3	12	response to oomycetes

3,14E-46	150	1037	response to organic substance
5,04E-12	47	388	response to osmotic stress
3,19E-11	55	528	response to other organism
2,71E-04	22	247	response to oxidative stress
1,64E-03	5	23	response to ozone
5,10E-16	60	471	response to radiation
1,07E-03	8	54	response to red light
5,21E-10	26	159	response to red or far red light
2,72E-07	20	135	response to salicylic acid stimulus
1,29E-12	46	360	response to salt stress
2,51E-08	11	33	response to steroid hormone stimulus
5,34E-55	304	3207	response to stimulus
2,13E-27	171	1853	response to stress
3,37E-04	7	35	response to sucrose stimulus
7,12E-10	41	359	response to temperature stimulus
3,57E-03	8	65	response to UV
4,18E-03	6	40	response to UV-B
1,51E-03	2	2	response to very low light intensity stimulus
6,84E-07	24	196	response to water
1,18E-06	23	188	response to water deprivation
7,92E-19	34	133	response to wounding
4,95E-04	6	27	salicylic acid mediated signaling pathway
7,62E-03	3	11	salicylic acid metabolic process
6,40E-03	5	31	systemic acquired resistance
8,70E-04	4	12	systemic acquired resistance, salicylic acid mediated signaling pathway
3,23E-03	8	64	two-component signal transduction system (phosphorelay)

**Tabelle Anhang T2: Übersicht der ausgewählten immunrelevanten biologischen Prozesse für *Pst* COR<sup>-</sup> vs. Mock.** Die Tabelle fasst die ausgewählten 82 immunrelevanten Funktionen zusammen. (x=Anzahl der korrespondierenden Proteine für den biologischen Prozess; n=Anzahl aller annotierten Proteine in BiNGO; p-Value=Wahrscheinlichkeit, ein identisches Ergebnis zu erzielen.)

p-Value	x	n	Biologischer Prozess (GO)
5,45E-04	4	21	auxin biosynthetic process
2,48E-03	4	31	auxin mediated signaling pathway
4,76E-03	4	37	auxin metabolic process
3,13E-03	4	33	cellular response to auxin stimulus
3,33E-04	5	33	cellular response to brassinosteroid stimulus
2,95E-13	32	361	cellular response to chemical stimulus
2,04E-12	26	254	cellular response to endogenous stimulus
2,74E-07	18	227	cellular response to hormone stimulus
3,10E-08	9	39	cellular response to jasmonic acid stimulus
8,36E-14	31	323	cellular response to organic substance
3,56E-03	3	17	cellular response to oxidative stress



6,79E-03	2	7	cellular response to oxygen radical
6,89E-03	4	41	cellular response to red or far red light
4,65E-07	7	27	cellular response to salicylic acid stimulus
3,33E-04	5	33	cellular response to steroid hormone stimulus
3,21E-10	41	729	cellular response to stimulus
6,79E-03	2	7	cellular response to superoxide
9,54E-06	29	637	defense response
2,32E-08	18	193	defense response to bacterium
1,46E-03	4	27	defense response to bacterium, incompatible interaction
1,18E-05	11	117	defense response to fungus
5,05E-04	5	36	defense response to fungus, incompatible interaction
7,66E-08	13	105	defense response, incompatible interaction
2,25E-07	18	224	hormone-mediated signaling pathway
6,04E-04	14	272	immune response
2,86E-04	15	283	immune system process
3,44E-04	14	257	innate immune response
3,10E-08	9	39	jasmonic acid mediated signaling pathway
3,56E-03	3	17	negative regulation of abscisic acid mediated signaling pathway
1,51E-05	5	18	negative regulation of defense response
3,39E-08	11	66	negative regulation of response to stimulus
2,48E-03	4	31	positive regulation of defense response
6,44E-04	6	57	positive regulation of response to stimulus
3,13E-03	4	33	regulation of abscisic acid mediated signaling pathway
2,87E-08	11	65	regulation of defense response
9,41E-06	7	41	regulation of immune response
9,41E-06	7	41	regulation of immune system process
6,59E-05	6	38	regulation of innate immune response
4,91E-03	2	6	regulation of plant-type hypersensitive response
9,38E-04	3	11	regulation of response to biotic stimulus
1,64E-12	21	157	regulation of response to stimulus
2,05E-08	12	78	regulation of response to stress
6,79E-03	2	7	regulation of systemic acquired resistance
7,24E-14	62	1168	response to abiotic stimulus
2,04E-07	20	272	response to abscisic acid stimulus
6,74E-03	12	282	response to auxin stimulus
9,47E-10	22	241	response to bacterium
1,40E-08	32	550	response to biotic stimulus
4,33E-04	6	53	response to brassinosteroid stimulus
4,91E-04	11	177	response to carbohydrate stimulus
3,31E-03	2	5	response to carbon dioxide
6,02E-21	92	1710	response to chemical stimulus
3,28E-03	7	104	response to chitin
3,00E-16	55	835	response to endogenous stimulus
1,16E-04	10	125	response to ethylene stimulus
6,89E-03	4	41	response to far red light
7,28E-06	13	156	response to fungus

1,73E-04	9	107	response to gibberellin stimulus
2,73E-12	46	767	response to hormone stimulus
4,93E-03	3	19	response to hypoxia
1,36E-03	18	434	response to inorganic substance
2,30E-04	4	17	response to insect
1,74E-08	16	148	response to jasmonic acid stimulus
4,65E-07	26	455	response to light stimulus
6,91E-04	3	10	response to molecule of bacterial origin
7,04E-20	68	1037	response to organic substance
1,31E-04	19	388	response to osmotic stress
5,30E-09	32	528	response to other organism
6,70E-03	11	247	response to oxidative stress
6,59E-03	3	21	response to oxygen levels
8,93E-07	26	471	response to radiation
4,37E-05	12	159	response to red or far red light
1,45E-06	13	135	response to salicylic acid stimulus
4,89E-05	19	360	response to salt stress
3,33E-04	5	33	response to steroid hormone stimulus
2,26E-26	145	3207	response to stimulus
9,33E-14	82	1853	response to stress
1,19E-03	16	359	response to temperature stimulus
6,31E-03	4	40	response to UV-B
1,91E-07	14	133	response to wounding
4,65E-07	7	27	salicylic acid mediated signaling pathway
1,23E-03	3	12	systemic acquired resistance, salicylic acid mediated signaling pathway

---

**Tabelle Anhang T3: Liste der funktionellen Hubproteine in den Immunnetzwerken.** Die Tabelle fasst die funktionellen Hubproteine ( $\geq 4$  IP) der Immunnetzwerke von *Pst* DC3000, *Pst* COR<sup>-</sup> und *Pst* COR<sup>-</sup> *hrpS* zusammen, anhand derer eine weitere Auswahl hinsichtlich einer immunrelevanten Funktion durchgeführt wurde.

### **Immunnetzwerk *Pst* DC3000**

Hubprotein	Name	Beschreibung
AT1G01010	NAC001	NAC domain-containing protein 1
AT1G01620	PIP1C	aquaporin PIP1-3
AT1G01910	AT1G01910	putative anion-transporting ATPase
AT1G02340	HFR1	transcription factor HFR1
AT1G03430	AHP5	histidine-containing phosphotransfer protein 5
AT1G04100	IAA10	auxin-responsive protein IAA10
AT1G04240	SHY2	auxin-responsive protein IAA3
AT1G04550	IAA12	AUX/IAA transcriptional regulator family protein
AT1G09530	PIF3	transcription factor PIF3
AT1G09570	PHYA	phytochrome A
AT1G10470	ARR4	two-component response regulator ARR4
AT1G11910	APA1	aspartic proteinase A1
AT1G12860	SCRM2	transcription factor SCREAM2
AT1G14920	GAI	DELLA protein GAI
AT1G15750	TPL	protein TOPLESS
AT1G17380	JAZ5	jasmonate-zim-domain protein 5
AT1G19180	JAZ1	jasmonate-zim-domain protein 1
AT1G19850	MP	auxin response factor 5
AT1G20950	AT1G20950	pyrophosphate--fructose 6-phosphate 1-phosphotransferase subunit alpha 1
AT1G22770	GI	gigantea protein (GI)
AT1G22920	CSN5A	COP9 signalosome complex subunit 5a
AT1G27300	AT1G27300	hypothetical protein
AT1G28480	GRX480	glutaredoxin-GRX480
AT1G30135	JAZ8	jasmonate-zim-domain protein 8
AT1G32230	RCD1	inactive poly [ADP-ribose] polymerase RCD1
AT1G32640	MYC2	transcription factor MYC2
AT1G48500	JAZ4	jasmonate-zim-domain protein 4
AT1G50420	SCL3	scarecrow-like protein 3
AT1G51950	IAA18	auxin-responsive protein IAA18
AT1G52830	IAA6	indole-3-acetic acid 6
AT1G52890	NAC019	NAC domain-containing protein 19
AT1G53510	MPK18	mitogen-activated protein kinase 18
AT1G56650	PAP1	transcription factor MYB75
AT1G58290	HEMA1	glutamyl-tRNA reductase 1
AT1G59940	ARR3	response regulator 3
AT1G63650	EGL3	transcription factor EGL1
AT1G64280	NPR1	Regulatory protein NPR1
AT1G66410	CAM1	calmodulin 1
AT1G69260	AFP1	ABI five binding protein
AT1G69600	ZFHD1	transcriptional factor ZFHD1
AT1G70410	BCA4	beta carbonic anhydrase 4
AT1G70700	TIFY7	protein TIFY 7
AT1G71230	CSN5B	COP9 signalosome complex subunit 5b
AT1G72450	JAZ6	jasmonate-zim-domain protein 6
AT1G73500	MKK9	MAP kinase kinase 9

AT1G74890	ARR15	two-component response regulator ARR15
AT1G74950	TIFY10B	protein TIFY 10B
AT1G75950	SKP1	S-phase kinase-associated protein 1
AT1G76670	AT1G76670	Nucleotide-sugar transporter family protein
AT1G77760	NIA1	nitrate reductase [NADH]
AT1G78300	GRF2	14-3-3-like protein GF14 omega
AT1G78570	RHM1	UDP-L-rhamnose synthase
AT1G80840	WRKY40	putative WRKY transcription factor 40
AT2G01570	RGA1	DELLA protein RGA
AT2G01760	RR14	two-component response regulator ARR14
AT2G01950	BRL2	serine/threonine-protein kinase BRI1-like 2
AT2G02810	UTR1	UDP-galactose transporter 1
AT2G02860	SUT2	sucrose transport protein SUC3
AT2G02950	PKS1	phytochrome kinase substrate 1
AT2G02990	RNS1	ribonuclease 1
AT2G18790	PHYB	phytochrome B
AT2G19480	NAP1;2	nucleosome assembly protein 1;2
AT2G20180	PIL5	transcription factor PIF1
AT2G22670	IAA8	auxin-responsive protein IAA8
AT2G22810	ACS4	1-aminocyclopropane-1-carboxylate synthase
AT2G30360	SIP4	CBL-interacting serine/threonine-protein kinase 11
AT2G30980	SKdZeta	Shaggy-related protein kinase zeta
AT2G31180	MYB14	myb domain protein 14
AT2G32950	COP1	E3 ubiquitin-protein ligase COP1
AT2G33310	IAA13	auxin-responsive protein IAA13
AT2G34650	PID	protein kinase PINOID
AT2G34690	ACD11	glycolipid transfer protein
AT2G35940	BLH1	BEL1-like homeodomain 1
AT2G37040	PAL1	phenylalanine ammonia-lyase 1
AT2G39800	P5CS1	delta1-pyrroline-5-carboxylate synthase 1
AT2G39940	COI1	coronatine-insensitive protein 1
AT2G40670	RR16	two-component response regulator ARR16
AT2G41070	EEL	ABI5-like transcription factor EEL
AT2G41090	AT2G41090	calmodulin-like protein 10
AT2G41100	TCH3	calmodulin-like protein 4
AT2G42580	TTL3	tetratricopeptide-repeat thioredoxin-like 3
AT2G42870	PAR1	phy rapidly regulated 1
AT2G43010	PIF4	transcription factor PIF4
AT2G43510	TI1	defensin-like protein 195
AT2G43790	MPK6	MAP kinase 6
AT2G44080	ARL	ARGOS-like protein
AT2G45960	PIP1B	aquaporin PIP1-2
AT2G46270	GBF3	G-box binding factor 3
AT2G47450	CAO	chloroplast signal recognition particle component CAO
AT3G02150	PTF1	transcription factor TCP13
AT3G04730	IAA16	auxin-responsive protein IAA16
AT3G05120	GID1A	putative gibberellin receptor GID1L1
AT3G07370	CHIP	E3 ubiquitin-protein ligase CHIP
AT3G13110	SERAT2;2	serine acetyltransferase 2;2
AT3G15360	TRX-M4	thioredoxin M4
AT3G15540	IAA19	auxin-responsive protein IAA19
AT3G16310	AT3G16310	mitotic phosphoprotein N' end (MPPN) family protein
AT3G16857	RR1	response regulator 1

AT3G17510	CIPK1	CBL-interacting serine/threonine-protein kinase 1
AT3G17860	JAZ3	jasmonate-zim-domain protein 3
AT3G20550	DDL	FHA domain-containing protein DDL
AT3G21510	AHP1	histidine-containing phosphotransmitter 1
AT3G23030	IAA2	auxin-responsive protein IAA2
AT3G23050	IAA7	auxin-responsive protein IAA7
AT3G27920	MYB0	trichome differentiation protein GL1
AT3G29350	AHP2	histidine-containing phosphotransfer protein 2
AT3G43810	CAM7	calmodulin 7
AT3G46090	ZAT7	zinc finger protein ZAT7
AT3G47620	TCP14	transcription factor TCP14
AT3G50060	MYB77	myb domain protein 77
AT3G51920	CAM9	calmodulin-like protein
AT3G53260	PAL2	phenylalanine ammonia-lyase 2
AT3G53900	UPP	uracil phosphoribosyltransferase
AT3G54220	SCR	protein scarecrow
AT3G57040	ARR9	two-component response regulator ARR9
AT3G57530	CPK32	calcium-dependent protein kinase 32
AT3G60600	VAP	vesicle associated protein
AT3G61060	PP2-A13	phloem protein 2-A13
AT3G62980	TIR1	protein TRANSPORT INHIBITOR RESPONSE 1
AT3G63010	GID1B	putative gibberellin receptor GID1L2
AT4G00430	PIP1;4	plasma membrane intrinsic protein 1;4
AT4G00480	ATMYC1	transcription factor MYC1
AT4G03190	GRH1	GRR1-like protein 1
AT4G09820	TT8	transcription factor TT8
AT4G14110	COP9	COP9 signalosome complex subunit 8
AT4G14560	IAA1	auxin-responsive protein IAA1
AT4G14640	CAM8	calmodulin-like protein 8
AT4G16760	ACX1	peroxisomal acyl-coenzyme A oxidase 1
AT4G17615	CBL1	calcineurin B-like protein 1
AT4G18710	BIN2	Shaggy-related protein kinase eta
AT4G19030	NLM1	aquaporin NIP1-1
AT4G19700	RING	SBP (S-ribonuclease binding protein) family protein
AT4G23650	CDPK6	calcium-dependent protein kinase 6
AT4G24210	SLY1	F-box protein GID2
AT4G26080	ABI1	protein phosphatase 2C 56
AT4G28640	IAA11	auxin-responsive protein IAA11
AT4G28720	YUC8	flavin-containing monooxygenase YUCCA8
AT4G29080	PAP2	phytochrome-associated protein 2
AT4G31800	WRKY18	WRKY DNA-binding protein 18
AT4G32880	HB-8	homeobox-leucine zipper protein ATHB-8
AT4G32980	ATH1	homeobox protein ATH1
AT4G33430	BAK1	Leu-rich receptor Serine/threonine protein kinase BAK1
AT4G34160	CYCD3;1	cyclin-D3-1
AT4G34590	GBF6	G-box binding factor 6
AT4G35450	AKR2	ankyrin repeat domain-containing protein 2
AT4G35580	NTL9	NAC transcription factor-like 9
AT4G36110	AT4G36110	SAUR-like auxin-responsive protein
AT4G37150	MES9	methyl esterase 9
AT4G37770	ACS8	1-aminocyclopropane-1-carboxylate synthase 8
AT4G37870	PCK1	phosphoenolpyruvate carboxykinase [ATP]
AT4G39400	BRI1	protein brassinosteroid insensitive 1

AT5G01810	CIPK15	CBL-interacting serine/threonine-protein kinase 15
AT5G01840	OFPI	ovate family protein 1
AT5G02200	FHL	far-red-elongated hypocotyl1-like protein
AT5G05410	DREB2A	dehydration-responsive element-binding protein 2A
AT5G10520	RBK1	receptor-like cytosolic serine/threonine-protein kinase RBK1
AT5G11060	KNAT4	homeobox protein knotted-1-like 4
AT5G11260	HY5	transcription factor HY5
AT5G13080	WRKY75	putative WRKY transcription factor 75
AT5G13220	JAZ10	jasmonate-zim-domain protein 10
AT5G13930	TT4	chalcone synthase
AT5G15330	SPX4	SPX domain-containing protein 4
AT5G15840	CO	zinc finger protein CONSTANS
AT5G17490	RGL3	DELLA protein RGL3
AT5G19000	BPM1	BTB-POZ and math domain-containing protein
AT5G20730	NPH4	auxin-regulated transcriptional activator NPH4
AT5G20900	JAZ12	jasmonate-zim-domain protein 12
AT5G21274	CAM6	calmodulin 6
AT5G22290	NAC089	membrane-tethered transcription factor ANAC089
AT5G24270	SOS3	calcineurin B-like protein 4
AT5G25220	KNAT3	homeobox protein knotted-1-like 3
AT5G25890	IAA28	auxin-responsive protein IAA28
AT5G35410	SOS2	CBL-interacting serine/threonine-protein kinase 24
AT5G35750	HK2	histidine kinase 2
AT5G37020	ARF8	auxin response factor 8
AT5G37670	AT5G37670	heat shock protein 15.7
AT5G39340	AHP3	histidine-containing phosphotransfer protein 3
AT5G42980	TRX3	thioredoxin H3
AT5G43700	ATAUX2-11	auxin-responsive protein IAA4
AT5G44790	RAN1	copper-transporting ATPase RAN1
AT5G47370	HAT2	homeobox-leucine zipper protein HAT2
AT5G57050	ABI2	protein phosphatase 2C 77
AT5G62920	ARR6	two-component response regulator ARR6
AT5G63310	NDPK2	nucleoside diphosphate kinase II
AT5G66190	FNR1	ferredoxin-NADP(+)-oxidoreductase 1

### **Immunnetzwerk Pst COR**

AT1G80840	WRKY40	putative WRKY transcription factor 40
AT2G32950	COP1	E3 ubiquitin-protein ligase COP1
AT1G77760	NIA1	nitrate reductase [NADH]
AT2G43790	ATMPK6	MAP kinase 6
AT1G52830	IAA6	indole-3-acetic acid 6
AT4G26070	MEK1	mitogen-activated protein kinase kinase 1
AT1G02340	HFR1	transcription factor HFR1
AT3G63010	ATGID1B	putative gibberellin receptor GID1L2
AT3G61060	AtPP2-A13	phloem protein 2-A13
AT3G53420	PIP2A	aquaporin PIP2-1
AT1G32640	ATMYC2	transcription factor MYC2
AT2G39940	COI1	coronatine-insensitive protein 1
AT3G12250	TGA6	transcription factor TGA6
AT3G17860	JAZ3	jasmonate-zim-domain protein 3
AT1G17380	JAZ5	jasmonate-zim-domain protein 5
AT1G21250	WAK1	wall-associated receptor kinase 1
AT3G20740	FIE	FERTILIZATION-INDEPENDENT ENDOSPERM

AT2G27100	SE	serrate RNA effector molecule
AT1G14920	GAI	DELLA protein GAI
AT3G02140	TMAC2	Ninja-family protein AFP4
AT3G11820	SYPI21	syntaxin-121
AT1G22920	CSN5A	COP9 signalosome complex subunit 5a
AT4G19700	RING	SBP (S-ribonuclease binding protein) family protein
AT5G17690	TFL2	chromo domain-containing protein LHP1
AT2G46830	CCA1	protein CCA1
AT1G15750	WSIP1	protein TOPLESS
AT2G41070	EEL	ABI5-like transcription factor EEL
AT1G01620	PIP1C	aquaporin PIP1-3
AT1G64280	NPR1	Regulatory protein NPR1
AT1G78570	RHM1	UDP-L-rhamnose synthase
AT5G13220	JAZ10	jasmonate-zim-domain protein 10
AT5G11060	KNAT4	homeobox protein knotted-1-like 4
AT3G17609	HYH	transcription factor HY5-like protein
AT1G73500	ATMKK9	MAP kinase kinase 9
AT4G14640	CAM8	calmodulin-like protein 8
AT5G42980	ATTRX3	thioredoxin H3
AT1G72450	JAZ6	jasmonate-zim-domain protein 6
AT1G70700	JAZ9	jasmonate-zim-domain protein 9
AT2G39010	PIP2E	plasma membrane intrinsic protein 2E
AT1G10210	ATMPK1	mitogen-activated protein kinase 1
AT4G28720	YUC8	flavin-containing monooxygenase YUCCA8
AT4G19030	NLM1	aquaporin NIP1-1
AT4G26080	ABI1	protein phosphatase 2C 56
AT1G02450	NIMIN1	protein NIM1-INTERACTING 1
AT3G50060	MYB77	myb domain protein 77
AT5G44790	RAN1	copper-transporting ATPase RAN1
AT1G20020	ATLFNR2	ferredoxin--NADP reductase, leaf isozyme 2
AT1G74950	JAZ2	jasmonate-zim-domain protein 2
AT5G15840	CO	zinc finger protein CONSTANS
AT1G28480	GRX480	glutaredoxin-GRX480
AT1G15080	ATPAP2	lipid phosphate phosphatase 2
AT1G30135	JAZ8	jasmonate-zim-domain protein 8
AT2G46270	GBF3	G-box binding factor 3
AT2G02860	SUT2	sucrose transport protein SUC3
AT1G19180	JAZ1	jasmonate-zim-domain protein 1
AT1G02090	FUS5	COP9 signalosome complex subunit 7
AT4G35580	NTL9	NAC transcription factor-like 9
AT2G42580	TTL3	tetratricopeptide-repeat thioredoxin-like 3
AT2G42870	PAR1	phy rapidly regulated 1
AT1G56650	PAP1	transcription factor MYB75
AT4G23650	CDPK6	calcium-dependent protein kinase 6
AT4G19660	NPR4	NPR1-like protein 4
AT1G51660	ATMKK4	mitogen-activated protein kinase kinase 4
AT1G70410	ATBCA4	beta carbonic anhydrase 4
AT3G62410	CP12-2	calvin cycle protein CP12-2
AT2G01760	ARR14	two-component response regulator ARR14
AT5G11260	HY5	transcription factor HY5
AT4G39400	BRI1	protein brassinosteroid insensitive 1
AT1G48500	JAZ4	jasmonate-zim-domain protein 4
AT1G78300	GRF2	14-3-3-like protein GF14 omega
AT2G35940	BLH1	BEL1-like homeodomain 1
AT2G22670	IAA8	auxin-responsive protein IAA8
AT5G06950	AHBP-1B	transcription factor TGA2

**Immunnetzwerk *Pst* COR<sup>-</sup> *hrpS***

AT1G56650	PAP1	production of anthocyanin pigment 1, MYB75
AT2G40220	ABI4	ethylene-responsive transcription factor ABI4
AT5G65420	CYCD4;1	cyclin-D4-1

---



## 9.4 Danksagung

Ich bedanke mich sehr herzlich bei meinem Betreuer Herrn Prof. Dr. Thomas Dandekar für die Überlassung der interessanten Projekte und die intensive Betreuung dieser Doktorarbeit. Einen besonderen Dank möchte ich an dieser Stelle auch für die großartige Unterstützung und Förderung aussprechen, aber auch die Möglichkeit, eigenständig zu Arbeiten und mich verwirklichen zu können.

Auch bedanke ich mich bei meiner Zweitgutachterin Frau Prof. Dr. Heike Walles für die gemeinsame Zusammenarbeit und die hilfreiche Unterstützung dieser Arbeit. Außerdem bedanke ich mich bei ihrer Arbeitsgruppe um Gudrun, Claudia und Sarah für die gelungene Zusammenarbeit, Unterstützung und hilfreichen Diskussionen. Ein besonderer Dank gilt vor allem Gudrun für die entgegengebrachte Unterstützung und wertvolle Hilfe.

Bei Janika, Jan und Prof. Dr. Dr. med. Thomas Thum bedanke ich mich ebenfalls für die gelungene Zusammenarbeit.

Auch bei der Arbeitsgruppe von Prof. Dr. Dandekar bedanke ich mich. Vor allem Liang und Naseem danke ich für die schöne Zeit und die vielen gemeinsamen Gespräche. Ein herzliches Dankeschön geht auch an Eva für die unkomplizierte Zusammenarbeit und die schönen Gespräche.

Ein besonderer Dank geht an meinen Bruder Danilo, aber auch Flo und Hagen für die sehr hilfreichen Korrekturen und nützlichen Ratschläge.

Insbesondere möchte ich mich sehr herzlich bei meinen Eltern und meinem Bruder Marco für die fortwährende Unterstützung und die schöne gemeinsame Zeit bedanken!

Mein besonderer Dank gilt meiner Frau Petra für die schönen gemeinsamen Jahre, aber insbesondere auch für die ständige Unterstützung und das fortwährende Verständnis!

## 9.5 Ehrenwörtliche Erklärung

### Erklärungen nach §4 Abs. 3 Satz 3, 5, 8 der Promotionsordnung der Fakultät für Biologie

Hiermit erkläre ich an Eides statt, die Dissertation

**„Systembiologische Analysen von Interaktionen: Zytokine (Pflanzenpathogene), 3D-Zellkulturen (Krebstherapie) und Drugtargets“**

eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

#### **Affidavit**

I hereby declare that my thesis entitled **„Systembiologische Analysen von Interaktionen: Zytokine (Pflanzenpathogene), 3D-Zellkulturen (Krebstherapie) und Drugtargets“** is the result of my own work.

I did not receive any help or support from commercial consultants. All sources and / or materials applied are listed and specified in the thesis.

Furthermore I verify that the thesis has not been submitted as part of another examination process neither in identical nor in similar form.

Würzburg, den

Unterschrift