

Aus der Klinik für Anaesthesiologie
Klinikum der Ludwig-Maximilians-Universität München
Direktor: Professor Dr. Bernhard Zwißler

Die Rolle intragenischer miRNA in der Regulation ihrer Host-Gene

Als kumulative Habilitationsschrift

Zur Erlangung des akademischen Grades eines habilitierten Doktors der Medizin an der
Ludwig-Maximilians-Universität München

Vorgelegt von
Ludwig Christian Giuseppe Hinske
(2017)

Hintergrund

Micro-RNAs (miRNAs) sind kleine, nicht-kodierende RNA-Sequenzen. Lee und Kollegen beschrieben in den 90er Jahren erstmals, dass das für die larvale Entwicklung notwendige Gen *lin-4* nicht in ein Protein translatiert wird, sondern dessen Transkript über basenkomplementäre Interaktion mit dem 3'-Ende des Transkripts des Gens *lin-14* dessen Translation epigenetisch negativ regulieren kann (R. C. Lee, Feinbaum, & Ambros, 1993). Die Entdeckung war revolutionär, da bis zu diesem Zeitpunkt davon ausgegangen wurde, dass nicht-translatierte RNA lediglich ein Abfallprodukt ohne relevante biologische Funktion sei. Erst einige Zeit später, im Jahr 2001, zeigten Lagos-Quintana und seine Kollegen, dass miRNAs in einer Vielzahl von Organismen nachweisbar sind, unter anderem in menschlichen Zellen. Außerdem beschrieben sie, dass miRNAs nicht nur organismus-, sondern auch gewebespezifisch exprimiert werden (Lagos-Quintana, Rauhut, Lendeckel, & Tuschl, 2001). Daraufhin stieg die Anzahl der Arbeiten, die sich mit der Rolle von miRNAs in der Pathogenese verschiedenster Erkrankungen beschäftigten, exponentiell an (Chan, Krichevsky, & Kosik, 2005; Hammond, 2006; Xie et al., 2005). Heutzutage ist man sich der zentralen Rolle von miRNAs als Regulatoren physiologischer Signalkaskaden bewusst (Ledderose et al., 2012; Martin et al., 2011; Tranter et al., 2011; Yan, Hao, Elton, Liu, & Ou, 2011).

Trotz intensiver Forschungsbemühungen sind allerdings viele Fragen im Bereich der miRNA-Forschung nicht ausreichend beantwortet. Bereits anhand der genomischen Lokalisation müssen drei Arten von miRNAs unterschieden werden. Erstens gibt es miRNAs, wie beispielsweise miR-21, die wie protein-kodierende Gene als solitäre

transkriptionelle Einheit vorliegen und somit eigenständig reguliert werden (Long et al., 2011). Weiterhin gibt es miRNAs, die als Polycistron vorliegen. Polycistrone sind Gencluster, die als ein Primärtranskript abgelesen werden. Dies ist insbesondere interessant, da Polycistrone im menschlichen Organismus untypisch sind (Baskerville & Bartel, 2005). Und als dritte Gruppe existieren intragenische miRNAs, d.h. miRNA-Gene, die innerhalb von proteinkodierenden, sogenannten Host-Genen, gelegen sind. Diese miRNA-Gene können zwar eigene regulatorische Elemente besitzen, wie beispielsweise miR-107 oder miR-126 (Monteys et al., 2010), sind aber häufig funktionell an die Expression ihrer Host-Gene gekoppelt (Rodriguez, Griffiths-Jones, Ashurst, & Bradley, 2004). Dieses Phänomen eröffnet einige wichtige Fragen: Ist diese Kopplung biologisch relevant? Wenn ja, wie kann diese untersucht und charakterisiert werden? Welche Rolle spielt eine solche Beziehung für mögliche pathogenetische Prozesse? Einige Arbeiten zeigen bereits, dass eine aufgehobene Kopplung intragenischer miRNAs und ihrer Host-Gene ein entscheidender Schritt in der Tumorphagenese sein könnte (Mayr & Bartel, 2009; Singh et al., 2009).

Das hier vorgestellte Habilitationsverfahren befasst sich deshalb mit der Untersuchung der funktionellen Beziehung intragenischer miRNAs zu ihren Host-Genen. Zur Bearbeitung dieser Thematik wurde ein breites Spektrum an bioinformatischen und molekularbiologischen Methoden etabliert und eingesetzt. Die Thematik wurde in drei Schritten bearbeitet:

1. Bioinformatische Grundlagen: Erstellung einer Datenbank und Evaluation bioinformatischer Methoden zur Untersuchung von Fragestellungen in Bezug auf intragenische miRNAs.

MiRNAs sind circa 20 Nukleotide kurze, einzelsträngige Nukleinsäuremoleküle. Sie binden an die nicht übersetzte 3'-RNA Region (3'-UTR) proteinkodierender Gene und reduzieren in Folge dessen deren Translation, entweder via translationaler Inhibition oder via mRNA Degradation (Baek et al., 2008; Lagos-Quintana et al., 2001; Lau, Lim, Weinstein, & Bartel, 2001; R. C. Lee & Ambros, 2001). Letzteres scheint dabei das dominierende Prinzip zu sein (Guo, Ingolia, Weissman, & Bartel, 2010). Der Mechanismus der Zielerkennung ist bis heute nicht vollständig verstanden und es existieren keine molekularbiologischen Methoden zur Large-Scale Detektion und Validierung von miRNA-Zielgen-Interaktionen. Dies ist nicht zuletzt dem Umstand geschuldet, dass eine miRNA mehrere hundert Zielgene haben und ein Gen von vielen miRNAs reguliert werden kann. Daher wurden diverse miRNA Zielvorhersage-Algorithmen entwickelt, die auf verschiedenen Prinzipien der Zielvorhersage beruhen. Einer der ältesten und am meisten benutzten Algorithmen ist TargetScan (Lewis, Burge, & Bartel, 2005; Lewis, Shih, Jones-Rhoades, Bartel, & Burge, 2003). TargetScan basiert auf dem Prinzip der sogenannten "Seed-Komplementarität". Als "Seed" einer miRNA bezeichnet man die Basen 2 - 8 (vom 5'-Ende der miRNA gezählt). Er ist der stärkste Prädiktor für die Zielgenerkennung einer miRNA (Brennecke, Stark, Russell, & Cohen, 2005). Andere Zielvorhersagealgorithmen nutzen die freie Bindungsenergie (Kertesz, Iovino, Unnerstall, Gaul, & Segal, 2007), Methoden der Mustererkennung

(Miranda et al., 2006) oder prädiktive Modelle basierend auf Transfektionsexperimenten (Krek et al., 2005; Wang & El Naqa, 2008), die jeweils in Sensitivität und Spezifität deutliche Unterschiede zeigen. Intragenische miRNAs können des Weiteren je nach Lokalisation in intronisch und exonisch untergliedert werden. Dabei muss allerdings die Strangspezifität, die Existenz möglicher alternativer Transkripte und das Vorhandensein möglicher miRNA-Promotoren berücksichtigt werden (Monteys et al., 2010). In einer ersten Arbeit entwickelten wir daher ein Datenbankmodell, in dem die verschiedenen Informationen aus unterschiedlichsten Quellen integriert wurden (L. C. Hinske, Heyn, Galante, Ohno-Machado, & Kreth, 2013). In diesem Modell werden die Informationen für proteinkodierende Gene des National Center for Biotechnology Information (NCBI) mit Transkript-Informationen aus der Reference Sequence Collection (RefSeq) verbunden (Pruitt, Tatusova, & Maglott, 2007). RefSeq ist eine kurierte, nicht-redundante Sammlung von Gentranskripten, die unter anderem durch den Genome Browser der University of California Santa Cruz zum Download bereitsteht. Das Mapping dieser Sequenzen auf das jeweils aktuelle menschliche Referenzgenom erlaubt insbesondere die Extraktion der 3'-UTR-Sequenz der jeweiligen Transkripte. Die jeweils aktuell registrierten miRNA-Gen-Koordinaten können von miRBase extrahiert werden (Griffiths-Jones, 2006; Griffiths-Jones, Grocock, van Dongen, Bateman, & Enright, 2006; Griffiths-Jones, Saini, van Dongen, & Enright, 2008). Dann müssen die miRNA-Gen-Koordinaten mit den RefSeq-Koordinaten verglichen werden, um miRNAs in intragenisch, und spezifischer in intronisch und exonisch klassifizieren zu können. Dabei definieren wir intragenische miRNAs als miRNAs, deren Genkoordinaten vollständig zwischen der Transkriptionsstartseite und dem Transkriptionsende liegen.

Exonisch sind miRNAs dann, wenn ein Teil des miRNA-Gens mit einem kodierenden Sequenzbereich überlappt. Zudem wurden Zielvorhersagealgorithmen integriert. Nach der entsprechenden Klassifikation der miRNAs wurden miRNA-Zielvorhersagen implementiert. Weiterhin wurden Vorhersagealgorithmen, Gen- und miRNA-Expressionsdaten und Protein-Interaktionsdaten integriert (L. C. Hinske et al., 2014; L. C. G. Hinske, Galante, Kuo, & Ohno-Machado, 2010). Abbildung 1 gibt eine grafische Übersicht über die verarbeiteten Datenquellen.

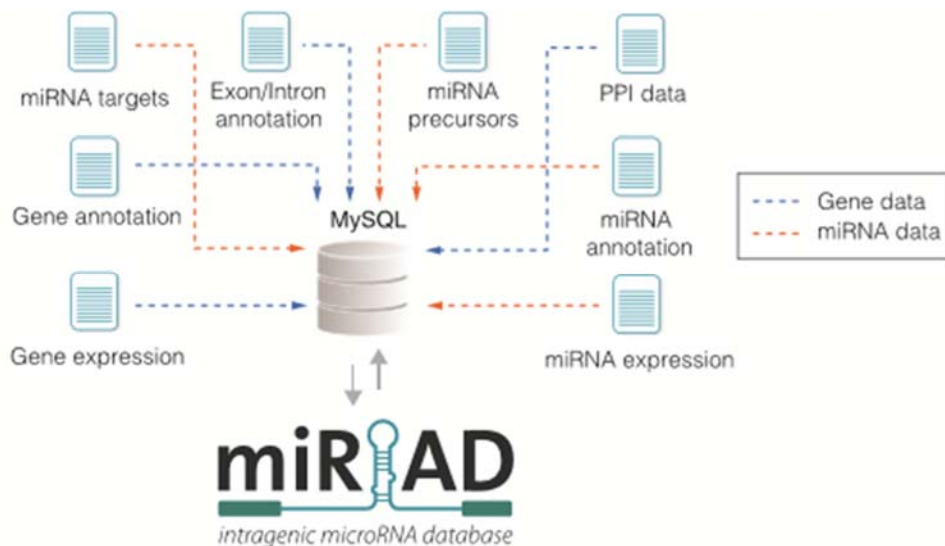


Abbildung 1. Übersicht über die verschiedenen Datenquellen, die in der von uns entwickelten Datenbank miRIAD integriert wurden (aus L.C. Hinske et al., 2014).

Zur besseren Benutzbarkeit dieser Datenbank haben wir in einer weiteren Studie eine Web-Applikation entwickelt (L. C. Hinske et al., 2014).

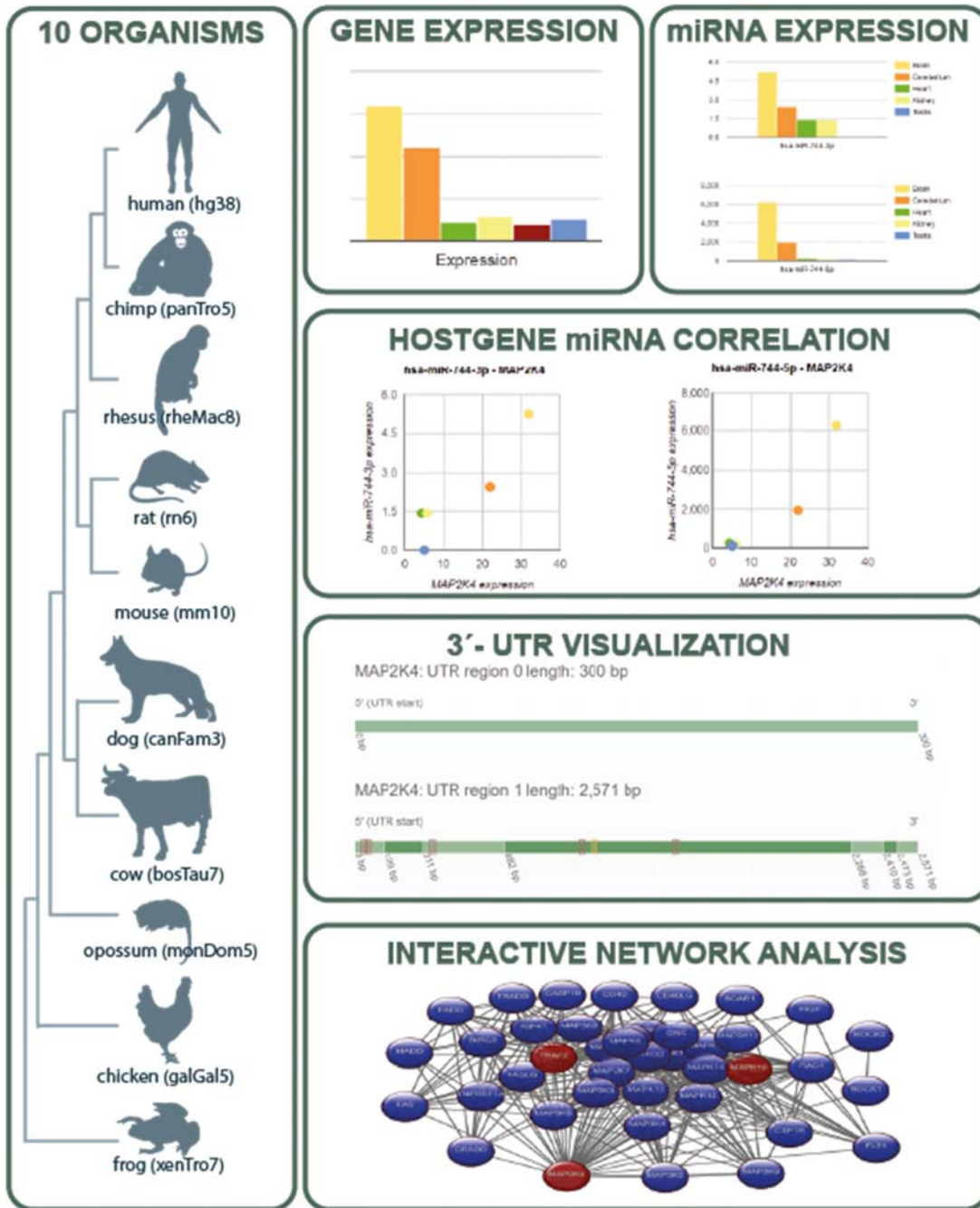


Abbildung 2. Übersicht über die Funktionalität der Web-Applikation zur effektiven Nutzung der miRIAD-Datenbank (aus L.C. Hinske et al., 2017).

Diese Software dient gezielt der Untersuchung der Rolle intronischer miRNAs und ihrer Host-Gene und erlaubte verschiedenste Analysen (Abbildung 2). Die Oberfläche ist

einfach aufgebaut und besteht vornehmlich aus einem Suchfeld. Dort kann sowohl nach Genen als auch nach miRNAs gesucht werden, einzeln oder als Liste. In dem resultierenden Suchergebnis werden übersichtlich alle proteinkodierenden Gene und/oder miRNAs dargestellt. Diejenigen Gene, die eine intragenische miRNA beinhalten beziehungsweise die miRNAs, die intragenisch gelegen sind, sind gekennzeichnet. Zudem kann auch die Liste aller intragenischen miRNAs oder aller Host-Gene aufgerufen werden.

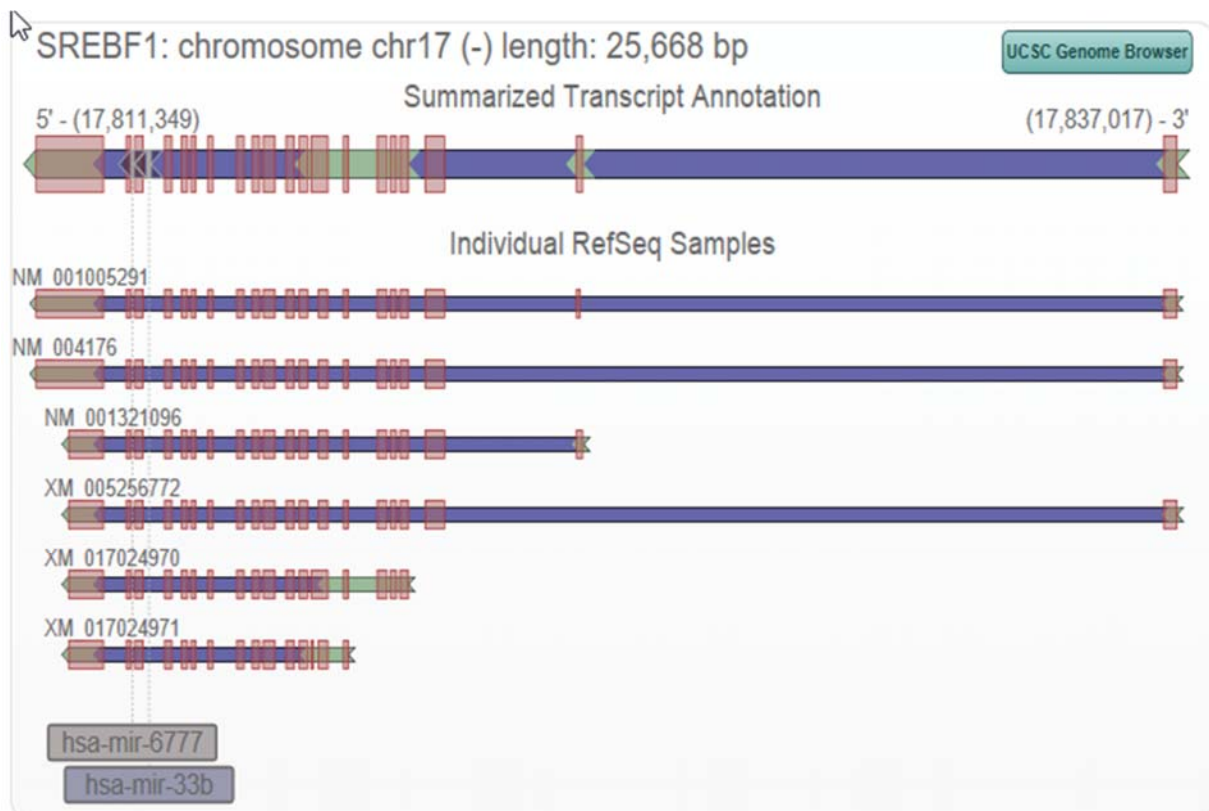


Abbildung 3. Visualisierung des Genmodells. Die Leserichtung ist durch die Pfeilrichtung gekennzeichnet, exonische Bereiche entsprechen den roten Kästchen, intronische Sequenzen den blauen und UTR-Sequenzen den grünen Bereichen (aus L.C. Hinske et al., 2017).

In der darauf folgenden Einzelansicht werden detaillierte Informationen über das Gen und die entsprechende miRNA angezeigt (Abbildung 3). Die Gen-Ansicht beginnt mit einer kurzen Beschreibung. Darunter befindet sich die Visualisierung des Genmodells, in der die Leserichtung, exonische, intronische und nicht-translatierte Bereiche gekennzeichnet sind. Auch mögliche intronische miRNAs werden mit Leserichtung dargestellt. Oben in der Grafik befindet sich die summative Darstellung, die die Informationen aus den darunter dargestellten Einzeltranskripten zusammenfasst. Diese Übersicht erlaubt die schnelle Erfassung von Informationen, wie der Distanz der intronischen miRNA zum nächsten Exon, des Vorhandenseins der miRNA in verschiedenen Isoformen des Gens sowie der Leserichtung der miRNA verglichen mit ihrem Host-Gen. MiRNAs, die laut Prädiktionsmodell ihr eigenes Host-Gen regulieren, werden blau dargestellt.

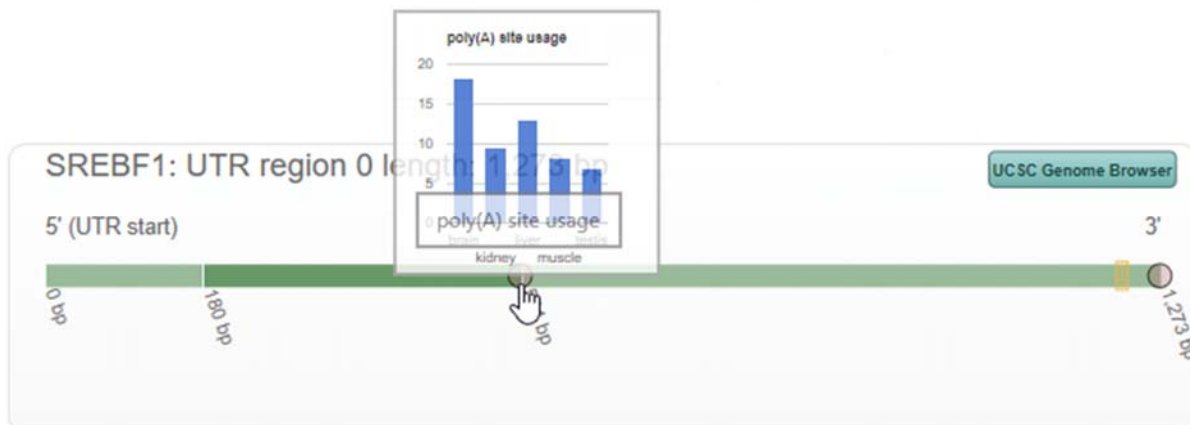


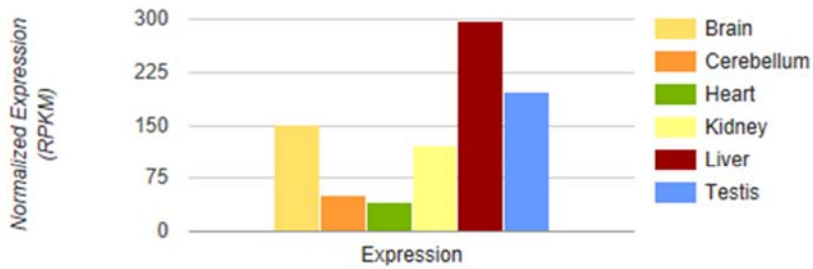
Abbildung 4. Visualisierung der 3'-UTR Region. Alternative Polyadenylierungsisoformen sind durch alternierende Grünbereiche gekennzeichnet, die mögliche miRNA-Bindungsstelle durch ein gelbes Kästchen, Informationen über die gewebeabhängige Verteilung der APA-Isoformen durch die Kreise (aus L.C. Hinske et al., 2017).

Der Genmodellgrafik folgt die gesonderte Darstellung des für die Genregulation durch miRNAs wichtigen 3'-UTRs (Abbildung 4). Diese für das Transkript besondere Region kann ebenfalls in verschiedenen Isoformen vorkommen, was man als alternative Polyadenylierung (APA) bezeichnet (Di Giammartino, Nishida, & Manley, 2011). Dieser Mechanismus ist mitunter für gewebespezifische miRNA-Regulation verantwortlich und scheint insbesondere bei der Aktivierung von Tumorgenen eine zentrale Rolle zu spielen (Mayr & Bartel, 2009; Zhang, Lee, & Tian, 2005).

Visuell werden diese alternativen Polyadenylierungsformen durch alternierende Farben gekennzeichnet. Um alternative Polyadenylierungsmuster darstellen zu können, haben wir einerseits die Daten von Derti et al. prozessiert und integriert, die für mehrere Spezies und Gewebe speziell poly(A)-Sequenzen isoliert und mithilfe von Next-Generation Sequencing sequenziert haben (Derti et al., 2012). Zum anderen haben wir einen Algorithmus entwickelt, um potentielle alternative poly(A) Varianten aus regulären RNA-Sequencing Daten zu extrahieren. Basierend auf dem Next-Generation Sequencing Datensatz von Brawand und Kollegen (Brawand et al., 2011) haben wir nach entsprechendem Mapping der Read-Sequenzen auf das jeweilige Referenzgenom nach Sequenzen gesucht, die mindestens vier aufeinander folgende Adenosin-Nukleotide ohne Korrelat im Referenzgenom aufweisen. Die entsprechende Stelle wurde nur dann als alternative Polyadenylierungsstelle in unsere Datenbank eingefügt, wenn sie von mindestens zwei nicht identischen Reads gestützt wurde. APA-Stellen, für die die Expression in verschiedenen Geweben quantifizierbar war, erscheinen in der Visualisierung als kleine Kreise. Durch Anklicken erscheint das Expressions-Balkendiagramm für diese Stelle. Zudem können vorhergesagte miRNA-

Bindungsstellen angezeigt werden, zusammen mit der dazugehörigen UTR-Sequenz und der relativen Position innerhalb des UTRs.

Gene Expression



gene	tissue	expression
SREBF1	Brain	151.3
SREBF1	Cerebellum	50.77
SREBF1	Heart	41.59
SREBF1	Kidney	121.53
SREBF1	Liver	296.38
SREBF1	Testis	196.6

Expression correlation between Host and miRNA

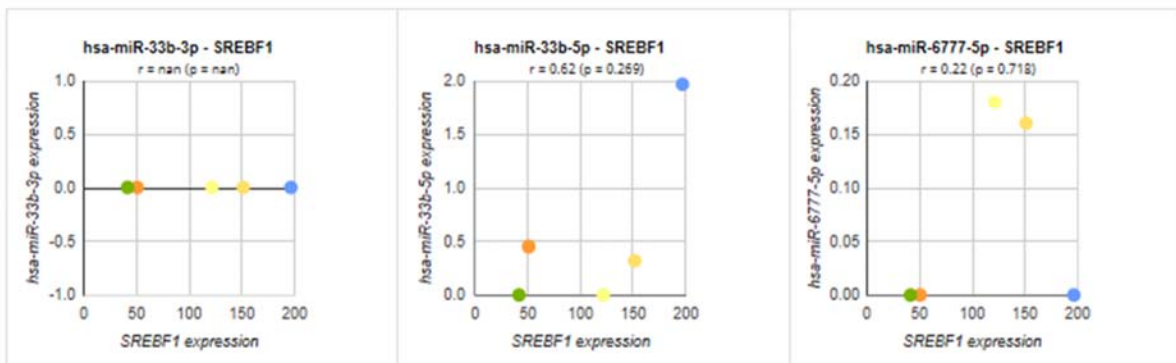


Abbildung 5. Darstellung der Genexpression in verschiedenen Geweben, sowie die Korrelation zwischen Host-Gen und den intronisch gelegenen miRNAs (aus L.C. Hinske et al., 2017).

Für einen Großteil der aktuell zehn in unserer Datenbank implementierten Spezies sind Expressionsdaten sowohl von miRNAs als auch von mRNA für die verschiedenen Gewebe enthalten. Insbesondere für die Beurteilung einer möglichen transkriptionellen Koregulation ist die Korrelation der einzelnen intragenischen miRNAs und ihrer Host-Gene relevant. Expressions- und Korrelationsdaten werden in unserer Applikation sowohl grafisch als auch tabellarisch dargestellt (Abbildung 5).

Da eine miRNA nicht nur ein einzelnes Gen reguliert, haben wir zudem einen Algorithmus entwickelt, um den Effekt einer miRNA auf ein Netzwerk von Genen zu quantifizieren. Dieses Netzwerk kann entweder vom Benutzer selbst definiert werden, beispielsweise als beobachtete Gen-Signatur für einen bestimmten Krankheitszustand. Da in unsere Datenbank aber auch Protein-Protein-Interaktionsdatenbanken eingebunden wurden, können miRNAs gesucht werden, die nicht nur das Zielgen selbst, sondern das mit dem Zielgen interagierende Netzwerk regulieren können. Für jede miRNA wird ein Score zu dem entsprechenden Netzwerk erstellt.

Zuerst wird die Wahrscheinlichkeit eines zufälligen Auftretens der Seed-Sequenz einer miRNA berechnet:

$$p(S) = \prod_{i=1}^n p(N_i|D), \text{ wobei}$$

S = Seed-Sequenz

n = Länge(S)

N_i = i-tes Nukleotid of S

D = Nukleotid-Verteilung.

Die Wahrscheinlichkeit, dass diese Sequenz mindestens r mal innerhalb einer zufälligen Sequenz der Länge N (UTR-Sequenz eines jeden Gens innerhalb eines Interaktionsnetzwerks) auftritt, ist gegeben durch:

$$p(x_t) = \left(1 - \sum_{i=0}^{r-1} \binom{L_x}{i} p(S)^i (1 - p(S))^{L_x-i}\right), \text{ wobei}$$

L_x = (Länge des 3'-UTRs von Element x_t) - (Länge der Seed-Sequenz n) + 1

r = erwartetes Auftreten der Sequenz (in unserer Anwendung wurde r=1 gesetzt)

Somit kann die zufällig erwartete Anzahl an Genen mit Seed-Site komplementärer Sequenz E(X_t) innerhalb des Netzwerks X durch die Summe der Einzelwahrscheinlichkeiten x errechnet werden:

$$E(X_t) = \sum_{x_t \in X} p(x_t)$$

Die erwartete Anzahl kann dann mit der beobachteten Anzahl verglichen und statistisch mit der Log-Odds Ratio quantifiziert werden:

$$\text{Score}(X|S) = \log\left(\frac{\frac{E(X_t) + O(X_t)}{E(X_t)} * \frac{E(X_n) + O(X_n)}{E(X_n)}}{\frac{|X|}{E(n)}}\right)$$

Zusammengefasst haben wir ein Webtool entwickelt, das eine Plattform zur Untersuchung miRNA-bezogener Fragestellungen bietet und die Visualisierung komplexer Zusammenhänge zwischen den verschiedenen Datensätzen erlaubt.

2. Die Beziehung zwischen Host-Gen und miRNA: Bioinformatische Evidenz für eine funktionelle Beziehung zwischen intronischen miRNAs und ihren Host-Genen

Intronische miRNAs werden zusammen mit ihrem Host-Gen als Primärtranskript exprimiert und vor der Splicing-Reaktion aus dem Transkript extrahiert (Kim & Kim, 2007). Danach erfolgt in mehreren Prozessierungsschritten die Reifung zur aktiven miRNA (Denli, Tops, Plasterk, Ketting, & Hannon, 2004; Han et al., 2004; Y. Lee et al., 2003). So reizvoll die Annahme einer funktionellen Beziehung zwischen Host-Gen und miRNA auch sein mag, könnte die Kolo-kalisation ebenso ein stochastisches Phänomen oder lediglich eine Informationskompression auf der DNA sein. Daher haben wir untersucht, ob es Evidenz für eine funktionelle Beziehung zwischen miRNAs und ihren

Host-Genen gibt (L. C. G. Hinske et al., 2010). Dazu haben wir die oben beschriebene Datenbank benutzt, um zuerst im Rahmen einer strukturellen Analyse genomische Charakteristika intronischer miRNAs zu extrahieren. Es zeigte sich, dass Host-Gene mit intronischen miRNAs insgesamt ca. dreimal länger als protein-kodierende Gene ohne miRNAs sind. Aber nicht nur die Gene an sich, sondern auch die entsprechenden 3'-UTR Sequenzen sind länger und enthalten mehr AU-reiche Regionen. AU-reiche Regionen wiederum sind maßgeblich für die Transkriptstabilität verantwortlich und eine Häufung in der UTR-Sequenz mit schnellerem Abbau und engmaschiger epigenetischer Kontrolle assoziiert (Jing et al., 2005). Danach untersuchten wir die Position intronischer miRNAs innerhalb der Gene. Unsere Hypothese war, dass im Falle eines funktionellen Zusammenhangs ein Positionsbias intronischer miRNAs zugunsten des 5'-Endes der Gene existieren müsste, damit möglichst viele alternative Transkripte mit der miRNA koexprimiert würden. Tatsächlich befinden sich ca. 60% aller intronischen miRNAs in den ersten fünf Introns und zeigen einen starken Leserichtungsbias: Die Leserichtung der intronischen miRNA und des Host-Gens ist deutlich häufiger dieselbe, als es durch Zufall erklärbar wäre. Im nächsten Schritt führten wir eine funktionelle Analyse durch, um die Hypothese eines funktionellen Zusammenhangs näher zu beleuchten. Zuerst zeigten wir, dass die Anzahl von Host-Genen, die laut Vorhersage von ihrer intragenischen miRNA reguliert werden, signifikant höher ist, als durch Zufall erklärbar (negative Regulation erster Ordnung). Dieses Ergebnis war konstant für alle in der Analyse benutzten Zielvorhersage-Algorithmen. Zur Beurteilung der Wahrscheinlichkeit einer Regulation auf höherer Ebene nutzten wir die Genkarten der Kyoto Encyclopedia of Genes and Genomes (KEGG), in der die Interaktionen zwischen Genen bezogen auf

bekannte Signalkaskaden kodiert sind (Kanehisa & Goto, 2000). Dann ermittelten wir für jedes Host-Gen, die intragenische miRNA und den dazugehörigen Signalweg die Anzahl der in der entsprechenden Signalkaskade vorhergesagten Ziel-Gene. Um die so ermittelte Anzahl statistisch bewerten zu können, führten wir pro Host-Gen 1000 Simulationen durch, in dem die Gene innerhalb der Kaskade durch zufällig ausgewählte Gene ersetzt wurden und erneut die Anzahl möglicher Ziel-Gene ermittelt wurde. Wir fanden eine hoch-signifikante Anreicherung von Zielgenen innerhalb der Host-Gen Signalkaskade. Basierend auf diesen Ergebnissen haben wir die Hypothese aufgestellt, dass ein möglicher Aspekt der Beziehung zwischen intronischen miRNAs und ihren Host-Genen die Verhinderung überschießender Expression im Sinne einer negativen Rückkopplung ist. Dabei haben wir ein mögliches Feedback erster Ordnung, das heißt die miRNA reguliert direkt ihr eigenes Host-Gen, von einem Feedback höherer Ordnung, bei dem die miRNA die Transkription (beispielsweise via Regulation eines Transkriptionsfaktors) oder den funktionellen Zustand (beispielsweise via Regulation einer Kinase) ihres Host-Gens beeinflusst, unterschieden (Abbildung 6).

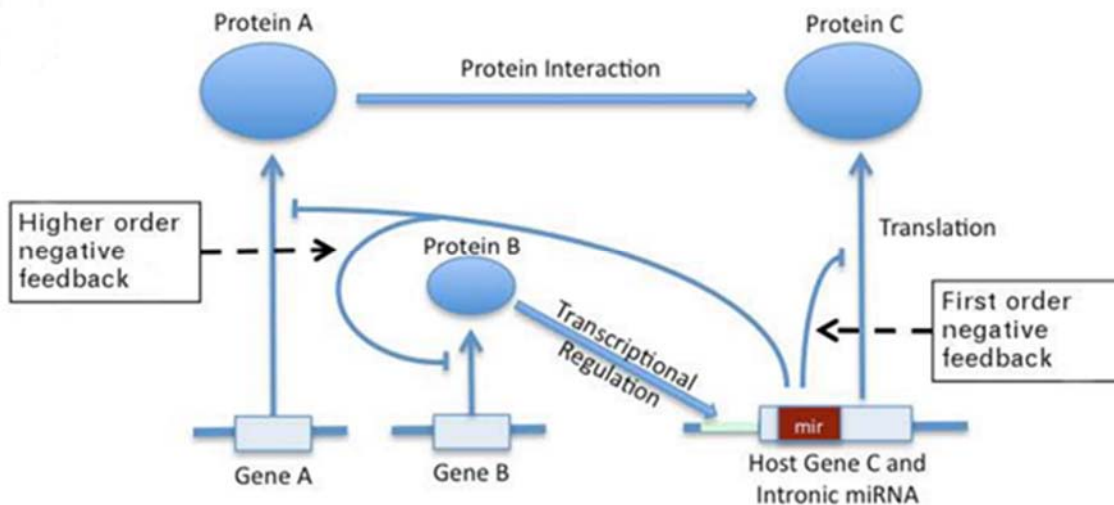


Abbildung 6. Modell der negativen Rückkopplung intronischer miRNAs auf die Signalkaskade der Host-Gene (aus L.C.G. Hinske et al., 2010).

3. Direkte und funktionelle negative Rückkopplung: Alternative Polyadenylierung als Mechanismus zur Rückkopplungsadjustierung

Eine Variante einer negativen Rückkopplung einer miRNA auf ihr Host-Gen ist die direkte Regulation (L. C. G. Hinske et al., 2010). Diese Form der Rückkopplung ist allerdings nur dann sinnvoll, wenn es Mechanismen gibt, mit Hilfe derer diese Beziehung an- und abgeschaltet werden kann. Dill und Kollegen haben erstmals anhand von miR-26b ein Beispiel einer intronischen miRNA zeigen können, die erst im Verlauf des zellulären Differenzierungsprozesses überhaupt in die biologisch aktive Form übersetzt wird und somit ihr Host-Gen reguliert (Dill, Linder, Fehr, & Fischer, 2012). Während die differenzielle Prozessierung intronischer miRNAs einen möglichen Weg der Regulation einer direkten Rückkopplung darstellt, wäre ein dynamischerer Prozess die Modifikation der Länge des 3'-UTR via alternativer Polyadenylierung.

Erreicht wird dies durch das Vorhandensein alternativer Polyadenylierungssignale. Diese wiederum werden unterteilt in "starke" und "schwache" Signale, die von einem Komplex verschiedener Proteine erkannt werden (Beaudoing, Freier, Wyatt, Claverie, & Gautheret, 2000; Di Giammartino et al., 2011). Wir haben daher die 3'-UTR-Sequenzen von Host-Genen mit Seed-komplementärer Sequenz für die eigene intronische miRNA mit denen ohne Seed-komplementäre Sequenz verglichen (L. C. Hinske et al., 2015). Dabei zeigt sich, dass erstere längere 3'-UTR Sequenzen mit mehr alternativen Polyadenylierungssignalen besitzen. Während sich bei Genen ohne intronische miRNA und Host-Genen ohne Seed-komplementäre Sequenz die "starken" Polyadenylierungssignale AAUAAA und AUUAAA zumeist am 3'-Ende des 3'-UTRs befanden, lagen diese bei Host-Genen mit Seed-komplementärer Sequenz präferenziell vor dieser Sequenz. Zudem zeigte sich, dass miRNAs, die potentiell ihr eigenes Host-Gen regulieren, auch vermehrt Transkripte von Genen regulieren, die für den Polyadenylierungsapparat wichtige Proteine kodieren, insbesondere *CPSF2* (Cleavage and Polyadenylation Specific Factor 2). *CPSF2* wiederum ist bereits mit der Veränderung der Signalerkennung in Verbindung gebracht worden (Herr, Molnàr, Jones, & Baulcombe, 2006; Kolev, Yario, Benson, & Steitz, 2008). Wir haben daher U87-Zellen nach Transfektion mit siCPSF2 sequenziert. Bei 97%-iger Reduktion von CPSF2-RNA zeigten sich, wie vorbeschrieben, grundsätzlich verlängerte UTR-Sequenzen, außer bei Host-Genen mit Seed-komplementärer Sequenz. Diese wurden tendenziell kürzer. Mittels Motiv-Detektionsanalyse konnten wir zudem zeigen, dass in den 3'-UTR Sequenzblöcken, die eine besonders starke Expressionszunahme verzeichneten, die "starken" Polyadenylierungssignale angereichert waren.

Zusammengefasst bedeutet dies, dass intronische miRNAs, die ihren eigenen Host regulieren, mitunter die eigene Wirkung auf das Host-Gen begrenzen.

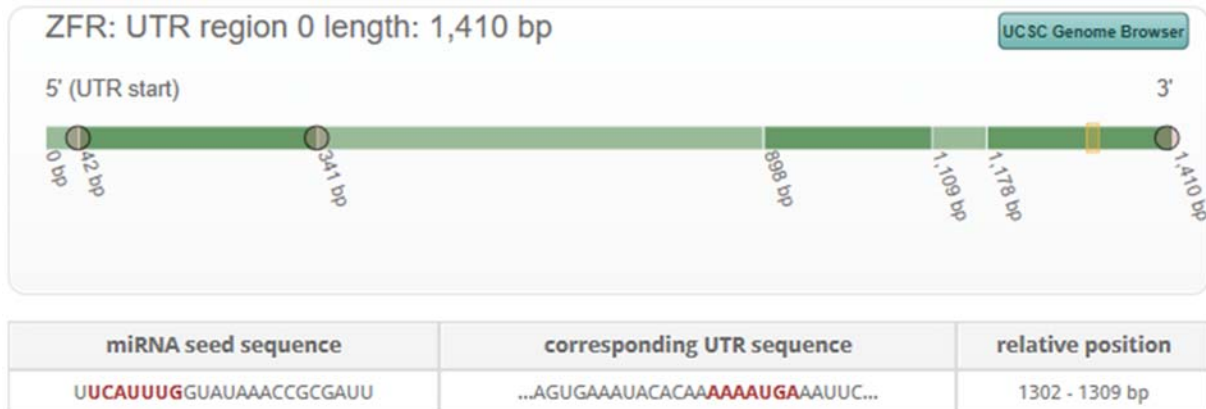


Abbildung 7. Die 3'-UTR von ZFR. Die Bindungsstelle für miR-579 ist als gelbes Kästchen dargestellt, die Bindungssequenz im unteren Bildbereich (aus L.C. Hinske et al., 2017).

Wir haben diese Hypothese an einem Beispiel evaluiert. Das Gen *ZFR* (Zinc Finger Recombinase) besitzt eine intronische miRNA, hsa-miR-579, sowie in der 3'-UTR Region eine Seed-komplementäre Sequenz für diese miRNA (Abbildung 7). Mittels Luciferase-Assay und im Western-Blot konnten wir zeigen, dass hsa-miR-579 sowohl *ZFR*, als auch CPSF2 direkt reguliert. Des Weiteren entdeckten wir mehrere potentielle Polyadenylierungsstellen in der 3'-UTR, die wir mittels 3'-RACE (Rapid Amplification of cDNA Ends) validierten. Wir konnten zeigen, dass von der Regulation lediglich die längste Transkriptvariante betroffen ist. Zusammenfassend konnten wir zeigen, dass eine direkte negative Rückkopplung durch intronische miRNAs nicht nur existiert, sondern durch alternative Polyadenylierung reguliert werden kann (Abbildung 8).

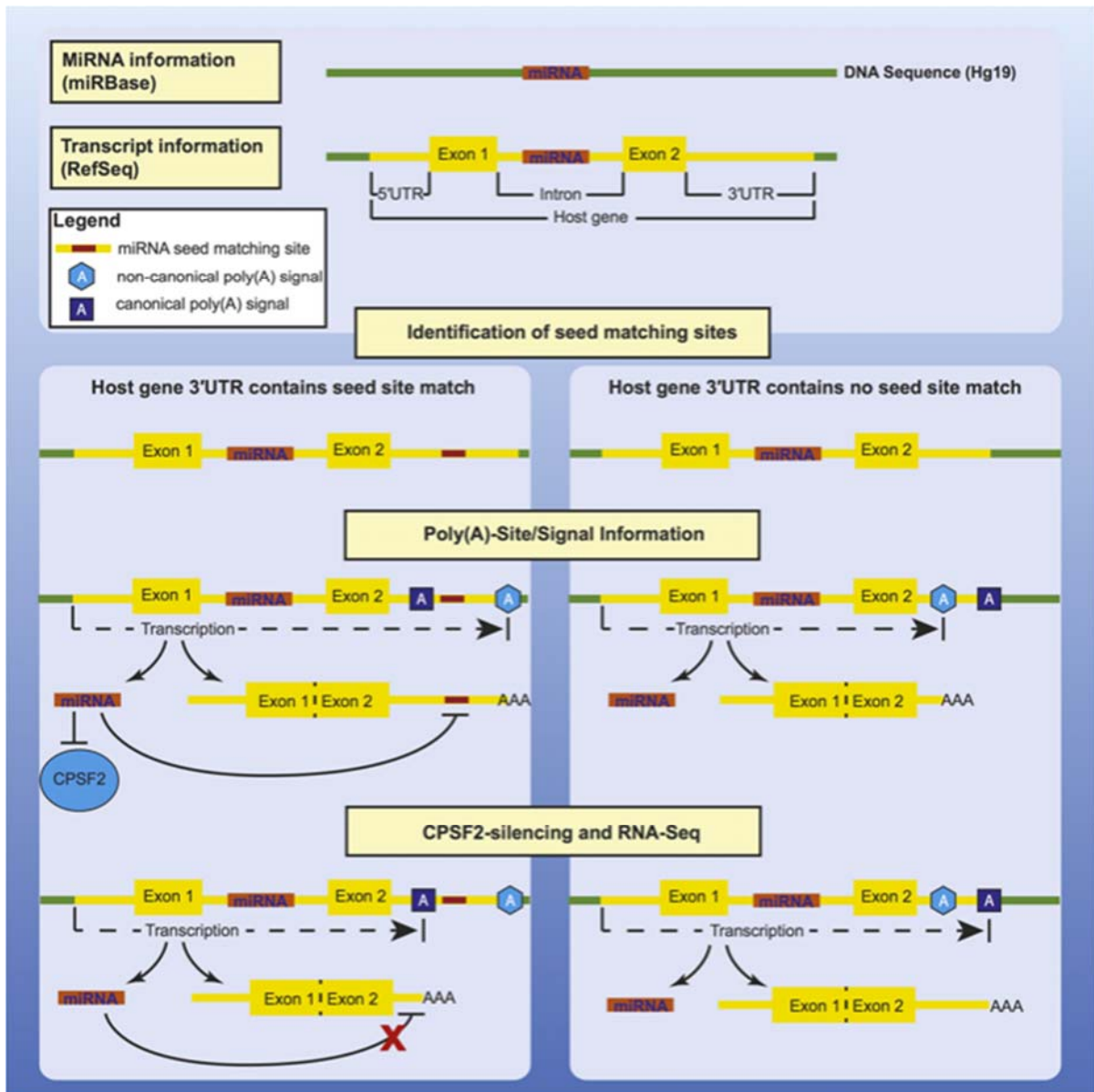


Abbildung 8. Modell der differentiellen Regulation der direkten Regulation von Host-Genen durch ihre intronischen miRNAs (aus L.C. Hinske et al., 2015).

Ein Großteil der Host-Genen besitzt keine seed-komplementäre Sequenz in der 3'-UTR, was eine mögliche indirekte Regulation nahe legen könnte. Diese Beziehung intragenischer miRNAs zu ihren Host-Genen scheint mitunter sogar von größerer klinischer Relevanz zu sein. Bereits im Jahr 2010 publizierten Tie und Kollegen, dass

die intronische miRNA hsa-miR-218, die im menschlichen Genom sowohl in *SLIT2* als auch in *SLIT3* intronisch vorkommt, mit ihrem Host-Gen *SLIT3* koexprimiert wird (Tie et al., 2010). Die Bindung von SLIT an den ROBO1-Rezeptor ist wiederum wichtig für die Invasion und Metastasierung kolorektaler Karzinome. Die Autoren zeigten, dass die ROBO1-expression von miR-218 reprimiert wird und dass in einer hochinvasiven Tumorzelllinie diese Regulation durch Silencing von *SLIT3* und miR-218 entfällt. Die Autoren konnten zudem zeigen, dass eine erniedrigte miR-218-Expression mit fortgeschrittenem Tumorstadium, lymphatischer Metastasierung, sowie schlechter Prognose korreliert (Tie et al., 2010). Kürzlich erst konnten Schmitt et al. zeigen, dass die miR-4728, die intronisch in dem insbesondere für das Mamma-Karzinom wichtigen Rezeptor ERBB2/HER2 des MAPK-Signalwegs gelegen ist, ein negativer Regulator dieser Signalkaskade ist (Schmitt et al., 2015). Die Autoren fanden auch hier einen ausgeprägten Zusammenhang zwischen der Expression der miRNA und der Überlebensraten der Patienten.

Ein weiteres, für viele Tumoren sehr zentrales Gen, ist AKT2 (AKT Serine/Threonine Kinase 2) (Agarwal, Brattain, & Chowdhury, 2013; Chautard, Ouédraogo, Biau, & Verrelle, 2014; Emdad, Hu, Das, Sarkar, & Fisher, 2015). AKT wird durch Phosphorylierung aktiviert, vermittelt Zellwachstum und -überleben und inhibiert Apoptose (Chautard et al., 2014; Cui et al., 2015; Emdad et al., 2015; Hu et al., 2014). Interessanterweise ist die intronisch gelegene miRNA hsa-miR-641 kaum untersucht. Wir haben die Hypothese aufgestellt, dass miR-641 ein negativer Regulator des PI3K/AKT-Signalweges ist und dass diese Beziehung im Rahmen der Glioblastompathogenese gestört sein könnte.

In einem ersten Schritt untersuchten wir daher die Expression von miR-641 in Gewebeproben von Glioblastom-Patienten und verglichen diese mit Normalhirn-Gewebe. Tatsächlich zeigte sich eine deutlich erniedrigte miR-641-Expression. Danach überprüften wir, ob miR-641 möglicherweise AKT2 direkt reguliert. Allerdings enthält die AKT2 3'-UTR keine Seed-komplementäre Sequenz für miR-641 und die AKT2-mRNA änderte sich nicht signifikant nach Transfektion dieser miRNA. Trotz unveränderter AKT2-Expression beobachteten wir aber eine deutliche Zunahme der Apoptoserate der transfizierten Zelllinie, vereinbar mit einer indirekten Regulation des PI3K/AKT-Signalwegs. Daher untersuchten wir den AKT2-Aktivierungszustand in diesen Zellen und fanden eine signifikant reduzierte AKT2-Phosphoryllierung. Basierend auf unserer Interaktionsdatenbank identifizierten wir drei Kinasen, die AKT2 aktivieren (Frias et al., 2006; Jacinto et al., 2006; Laplante & Sabatini, 2012; Scheid, Marignani, & Woodgett, 2002) und deren Expressionslevel durch miR-641-Transfektion deutlich reduziert wurden: PIK3R3, PDK2 und MAPKAP1. Für zwei der drei Kinasen (PIK3R3 und MAPKAP1) wiesen wir eine direkte Regulation durch miR-641 nach, die allerdings die ausgeprägten Expressionsänderungen insbesondere von PIK3R3 nicht erklären konnte. Daraufhin suchten wir nach Transkriptionsfaktoren, die möglicherweise durch miR-641 reguliert werden und sowohl mit PIK3R3 als auch PDK2 interagieren. NFAT5 wurde als wahrscheinlicher Kandidat identifiziert und in der Folge von uns validiert. Zusammenfassend konnten wir am Beispiel von AKT2 und miR-641 zeigen, dass intronische miRNAs zentral für die Regulation ihrer Host-Gen Signalwege sein können und dass ein Wegfall der Regulation mit direkten Implikationen für die Tumorphathogenese verbunden sein kann.

Das Verständnis der Beziehung intronischer miRNAs zu deren Host-Genen ist aber nicht nur von pathogenetischer Bedeutung, sondern könnte auch klinische Anwendung finden. Eine kürzlich veröffentlichte Arbeit befasst sich beispielsweise mit der intronisch gelegenen miR-4722, die spezifisch für den menschlichen Organismus ist. Das Gen dieser miRNA liegt in Intron 5 von IL-18RAP. Sowohl intronische miRNA als auch Host-Gen sind koreguliert und in der Sepsis verstärkt exprimiert. Die Autoren konnten zeigen, dass die Bestimmung der Expression von miR-4722 eine Unterscheidung von Patienten mit Systemic Inflammatory Response Syndrome von Patienten mit Sepsis erlaubt (Ma et al., 2013).

In dieser Habilitationsarbeit wurden die nötigen Methoden entwickelt, um die Beziehung intragenischer miRNAs zu ihren Host-Genen und die Bedeutung für die Klinik zu bearbeiten. Diese Methoden wurden erfolgreich eingesetzt, um die zentrale Hypothese einer negativen Rückkopplung der miRNAs innerhalb der Signalwege ihrer Host-Gene zu etablieren sowie Beispiele sowohl für direkte als auch indirekte Rückkopplung zu validieren und regulative Mechanismen zu identifizieren. Abschließend konnte die klinische Relevanz dieser Mechanismen am Beispiel der Glioblastompathogenese gezeigt werden. Es ist davon auszugehen, dass die Beziehung intronischer miRNAs und ihrer Host-Gene eine zunehmend zentrale Rolle in vielen klinischen Bereichen einnehmen wird und möglicherweise interessante und neue diagnostische wie therapeutische Optionen bietet.

Literaturverzeichnis

- Agarwal, E., Brattain, M. G., & Chowdhury, S. (2013). Cell survival and metastasis regulation by Akt signaling in colorectal cancer. *Cellular Signalling*, *25*(8), 1711–1719.
- Baek, D., Villén, J., Shin, C., Camargo, F. D., Gygi, S. P., & Bartel, D. P. (2008). The impact of microRNAs on protein output. *Nature*, *455*(7209), 64–71.
- Baskerville, S., & Bartel, D. P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA*, *11*(3), 241–247.
- Beaudoing, E., Freier, S., Wyatt, J. R., Claverie, J. M., & Gautheret, D. (2000). Patterns of variant polyadenylation signal usage in human genes. *Genome Research*, *10*(7), 1001–1010.
- Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csárdi, G., Harrigan, P., ... Kaessmann, H. (2011). The evolution of gene expression levels in mammalian organs. *Nature*, *478*(7369), 343–348.
- Brennecke, J., Stark, A., Russell, R. B., & Cohen, S. M. (2005). Principles of MicroRNA–Target Recognition. *PLoS Biology*, *3*(3), e85.
- Chan, J. A., Krichevsky, A. M., & Kosik, K. S. (2005). MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Research*, *65*(14), 6029–6033.
- Chautard, E., Ouédraogo, Z. G., Biau, J., & Verrelle, P. (2014). Role of Akt in human malignant glioma: from oncogenesis to tumor aggressiveness. *Journal of Neuro-Oncology*, *117*(2), 205–215.
- Cui, Y., Lin, J., Zuo, J., Zhang, L., Dong, Y., Hu, G., ... Lu, Y. (2015). AKT2-knockdown

- suppressed viability with enhanced apoptosis, and attenuated chemoresistance to temozolomide of human glioblastoma cells in vitro and in vivo. *OncoTargets and Therapy*, 8, 1681–1690.
- Denli, A. M., Tops, B. B. J., Plasterk, R. H. A., Ketting, R. F., & Hannon, G. J. (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature*, 432(7014), 231–235.
- Derti, A., Garrett-Engele, P., Macisaac, K. D., Stevens, R. C., Sriram, S., Chen, R., ... Babak, T. (2012). A quantitative atlas of polyadenylation in five mammals. *Genome Research*, 22(6), 1173–1183.
- Di Giammartino, D. C., Nishida, K., & Manley, J. L. (2011). Mechanisms and consequences of alternative polyadenylation. *Molecular Cell*, 43(6), 853–866.
- Dill, H., Linder, B., Fehr, A., & Fischer, U. (2012). Intronic miR-26b controls neuronal differentiation by repressing its host transcript, *ctdsp2*. *Genes & Development*, 26(1), 25–30.
- Emdad, L., Hu, B., Das, S. K., Sarkar, D., & Fisher, P. B. (2015). AEG-1-AKT2: A novel complex controlling the aggressiveness of glioblastoma. *Molecular & Cellular Oncology*, 2(3), e995008.
- Frias, M. A., Thoreen, C. C., Jaffe, J. D., Schroder, W., Sculley, T., Carr, S. A., & Sabatini, D. M. (2006). mSin1 is necessary for Akt/PKB phosphorylation, and its isoforms define three distinct mTORC2s. *Current Biology: CB*, 16(18), 1865–1870.
- Griffiths-Jones, S. (2006). miRBase: the microRNA sequence database. *Methods in Molecular Biology*, 342, 129–138.
- Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A., & Enright, A. J.

- (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research*, 34(Database issue), D140–4.
- Griffiths-Jones, S., Saini, H. K., van Dongen, S., & Enright, A. J. (2008). miRBase: tools for microRNA genomics. *Nucleic Acids Research*, 36(Database issue), D154–8.
- Guo, H., Ingolia, N. T., Weissman, J. S., & Bartel, D. P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, 466(7308), 835–840.
- Hammond, S. M. (2006). MicroRNAs as oncogenes. *Current Opinion in Genetics & Development*, 16(1), 4–9.
- Han, J., Lee, Y., Yeom, K.-H., Kim, Y.-K., Jin, H., & Kim, V. N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. *Genes & Development*, 18(24), 3016–3027.
- Herr, A. J., Molnár, A., Jones, A., & Baulcombe, D. C. (2006). Defective RNA processing enhances RNA silencing and influences flowering of Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, 103(41), 14994–15001.
- Hinske, L. C., dos Santos, F. R. C., Ohara, D. T., Ohno-Machado, L., Kreth, S., & Galante, P. A. F. (2017). miRIAD update: using alternative polyadenylation, protein interaction network analysis and additional species to enhance exploration of the role of intragenic miRNAs and their host genes. *Database: The Journal of Biological Databases and Curation*, 2017,1-8. <https://doi.org/10.1093/database/bax053>
- Hinske, L. C., França, G. S., Torres, H. A. M., Ohara, D. T., Lopes-Ramos, C. M., Heyn, J., ... Galante, P. A. F. (2014). miRIAD-integrating microRNA inter- and intragenic data. *Database: The Journal of Biological Databases and Curation*, 2014.

<https://doi.org/10.1093/database/bau099>

- Hinske, L. C., Galante, P. A. F., Limbeck, E., Möhnle, P., Parmigiani, R. B., Ohno-Machado, L., ... Kreth, S. (2015). Alternative Polyadenylation Allows Differential Negative Feedback of Human miRNA miR-579 on Its Host Gene ZFR. *PloS One*, *10*(3), e0121507.
- Hinske, L. C. G., Galante, P. A. F., Kuo, W. P., & Ohno-Machado, L. (2010). A potential role for intragenic miRNAs on their hosts' interactome. *BMC Genomics*, *11*, 533.
- Hinske, L. C., Heyn, J., Galante, P. A. F., Ohno-Machado, L., & Kreth, S. (2013). Setting up an intronic miRNA database. *Methods in Molecular Biology*, *936*, 69–76.
- Hu, B., Emdad, L., Bacolod, M. D., Kegelman, T. P., Shen, X.-N., Alzubi, M. A., ... Fisher, P. B. (2014). Astrocyte elevated gene-1 interacts with Akt isoform 2 to control glioma growth, survival, and pathogenesis. *Cancer Research*, *74*(24), 7321–7332.
- Jacinto, E., Facchinetti, V., Liu, D., Soto, N., Wei, S., Jung, S. Y., ... Su, B. (2006). SIN1/MIP1 maintains rictor-mTOR complex integrity and regulates Akt phosphorylation and substrate specificity. *Cell*, *127*(1), 125–137.
- Jing, Q., Huang, S., Guth, S., Zarubin, T., Motoyama, A., Chen, J., ... Han, J. (2005). Involvement of microRNA in AU-rich element-mediated mRNA instability. *Cell*, *120*(5), 623–634.
- Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30.
- Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., & Segal, E. (2007). The role of site accessibility in microRNA target recognition. *Nature Genetics*, *39*(10), 1278–1284.

- Kim, Y.-K., & Kim, V. N. (2007). Processing of intronic microRNAs. *The EMBO Journal*, 26(3), 775–783.
- Kolev, N. G., Yario, T. A., Benson, E., & Steitz, J. A. (2008). Conserved motifs in both CPSF73 and CPSF100 are required to assemble the active endonuclease for histone mRNA 3'-end maturation. *EMBO Reports*, 9(10), 1013–1018.
- Krek, A., Grün, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., ... Rajewsky, N. (2005). Combinatorial microRNA target predictions. *Nature Genetics*, 37(5), 495–500.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., & Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science*, 294(5543), 853–858.
- Laplanche, M., & Sabatini, D. M. (2012). mTOR signaling in growth control and disease. *Cell*, 149(2), 274–293.
- Lau, N. C., Lim, L. P., Weinstein, E. G., & Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science*, 294(5543), 858–862.
- Ledderose, C., Möhnle, P., Limbeck, E., Schütz, S., Weis, F., Rink, J., ... Kreth, S. (2012). Corticosteroid resistance in sepsis is influenced by microRNA-124–induced downregulation of glucocorticoid receptor- α^* . *Critical Care Medicine*, 40(10), 2745.
- Lee, R. C., & Ambros, V. (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. *Science*, 294(5543), 862–864.
- Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, 75(5), 843–854.

- Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., ... Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature*, *425*(6956), 415–419.
- Lewis, B. P., Burge, C. B., & Bartel, D. P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, *120*(1), 15–20.
- Lewis, B. P., Shih, I.-H., Jones-Rhoades, M. W., Bartel, D. P., & Burge, C. B. (2003). Prediction of mammalian microRNA targets. *Cell*, *115*(7), 787–798.
- Long, Y.-S., Deng, G.-F., Sun, X.-S., Yi, Y.-H., Su, T., Zhao, Q.-H., & Liao, W.-P. (2011). Identification of the transcriptional promoters in the proximal regions of human microRNA genes. *Molecular Biology Reports*, *38*(6), 4153–4157.
- Martin, J., Jenkins, R. H., Bennagi, R., Krupa, A., Phillips, A. O., Bowen, T., & Fraser, D. J. (2011). Post-transcriptional regulation of Transforming Growth Factor Beta-1 by microRNA-744. *PloS One*, *6*(10), e25044.
- Mayr, C., & Bartel, D. P. (2009). Widespread shortening of 3' UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell*, *138*(4), 673–684.
- Ma, Y., Vilanova, D., Atalar, K., Delfour, O., Edgeworth, J., Ostermann, M., ... Lord, G. M. (2013). Genome-wide sequencing of cellular microRNAs identifies a combinatorial expression signature diagnostic of sepsis. *PloS One*, *8*(10), e75918.
- Miranda, K. C., Huynh, T., Tay, Y., Ang, Y.-S., Tam, W.-L., Thomson, A. M., ... Rigoutsos, I. (2006). A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell*, *126*(6), 1203–1217.
- Monteys, A. M., Spengler, R. M., Wan, J., Tecedor, L., Lennox, K. A., Xing, Y., &

- Davidson, B. L. (2010). Structure and activity of putative intronic miRNA promoters. *RNA*, *16*(3), 495–505.
- Pruitt, K. D., Tatusova, T., & Maglott, D. R. (2007). NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research*, *35*(Database issue), D61–5.
- Rodriguez, A., Griffiths-Jones, S., Ashurst, J. L., & Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. *Genome Research*, *14*(10A), 1902–1910.
- Scheid, M. P., Marignani, P. A., & Woodgett, J. R. (2002). Multiple phosphoinositide 3-kinase-dependent steps in activation of protein kinase B. *Molecular and Cellular Biology*, *22*(17), 6247–6260.
- Schmitt, D. C., Madeira da Silva, L., Zhang, W., Liu, Z., Arora, R., Lim, S., ... Tan, M. (2015). ErbB2-intronic microRNA-4728: a novel tumor suppressor and antagonist of oncogenic MAPK signaling. *Cell Death & Disease*, *6*, e1742.
- Singh, P., Alley, T. L., Wright, S. M., Kamdar, S., Schott, W., Wilpan, R. Y., ... Graber, J. H. (2009). Global changes in processing of mRNA 3' untranslated regions characterize clinically distinct cancer subtypes. *Cancer Research*, *69*(24), 9422–9430.
- Tie, J., Pan, Y., Zhao, L., Wu, K., Liu, J., Sun, S., ... Fan, D. (2010). MiR-218 inhibits invasion and metastasis of gastric cancer by targeting the Robo1 receptor. *PLoS Genetics*, *6*(3), e1000879.
- Tranter, M., Helsley, R. N., Paulding, W. R., McGuinness, M., Brokamp, C., Haar, L., ... Jones, W. K. (2011). Coordinated post-transcriptional regulation of Hsp70.3 gene

- expression by microRNA and alternative polyadenylation. *The Journal of Biological Chemistry*, 286(34), 29828–29837.
- Wang, X., & El Naqa, I. M. (2008). Prediction of both conserved and nonconserved microRNA targets in animals. *Bioinformatics*, 24(3), 325–332.
- Xie, X., Lu, J., Kulbokas, E. J., Golub, T. R., Mootha, V., Lindblad-Toh, K., ... Kellis, M. (2005). Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature*, 434(7031), 338–345.
- Yan, L., Hao, H., Elton, T. S., Liu, Z., & Ou, H. (2011). Intronic microRNA suppresses endothelial nitric oxide synthase expression and endothelial cell proliferation via inhibition of STAT3 signaling. *Molecular and Cellular Biochemistry*, 357(1-2), 9–19.
- Zhang, H., Lee, J. Y., & Tian, B. (2005). Biased alternative polyadenylation in human tissues. *Genome Biology*, 6(12), R100.

Manuskripte der kumulativen Habilitationsschrift

Folgende Manuskripte sind aufgrund von Copyright-Beschränkungen lediglich online einsehbar.

Hinske, L. C., Heyn, J., Galante, P. A. F., Ohno-Machado, L., & Kreth, S. (2013). Setting up an intronic miRNA database. *Methods in Molecular Biology* , 936, 69–76.
DOI: 10.1007/978-1-62703-083-0_5

Hinske, L. C., Heyn, J., Hübner, M., Rink, J., Hirschberger, S., & Kreth, S. (2017). Intronic miRNA-641 controls its host Gene's pathway PI3K/AKT and this relationship is dysfunctional in glioblastoma multiforme. *Biochemical and Biophysical Research Communications*, 2017 Aug 5;489(4):477-483.
DOI: 10.1016/j.bbrc.2017.05.175