

**MicroRNA-149 ist assoziiert mit der Bildung von  
orofazialen Spalten und veränderter  
Neuralleistenzell-Migration**

Inaugural-Dissertation  
zur Erlangung des Doktorgrades  
der Hohen Medizinischen Fakultät  
der Rheinischen Friedrich-Wilhelms-Universität  
Bonn

**Laura Gesine Stüssel**

aus Dresden

2024

Angefertigt mit Genehmigung  
der Medizinischen Fakultät der Universität Bonn

1. Gutachterin: Dr. rer. nat. Kerstin Ludwig
2. Gutachter: Prof. Dr. Rayk Behrendt

Tag der Mündlichen Prüfung: 14.02.2024

Aus dem Institut für Humangenetik

Direktor: Prof. Dr. med. Markus M. Nöthen

Für meine Eltern



## Inhaltsverzeichnis

	<b>Abkürzungsverzeichnis</b>	6
<b>1.</b>	<b>Deutsche Zusammenfassung</b>	8
1.1	Einleitung	8
	1.1.2 Embryologische Grundlagen	8
	1.1.2 Epidemiologie und Ätiologie von orofazialen Spalten	8
	1.1.3 MicroRNAs und orofaziale Spalten	10
	1.1.4 Zielsetzung	12
1.2	Material und Methoden	12
	1.2.1 Erzeugung und Qualitätskontrolle von hNCC	12
	1.2.2 MicroRNA-Expressionsanalyse	13
	1.2.3 Target Prediction	13
	1.2.4 Integration mit genetischen Daten	14
	1.2.5 Replikationsanalysen	14
	1.2.6 In-vitro-Analyse von miR-149	15
1.3	Ergebnisse	15
	1.3.1 Identifikation von Kandidaten-miRNAs	15
	1.3.2 Priorisierung von miR-149-3p	16
	1.3.3 In-vitro- und in-silico-Untersuchungen zu miR-149	20
1.4	Diskussion	22
1.5	Zusammenfassung	30
1.6	Literaturverzeichnis der deutschen Zusammenfassung	32
<b>2.</b>	<b>Veröffentlichung</b>	44
<b>3.</b>	<b>Danksagung</b>	52

## Abkürzungsverzeichnis

ACTC1	Actin Alpha Cardiac Muscle
ANKRD1	Ankyrin Repeat Domain-Containing Protein 1
BMP4	Bone Morphogenic Protein 4
CDS	Coding sequence, kodierende Sequenz
ChIP	Chromatin Immunpräzipitation
CL/P	Cleft lip with or without cleft palate
CLASH	Cross-linking, Ligation And Sequencing of Hybrids
CPO	Cleft palate only
CRYAB	Crystallin Alpha B
DE	Differentiell exprimiert
DNA	Desoxyribonucleic acid, Desoxyribonukleinsäure
eQTL	Expression quantitative trait loci
FNP	Frontonasale Prominenz
FGF	Fibroblast Growth Factor
GO	Gene Ontology
GPC1	Glypican 1
GREM1	Gremlin 1
GWAS	Genomweite Assoziationsstudie(n)
hiPSC	Human induced pluripotent stem cells, humane induzierte pluripotente Stammzellen
(h)NCC	(Human) neural crest cells, (humane) Neuralleistenzellen
kb	Kilobase
LKG	Lippen-Kiefer-Gaumenspalte
LNP	Laterale nasale Prominenz
MAF	Minore Allelfrequenz
miRNA(s), miR-	MicroRNA(s)
MNP	Mediale nasale Prominenz
MSX1	Homöobox-Protein 1

NOG	Noggin
(ns)CL/P	(Non-syndromic) cleft lip with or without cleft palate
nt	Nukleotid(e)
OFC	Orofacial clefts, orofaziale Spalten
PAX9	Paired box 9
RNA	Ribonucleic acid, Ribonukleinsäure
SNV	Single nucleotide variant, Einzelnukleotidvariante
TAD	Topographically associated domain, topografisch assoziierte Domäne
TAGLN	Transgelin
TGF(B2)	Transforming growth factor (beta 2)
UTR	Untranslated region

# 1. Deutsche Zusammenfassung

## 1.1 Einleitung

### 1.1.1 Embryologische Grundlagen

Die Entwicklung des menschlichen Gesichts findet im frühen Schwangerschaftsverlauf, d.h. zwischen der vierten und siebten Woche bzw. in den Carnegie Stadien 10-18, statt (Ji et al., 2020; O’Rahilly, 1972). Ausgehend von Neuralleistenzellen (NCC), die bilateral nach ventral zur frontonasalen Prominenz (FNP) und in den ersten Kiemenbogen migrieren, bilden sich unter anderem die Strukturen des Ober- und Unterkiefers (Cordero et al., 2011). Die Entwicklung von Oberkiefer und -lippe laufen wie folgt ab: Die zunächst getrennten medialen und lateralen nasalen Prominenz (MNP, LNP) und die sogenannten Oberkieferwülste fusionieren in beiden Hälften des Gesichts (Ji et al., 2020) miteinander und formen die Oberlippe. Im Bereich der Mittellinie entsteht aus den zusammenwachsenden bilateralen MNPs außerdem das Philtrum und der anteriore (primäre) Gaumen; aus den zusammenwachsenden Oberkieferwülsten entwickelt sich der posteriore (sekundäre) Gaumen. Zeitlich betrachtet entsteht dabei zuerst die Lippe, dann der Gaumen. Im Verlauf fusionieren zudem primärer und sekundärer Gaumen. Auf zellulärer Ebene sind dabei neben der NCC-Differenzierung und -migration (Trainor, 2010) auch die epithelial-mesenchymale Transition, die kontrollierte Apoptose von Epithelzellen und die Interaktion mit der extrazellulären Matrix an der Entstehung der Gesichtsstrukturen beteiligt (Ji et al., 2020). Fehler bei diesen eng aufeinander abgestimmten Prozessen können zu Deformitäten im Gesicht, unter anderem zu Spaltbildungen, führen. Die Erforschung menschlicher Spaltbildungen ist aufgrund des frühen Ablaufs der Gesichtsentwicklung in der Schwangerschaft und dem damit einhergehenden limitierten Zugang zu relevanten Geweben leider erschwert.

### 1.1.2 Epidemiologie und Ätiologie von orofazialen Spalten

Orofaziale Spalten (orofacial clefts, OFC) sind mit einer geschätzten Prävalenz von 1:700 Neugeborenen häufige angeborene Fehlbildungen, die eine große phänotypische Breite aufweisen (Mangold et al., 2017, Mossey et al., 2009). Die Behandlung der Betroffenen ist komplex und erfordert eine interdisziplinäre Betreuung, welche bis ins Erwachsenenalter andauern kann (Lewis et al., 2017; Worley et al., 2018). Dabei können OFC zu diversen Folgebeschwerden, wie z.B. erschwerte Nahrungsaufnahme (Mossey



et al., 2009), rezidivierenden Mittelohrentzündungen (Karanth und Whittmore, 2018) und Schlafapnoe (Tan et al., 2016), führen. Solche Spaltbildungen gehen zudem mit einer erhöhten Mortalität einher (Christensen et al., 2004).

Innerhalb der OFC kann bezüglich der betroffenen anatomischen Strukturen zwischen einseitigen und beidseitigen Spalten, medianen und lateralen Formen, separaten Spaltbildungen der Lippe bzw. des Gaumens und kombinierten Formen unterschieden werden (Mangold et al., 2017). Der größte Anteil entfällt auf Lippenspalten mit oder ohne Beteiligung des Gaumens, wobei hier der Alveolarfortsatz des Oberkiefers mit gespalten sein kann (cleft lip with or without palate, CL/P, im Deutschen Lippen-Kiefer-Gaumenspalten, LKG) (Mossey et al., 2009). Seltener sind reine Gaumenspalten (cleft palate only, CPO), die etwa ein Drittel der Fälle von orofazialen Spalten ausmachen (Martinelli et al., 2020). Jede Form der OFC kann in Hinblick auf ihre Ätiologie als syndromale oder nicht-syndromale Form auftreten. LKG liegen dabei seltener syndromal vor als CPO (Watkins et al., 2014). Syndromale OFC sind assoziiert mit Fehlbildungen der Haut und der Zähne, des Herzens oder mit malignen Erkrankungen (Setó-Salvia und Stanier 2014). Außerdem treten sie im Rahmen von komplexen Krankheitsbildern, z.B. beim Down-Syndrom (Stoll et al., 2015) auf. Liegen OFC isoliert vor, also primär ohne begleitende phänotypische Veränderungen, bezeichnet man diese als nicht-syndromale (non-syndromic, ns) Formen. Ihre Genese ist als multifaktoriell anzusehen, d.h. sowohl genetische Faktoren als auch Umwelteinflüsse sind an ihrer Entstehung beteiligt. Als relevante Umweltfaktoren werden unter anderem Alkohol- oder Nikotinkonsum der Mutter in der Schwangerschaft diskutiert (Beaty et al., 2013; Carlson et al., 2022). Allerdings ist dieser Zusammenhang noch nicht abschließend geklärt.

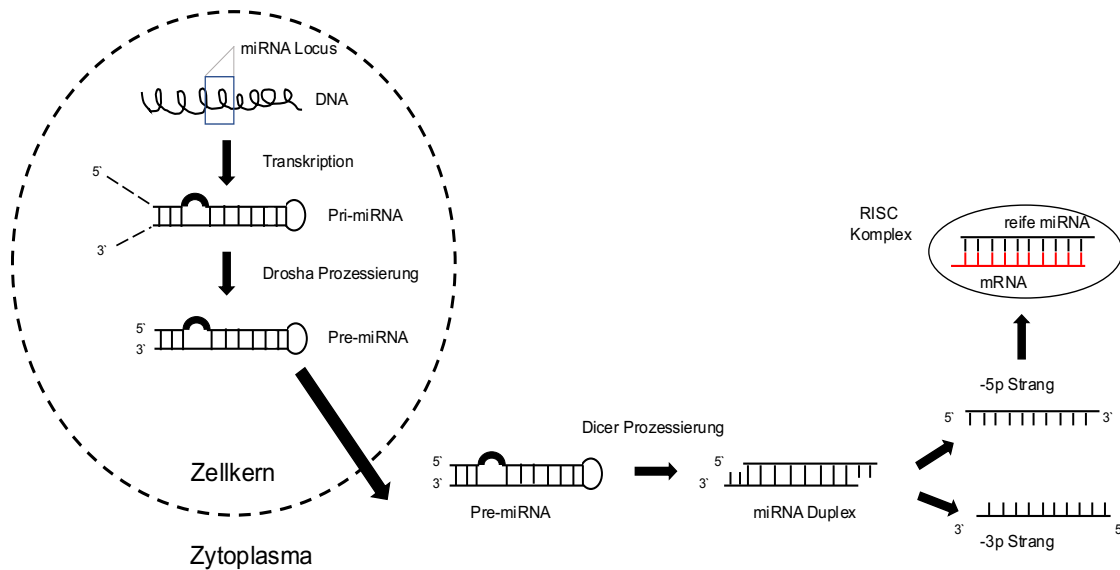
Die grundsätzliche Anwesenheit einer genetischen Komponente bei der Entwicklung von OFC konnte durch Zwillingsstudien und Analysen von großen Fehlbildungsregistern gesichert werden (Carter et al., 1982; Christensen und Fogh-Andersen, 1993; Grosen et al., 2010; Shields et al., 1979; Sivertsen et al., 2008). Diese zeigten, dass insbesondere monozygote Zwillinge, aber auch dizygote Zwillinge und erstgradige Verwandte von Spaltpatienten, ebenfalls eine höhere Rate an OFC aufwiesen, ohne jedoch die dafür verantwortlichen Gene zu identifizieren.

Der weitere Fokus dieser Arbeit soll auf nicht-syndromalen Lippenpalten mit oder ohne Kieferbeteiligung (nsCL/P) liegen, weswegen im Folgenden nur die für diesen Phänotyp relevanten genetischen Erkenntnisse zusammengefasst werden. Erste Kandidatengene für nsCL/P, wie das Homöobox-Protein MSX1, wurden z.B. durch Daten aus Mausexperimenten und Linkage-Disequilibrium-Analysen in betroffenen Familien postuliert (van den Boogaard et al., 2000). Dabei konnten jedoch meist nur kleinere Kollektive und einzelne Gene bzw. Loci beleuchtet werden. Große Fortschritte im Verständnis von nsCL/P konnten daher in den letzten Jahren vor allem durch die zunehmende Verfügbarkeit von Hochdurchsatzmethoden wie genomweiten Assoziationsstudien (GWAS), gemacht werden. Mittels systematischer GWAS wurden mindestens 40 Risiko-Loci für nsCL/P identifiziert (Reynolds et al., 2020). Dazu gehören beispielsweise Abschnitte auf den chromosomalen Bereichen 9q21 (Beaty et al., 2013; Leslie et al., 2017; Marazita et al., 2004), 1q32 und 8q24.21 (Beaty et al., 2010; Birnbaum et al., 2009). Ein Großteil dieser bisher identifizierten Varianten liegt in sogenannten nicht-kodierenden Regionen des Genoms (Thieme und Ludwig 2017) und kann keinen einzelnen Genen zugeordnet werden. Ihre Effekte sind dementsprechend schwieriger hervorzusagen. Nicht-kodierende Regionen sind aber durchaus an der Ausprägung von Phänotypen beteiligt. Sie tragen zur Genregulation bei, indem sie z.B. als Enhancer als Bindungsstelle für Transkriptionsfaktoren agieren und somit die Transkription ihrer Ziel-Gene verstärken. Auf posttranskriptioneller Ebene können nicht-kodierende Ribonukleinsäuren (RNA) an der Genregulation teilnehmen (Maston et al. 2006; Mattick und Makunin 2006). Die Identifikation von Regulations-Mechanismen solcher intergenischer Regionen spielt somit eine essenzielle Rolle im Verständnis komplexen Erkrankungen wie nsCL/P (Thieme und Ludwig 2017).

### 1.1.3 MicroRNAs und orofaziale Spalten

MicroRNAs (miRNAs) zählen zu den nicht-kodierenden RNAs und weisen eine vergleichsweise kurze Länge von 21-23 Nukleotiden (nt) auf. Ihre Biosynthese ist ein komplexer Prozess mit diversen Vorstufen (Abb. 1, Seite 11). Die reife miRNA kann an der 3' untranslatierten Region (3'UTR) ihrer komplementären Ziel-mRNA binden (Eulalio et al., 2008), was zum Abbau der mRNA führt und den miRNAs somit die Rolle von Gen-Silencern zukommen lässt. In den letzten Jahren wurden zunehmend auch Interaktionen

von miRNAs mit der 5'UTR und den kodierenden Bereichen von Genen beschrieben (Broughton et al., 2016).



### Abb.1: Kanonische Biosynthese von miRNAs

Nach Transkription des miRNA-Gens durch die RNA-Polymerase II entsteht die Stem-loop-förmige pri-miRNA, welche ein 5'Cap und einen Poly-A-Schwanz enthält. Mithilfe der RNase DROSHA werden diese Enden abgespalten, die pri-miRNA gekürzt und es entsteht die etwa 70nt lange pre-miRNA. Diese wird dann aus dem Zellkern ins Zytoplasma exportiert, wo sie durch eine weitere RNase (Dicer) zum miRNA-Duplex (ca. 22nt) gekürzt wird. Im Folgenden entstehen reife Einzelstrang-miRNAs, die im RISC-Komplex bei Komplementarität zum Abbau von mRNAs führen können (Colpaert und Calore, 2019).

Ziel-Gene von miRNAs können entweder durch in vitro-Experimente (z.B. qPCRs, Next Generation Sequencing) bestimmt werden oder mithilfe von bioinformatischen Tools über die Sequenz der miRNA vorhergesagt werden (Grigaitis et al., 2020); dies bezeichnet man als Target Prediction.

Die Beteiligung von miRNAs an der Entstehung und am Fortschreiten menschlicher Erkrankungen wurde bereits für diverse hochprävalente Krankheitsbilder beschrieben: zum Beispiel für das Mamma- und Prostatakarzinom (Loh et al., 2019; Sekhon et al., 2016) oder die nicht-alkoholische Fettleber (Becker et al., 2015). Auch im Bereich der OFC konnten in den letzten Jahren einige miRNAs identifiziert werden, die Einfluss auf deren Entstehung haben könnten (Seelan et al., 2022). So konnte für miR-140, eine an der Regulation der Palatogenese von Mäusen beteiligte miRNA (Eberhart et al., 2008), gezeigt werden, dass deren Prozessierung durch eine nsCL/P assoziierte Einzelnukleo-

tidvariante (single nucleotide variant, SNV) verändert wurde. Die Ausbildung eines Spalt-Phänotypes unter erhöhter Expression von miR-140 bei Zebrafischen ließ sich auf eine reduzierte Migration der kranialen NCC zurückführen (Dougherty et al. 2012; Eberhart et al. 2008). Weitere Beispiele sind miR-193a-3p und let-7c-5p, welche an der Regulation von Proliferation und Apoptose von mesenchymalen und epithelialen Zellen des menschlichen Gaumens und des Mundes beteiligt sind (Fu et al., 2021).

#### 1.1.4 Zielsetzung

Ziel der vorliegenden Arbeit war, die Rolle von miRNAs bei der genetisch vermittelten Entstehung von nsCL/P genauer zu beleuchten. Vorangegangene Studien hierzu waren meist basierend auf miRNA-Expressionsmustern aus Tiermodellen erfolgt (Ding et al., 2016; Eberhart et al., 2008; Li et al., 2010), was die Integration mit genetischen Daten großer Kohorten einschränkt. In dieser Arbeit sollte daher zunächst ein miRNA-Datensatz aus humanen NCC (hNCC), welche aus pluripotenten Zelllinien generiert werden können, erstellt werden. Basierend auf den Expressionsdaten und Informationen aus der Literatur sollten anschließend potenziell an der frühen Gesichtsentwicklung beteiligte miRNAs identifiziert werden. Ausgehend von der Annahme, dass sich an Loci der Kandidaten-miRNAs nsCL/P assoziierte genetischen Varianten finden können, sollten anschließend vorab erhobene genetische Daten (GWAS) von nsCL/P Patienten integriert werden. Weitere Evidenz sollte durch Replikation in einem unabhängigen Sample sowie in vitro-Analysen ausgewählter Kandidaten-miRNAs generiert werden.

## 1.2 Material und Methoden

### 1.2.1 Erzeugung und Qualitätskontrolle von hNCC

Die Herstellung und Weiterverarbeitung der hNCC erfolgte nach etablierten Protokollen durch unsere Kooperationspartnerin Dr. rer. nat. Magdalena Laugsch (Laugsch et al., 2019; Prescott et al., 2015). Eine ausführliche Beschreibung ist in den Supplementary Files der Originalpublikation zu finden. Die Fibroblasten- oder Keratinozytenproben von vier gesunden Individuen (Zelllinie DPEDi001-A, Ethikvotum EK153042016 der Uniklinik Köln) wurden über pluripotente Stammzellen (human induced pluripotent stem cells, hiPSC) zu hNCC reprogrammiert. Für jede hNCC-Probe wurden Gesamt-RNA und -DNA extrahiert und es erfolgte eine Qualitätskontrolle. 1.) Die DNA-Proben wurden mit-

tels NanoDrop quantifiziert und die genomische Integrität wurde durch Karyotypisierung bestätigt. 2.) Die RNA-Integrität wurde mittels Bioanalyzer validiert und im Folgenden mittels 3'RNA-Seq auf die Expression von Neuralleisten-, iPSC- und mesenchymalen Markern untersucht.

### 1.2.2 MicroRNA-Expressionsanalyse

Für die miRNA-Expressionsanalyse aus den vier hNCC-Proben wurde der Affymetrix GeneChip® (Affymetrix, Santa Clara, CA) genutzt. Die Proben wurden in Übereinstimmung mit den Herstellerinformationen vorbereitet und verarbeitet. Die Expressionsdaten wurden dann mithilfe der Affymetrix® Expression Console™ Software analysiert. Dabei wurden entsprechend eines etablierten Protokolls (Hochfeld et al., 2017) die initialen Expressionswerte nach Quantilnormalisierung und log<sub>2</sub>-Transformierung (mittels robuster multi-array Analyse) einer „detection above background“-Analyse unterzogen. Dabei wird die Signalintensität jeder einzelnen (miRNA)-Probe mit der Signalintensität der negativen Kontrollen verglichen und ein P-Wert generiert, der die Wahrscheinlichkeit der Anwesenheit der jeweiligen miRNA repräsentiert. MiRNAs wurden als exprimiert definiert, wenn sich ein P-Wert von kleiner 0,01 in mindestens einem der vier hNCC-Samples ergab. MiRNAs, die in allen vier Samples exprimiert waren, wurden als Kandidaten-miRNAs angesehen („hNCC batch“). Weitere Kandidaten-miRNAs wurden in einem „literature batch“ zusammengefasst. Diese miRNAs mussten zwei Kriterien erfüllen: Zum einen musste ihre Rolle in der kraniofazialen Entwicklung durch bereits veröffentlichte Studien beschrieben sein. Zweitens mussten die entsprechenden miRNAs in mindestens einer der vier hNCC-Proben ebenfalls unter einem P-Wert < 0,01 exprimiert sein.

### 1.2.3 Target Prediction

Zur Target Prediction von hsa-miR-149-3p wurde das Online-Tool MirWalk3.0 (Sticht et al., 2018) benutzt. Der dort verwendete Algorithmus nutzt Cross-linking, Ligation And Sequencing of Hybrids (CLASH)-Daten und einen Random-Forest-Test-Ansatz zur Erstellung eines Scores. Dieser dient als Maß für die Wahrscheinlichkeit einer miRNA-Target-Interaktion. Der Cut-off für den Score wurde bei 0,95 gesetzt. Es wurden Ziel-Gene sowohl in den 3'- und 5'UTRs als auch kodierenden Bereichen von Genen analy-

siert. Um Redundanz zu verhindern, wurde dabei jedes Gen nur einmal pro Region gezählt (3'UTR, 5'UTR und kodierender Bereich (CDS)).

#### 1.2.4 Integration mit genetischen Daten

Die benutzten GWAS-Daten für nsCL/P stammten aus einer zuvor veröffentlichten Studie zu nsCL/P mit europäischen und asiatischen Patienten (399 Patienten, 1318 Kontrollen, Ludwig et al., 2017). Die chromosomalen Positionen der Kandidaten-miRNAs wurden mit einem Fenster von  $\pm$  einer Kilobase (kb) auf dem Genom Browser der University of California Santa Cruz (UCSC, Genome Reference Consortium Human Build 37 (GRCh37, hg 19)) abgerufen. Innerhalb dieser chromosomalen Positionen wurden dann die Varianten aus den nsCL/P Daten sowie deren P-Werte extrahiert. Es wurden lediglich SNV mit einer ausreichenden Imputationsqualität (Infoscore  $> 0,4$ ) und einer minderen Allelfrequenz (MAF) von größer 1% berücksichtigt. Das Signifikanzlevel nach multipler Testung wurde bei  $P = 0,05/n_{\text{SNVs}}$  gesetzt. SNV mit einem P-Wert von  $< 0,01$  entweder in der rein europäischen Kohorte ( $\text{meta}_{\text{Euro}}$ ) oder in der Kohorte europäischer und asiatischer Patienten ( $\text{meta}_{\text{all}}$ , (Ludwig et al., 2017)) wurden als Kandidatenvarianten definiert. Für einzelne Loci wurden regionale Assoziationsplots mithilfe des Softwaretools LocusZoom (Pruim et al., 2011) generiert.

#### 1.2.5 Replikations-Analysen

Zur Bestätigung der durch die Analyse der GWAS-Daten identifizierten Kandidatenvarianten erfolgte eine Genotypisierung in einer unabhängigen Kohorte aus 223 nsCL/P Patienten und 968 Kontrollen europäischer Herkunft (Bonn Kohorte, (Mangold et al., 2011)). Dabei wurden 21 der 25 Kandidatenvarianten mittels MALDI-TOF (matrix-assisted laser-desorption-ionisation – time of flight mass spectrometry) basierten Genotypisierung auf einem MassARRAY System (Agena Bioscience) analysiert. Das Primer- und Plex-Design erfolgte mithilfe der AgenaCx Plattform. Für vier der ursprünglich 25 Varianten konnten keine eindeutigen (uniquen) Primer designt werden. Im Anschluss an die Genotypisierung erfolgte eine manuelle Kontrolle der Ergebnisse mithilfe der iSeq™ und PlateEditor Software (jeweils Agena Bioscience). Im Rahmen dieser Qualitätskontrolle wurden sechs weitere Varianten (nachdem sie in weniger als 95% der Individuen genotypisiert werden konnten) und 27 Individuen (2 Patienten, 25 Kontrollen) aus den

Analysen ausgeschlossen. Die statistischen Analysen beinhalteten Armitage-Trend- und Genotyp-basierte Tests. Eine Kandidatenvariante (rs149251515) mit unabhängiger Evidenz aus dieser Replikationsanalyse wurde in zwei unabhängigen Kohorten analysiert: 1.) in silico Analyse und Abruf eines P-Werts aus einer veröffentlichten nsCL/P GWAS mit Individuen gemischter Herkunft (Leslie et al., 2016; Welzenbach et al., 2021). 2.) Restriktions-Fragment-Längen-Polymorphismus-Analysen und logistische Regression in einer polnischen Kohorte (471 Patienten und 832 Kontrollen, (Mostowska et al., 2018)).

### 1.2.6 In-vitro-Analyse von miR-149

Diese Untersuchungen erfolgten ebenfalls durch unsere Kooperationspartnerin Frau Dr. Lausch. Zur Unter- und Überexpression von miRNA-149-3p wurden miRVana-Inhibitor und miRVana-Mimic genutzt. Im Folgenden wurden zwei Untersuchungen in Bezug auf die veränderten miRNA-Expressionslevels durchgeführt: Zum einen das Migrationsverhalten der hNCC (Scratch Assay), zum anderen erfolgte eine 3'RNA-Seq (Lexogen). Die Analyse dieser Daten erfolgte durch die Ko-Autorin Frau Dr. Hollstein. Mithilfe der 3'RNA-Seq wurden differentiell exprimierte (DE) Gene, also Gene mit verändertem Expressionslevel nach mirVana-Transfektion (höher oder niedriger im Vergleich zu den nicht-transfizierten Kontrollen) bestimmt. Dabei wurden je zwei Replikate von Kontrollen, miRNA-Mimic und miRNA-Inhibitor analysiert. Zur bioinformatischen Analyse wurde ein etablierter limma/edgeR-Workflow verwendet (Smyth, 2004). Für die DE Gene ( $P_{adj.} < 0.05$ ,  $\log_2\text{fold change} > 1.5$ ) erfolgte außerdem eine GO-Term Analyse mithilfe von ConsensusPathDB (Kamburov et al., 2009).

## 1.3 Ergebnisse

### 1.3.1 Identifikation von Kandidaten miRNAs

Die korrekte Differenzierung und hNCC-Identität konnten mithilfe der RNA-Seq Daten bestätigt werden: Es zeigte sich eine niedrige Expression von hiPSC-Markergenen und eine fehlende Expression von mesenchymalen Markergenen bei deutlicher Expression bekannter NCC-Marker (z.B. NRF2F112, SOX9, TFAP2A, Abb. 1A der Veröffentlichung, Seite 43, (Rada-Iglesias et al., 2012)). Begleitend konnte eine hohe Korrelation zwischen den einzelnen hNCC-Proben gezeigt werden (durchschnittlicher Pearson's  $r^2 = 0,896$ , Abb. 1B der Veröffentlichung, Seite 46). Ergänzend wurden die RNA-

Expressions-Daten mit öffentlich zugänglichen Transkriptom-Daten anderer hNCC-Proben verglichen, hier zeigte sich eine Korrelation ( $r^2$ ) von 0,674 (Abb. 1 der Supplementary Files). Insgesamt wurden 151 Kandidaten-miRNAs definiert. Diese setzen sich zusammen aus 1.) dem „hNCC batch“ mit 111 reifen miRNAs und drei Stemloop Sequenzen (entsprechend sechs reifen miRNAs), die jeweils in allen vier hNCC als exprimiert detektiert wurden sowie aus 2.) 47 miRNAs des „literature batches“ (Expression in mindestens einer hNCC Probe und Anhalt für kraniofaziale Relevanz in der Literaturrecherche, Tab. 1). Insgesamt wurden 13 miRNAs identifiziert, die beiden „batches“ zuzuordnen waren (Details s. Tab. 1 sowie Figure 1C und 1D der Veröffentlichung, Seite 46 und Tab. 1 der Supplementary Files).

**Tab. 1: Expression von miRNAs in den vier hNCC Samples und durch die Literaturrecherche identifizierte miRNAs.**

	hNCC-Proben					Literaturrecherche („literature batch“)
	hNCC_1	hNCC_2	hNCC_3	hNCC_4	Überschneidung („hNCC batch“)	
Anzahl miRNAs	382	293	416	258	111 (+6 aus Stem-loops)	47

### 1.3.2 Priorisierung von miR-149-3p

Innerhalb der uns vorliegenden GWAS-Daten zu nsCL/P (Ludwig et al., 2017) wurde nach SNVs gesucht, die an Loci der definierten Kandidaten-miRNAs ( $n=151$ ) lagen. Insgesamt konnten 13.002 SNVs identifiziert werden, welche innerhalb des definierten Fensters von 1kb um die miRNA-Loci lagen. Tabelle 2 (Seite 17) stellt die 25 SNVs an 13 miRNA Loci dar, die einen Assoziations-P-Wert von  $< 0,01$  in einer der beiden Analysen ( $meta_{all}$  – europäische und asiatische Patienten oder  $meta_{euro}$  – nur europäische Patienten) aufwiesen. Für keine der Varianten waren die Assoziations-Werte nach der Korrektur für multiples Testen weiterhin signifikant.

Ebenfalls in Tabelle 2 ist zu sehen, dass nach dem initialen Primerdesign zunächst 21 SNV zur Genotypisierung im unabhängigen Sample verblieben. Im Rahmen der Qualitätskontrolle wurden weitere sechs SNV exkludiert. Von den übrigen 15 Varianten zeigte eine SNV auf Chromosom 2 ( $rs149251515$ ,  $P_{genotype} = 0,049$ ,  $P_{Armitage} = 0,08$ , odds ratio = 1,57 [95% Konfidenzintervall: 0,94-2,61], markiert in Tab. 2) eine nominelle Signifikanz in der Replikationskohorte.



**Tab. 2: Übersicht zu Kandidaten-SNVs an Kandidaten-miRNA Loci – Assoziationsstatistiken in nsCL/P GWAS (Ludwig et al., 2017) und Replikation in unabhängigem Sample**

Pmeta – niedrigster Assoziations-P-Wert (aus entweder der europäischen Kohorte – <sup>a</sup> oder aus der europäischen + asiatischen Kohorte – <sup>b</sup>)

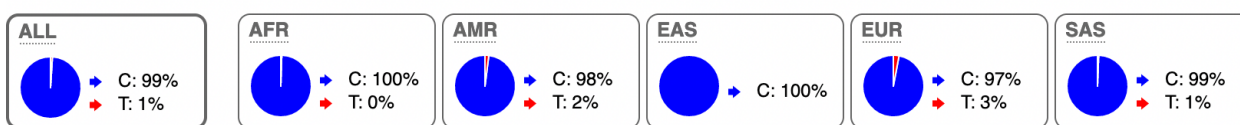
Chr. – Chromosom

P\_Ge/ P\_Tr – Genotyp bzw. Armitage Trend basierter P-Wert

QC – Qualitätskontrolle

SNV	miRNA	Chr.	Assoziationsstatistik nach Ludwig et al.			Assoziationsstatistik Replikationskohorte		
			Allele	Pmeta	Risiko Allel	MAF	P_Ge	P_Tr
rs149251515	hsa-mir-149	2	C/T	1,47E-03 <sup>a</sup>	T	0,048	4,94E-02	8,00E-02
rs9854333	hsa-mir-885	3	A/G	9,75E-03 <sup>b</sup>	G	0,368	7,78E-01	4,86E-01
rs3774159	hsa-mir-885	3	C/G	9,33E-03 <sup>b</sup>	G	0,373	7,61E-01	4,60E-01
rs7372209	hsa-mir-26a	3	C/T	5,78E-03 <sup>a</sup>	C	Exkludiert im Rahmen der QC		
rs13202946	hsa-mir-30a	6	C/T	6,92E-03 <sup>b</sup>	T	Exkludiert im Rahmen der QC		
rs45560031	hsa-mir-4649	7	C/G	9,68E-03 <sup>b</sup>	C	0,050	3,88E-01	6,19E-01
rs41280157	hsa-mir-4668	9	A/C	7,32E-05 <sup>a</sup>	C	0,014	9,99E-01	9,58E-01
rs2484293	hsa-mir-7162	10	A/G	8,89E-03 <sup>a</sup>	G	0,156	9,04E-01	6,53E-01
rs7946008	hsa-mir-6744	11	C/T	9,20E-03 <sup>b</sup>	T	Exkludiert im Rahmen der QC		
rs7134440	hsa-mir-6757	12	C/T	8,80E-03 <sup>b</sup>	T	Exkludiert im Rahmen der QC		
rs113552903	hsa-mir-320d	13	C/T	3,45E-04 <sup>b</sup>	C	Exkludiert aufgrund des Primer-Designs		
rs113944487	hsa-mir-320d	13	A/G	1,85E-03 <sup>b</sup>	G	Exkludiert aufgrund des Primer-Designs		
rs7332128	hsa-mir-320d	13	G/T	1,88E-03 <sup>b</sup>	T	Exkludiert im Rahmen der QC		
rs10048129	hsa-mir-140	16	C/T	1,11E-03 <sup>a</sup>	C	Exkludiert im Rahmen der QC		
rs10048088	hsa-mir-140	16	A/G	3,33E-03 <sup>a</sup>	G	0,278	8,24E-01	8,04E-01
rs10048090	hsa-mir-140	16	A/G	3,25E-03 <sup>a</sup>	A	0,277	8,38E-01	8,67E-01
rs2102066	hsa-mir-140	16	A/G	3,19E-03 <sup>a</sup>	A	0,278	8,18E-01	8,12E-01
rs7206222	hsa-mir-140	16	A/G	3,10E-03 <sup>a</sup>	A	0,278	5,34E-01	5,46E-01
rs8047818	hsa-mir-140	16	A/G	3,28E-03 <sup>a</sup>	A	0,281	7,00E-01	8,94E-01
rs2270841	hsa-mir-140	16	C/T	2,65E-03 <sup>a</sup>	C	0,277	8,28E-01	8,45E-01
rs7239622	hsa-mir-320c	18	C/T	9,00E-03 <sup>a</sup>	C	0,462	5,54E-01	2,93E-01
rs113687432	hsa-mir-320c	18	C/T	4,75E-04 <sup>a</sup>	T	Exkludiert aufgrund des Primer-Designs		
rs112021519	hsa-mir-320c	18	A/C	4,76E-04 <sup>a</sup>	A	Exkludiert aufgrund des Primer-Designs		
rs75750038	hsa-mir-6088	19	A/G	4,96E-03 <sup>a</sup>	A	0,027	6,53E-01	3,55E-01
rs8111742	hsa-mir-let-7e*-cluster	19	A/G	8,81E-03 <sup>a</sup>	A	0,247	6,35E-01	3,42E-01

Im Folgenden wurden weitere Informationen zu rs149251515 eingeholt, um ihre Bedeutung bei - wie beschrieben - lediglich nomineller Signifikanz in der Genotypisierung besser einzuschätzen. Assoziationsdaten von rs149251515 wurden in den GWAS zweier unabhängiger nsCL/P Kohorten abgerufen. In einer Kohorte mit durchmischter Herkunft ergab sich ebenfalls nominelle Signifikanz ( $P = 0,033$  (Leslie et al., 2016)), während sich jedoch keine signifikante Assoziation in einer polnischen Kohorte ( $P = 0,71$  (Mostowska et al., 2018)) zeigte. Die korrekte Imputation von rs149251515 in den Array-Daten wurde überprüft, indem zum einen der Infoscore abgerufen wurde, dieser betrug  $> 0,6$ . Zum anderen wurde die MAF der Variante mit Genomdaten auf gnomAD und Ensembl abgeglichen; diese zeigte für europäische Kohorten ein vergleichbares Ergebnis (Abb. 2). Die MAF dieser nominell signifikanten Variante in der europäischen Vergleichspopulation ist 0,03. Sie stellt damit eine eher seltenere Variante am miR-149-Locus dar.

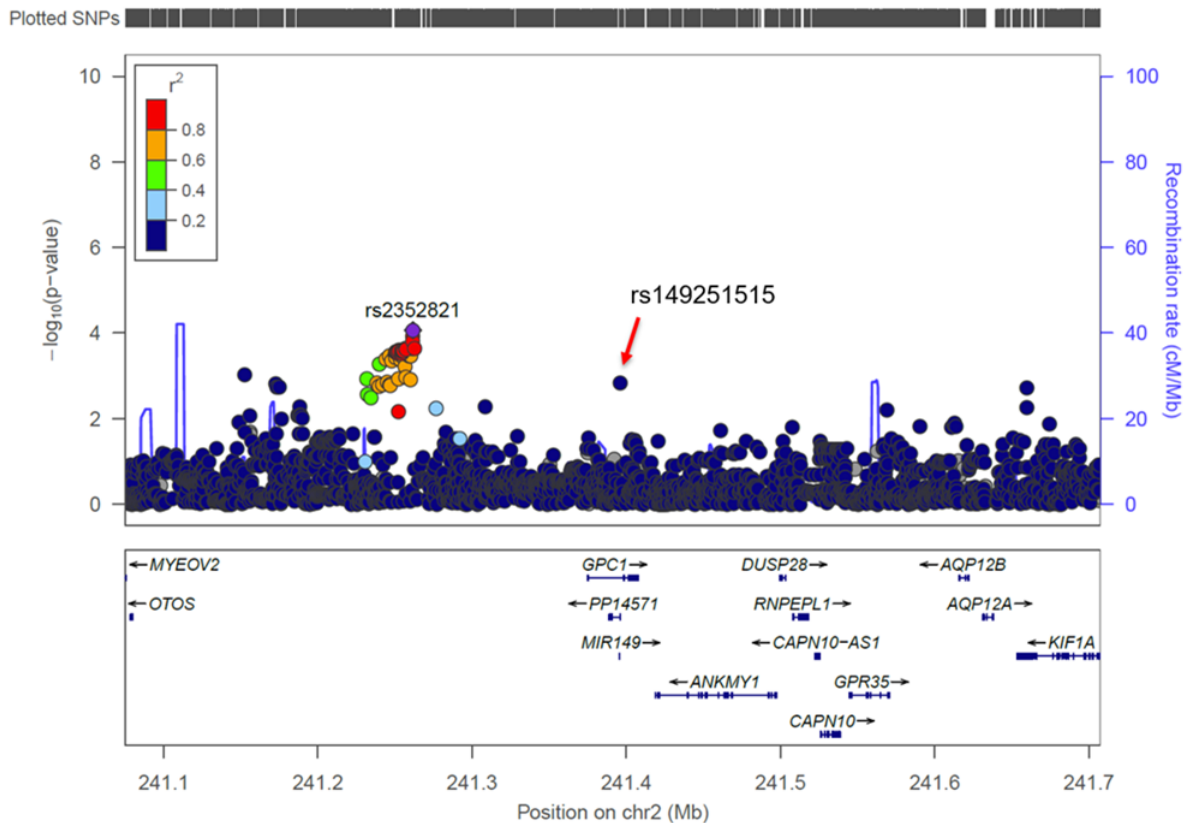


**Abb. 2: Allelfrequenzen für rs149251515 nach Ensembl (Zerbino et al., 2017)**

Die Variante kommt vor allem in der europäischen Population vor (MAF = 3%).

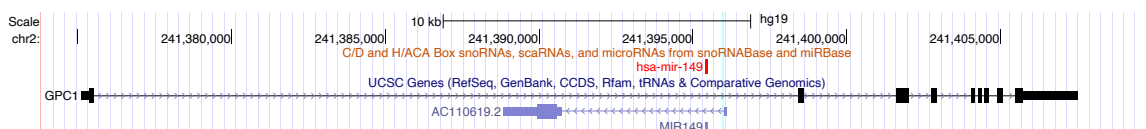
ALL – gesamte Population. AFR – Afrikanische Population. AMR – Amerikanische Population. EAS – Ostasiatische Population. EUR – Europäische Population. SAS – Südasiatische Population.

Um mögliche assoziierte Varianten zu identifizieren, die durch das Setzen des 1kb-Fensters den Analysen entgangen waren, wurde die gesamte assoziierte Region in meta<sub>all</sub> untersucht. Es konnten keine weiteren Varianten gefunden werden, die in Linkage Disequilibrium ( $r^2 > 0,8$ ) zu rs149251515 standen. Downstream konnten mehrere häufige Varianten mit P-Werten  $< 0,001$  identifiziert werden (Abb. 3, Seite 19, entsprechend Appendix Fig. S3, Stüssel et al., 2022).

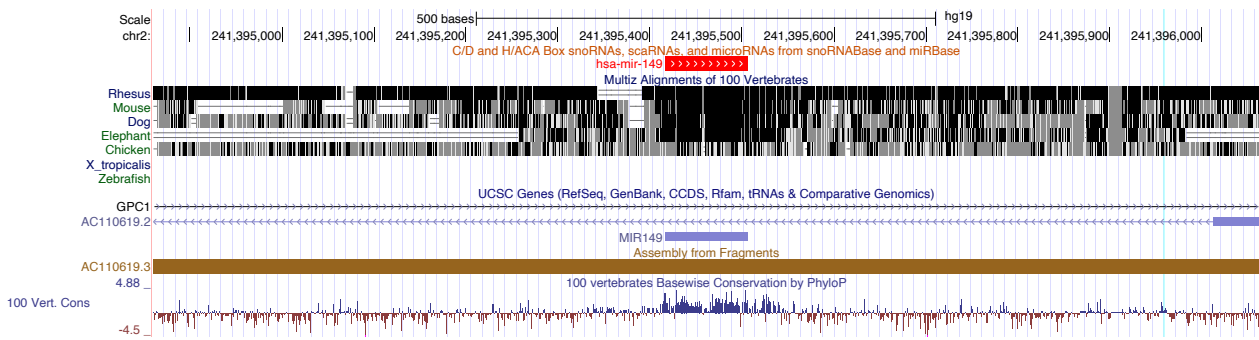


**Abb. 3: Regionaler Assoziationsplot des rs149251515/miR-149 Locus (entsprechend Appendix Fig. S3, Stüssel et al., 2022) mit einem Fenster von etwa 600kb um rs149251515 in der metaall GWAS Kohorte (Ludwig et al., 2017) mit Darstellung häufiger Varianten 250kb downstream. Auf der x-Achse dargestellt ist die chromosomale Position mit den in diesem Intervall gelegenen Genen. Die y-Achse zeigt die entsprechenden (logarithmischen) P-Werte der Varianten und die Rekombinationsrate. Farblich kodiert sind die Korrelationskoeffizienten der Varianten ( $r^2$ ) als Maß des Linkage Disequilibriums.**

Zur Evaluation möglicher Interaktionen und Funktionalität der intergenisch gelegenen Variante wurde der Locus außerdem in Hinblick auf nahegelegene Gene und Konservierungsgrad und untersucht (Abb. 4 und 5, Seite 19/20).



**Abb. 4: miR-149 Locus entsprechend des UCSC Genom Browsers (modifiziert nach Stüssel et al., 2022 (Appendix Fig. S2), Kent et al., 2002). Dargestellt ist der chromosomale Abschnitt auf Chromosom 2, mit intronischer Lage der miR-149 innerhalb von GPC. Mit freundlicher Genehmigung des UCSC Genome Browsers.**



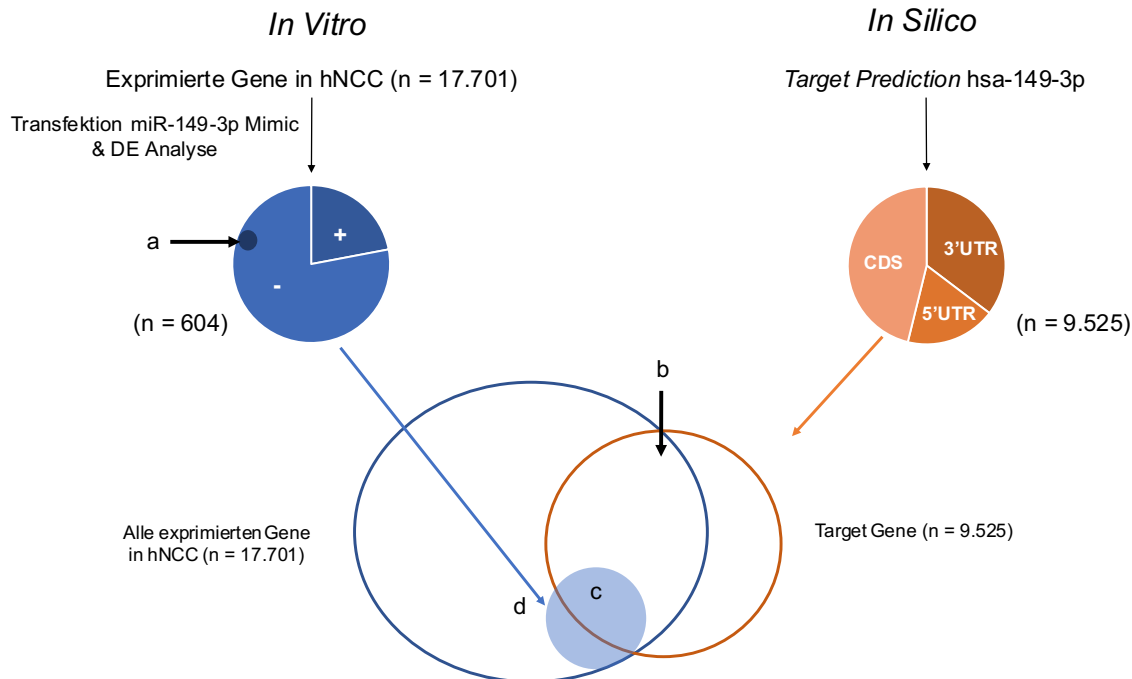
**Abb. 5: miR-149 Locus im UCSC Browser in höherer Auflösung mit Darstellung der Konservierung (modifiziert nach Stüssel et al., 2022 (Appendix Fig. S2), Kent et al., 2002).**

Dargestellt sind neben der humanen Sequenz mehrere Wirbeltierspezies (Rhesusaffe, Maus, Hund, Elefant, Huhn) und es zeigt sich anhand des PhyloP Scores eine hohe Konservierung des Locus. Mit freundlicher Genehmigung des UCSC Genome Browsers. [https://genome.ucsc.edu/s/laurastuessel/Abbildung%205%](https://genome.ucsc.edu/s/laurastuessel/Abbildung%205%20)

### 1.3.3 In-vitro- und in-silico-Untersuchungen zu miR-149

Nach Priorisierung der miR-149-3p (die 5p-Form war in den hNCC nicht exprimiert) stellte sich die Frage nach möglichen Pathomechanismen, durch die sie an der kraniofazialen Entwicklung teilnimmt. Die miR-149-3p-abhängige Veränderungen der hNCC-Migrationsfähigkeit wurde mithilfe eines Scratch Assays untersucht. In hNCC, die mit einem miR-149-3p mirVana Mimic transfiziert worden waren, zeigte sich eine signifikante Zunahme der Migration im Vergleich zu nicht behandelten Kontrollen (Abb. 2A-C der Veröffentlichung, Seite 48). Dies zeigte sich sowohl nach 30 Stunden ( $P = 0,0458$ , T-Test) als auch nach 40 Stunden ( $P = 0,0085$ ). Nach miR-149-3p Inhibierung ergaben sich keine signifikanten Unterschiede zwischen transfizierten Zellen und Kontrollen.

Um mögliche regulatorische Pfade zu identifizieren, die an dieser Veränderung der Migrationsfähigkeit der hNCC beteiligt sein könnten, wurden potenzielle Ziel-Gene von miR-149-3p in vitro und in silico betrachtet: In der 3'RNA-Seq der mit miR-149-3p transfizierten hNCC zeigten sich im Vergleich zu unbehandelten hNCC 604 differentiell exprimierte Gene (Abb. 6, Seite 21 und Abb. 2D der Veröffentlichung, Seite 48). Die Target Prediction für miR-149-3p konnte insgesamt 9.525 Gene als Ziel-Gene identifizieren, in Abgleich mit den RNA-Seq-Daten waren 189 dieser Gene nach miR-149-3p Überexpression herunterreguliert worden (negativer foldChange, Details s. Abb. 6, modifiziert nach Stüssel et al., 2022, Figure 2E, Seite 48).



**Abb. 6: Ergebnisse der in vitro und in silico Untersuchungen zu miR-149-3p (modifiziert nach Stüssel et al., 2022)**

**In vitro.** In der RNA-Seq der hNCC-Proben konnten insgesamt 17.701 Gene als exprimiert identifiziert werden. Nach Transfektion der Zellen mit miR-149-3p Mimic wurden insgesamt 604 Gene als differentiell exprimiert (DE) identifiziert. Davon waren 133 heraufreguliert (+) und 471 herabreguliert (-). In den herabregulierten Genen waren 17 enthalten, die zuvor mit der kraniofazialen Entwicklung in Verbindung gebracht wurden (a). **In silico.** Die Target Prediction ergab 4.084 Ziel-Genen in der 3'UTR, 2.142 in der 5'UTR und 5.337 in der kodierenden Sequenz (CDS). Von diesen insgesamt 9.525 Genen waren 6.497 auch in den hNCC exprimiert (b). Von den 604 DE Genen (d) waren 236 (c) unter den vorhergesagten Genen.

Zur funktionellen Einordnung der differentiell exprimierten Gene wurde eine Gene Ontology (GO) Term Analyse durchgeführt. Diese ergab signifikante Ergebnisse für verschiedene embryologische Entwicklungen (z.B. „anatomical structure morphogenesis“). Die 604 DE Gene wurden für 924 GO-Gen-Sets als signifikant angereichert identifiziert. Nach Einschränkung auf spezifischere molekulare Pathways konnten 138 Gen-Sets mit signifikanter Überrepräsentation der differentiell exprimierten Gene beobachtet werden. Zu diesen Pathways zählten unter anderem gesichtsentwicklungsrelevante GO-Terms wie „Regulation der Knorpelentwicklung“, „NCC Differenzierung“ und „NCC Entwicklung“ (Abb. 2E der Veröffentlichung, Seite 45).

#### 1.4 Diskussion

Dieser Arbeit liegt die Hypothese zugrunde, dass miRNAs, die in humanen NCC exprimiert sind, eine Funktion während der Gesichtsentwicklung übernehmen und dass Varianten an Loci solcher miRNAs zu kraniofazialen Fehlbildungen wie nsCL/P beitragen können. Es war das Ziel, solche miRNAs bzw. Loci zu finden und mögliche zugrundeliegende Mechanismen für die Ausbildung des nsCL/P Phänotypes zu beleuchten. Durch unsere Untersuchungen ist es gelungen, eine größere Gruppe an humanen miRNAs (n=151) zu identifizieren, die in hNCC exprimiert sind. Mit miR-149 wurde eine Kandidaten-miRNA mit Einfluss auf die Migrationsfähigkeit von hNCC identifiziert, für deren Locus es erste Anhaltspunkte für die Beteiligung an nsCL/P gibt.

Zunächst wurde eine Microarray-basierte Expressionsanalyse durchgeführt. Diese konnte in glaubhaft richtig differenzierten NCC erhoben werden. Allerdings ist zu bemerken, dass die benutzten Proben einen unerwartet niedrigen Übereinstimmungsgrad mit den öffentlich zugänglichen Transkriptomdaten anderer NCC-Samples aufwiesen. Dies kann zum einen an der geringen Größe unseres Samples (n=4), dem ähnlichen genetischem Background der Donoren (kaukasisch) und dem spezifischen genutzten Protokoll, bei dem die NCC aus hiPSC und nicht direkt aus embryonalem Gewebe gewonnen wurden, zusammenhängen. Es liegt nahe, dass es in dieser kleinen und nicht vollständig repräsentativen Gruppe zu falsch-negativen Ergebnissen gekommen ist, was in Teilen durch die Hinzunahme des „literature batches“ ausgeglichen werden sollte. Auffällig war jedoch der geringe Überschneidungsgrad zwischen beiden „batches“ (13 miRNAs, 12% des „hNCC“ batches, siehe Tab. 1 der Veröffentlichung, Seite 47). Dies lässt sich zum einen dadurch erklären, dass die Evidenz zu miRNAs aus der Literaturrecherche vorrangig, wenn auch nicht ausschließlich, in nicht-menschlichen Spezies (Zebrafisch, Maus, Vogelspezies) generiert wurde. Weiterhin war zum Zeitpunkt der Literaturrecherche nur eine geringe Zahl an Studien zu nsCL/P und miRNAs durchgeführt worden, wodurch die Suchbegriffe nur ein unvollständiges Bild abgreifen konnten. In der Zwischenzeit konnten neuere Studien, zunehmend auch mithilfe von humanen Bioproben, neue Kandidaten-miRNAs identifizieren, die unserer Literaturrecherche noch entgangen

sind. Es wurde versucht, die Anzahl falsch-positiver Kandidaten-miRNAs durch die Anlage strenger „Aufnahmekriterien“ für beide „batches“ zu reduzieren. Dennoch ist davon auszugehen, dass nicht alle miRNAs, die als exprimiert identifiziert wurden, an der Gesichtsentwicklung beteiligt sein müssen. NCC sind ausgesprochen multipotent und werden während der embryonalen Entwicklung u.a. auch zu Zellen des peripheren Nervensystems, zu Melanozyten oder zu kardialen Bindegewebszellen differenziert (Dupin et al., 2018). Es ist durch die hier durchgeführten Experimente nicht klar zu unterscheiden, ob manche der exprimierten miRNAs nicht eher an diesen Prozessen beteiligt sind, worauf es im weiteren Verlauf auch bei den Ergebnissen der DE- und GO-Term-Analyse Hinweise gab. Die Aussagekraft dieses Teils der Untersuchungen ließe sich zum Beispiel durch Replikation in einem größeren, heterogeneren Sample sowie durch die Nutzung von kranialen NCC erhöhen. Diese Subgruppe von NCC ist spezifisch für die Gesichtsentwicklung und kann ebenfalls aus pluripotenten Zellen differenziert werden (Barembaum und Bronner, 2013; Kobayashi et al., 2022).

Für die weiteren Untersuchungen sind wir davon ausgegangen, dass einige der potenziell an der Gesichtsentwicklung beteiligten miRNAs, auch an Fehlbildungen des Gesichts, konkret an nsCL/P, beteiligt sein können. Dies kann z.B. durch genetische Variation am Locus der miRNA vermittelt sein. Um solche Varianten aufzudecken, ist es erforderlich, die genetischen Daten von nsCL/P Patienten zu betrachten. Uns lagen diese in Form einer bereits veröffentlichte GWAS zu nsCL/P vor (Ludwig et al., 2017). GWAS ermöglichen es, über das gesamte Genom verteilte Allele zu identifizieren, die signifikant häufiger bei Betroffenen einer Erkrankung auftreten als bei Nicht-Betroffenen. Dabei können arraybasiert sowohl sehr große Patientenkollektive als auch mehrere Hunderttausend SNV untersucht werden. Durch Integration mit bekannten Haplotypen aus Referenzgruppen, also Allelgruppen, die gemeinsam vererbt werden, kann die Menge an identifizierten SNVs weiter erhöht werden (Imputation, Marchini et al., 2010). Die große Menge an Vergleichspunkten macht es wahrscheinlich, dass es unter standardmäßig angelegten Signifikanzniveaus ( $p < 0,05$  oder  $p < 0,01$ ) zu einer hohen Rate an falsch-positiven Ergebnissen kommt. Daher wird bei GWAS ein strengeres und für das multiple Testen korrigierte Signifikanzlevel, die sogenannte genomweite Signifikanz ( $p < 5 \times 10^{-8}$ ), angewandt (Chen et al., 2021). Bei der Untersuchung der Loci der Kandidaten-

miRNAs konnte nur eine sehr geringe Anzahl an nominell, also nicht genomweit, signifikanten SNVs (25 von 13 002 Varianten an Kandidaten-Loci) gefunden werden. Es kommen dafür mehrere mögliche Erklärungen in Frage: Zum einen kann vermutet werden, dass miRNAs im Allgemeinen oder die von uns identifizierten Kandidaten keinen signifikanten Einfluss auf die Entstehung von nsCL/P haben. Zum anderen ist es möglich, dass Varianten im Bereich von miRNAs aufgrund des hohen Konservierungsgrades und der hohen Spezifität der miRNAs selten vorkommen. Seltene Varianten entgehen häufig GWAS, da sie auch in den untersuchten Kollektiven nur selten auftreten und somit schwerer statistische Signifikanz erreichen. Zudem basieren GWAS auf Arrays mit vorgegeben SNV und Referenzgenomen zur Imputation, die nicht alle seltenen Allele bzw. Haplotypen abdecken können (Lee et al., 2014). Das genutzte Referenzgenom, 1000 Genomes, bildet die genetische Variation von hunderten Europäern ab, aber natürlich nicht alle potenziell existierenden Allele. Dennoch spielen seltene Varianten eine wichtige Rolle bei der Entstehung häufiger, multifaktorieller Erkrankungen wie nsCL/P. Sie können bei betroffenen Individuen entscheidend zur Ausprägung des Phänotyps beitragen (Ludwig et al., 2019). Ein Weg seltene Varianten besser zu identifizieren, wäre beispielsweise die Zuhilfenahme von Next Generation Sequencing, bei dem komplette Exome oder Genome bzw. die Kandidatenregionen sequenziert werden können oder die Verbesserung der Imputation mithilfe größerer Referenzpanels, die zunehmend verfügbar sind (Das et al., 2018). Zuletzt erscheint es auch denkbar, dass mögliche Effekte von NCC-exprimierten miRNAs auf nsCL/P nicht durch Variation am miRNA-Locus selbst, sondern durch Variation im Bereich regulierender Elemente vermittelt sind. Während manche (cis-) regulierende Elemente wie Promotoren in direkter Umgebung des miRNA-Locus zu finden sind, können z.B. Enhancer mehrere Tausend bp entfernt sein. Sogenannte trans-regulierende Elemente, wie die Loci von Transkriptionsfaktoren, können sogar auf anderen Chromosomen liegen (Fang et al., 2020). Transkriptionsfaktoren, die dann DNA im Bereich der miRNA binden, können ihre Transkription hoch- oder herunterregulieren. Varianten dieser können über eine veränderte DNA-Bindung zu veränderten miRNA-Leveln führen (Gulyaeva and Kushlinskiy, 2016). Durch das gesetzte Fenster von 1kb um den miRNA-Locus besteht die Gefahr, dass Varianten in cis-regulierenden Elementen nicht detektiert wurden. Gleichzeitig nimmt die Wahrscheinlichkeit zu, dass Signale in einem größeren Fenster anderen Genen oder regulierenden



Elementen zuzuordnen sind. Daher haben wir uns zur Vermeidung falsch-positiver Ergebnisse für dieses konservative Fenster entschieden. Trans-regulatorische Effekte wären auch einem größeren Fenster entgangen und ihre Detektion würde weitere ergänzende Untersuchungen, wie beispielsweise Chromatin Immunpräzipitation (ChIP) erfordern.

Für lediglich eine SNV ergab sich in den Replikationsuntersuchungen Anhalt für Signifikanz. Dabei ist zu bemerken, dass ein großer Anteil der Kandidaten-SNVs aufgrund des Designs und der Qualitätskontrolle des Experiments nicht untersucht werden konnte. Ein optimiertes Design bzw. die Replikation in einem anderen Kollektiv könnte eventuell für diese Varianten Signifikanz erbringen.

Für die weiter verfolgte Variante rs149251515 ergaben sich nur nominell signifikante P-Werte. Auch der Abgleich mit zwei weiteren nsCL/P Kollektiven konnte keine eindeutige Signifikanz erbringen. Es kommen mehrere Gründe für diese Ergebnisse in Frage:

1.) erscheint es möglich, dass der miR-149-Befund aus der Expressionsanalyse ein falsch positiver ist. Dies konnte jedoch durch Hinzunahme von anderen öffentlich zugänglichen miRNA-Expressionsdaten entkräftet werden. Die miRNA war ebenfalls in murinen Gesichtsgeweben in entsprechend frühen Entwicklungsstadien exprimiert (Brinkley et al., 2016; Ding et al., 2016; Mukhopadhyay et al., 2011).

2.) Die eher geringe Allelfrequenz (3%) und begrenzte Samplegröße der Replikationsstudie kann ein Grund für die geringe Signifikanz der Variante sein. Wie oben diskutiert, ist es nicht überraschend, dass Varianten im Bereich von miRNAs eher selten und somit Untersuchungen in kleineren Kollektiven entgehen. Eine Replikation in einem größeren Sample und die direkte Sequenzierung der Variante würden helfen, die Ergebnisse besser einzuordnen. Um wirkliche Aussagen über mögliche Effekte der Variante zu treffen, wäre es notwendig, Studien zu sogenannten expression quantitative trait loci (eQTL) durchzuführen. Dabei wird versucht, den entsprechenden Genotyp mit Veränderungen von mRNA-Leveln in für den Phänotyp-relevanten Geweben zu korrelieren und dabei Patienten und Kontrollen zu vergleichen (Li et al., 2022; Westra und Franke, 2014). Da-

für wären in diesem Falle insbesondere wegen der niedrigen Allelfrequenz von rs14925151 sehr große Patientenkohorten erforderlich (Baccarella et al., 2018).

3.) Kann aufgrund der oben beschriebenen Einschränkungen, die mit dem engen Fenster von 1kb um den miRNA Locus einhergehen, gemutmaßt werden, dass rs14925151 nicht unbedingt die relevante Variante ist, sondern nur ein „Marker“ für andere Effekte im Bereich miR-149-Locus. Es konnte downstream der Variante ein assoziiertes Cluster häufiger Varianten identifiziert werden (Abb. 3, Seite 19). Diese Varianten liegen innerhalb derselben topologically associating domain (TAD) wie miR-149. TADs sind 3D Strukturen der DNA, innerhalb derer die enthaltenen Sequenzen stärker miteinander interagieren (z.B. in Form von Enhancern und Promotern) als mit Regionen, die nicht in der TAD liegen (Tena und Santos-Pereira, 2021). Somit wäre eine Beeinflussung von miR-149 oder Glypican 1 (GPC1), zu dem miR-149 intronisch liegt, durch diese häufigen Varianten möglich.

In Zusammenschau dieser Erklärungsmöglichkeiten für die ambivalenten Ergebnisse zu rs14925151 wurde sich dafür entschieden, dennoch weitere Untersuchungen zur möglichen Funktionalität von miR-149 in Bezug auf die kraniofaziale Entwicklung anzustellen. Aus der Target Prediction ging zunächst eine sehr hohe Anzahl potenzieller Ziel-Gene hervor, z.B. 4.084 für den kanonischen Weg. Die geringe Spezifität von bioinformatischen Tools zur Target Prediction ist ein bekanntes Problem (Fridrich et al., 2019), und es ist somit davon auszugehen, dass unter den vorausgesagten Ziel-Genen eine große Anzahl an falsch Positiven ist. Strengere Kriterien und Optimierung der Algorithmen z.B. durch Kombination mit validierten Ziel-Genen aus in vitro Analysen können die Anzahl der falsch Positiven weiter reduzieren. Im Rahmen unserer Untersuchungen wurden DE-Analysen in NCC durchgeführt, um mögliche Targets zu validieren. 88% der DE-exprimierten Gene waren herabreguliert, übereinstimmend mit der Annahme, dass miRNAs in der Regel ihre Ziel-Gene durch Abbau der exprimierten miRNAs silencen. Es zeigte sich ein plausibler Überschneidungsgrad zwischen herabregulierten und predicteten Genen (n=189). In der Vergangenheit konnte gezeigt werden konnte, dass intronisch gelegene miRNAs häufig an der Regulation ihrer Host-Gene beteiligt sind (Chamorro-Jorganes et al., 2014; Zeidler et al., 2020). Daher lag es nahe, zunächst diese Interaktion genauer zu untersuchen. Während GPC1 in der Target Prediction als Ziel-Gen identi-

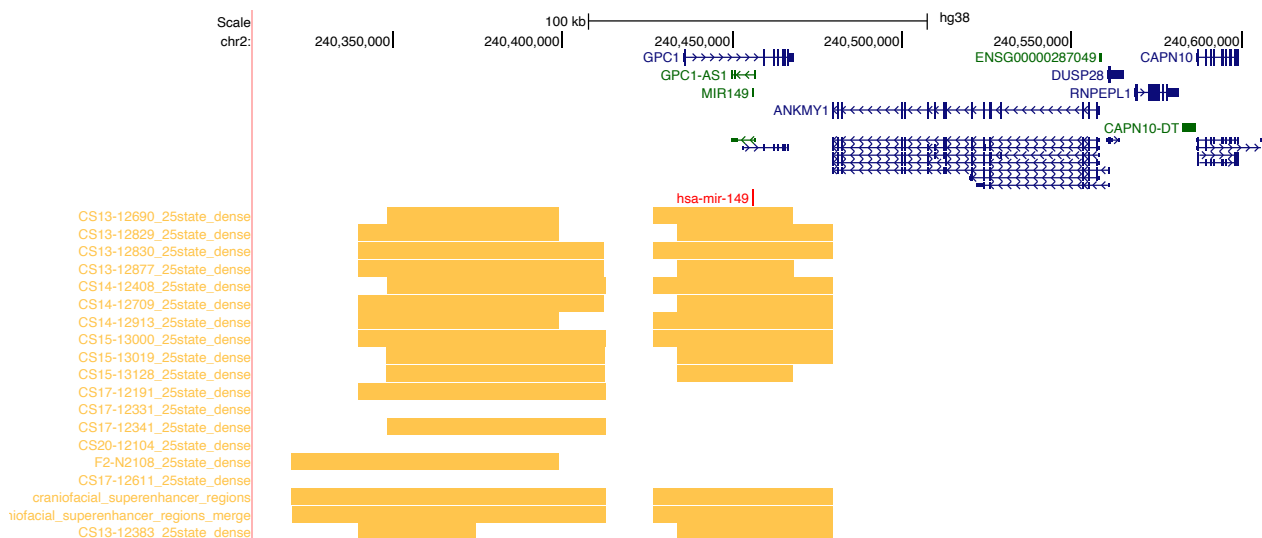
fiziert wurde, war die beobachtete Veränderung der Menge an GPC1 mRNA nach miR-149 Überexpression nicht signifikant. Expressionsdaten aus Mäusen und hNCC zeigten für alle Gene, die der TAD von rs14921515 zuzuordnen sind, nur geringe Expressionen für alle Gene außer *GPC1*. Von allen Genen, die der TAD zuzuordnen waren, wurden etwa 10% auch in der Target Prediction vorhergesehen, keines davon war DE expriert. MiRNAs können natürlich auch mit Ziel-Genen außerhalb einer TAD interagieren. Es ist daher wahrscheinlich, dass die Effekte von miR-149, zumindest im Kontext der genutzten NCC, nicht auf einer Interaktion mit GPC1 oder Genen innerhalb der TAD beruhen, sondern auf Interaktionen mit Genen außerhalb dieses Bereichs.

Unter den 471 DE herabregulierten Genen waren 17 Gene enthalten, die bereits im Rahmen genetischer bzw. funktioneller Untersuchungen als relevant für kraniofaziale Pathologien beschreiben worden sind: Dazu gehören beispielsweise Noggin (*NOG*) und Gremlin 1 (*GREM1*), welche an nsCL/P Suszeptibilitäts-Loci liegen (Ludwig et al., 2016; Mangold et al., 2009), sowie Bone morphogenic protein 4 (*BMP4*). Frühere Studien in Mäusen konnten zeigen, dass *Bmp4* an einem Netzwerk mit *Msx1/2*, *Pax9* (paired box 9) und *Grem1* beteiligt ist, welches das Wachstum der MNP im Rahmen der Lippenentwicklung beeinflusst und zu Spaltbildungen führen kann. Mitglieder dieses Netzwerkes (*MSX1/2*, *PAX* und *GREM1*) konnten durch Integration mit humanen GWAS-Daten mit nsCL/P in Verbindung gebracht werden, *PAX9* im Speziellen mit dem Subphänotypen nsCLP, also einer Spalte der Lippe und des Gaumens (Nakatomi et al., 2020). Weiterhin wurden z.B. zwei Mitglieder der Fibroblast Growth Factor (FGF)-Familie, *FGF1* und *FGF23*, als herabreguliert detektiert. Die Rolle von FGF und ihren Rezeptoren ist sowohl für die NCC-Entwicklung (Zhang et al., 2018) als auch für Gesichtsentwicklung und Spaltbildungen im Speziellen bekannt (Stanier und Pauws, 2012). Zudem wurde Transforming Growth Factor Beta 2 (*TGFB2*) identifiziert. Die Rolle von Mitgliedern der TGF-Beta-Familie ist für die Gesichtsentwicklung schon lange beschrieben (Dudas und Kaartinen, 2005), kürzlich wurde *TGFB2* als Target von miR-193-3p eine miRNA-vermittelte Rolle in der kraniofazialen Entwicklung von Mäusen attestiert (Fu et al., 2021). Im Gegensatz dazu sind die DE Gene mit der höchsten Signifikanz (Transgelin (*TAGLN*) und Ankyrin Repeat Domain-Containing Protein 1 (*ANKRD1*)) sowie die Gene mit der deutlichsten Herabregulation unter miR-149-3p Überexpression (negativer FoldChange, Ac-

tin Alpha Cardiac Muscle 1 (ACTC1) und Crystallin Alpha B (CRYAB)) bisher nicht im Zusammenhang mit kraniofazialen Prozessen beschrieben worden. TAGLN und ANKRD1 wurden auch in der Target Prediction identifiziert, die beiden anderen Gene nicht. TAGLN kodiert für einen mesenchymalen Marker und beeinflusst unter anderem die epithelial-mesenchymale Transition und Zellmigration (Chen et al., 2019; Daniel et al., 2012). Interessanterweise wurde seine Beteiligung auch am Verschluss der ventralen Mittellinie in der Embryonalzeit (Aldeiri et al., 2018) und an der Proliferation sowie Migration enterischer NCC im Rahmen des Morbus Hirschsprung (Wang et al., 2021) beschrieben. ANKRD1 wird eine Rolle vor allem im Bereich der Herzentwicklung und diverser Kardiomyopathien und Herzinsuffizienz zugeschrieben (Ling et al., 2017; Mikhailov und Torrado, 2008). Untersuchungen zu ANKRD1 in NCC liegen bisher nicht vor, allerdings wurde für ein weiteres Mitglied der ANKRD-Gengruppe, ANKRD13, eine miR-204-abhängige Dysregulation der NCC-Migration in Medeka-Fischen festgestellt (Avellino et al., 2013). Für ACTC1 und CRYAB wurden Rollen für Kardiomyopathien und verschiedene Malignome identifiziert (Jordan et al., 2021; Zhang et al., 2019).

Die Ergebnisse der DE- und GO-Term-Analyse sowie die unter miR-149-3p Überexpression beobachtete vermehrte Migrationsfähigkeit der NCC können keinen eindeutigen Hinweis auf ihre Mitwirkung an kraniofazialen Prozessen geben. Es ist zweifellos zu diskutieren, ob die beobachteten Effekte von miR-149 weniger spezifisch auf die Gesichtsentwicklung, sondern allgemeiner gesehen auf NCC-vermittelten Entwicklungsprozesse zu beziehen sind. Zwar führten veränderte migratorische Eigenschaften von NCC in früheren Studien zur Ausbildung von kraniofazialen Phänotypen (Chiquet et al., 2018; Dougherty et al., 2012; Eberhart et al., 2008). Konkret für miR-149 konnte aber auch in Endothelzellen bzw. in humanen aortalen glatten Muskelzellen, einem NCC-Derivat, eine miR-149 abhängige gesteigerte Migrationsfähigkeit gezeigt werden (Chamorro-Jorganes et al., 2014, Lin et al., 2021). Im Zusammenhang mit kraniofazialen Prozessen existieren für miR-149 bisher keine weiteren Studien. Aufgrund der bereits diskutierten Multipotenz der von uns genutzten NCC kann dies unter unserem experimentellen Design nicht klar differenziert werden. Um dieses Problem anzugehen, kommen für weiterführende Untersuchungen verschiedene Ansätze in Frage: Beispielsweise wären in vit-

ro-Untersuchungen in spezifischeren humanen Geweben wie kranialen NCC oder in frühem embryonalen Gesichtsgewebe oder in vivo-Versuche mit Modellorganismen erforderlich. Uns lagen öffentlich zugängliche ChIP-Seq-Daten aus frühen menschlichen Embryonalstadien vor (Wilderman et al., 2018). Bei ChIP-Seq wird die Modifikation von Histonen, DNA-bindenden Proteinen, untersucht. Bestimmte Modifikationsmuster an einem Locus können u.a. dafür sprechen, dass diese Region an der „Aktivierung“ der Transkription weiter entfernter Gene beteiligt ist, und somit als Enhancer fungiert (Pott und Lieb, 2015). Diese Muster sind häufig spezifisch für bestimmte Gewebe oder Zeitpunkte der Entwicklung und dementsprechend ggf. spezifisch für Entwicklungsprozesse an diesem Ort oder zu dieser Zeit (Wilderman et al., 2018). Der Region um den miR-149-Locus wird eine starke regulatorische Aktivität zugeschrieben und sie wird als kraniofaziale Superenhancer-Region klassifiziert (Abb. 7).



**Abb. 7: UCSC-Track des miR-149 Locus zeigt die Lage innerhalb einer kraniofazialen Superenhancer Region (Kent et al., 2002, Wilderman et al., 2018).** <https://genome.ucsc.edu/s/laurastuessel/Abbildung%207%20St%C3%BCssel>

Superenhancer sind dabei als Regionen oft von mehreren nah gelegenen Enhancern definiert, die eine ungewöhnlich starke Bindung bestimmter Transkriptionsfaktoren zeigten (Pott and Lieb, 2015; Whyte et al., 2013). Neben miR-149 liegt auch GPC1 in der Superenhancer-Region, die miRNA liegt dabei intronisch zu GPC1. Innerhalb des etwa 40kb langen Superenhancers liegt ansonsten nur noch eine lange nicht kodierende RNA (lncRNA). Die Lage von miR-149 innerhalb einer solchen Region kann somit für eine

kraniofaziale Rolle sprechen, beweist sie aber nicht. Dazu wäre es weiterführend wichtig, die in vivo-Effekte einer veränderten miR-149 Expression in Modelltieren wie Zebrafischen oder Mäusen auf kraniofaziale Phänotypen zu untersuchen.

Zusammenfassend bleiben somit insbesondere in Bezug auf die genetischen Daten zu rs14921515 und die funktionelle Relevanz von miR-149 bei der Gesichtsentwicklung noch wesentliche Fragen offen: Sind die bisher erhobenen genetischen Befunde der GWAS- und Replikationsuntersuchungen auf die geringe Allelfrequenz und die vergleichsweise kleinen Versuchskohorten zurückzuführen oder tatsächlich nicht ausreichend signifikant? Können die beobachteten Effekte von miR-149 in NCC auf die Gesichtsentwicklung bezogen werden oder eher auf allgemeine NCC-vermittelte Prozesse? Wie genau interagieren rs14921515 und miR-149 und durch welche molekularen Prozesse greift miR-149 in die Migration der NCC ein? Führt die veränderte Migrationsfähigkeit der NCC zu einem kraniofazialen Phänotyp? Um diese Fragen zu beantworten, sind weiterführende Studien wie die direkte Sequenzierung des Locus, eQTL-Studien, in vivo-Untersuchungen und die Zuhilfenahme anderer Gewebe, wie z.B. kraniale NCC, erforderlich. Dennoch stellt diese Arbeit ein wichtiges Puzzleteil bei der Untersuchung miRNA-vermittelter kraniofazialer Prozesse dar. Sie zeichnet sich insbesondere durch die Nutzung humanen Gewebes und die Integration mit genetischen Daten zu nsCL/P aus. Die grundsätzliche Herangehensweise ist auch für andere Pathologien, für die relevante Gewebe und GWAS-Daten vorliegen, denkbar, und die generierte Liste an Kandidaten-miRNAs kann auch für andere NCC-vermittelte Pathologien genutzt werden.

## 1.5 Zusammenfassung

Die vorliegende Studie befasst sich mit der Untersuchung von miRNAs im Zusammenhang mit der Gesichtsentwicklung und fragt, inwieweit diese eine Rolle bei der Ausbildung der nsCL/P (non-syndromic cleft lip with or without cleft palate), einer häufigen angeborenen Fehlbildung des Menschen, spielen. Besonderer Wert wurde bei der Identifizierung relevanter miRNAs auf die Nutzung humaner Gewebe gelegt, um die Integration mit genetischen Daten zu nsCL/P zu ermöglichen. Mittels systematischer Expressionsanalysen in Neuralleistenzellen (neural crest cells, NCC) wurde miR-149 als mögliche Kandidaten-miRNA in NCC-vermittelten molekularen Prozessen identifiziert. Durch die Analyse von umfangreichen genetischen Daten von nsC/P-Patienten und Kontrollkohor-

ten konnte Evidenz für einen Zusammenhang zwischen einer nah am miRNA149-liegenden genetischen Variante (rs4921515) und dem erhöhten Risiko für eine Spaltbildung im Gesicht erhalten werden. Als potenzieller Pathomechanismus wurde dabei durch in vitro-Manipulation von NCC die Beeinflussung der Zellmigration nahegelegt. Die Studie, welche durch den Einsatz verschiedener state of the art-Technologien wie z.B. in vitro-Manipulation induzierter pluripotenter Stammzellen und molekularer Hochdurchsatz-Analysen ermöglicht wurde, gibt somit Hinweise auf die Rolle der miRNA-149 in der humanen Gesichtsentwicklung. Die Ergebnisse bedürfen jedoch weiterer unabhängiger Untersuchungen, um die erhobenen genetischen Befunde zu validieren und die Rolle von miR-149 in Bezug auf die kraniofaziale Entwicklung genauer zu verstehen. Der generierte Datensatz zur miRNA-Expression in NCC kann für weitere explorative Analysen, auch in anderen NCC-vermittelten Phänotypen, verwendet werden.

## 1.6 Literaturverzeichnis der deutschen Zusammenfassung

Aldeiri B, Roostalu U, Albertini A, Behnsen J, Wong J, Morabito A, Cossu G. Abrogation of TGF-beta signalling in TAGLN expressing cells recapitulates Pentalogy of Cantrell in the mouse. *Sci Rep.* 2018;8(1):3658

Avellino R, Carrella S, Pirozzi M, Risolino M, Salierno FG, Franco P, Stoppelli P, Verde P, Banfi S, Conte I. miR-204 Targeting of Ankrd13A Controls Both Mesenchymal Neural Crest and Lens Cell Migration. *PLoS One* 2013;8(4):e61099

Baccarella A, Williams CR, Parrish JZ, Kim CC. Empirical assessment of the impact of sample number and read depth on RNA-Seq analysis workflow performance. *BMC Bioinformatics.*2018;19(1):423

Barembaum M, Bronner ME. Identification and dissection of a key enhancer mediating cranial neural crest specific expression of transcription factor, Ets-1. *Dev Biol* 2013;382(2):567–575

Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, Liang KY, Wu T, Murray T, Fallin MD, Redett RA, Raymond G, Schwender H, Jin SC, Cooper ME, Dunnwald M, Mansilla MA, Leslie E, Bullard S, Lidral AC, Moreno LM, Menezes R, Vieira AR, Petrin A, Wilcox AJ, Lie RT, Jabs EW, Wu-Chou YH, Chen PK, Wang H, Ye X, Huang S, Yeow V, Chong SS, Jee SH, Shi B, Christensen K, Melbye M, Doheny KF, Pugh EW, Ling H, Castilla EE, Czeizel AE, Ma L, Field LL, Brody L, Pangilinan F, Mills JL, Molloy AM, Kirke PN, Scott JM, Arcos-Burgos M, Scott AF. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. *Nat Genet* 2010;42(6):525-529

Beaty TH, Taub MA, Scott AF, Murray JC, Marazita ML, Schwender H, et al. Confirming genes influencing risk to cleft lip with/without cleft palate in a case-parent trio study. *Hum Genet* 2013; 132(7):771–781



Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, Rubini M, Baluardo C, Ferrian M, Almeida de Assis N, Alblas MA, Barth S, Freudenberg J, Lauster C, Schmidt G, Scheer M, Braumann B, Bergé SJ, Reich RH, Schiefke F, Hemprich A, Pötzsch S, Steegers-Theunissen RP, Pötzsch B, Moebus S, Horsthemke B, Kramer FJ, Wienker TF, Mossey PA, Propping P, Cichon S, Hoffmann P, Knapp M, Nöthen MM, Mangold E. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. *Nat Genet* 2009;41(4):473–477

Brinkley JF, Fisher S, Harris MP, Holmes G, Hooper JE, Jabs EW, et al. The FaceBase Consortium: a comprehensive resource for craniofacial researchers. *Development*. 2016;143(14):2677–2688

Broughton JP, Lovci MT, Huang JL, Yeo GW, Pasquinelli AE. Pairing beyond the Seed Supports MicroRNA Targeting Specificity. *Mol Cell* 2016 Oct;64(2):320–33

Carlson JC, Shaffer JR, Deleyiannis F, Hecht JT, Wehby GL, Christensen K, Feingold E, Weinberg SM, Marazita ML, Leslie EJ. Genome-wide Interaction Study Implicates VGLL2 and Alcohol Exposure and PRL and Smoking in Orofacial Cleft Risk. *Front Cell Dev Biol* 2022;10:621261

Carter CO, Evans K, Coffey R, Roberts JA, Buck A, Roberts MF. A three generation family study of cleft lip with or without cleft palate. *J Med Genet* 1982;19(4):246–261

Chamorro-Jorganes A, Araldi E, Rotllan N, Cirera-Salinas D, Suarez Y. Autoregulation of glypican-1 by intronic microRNA-149 fine tunes the angiogenic response to FGF2 in human endothelial cells. *J Cell Sci*. 2014;127(Pt 6):1169-1177

Chen Z, Boehnke M, Wen X, Mukherjee B. Revisiting the genome-wide significance threshold for common variant GWAS. *G3 (Bethesda)* 2021;11(2):jkaa056

Chiquet BT, Yuan Q, Swindell EC, Maili L, Plant R, Dyke J, Boyer R, Teichgraeber JF, Greives MR, Mulliken JB, Letra A, Blanton SH, Hecht JT. Knockdown of *Crispld2* in

zebrafish identifies a novel network for nonsyndromic cleft lip with or without cleft palate candidate genes. *Eur J Hum Genet* 2018;26(10):1441-1450

Christensen K, Juel K, Herskind AM, Murray JC. Long term follow up study of survival associated with cleft lip and palate at birth. *BMJ* 2004; 328:1405

Christensen K, Fogh-Andersen P. Cleft lip (+/- cleft palate) in Danish twins, 1970-1990. *Am J Med Genet* 1993;47(6):910–916

Colpaert RMW, Calore M. MicroRNAs in Cardiac Diseases. *Cells* 2019;8(7):737

Cordero DR, Brugmann S, Chu Y, Bajpai R, Jame M, Helms JA. Cranial neural crest cells on the move: Their roles in craniofacial development. *Am J Med Genet* 2011; Part A 155: 270–279

Daniel C, Lüdke A, Wagner A, Todorov VT, Hohenstein B, Hugo C. Transgelin is a marker of repopulating mesangial cells after injury and promotes their proliferation and migration. *Lab Invest* 2012;92(6):812–826

Das S, Abecasis GR, Browning BL. Genotype Imputation from Large Reference Panels. *Annu Rev Genomics Hum Genet* 2018;19(1):73–96

Ding HL, Hooper JE, Batzel P, Eames BF, Postlethwait JH, Artinger KB, Clouthier DE. MicroRNA Profiling during Craniofacial Development: Potential Roles for Mir23b and Mir133b. *Front Physiol.* 2016;7:281

Dougherty M, Kamel G, Shubinets V, Hickey G, Grimaldi M, Liao EC. Embryonic fate map of first pharyngeal arch structures in the *sox10*: kaede zebrafish transgenic model. *J Craniofac Surg* 2012;23(5):1333–1337

Dudas M, Kaartinen VBT-CT in DB. TGF- $\beta$  Superfamily and Mouse Craniofacial Development: Interplay of Morphogenetic Proteins and Receptor Signaling Controls Normal Formation of the Face. In Academic Press 2005; 65–133

Dupin E, Calloni GW, Coelho-Aguiar JM, Le Douarin NM. The issue of the multipotency of the neural crest cells. *Dev Biol* 2018;444 Suppl:S47–59

Eberhart JK, He X, Swartz ME, Yan YL, Song H, Boling TC, Kunerth AK, Walker MB, Kimmel CB, Postlethwait JH. MicroRNA Mirn140 modulates Pdgf signaling during palatogenesis. *Nat Genet* 2008;40(3):290–298

Eulalio A, Huntzinger E, Izaurralde E. Getting to the Root of miRNA-Mediated Gene Silencing. *Cell* 2008;132(1):9–14

Fang W, Wen Y, Wei X. Identification and Characterization of Cis-Regulatory Elements for Photoreceptor-Type-Specific Transcription in ZebraFish. *Methods Mol Biol* 2020;2092:123–145

Fridrich A, Hazan Y, Moran Y. Too Many False Targets for MicroRNAs: Challenges and Pitfalls in Prediction of miRNA Targets and Their Gene Ontology in Model and Non-model Organisms. *Bioessays* 2019;41(4):e1800169

Fu C, Lou S, Zhu G, Fan L, Yu X, Zhu W, Ma L, Wang L, Pan Y. Identification of New miRNA-mRNA Networks in the Development of Non-syndromic Cleft Lip With or Without Cleft Palate. *Front cell Dev Biol* 2021;9:631057

Grigaitis P, Starkuviene V, Rost U, Serva A, Pucholt P, Kummer U. miRNA target identification and prediction as a function of time in gene expression data. *RNA Biol* 2020;17(7):990–1000

Grosen D, Chevrier C, Skytthe A, Bille C, Mølsted K, Sivertsen A, Murray JC, Christensen K. A cohort study of recurrence patterns among more than 54,000 relatives of oral

cleft cases in Denmark: support for the multifactorial threshold model of inheritance. *J Med Genet* 2010; 47(3):162-168

Gulyaeva LF, Kushlinskiy NE. Regulatory mechanisms of microRNA expression. *J Transl Med* 2016; 14(1):143

Hochfeld LM, Anhalt T, Reinbold CS, Herrera-Rivero M, Fricker N, Nöthen MM, Heilmann-Heimbach S. Expression profiling and bioinformatic analyses suggest new target genes and pathways for human hair follicle related microRNAs. *BMC Dermatol.* 2017;17(1):3

Ji Y, Garland MA, Sun B, Zhang S, Reynolds K, McMahon M, Rajakumar R, Islam MS, Liu Y, Chen Y, Zhou CJ. Cellular and developmental basis of orofacial clefts. *Birth Defects Res* 2020; 112: 1558–1587

Jordan E, Peterson L, Ai T, Asatryan B, Bronicki L, Brown E, Celeghein R, Edwards M, Fan J, Ingles J, James CA, Jarinova O, Johnson R, Judge DP, Lahrouchi N, Lekanne Deprez RH, Lumbers RT, Mazzarotto F, Medeiros Domingo A, Miller RL, Morales A, Murray B, Peters S, Pilichou K, Protonotarios A, Semsarian C, Shah P, Syrris P, Thaxton C, van Tintelen JP, Walsh R, Wang J, Ware J, Hershberger RE. Evidence-Based Assessment of Genes in Dilated Cardiomyopathy. *Circulation* 2021;144(1):7-19

Kamburov A, Wierling C, Lehrach H, Herwig R. ConsensusPathDB--a database for integrating human functional interaction networks. *Nucleic Acids Res* 2009;37(Database issue):D623–628

Karanth TK, Whittlemore KR. Middle-ear disease in children with cleft palate. *Auris Nasus Larynx* 2018; 45: 1143–1151

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res* 2002;12(6):996–1006

Kobayashi Y, Nouet J, Baljinnyam E, Siddiqui Z, Fine DH, Fraidenraich D, Kumar VA, Shimizu E. iPSC-derived cranial neural crest-like cells can replicate dental pulp tissue with the aid of angiogenic hydrogel. *Bioact Mater* 2022;14:290–301

Laugsch M, Bartusel M, Rehim R, Alirzayeva H, Karaolidou A, Crispatzu G, Zentis P, Nikolic M, Bleckwehl T, Kolovos P, van Ijcken WFJ, Šarić T, Koehler K, Frommolt P, Lachlan K, Baptista J, Rada-Iglesias A. Modeling the Pathological Long-Range Regulatory Effects of Human Structural Variation with Patient-Specific hiPSCs. *Cell Stem Cell*. 2019;24(5):736-752.e12

Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. *Am J Hum Genet* 2014;95(1):5–23

Leslie EJ, Carlson JC, Shaffer JR, Buxó CJ, Castilla EE, Christensen K, Deleyiannis FWB, Field LL, Hecht JT, Moreno L, Orioli IM, Padilla C, Vieira AR, Wehby GL, Feingold E, Weinberg SM, Murray JC, Marazita ML. Association studies of low-frequency coding variants in nonsyndromic cleft lip with or without cleft palate. *Am J Med Genet A* 2017;173(6):1531-1538

Lewis CW, Jacob LS, Lehmann CU, Krol D, Gereige R, Karp J, Fisher-Owens S, Braun P, Jacob L, Segura A. The Primary Care Pediatrician and the Care of Children With Cleft Lip and/or Cleft Palate. *Pediatrics* 2017; 139: e20170628

Li L, Meng T, Jia Z, Zhu G, Shi B. Single nucleotide polymorphism associated with non-syndromic cleft palate influences the processing of miR-140. *Am J Med Genet A* 2010;152A(4):856–862

Lin D-S, Zhang C-Y, Li L, Ye G-H, Jiang L-P, Jin Q. Circ\_ROBO2/miR-149 Axis Promotes the Proliferation and Migration of Human Aortic Smooth Muscle Cells by Activating NF- $\kappa$ B Signaling. *Cytogenet Genome Res* 2021;161(8–9):414–424

Ling SSM, Chen Y-T, Wang J, Richards AM, Liew OW. Ankyrin Repeat Domain 1 Protein: A Functionally Pleiotropic Protein with Cardiac Biomarker Potential. *Int J Mol Sci* 2017;18(7)

Loh HY, Norman BP, Lai KS, Rahman NMANA, Alitheen NBM, Osman MA. The Regulatory Role of MicroRNAs in Breast Cancer. *Int J Mol Sci* 2019;20(19)

Ludwig KU, Ahmed ST, Böhmer AC, Sangani NB, Varghese S, Klamt J, Schuenke H, Gültepe P, Hofmann A, Rubini M, Aldhorae KA, Steegers-Theunissen RP, Rojas-Martinez A, Reiter R, Borck G, Knapp M, Nakatomi M, Graf D, Mangold E, Peters H. Meta-analysis Reveals Genome-Wide Significance at 15q13 for Nonsyndromic Clefting of Both the Lip and the Palate, and Functional Analyses Implicate *GREM1* As a Plausible Causative Gene. *PLoS Genet* 2016;12(3):e1005914

Ludwig KU, Böhmer AC, Bowes J, Nikolic M, Ishorst N, Wyatt N, Hammond NL, Gözl L, Thieme F, Barth S, Schuenke H, Klamt J, Spielmann M, Aldhorae K, Rojas-Martinez A, Nöthen MM, Rada-Iglesias A, Dixon MJ, Knapp M, Mangold E. Imputation of orofacial clefting data identifies novel risk loci and sheds light on the genetic background of cleft lip ± cleft palate and cleft palate only. *Hum Mol Genet* 2017; 26(4):829–842

Ludwig KU, Degenhardt F, Nöthen MM. Die Rolle seltener Varianten bei häufigen Krankheiten. *Med Genet* 2019;31(2):212–221

Mangold E, Reutter H, Birnbaum S, Walier M, Mattheisen M, Henschke H, Lauster C, Schmidt G, Schiefke F, Reich RH, Scheer M, Hemprich A, Martini M, Braumann B, Krimmel M, Opitz C, Lenz JH, Kramer FJ, Wienker TF, Nöthen MM, Diaz Lacava A. Genome-wide linkage scan of nonsyndromic orofacial clefting in 91 families of central European origin. *Am J Med Genet A* 2009;149A(12):2680-2694

Mangold E, Ludwig KU, Birnbaum S, Baluardo C, Ferrian M, Herms S, Reutter H, de Assis NA, Chawa TA, Mattheisen M, Steffens M, Barth S, Kluck N, Paul A, Becker J, Lauster C, Schmidt G, Braumann B, Scheer M, Reich RH, Hemprich A, Pötzsch S,

Blaumeiser B, Moebus S, Krawczak M, Schreiber S, Meitinger T, Wichmann HE, Steegers-Theunissen RP, Kramer FJ, Cichon S, Propping P, Wienker TF, Knapp M, Rubini M, Mossey PA, Hoffmann P, Nöthen MM. Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. *Nat Genet* 2010;42(1):24–26

Mangold E, Kreiß M, Nöthen MM. Syndromale und nichtsyndromale orofaziale Spalten. *Medizinische Genetik* 2017; 29: 397–441

Marazita ML, Murray JC, Lidral AC, Arcos-Burgos M, Cooper ME, Goldstein T, Maher BS, Daack-Hirsch S, Schultz R, Mansilla MA, Field LL, Liu YE, Prescott N, Malcolm S, Winter R, Ray A, Moreno L, Valencia C, Neiswanger K, Wyszynski DF, Bailey-Wilson JE, Albacha-Hejazi H, Beaty TH, McIntosh I, Hetmanski JB, Tunçbilek G, Edwards M, Harkin L, Scott R, Roddick LG. Meta-analysis of 13 genome scans reveals multiple cleft lip/palate genes with novel loci on 9q21 and 2q32-35. *Am J Hum Genet* 2004;75(2):161-173

Martinelli M, Palmieri A, Carinci F, Scapoli L. Non-syndromic Cleft Palate: An Overview on Human Genetic and Environmental Risk Factors. *Front Cell Dev Biol* 2020. 8:592271  
Maston GA, Evans SK, Green MR. Transcriptional Regulatory Elements in the Human Genome. *Annu Rev Genomics Hum Genet* 2006;7(1):29–59

Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet* 11 2010, 499–511

Mattick JS, Makunin IV. Non-coding RNA. *Hum Mol Genet* 2006; 15 Spec No 1:R17-29  
Mikhailov AT, Torrado M. The enigmatic role of the ankyrin repeat domain 1 gene in heart development and disease. *Int J Dev Biol* 2008;52(7):811–821

Mossey PA, Little J, Munger RG, Dixon MJ, Shaw WC. Cleft lip and palate. *Lancet* 2009; 374: 1773–1785

Mostowska A, Gaczkowska A, Żukowski K, Ludwig KU, Hozyasz KK, Wójcicki P, Mangold E, Böhmer AC, Heilmann-Heimbach S, Knapp M, Zadurska M, Biedziak B, Budner M, Lasota A, Daktera-Micker A, Jagodziński PP. Common variants in DLG1 locus are associated with non-syndromic cleft lip with or without cleft palate. *Clin Genet* 2018;93(4):784–793

Mukhopadhyay P, Brock G, Appana S, Webb C, Greene RM, Pisano MM. MicroRNA gene expression signatures in the developing neural tube. *Birth Defects Res A Clin Mol Teratol* 2011;91(8):744-762

Nakatomi M, Ludwig KU, Knapp M, Kist R, Lisgo S, Ohshima H, Mangold E, Peters H. Msx1 deficiency interacts with hypoxia and induces a morphogenetic regulation during mouse lip development. *Development* 2020;147(21):dev189175

O’Rahilly, R. Guide to the staging of human embryos. *Anat. Anz* , 1972; 130: 556–559

Polychronopoulos D, King JWD, Nash AJ, Tan G, Lenhard B. Conserved non-coding elements: developmental gene regulation meets genome organization. *Nucleic Acids Res* 2017; 45(22):12611–12624

Pott S, Lieb JD. What are super-enhancers? *Nat Genet* 2015;47(1):8–12

Prescott SL, Srinivasan R, Marchetto MC, Grishina I, Narvaiza I, Selleri L, Gage FH, Swigut T, Wysocka J. Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest. *Cell* 2015;163(1):68–83

Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26(18):2336–2337



Rada-Iglesias A, Bajpai R, Prescott S, Brugmann SA, Swigut T, Wysocka J. Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. *Cell Stem Cell* 2012;11(5):633-648

Reynolds K, Zhang S, Sun B, Garland MA, Ji Y, Zhou CJ. Genetics and signaling mechanisms of orofacial clefts. *Birth defects Res* 2020; 112(19):1588–1634

Seelan RS, Pisano MM, Greene RM. MicroRNAs as epigenetic regulators of orofacial development. *Differentiation* 2022; 124:1–16

Sekhon K, Bucay N, Majid S, Dahiya R, Saini S. MicroRNAs and epithelial-mesenchymal transition in prostate cancer. *Oncotarget* 2016;7(41):67597–67611

Setó-Salvia N, Stanier P. Genetics of cleft lip and/or cleft palate: Association with other common anomalies. *Eur J Med Genet* 2014; 57(8):381–393

Shields ED, Bixler D, Fogh-Andersen P. Facial clefts in Danish twins. *Cleft Palate J* 1979;16(1):1–6

Sivertsen A, Wilcox AJ, Skjaerven R, Vindenes HA, Abyholm F, Harville E, Lie RT. Familial risk of oral clefts by morphological type and severity: population based cohort study of first degree relatives. *BMJ* 2008;336(7641):432-434

Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;3:Article3

Stanier P, Pauws E. Development of the lip and palate: FGF signalling. *Cleft Lip Palate Epidemiol Aetiol Treat* 2012;16:71–80

Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: An online resource for prediction of microRNA binding sites. *PLoS One* 2018;13(10):e0206239–e0206239

Stoll C, Dott B, Alembik Y, Roth M-P. Associated congenital anomalies among cases with Down syndrome. *Eur J Med Genet* 2015; 58(12):674–680

Tan HL, Kheirandish-Gozal L, Abel F, Gozal D. Craniofacial syndromes and sleep-related breathing disorders. *Sleep Med Rev* 2016; 27: 74–88

Tena JJ, Santos-Pereira JM. Topologically Associating Domains and Regulatory Landscapes in Development, Evolution and Disease. *Front Cell Dev Biol* 2021;9:702787

Thieme F, Ludwig KU. The Role of Noncoding Genetic Variation in Isolated Orofacial Clefts. *J Dent Res* 2017; 96(11):1238–1247

Trainor PA. Craniofacial birth defects: The role of neural crest cells in the etiology and pathogenesis of Treacher Collins syndrome and the potential for prevention. *Am J Med Genet* 2010; 152A: 2984–2994

Van den Boogaard MJ, Dorland M, Beemer FA, van Amstel HK. MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans. *Nat Genet* 2000; 24(4):342–343

Wang B, Fang X, Sun X, Du C, Zhou L, Lv X, Li Y, Li H, Tang W. m(6)A demethylase ALKBH5 suppresses proliferation and migration of enteric neural crest cells by regulating TAGLN in Hirschsprung's disease. *Life Sci* 2021;278:119577

Watkins SE, Meyer RE, Strauss RP, Aylsworth AS. Classification, epidemiology, and genetics of orofacial clefts. *Clin Plast Surg* 2014; 41(2):149–163

Welzenbach J, Hammond NL, Nikolić M, Thieme F, Ishorst N, Leslie EJ, Weinberg SM, Beaty TH, Marazita ML, Mangold E, Knapp M, Cotney J, Rada-Iglesias A, Dixon MJ, Ludwig KU. Integrative approaches generate insights into the architecture of non-syndromic cleft lip with or without cleft palate. *HGG Adv* 2021;2(3):100038

Westra H-J, Franke L. From genome to function by studying eQTLs. *Biochim Biophys Acta - Mol Basis Dis* 2014;1842(10):1896–1902

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 2013;153(2):307-319

Wilderman A, VanOudenhove J, Kron J, Noonan JP, Cotney J. High-Resolution Epigenomic Atlas of Human Embryonic Craniofacial Development. *Cell Rep.* 2018;23(5):1581–1597

Worley ML, Patel KG, Kilpatrick LA. Cleft Lip and Palate. *Clin Perinatol* 2018; 45: 661–678

Zeidler M, Hüttenhofer A, Kress M, Kummer KK. Intragenic MicroRNAs Autoregulate Their Host Genes in Both Direct and Indirect Ways-A Cross-Species Analysis. *Cells* 2020 Jan;9(1):232

Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. *Nucleic Acids Res* 2017;46(D1):D754–761

Zhang J, Liu J, Wu J, Li W, Chen Z, Yang L. Progression of the role of CRYAB in signaling pathways and cancers. *Onco Targets Ther* 2019;12:4129–39

Zhang X-T, Wang G, Li Y, Chuai M, Lee KKH, Yang X. Role of FGF signalling in neural crest cell migration during early chick embryo development. *Zygote* 2018;26(6):457–464

## 2. Veröffentlichung



### Research Reports: Biological

# MiRNA-149 as a Candidate for Facial Clefting and Neural Crest Cell Migration

Journal of Dental Research  
2022, Vol. 101(3) 323–330  
© International Association for Dental Research and American Association for Dental, Oral, and Craniofacial Research 2021  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/00220345211038203  
journals.sagepub.com/home/jdr

L.G. Stüssel<sup>1\*</sup>, R. Hollstein<sup>1\*</sup> , M. Laugsch<sup>2,3,4</sup>, L.M. Hochfeld<sup>1</sup> , J. Welzenbach<sup>1</sup>, J. Schröder<sup>1</sup>, F. Thieme<sup>1</sup>, N. Ishorst<sup>1</sup>, R. Olmos Romero<sup>3,4</sup>, L. Weinhold<sup>5</sup>, T. Hess<sup>1,6</sup>, J. Gehlen<sup>1,6</sup>, A. Mostowska<sup>7</sup>, S. Heilmann-Heimbach<sup>1</sup>, E. Mangold<sup>1</sup>, A. Rada-Iglesias<sup>2,8,9</sup>, M. Knapp<sup>5</sup>, C.P. Schaaf<sup>3,4</sup>, and K.U. Ludwig<sup>1</sup>

### Abstract

Nonsyndromic cleft lip with or without palate (nsCL/P) ranks among the most common human birth defects and has a multifactorial etiology. Human neural crest cells (hNCC) make a substantial contribution to the formation of facial bone and cartilage and are a key cell type in terms of nsCL/P etiology. Based on increasing evidence for the role of noncoding regulatory mechanisms in nsCL/P, we investigated the role of hNCC-expressed microRNAs (miRNA) in cleft development. First, we conducted a systematic analysis of miRNAs expressed in human-induced pluripotent stem cell-derived hNCC using Affymetrix microarrays on cell lines established from 4 unaffected donors. These analyses identified 152 candidate miRNAs. Based on the hypothesis that candidate miRNA loci harbor genetic variation associated with nsCL/P risk, the genomic locations of these candidates were cross-referenced with data from a previous genome-wide association study of nsCL/P. Associated variants were reanalyzed in independent nsCL/P study populations. Jointly, the results suggest that miR-149 is implicated in nsCL/P etiology. Second, functional follow-up included in vitro overexpression and inhibition of miR-149 in hNCC and subsequent analyses at the molecular and phenotypic level. Using 3'RNA-Seq, we identified 604 differentially expressed (DE) genes in hNCC overexpressing miR-149 compared with untreated cells. These included *TLR4* and *JUNB*, which are established targets of miR-149, and *NOG*, *BMP4*, and *PAX6*, which are reported nsCL/P candidate genes. Pathway analyses revealed that DE genes were enriched in pathways including regulation of cartilage development and NCC differentiation. At the cellular level, distinct hNCC migration patterns were observed in response to miR-149 overexpression. Our data suggest that miR-149 is involved in the etiology of nsCL/P via its role in hNCC migration.

**Keywords:** cleft lip, cleft palate, small RNAs, gene regulation, noncoding, embryonic development

### Introduction

Nonsyndromic cleft lip with or without cleft palate (nsCL/P) represents one of the most common human birth defects, with a prevalence of about 1 in 1,000 live births (Mossey and Modell 2012). NsCL/P has a multifactorial etiology, comprising environmental influences and a strong genetic component (Grosen et al. 2010). To date, systematic genetic studies (mainly genome-wide association studies [GWAS]) have identified at least 40 genetic risk loci for nsCL/P. Their predominant location in noncoding regions of the human genome poses a substantial challenge for biological interpretation (Thieme and Ludwig 2017).

One potential mechanism through which noncoding variation may lead to disease is the regulatory impact of noncoding RNA molecules, such as microRNAs (miRNAs). miRNAs are small molecules of about 22 nucleotides (nt) that bind to complementary sequences within the mRNAs of their respective target genes. This leads to a downregulation of the target genes at the posttranscriptional level (Schoen et al. 2017). Accordingly, any form of genetic variability with an influence on miRNA expression levels and/or that modifies miRNA target recognition sites may result in interindividual variation in

<sup>1</sup>Institute of Human Genetics, University of Bonn, School of Medicine and University Hospital Bonn, Bonn, Germany

<sup>2</sup>Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany

<sup>3</sup>Institute of Human Genetics, University of Heidelberg, Heidelberg, Germany

<sup>4</sup>Institute of Human Genetics, CMMC, University Hospital Cologne, Cologne, Germany

<sup>5</sup>Institute of Medical Biometry Informatics and Epidemiology, University of Bonn, Bonn, Germany

<sup>6</sup>Center of Human Genetics, University Hospital of Marburg, Marburg, Germany

<sup>7</sup>Department of Biochemistry and Molecular Biology, Poznan University of Medical Sciences, Poznan, Poland

<sup>8</sup>Cologne Excellence Cluster for Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, Cologne, Germany

<sup>9</sup>Institute of Biomedicine and Biotechnology, University of Cantabria, Santander, Spain

\*Authors contributing equally to this article.

A supplemental appendix to this article is available online.

### Corresponding Author:

K.U. Ludwig, Institute of Human Genetics, University of Bonn, School of Medicine and University Hospital Bonn, Venusberg Campus 1, Gebäude 76, Bonn, 53127, Germany.

Email: Kerstin.ludwig@uni-bonn.de

downstream effects on the cellular gene regulatory network. As one example, a single nucleotide variant (SNV), which was associated with nsCL/P in the Chinese population, was shown to alter processing of miR-140 (Li et al. 2010), which in turn had been previously identified as a regulator of palatogenesis in mice (Eberhart et al. 2008). Further examples include the creation of a novel noncoding mRNA by an nsCL/P-associated SNV located in intron 2 of *CDH2* (Kumari et al. 2018); the recent identification of miRNAs let-7c-5p and miR-193a-3p, which were shown to regulate proliferation and apoptosis in human embryonic palatal mesenchyme cells and oral epithelial cells (Fu et al. 2021); and the establishment of miRNA cluster miR-17-92 as a component of the fine-tuning complex involved in murine palatal shelf outgrowth (Wang et al. 2013). To date, systematic analyses, such as array-based miRNA assessments, have been restricted to nonhuman or adult human specimens, which limits their utility in terms of the interpretation of nsCL/P genetic association findings from GWAS.

The lack of systematic miRNA analyses in human cell types of relevance to nsCL/P is attributable in part to a limited access to human embryonic facial tissue. It has been shown that specific mesenchymal precursor cells that are present only during embryonic development, namely, human neural crest cells (hNCC), play an important role in craniofacial development. Genetic variants that modulate enhancer activity in hNCC affect human phenotypic diversity (Rada-Iglesias et al. 2013), and different genetic loci have been shown to affect both neural crest development and craniofacial morphology, often in a dosage-sensitive manner (Prescott et al. 2015). hNCC give rise to most of the craniofacial tissue and cartilage (Trainor and Tam 1995; Adameyko and Fried 2016), and there is increasing evidence that alterations in hNCC development contribute to nsCL/P etiology (Ludwig et al. 2017). Recently, it was shown that regions associated with normal-range facial variation are enriched for enhancer activity in cranial neural crest cells and craniofacial tissues (White et al. 2021).

The aim of the present study was the identification of novel candidate miRNAs for nsCL/P. First, we profiled miRNA expression in hNCC derived from human-induced pluripotent stem cells (hiPSC) and prioritized miRNAs by integrating genetic findings with genetic data from our recently generated nsCL/P GWAS data (Ludwig et al. 2017). Second, we performed replication studies, which identified miRNA-149 as a plausible candidate miRNA for nsCL/P. Finally, we performed in vitro characterization in hNCC at both the phenotype and the molecular level. Our results implicate miRNA-149 as a molecular candidate in nsCL/P pathobiology and suggest that impaired NCC migration might be a potential functional mechanism at this genetic risk locus.

## Materials and Methods

Details of the methods can be found in the Appendix. The Appendix tables have been deposited at Zenodo, under doi:10.5281/zenodo.5036358.

## Generation of hNCC

The fibroblast or keratinocyte samples of 4 healthy donors were reprogrammed into hiPSC, as described elsewhere (Laugsch et al. 2019). For each hNCC sample, total RNA and DNA were extracted using the innuPREP RNA-Mini-Kit (Analytic Jena) as described elsewhere (Laugsch et al. 2019). Quality control included molecular karyotyping and systematic expression analysis of specific marker genes (i.e., neural crest markers, iPSC and mesenchymal markers) using 3'RNA-seq.

Data analysis was performed using an in-house pipeline (see the methods in the Appendix).

## MiRNA Profiling

### MiRNA Analysis

MiRNA expression analysis was performed using the Affymetrix GeneChip miRNA 4.0 and Affymetrix Expression Console software (Affymetrix). These analyses were based on robust multiarray analysis and detection above background, in accordance with established, in-house protocols (Hochfeld et al. 2017). miRNAs were classified as “expressed” if identified with a detection *P* value <0.01 in any of the 4 individual hNCC samples. miRNAs expressed in all 4 donors were considered candidate miRNAs (“hNCC batch”). We further prioritized miRNAs that 1) have been reported to be involved in craniofacial development in previous studies and 2) were expressed in at least 1 of the 4 hNCC samples (“literature batch”).

### Target Prediction

Target prediction for hsa-miR-149-3p was performed using MirWalk3.0 (Sticht et al. 2018). For further details, see the methods in the Appendix.

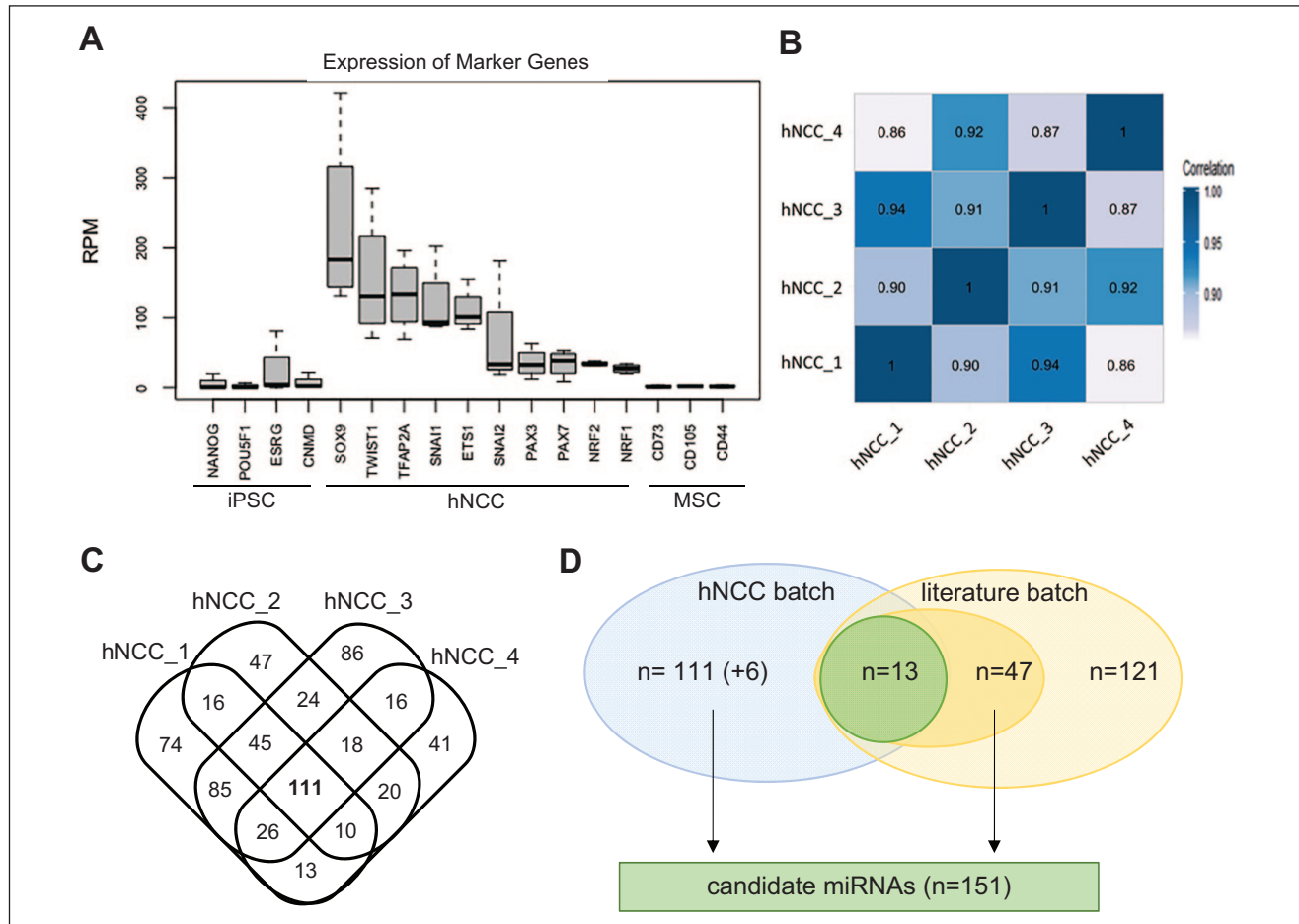
## Integration with Genetic Data

### GWAS Data

GWAS summary statistics for nsCL/P were retrieved from a previously published study of nsCL/P (Ludwig et al. 2017). For each candidate miRNA±1 kb, common variants with *P*-values <0.01 were defined as “candidate variants.”

### Replication Analyses

We first analyzed 21 of 25 candidate variants using targeted genotyping in an independent cohort of 223 nsCL/P cases and 968 controls of European descent. Statistical analyses were performed using Armitage trend and genotypic tests (replication I). One SNV with supportive evidence (rs149251515) was subsequently analyzed in 2 additional cohorts (replication II), that is, from Poland (471 cases and 832 controls; Mostowska et al. 2018) and from a multiethnic GWAS (phs000774, retrieved upon approved data access).



**Figure 1.** Identification of microRNAs expressed in human neural crest cells (hNCC). **(A)** To show full differentiation, the expression of induced pluripotent stem cell marker genes (*NANOG*, *POU5F1*, *ESRG*, *CNMD*) was analyzed (Sekine et al. 2020). The hNCC status of 4 independent cell lines was confirmed by the detection of strong marker gene expression in 3'RNA-Seq analyses (*SOX9*, *TFAP2A*, *SNAI1/2*, *TWIST1*, *PAX3/7*, *NRF1/2*, and *ETS1*), described as marker genes in Betancur et al. 2010; Rada-Iglesias et al. 2012). The expression of genes (*CD73*, *CD105*, *CD44*) described as mesenchymal specifiers served as a negative control (Gomez-Salazar et al. 2020). Counts are shown in reads per million (RPM). **(B)** High intersample correlation confirmed the reproducibility of results across all donors (average Pearson's correlation 0.896). **(C)** Venn diagram showing overlapping expression of miRNAs between the 4 hNCC samples. A total of 111 miRNAs were expressed in all 4 samples, subsequently referred to as the "hNCC batch." A total of 111 miRNAs were expressed in all 4 samples. **(D)** Together with three stem-loop sequences (each having a -3p and -5p version, as indicated in brackets), the 111 miRNAs are referred to as "hNCC batch". This batch was complemented with a set of 47 miRNAs that were identified as candidate miRNAs based on previous reports of their role in orofacial development. Thirteen miRNAs overlapped between both batches, resulting in a list of 151 candidate miRNAs.

## In Vitro Analysis of miR-149

miRNA-149-3p abundance was modified in hNCC using either mirVana Inhibitor or mirVana miRNA Mimic (assay hsa-mir-149-3p, Thermo Fisher Scientific). Briefly, hNCC were transfected with 50 nM mirVana Inhibitor/Mimic. Forty-eight hours after transfection, hNCCs were characterized as follows: 1) at the phenotypic level using scratch assays to analyze migration behavior and 2) at the molecular level to identify differentially expressed (DE) genes compared with untreated control.

## Results

### Identification of Candidate miRNAs Expressed in hNCC

Four hiPSC cell lines derived from healthy donors were differentiated into hNCC according to previously established

protocols (Laugsch et al. 2019). Quality control based on RNA-Seq data confirmed low expression of iPSC marker genes (Sekine et al. 2020), equal expression of established hNCC marker genes (e.g., *NR2F1/2*, *SOX9*, and *TFAP2A*; Rada-Iglesias et al. 2012), as well as the absence of mesenchymal marker genes (Gomez-Salazar et al. 2020; Fig. 1A), with robust intersample correlations (average Pearson's  $r^2=0.896$ ; Fig. 1B). For further validation of the cellular identity, we compared these RNA-Seq data with publicly available hNCC transcriptome data ( $r^2=0.674$ ; Appendix Fig. S1).

miRNA profiling revealed an abundance of 382, 293, 416, and 258 human miRNAs in samples hNCC\_1 to hNCC\_4, respectively (Fig. 1C). The list of 633 miRNAs included 111 miRNAs and three stem-loop sequences that were detected at  $P_{\text{det}} < 0.01$  across all 4 hNCC samples ("hNCC batch"). Of 121 miRNAs that have been implicated in craniofacial development, 47 miRNAs were expressed ( $P_{\text{det}} < 0.01$ ) in at least 1

**Table.** Candidate miRNAs after Filtering.<sup>a</sup>

Hairpin miRNA	Mature miRNA	Highest Detection <i>P</i> Value <sup>b</sup>	Evidence from the Literature	Reference (PMID)
hsa-mir-16	hsa-miR-16-2-3p	0.0051	Expression elevated in plasma samples of patients with nonsyndromic cleft lip	27470385
hsa-mir-17~92 cluster	hsa-miR-17-5p	$8.72 \times 10^{-10}$	Severe craniofacial phenotypes in miR-17-92 mutant mice embryos	24068957; 24068957
	hsa-miR-19b-3p	0.0044		
	hsa-miR-20a-5p	$8.72 \times 10^{-10}$		
	hsa-miR-92a-3p	$8.72 \times 10^{-10}$		
	hsa-miR-92a-1-5p	0.0051		
	hsa-miR-92a-2-5p	$8.60 \times 10^{-04}$		
hsa-mir-20b	hsa-miR-20b-5p	$8.72 \times 10^{-10}$	Paralogue to hsa-mir-17~92 cluster	
hsa-mir-23b	hsa-miR-23b-5p	$7.25 \times 10^{-06}$	Differential expression in avian facial tissues during early development	27471470
hsa-mir-92b	hsa-miR-92b-3p	0.0069	Paralogue to hsa-mir-17~92 cluster	
hsa-mir-106a	hsa-miR-106a-5p	$8.72 \times 10^{-10}$	Paralogue to hsa-mir-17~92 cluster	
hsa-mir-199a	hsa-miR-199a-3p	0.0072	Involved in chondrogenic differentiation of stem cells, expressed in developing murine facial structures	19787082; 19251704; 24849136; 27471470
hsa-mir-363	hsa-miR-363-3p	0.0057	Differential expression in murine facial tissue and involvement in regulation of Hand2 and Tbx3 signaling	24849136

<sup>a</sup>Thirteen microRNAs were robustly expressed at a detection *P* value of <0.01 and have been implicated previously in craniofacial development and/or disease. miRNA, microRNA; PMID, identification number in PubMed.

<sup>b</sup>Among the 4 individual human neural crest cell lines.

hNCC sample (“literature batch”). Thirteen miRNAs were present in both batches (Table). Overall, 151 candidate miRNAs were expressed in hNCC (Fig. 1D, full list provided as Appendix Table S1, available at doi:10.5281/zenodo.5036358).

### Prioritization of miR-149

GWAS summary statistics from a published study of nsCL/P, which comprised European (meta<sub>Euro</sub>) and European+Asian samples (meta<sub>all</sub>; Ludwig et al. 2017), were screened for co-localizing SNVs ( $\pm 1$  kb). A total of 13,002 SNVs mapped with any of the 151 candidate miRNAs. Of these, 25 variants at 13 loci showed an association *P*-value of <0.01 in either meta<sub>Euro</sub> or meta<sub>all</sub> (Appendix Table S2, available at doi:10.5281/zenodo.5036358). None of the variants mapped directly within a miRNA sequence, and none of the association *P*-values withstood correction for multiple testing.

Targeted association analysis of 15 candidate variants in our in-house European replication cohort (223 nsCL/P patients, 968 population-based controls) revealed 1 nominally significant variant on chromosome 2 (rs149251515,  $P_{\text{genotype}}=0.049$ ,  $P_{\text{Armitage}}=0.08$ , odds ratio=1.57 [95% confidence interval: 0.94–2.61]; Appendix Table S2, available at doi:10.5281/zenodo.5036358). Further analyses of rs149251515 in 2 independent nsCL/P cohorts showed a nominally significant association in an independent GWAS (Leslie et al. 2016,  $P=0.033$ ), while no significant association was detected in a Polish cohort (Mostowska et al. 2018),  $P=0.71$ ).

Variant rs149251515 represents a low-frequency variant (MAF=0.03) at the miR-149 locus and is observed primarily in the European population. To validate imputation accuracy for this variant in the array data, we confirmed the consistency of MAF with gnomAD data (European population)

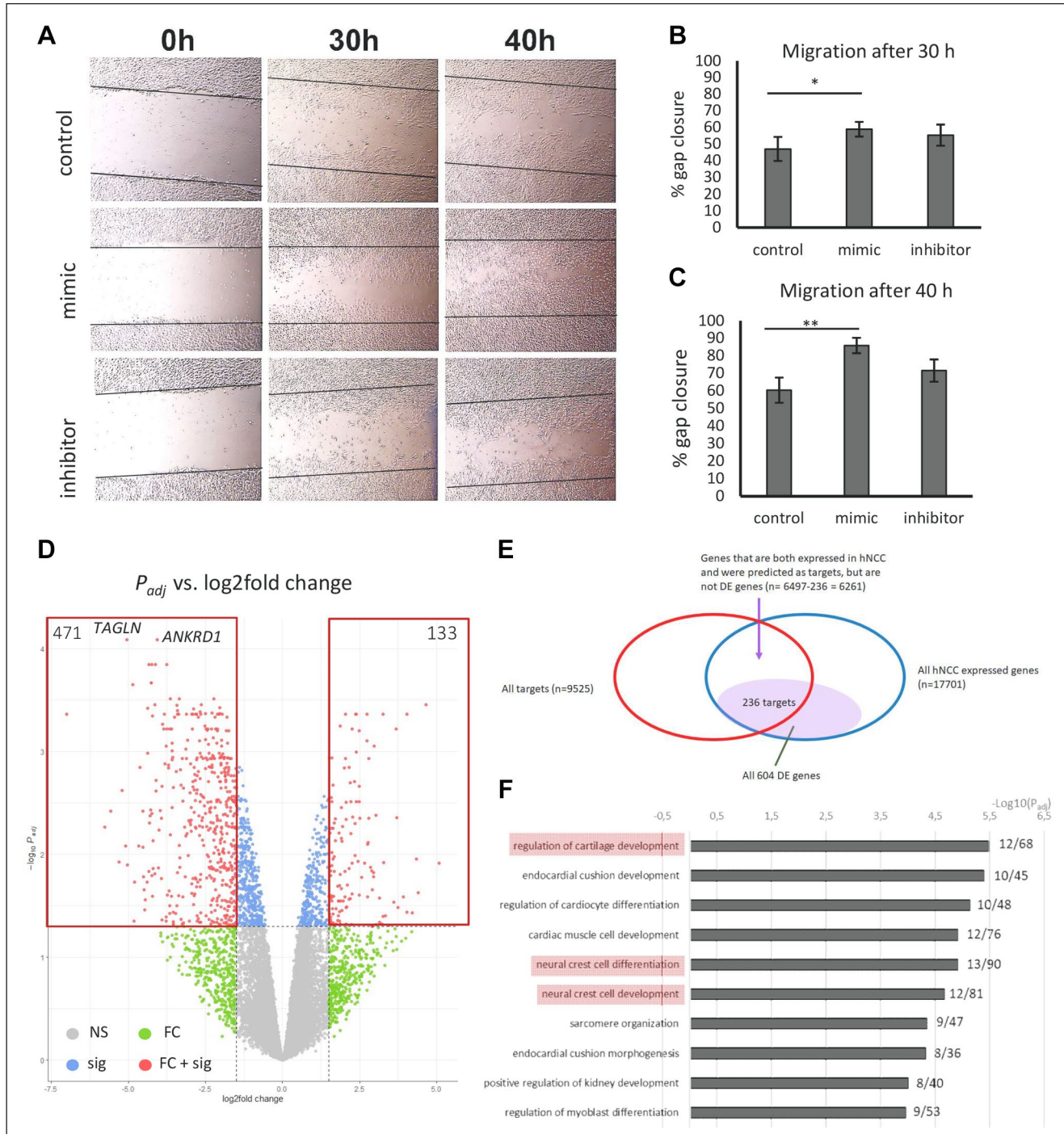
and imputation quality (info score >0.6). MiR-149 is highly conserved throughout vertebrate species (Appendix Fig. S2) and is located intronically in *GPC1*. Inspection of the entire association region in meta<sub>all</sub> revealed no other variant in linkage disequilibrium to rs149251515 (at  $r^2 > 0.8$ ). However, a set of common risk variants with association *P*-values <0.001 was located 250 kb downstream of miR-149 (Appendix Fig. S3).

### In Vitro Scratch Assay in hNCC

Detailed inspection of miR-149 in the hNCC expression data revealed that the processed form miR-149-3p was present in all 4 samples, whereas miR-149-5p was not expressed. Functional analyses were thus performed for miR-149-3p. Based on previous reports that miR-149-3p regulates proliferation and migration in diverse cell types (Chamorro-Jorganes et al. 2014; Shen et al. 2017; Yao and Wu 2019), we performed in vitro scratch assays in hNCC that involved the manipulation of miR-149-3p abundance using hsa-miR-149-3p mirVana miRNA inhibitor (i.e., decreased miR-level) or mirVana miRNA mimic (i.e., increased miR-level). hNCC transfected with miR-149-3p mirVana mimic showed a significant increase in migration compared with untreated hNCC at both 30 h ( $P=0.0458$ , *t* test) and 40 h ( $P=0.0085$ ) postscratching (Fig. 2A, B), respectively. No significant difference between the conditions was observed for miRNA-149 inhibition.

### Identification of miR-149-3p Targets in hNCC

To systematically identify target genes of miR-149-3p, 3'RNA-Seq was performed on hNCC (p2 stage) treated with miR-149-3p mimic and controls 48 h posttransfection. A total of 604 DE genes were observed (133 up- and 471 downregulated; Fig.



**Figure 2.** In vitro investigation of the functional effects of miR-149-3p. **(A–C)** Scratch assays were performed after the transient transfection of Human neural crest cells (hNCC) with hsa-miR-149-3p mimic (increasing level) or inhibitor (decreasing level). Quantification of the migration pattern revealed a significant increase in migration at **(B)** 30h (\* $P = 0.046$ ,  $t$  test,  $n = 3$ ) and **(C)** 40h (\*\* $P = 0.0085$ ,  $t$  test,  $n = 3$ ) after scratching. **(D)** Analysis of differentially expressed (DE) genes in hNCC. Expression data were generated using 3'-mRNA-Seq 48h after transient transfection with hsa-miR-149-3p mimic. Volcano plot shows the results of the DE analysis, with each dot representing a DE gene compared with untreated control cells. NS, not significant; FC,  $\log_2$  fold change  $> 1.5$ , sig, genes with adjusted  $P < 0.05$ . Significant downregulation was observed in 471 genes (red dots, left panel,  $P_{adj} < 0.05$ ,  $\log_2$ fold change  $> -1.5$ ). Significant upregulation was observed in 133 genes (red dots, right panel,  $P_{adj} < 0.05$ ,  $\log_2$ fold change  $> 1.5$ ). **(E)** Comparison of in silico prediction of miRNA targets with results of the in vitro DE analysis in **(D)**. **(F)** Gene Ontology (GO) analysis of DE genes. Shown are the top 10 GO terms of term level 5 with a group size of  $< 100$ . NCC-related pathways are marked in red.



2C; Appendix Table S3, doi:10.5281/zenodo.5036358). Of the latter, 17 have been implicated previously in orofacial clefting and/or face morphology at the functional and/or genetic level. The results were compared with in silico target prediction data for miR-149-3p, which had revealed 4,084 unique 3'UTR targets, 2,142 unique 5'UTR targets, and 5,337 targets within the coding sequences. Overall, 9,525 genes were predicted as targets at least once, and 6,497 of these were also expressed in hNCC (Fig. 2D). Of the 604 DE genes, 236 were among the predicted targets of miR-149-3p, and of those, 189 were downregulated by miR-149-3p overexpression (indicated by a negative fold change).

### Pathway Enrichment Analysis

Gene Ontology (GO) Term analysis of the set of 604 DE genes was performed using ConsensusDB. To enable detection of downstream effects in gene-regulatory networks, we included genes up- and downregulated after miR-149 overexpression. This revealed a significant enrichment in developmental pathways, such as anatomical structure morphogenesis, anatomical structure development, and system development (Appendix Table S4, doi:10.5281/zenodo.5036358). DE genes were significantly enriched in a total of 924 GO gene sets ( $P < 0.05$ ), 146 of which represent developmental processes. Among the more specific molecular pathways (defined as term-level 5 and gene sets size  $< 100$ ), 138 gene sets with a significant overrepresentation of DE genes were observed (Appendix Table S5, doi:10.5281/zenodo.5036358). These pathways included “regulation of cartilage development,” “neural crest cell differentiation,” and “neural crest cell development” (Fig. 2E).

### Discussion

Although altered miRNA regulation has been proposed as a pathomechanism in the complex etiology of nsCL/P (Schoen et al. 2017), the vast majority of molecular miRNA-studies of nsCL/P to date have been limited by the investigation of non-human samples (Eberhart et al. 2008; Li et al. 2010; Ding et al. 2016). The present study leveraged recent developments in induced pluripotent stem cell technology, high-throughput molecular analyses, and in vitro manipulations to generate a novel miRNA data set of relevance to human craniofacial development and disease.

The central aim of our study was the identification of hNCC-expressed miRNAs that might be involved in nsCL/P etiology. Through array-based miRNA profiling, we generated a candidate list of miRNAs expressed in hiPSC-derived hNCC. Of the 111 miRNAs that were expressed in all 4 donors, 13 had been identified in the context of craniofacial development by previous independent studies (e.g., the mir-17~92 cluster and mir-23b; Ding et al. 2013, 2016). Overall, 151 miRNAs represent candidate miRNAs for future functional analyses into hNCC-related development and disease.

Based on our hypothesis that candidate miRNA loci harbor genetic variation associated with nsCL/P risk, we integrated previously generated nsCL/P GWAS data to search

for associated risk variants at the selected miRNA loci. The analyses revealed a low number of associated SNVs within the miRNA genomic loci (25 of 13,002 SNVs). Possible explanations are that 1) in general, miRNAs do not contribute to nsCL/P etiology at significant extent; 2) because of their high specificity and conservations, associated variants within miRNAs are predominantly from the rare allele spectrum and, thus, these variants have escaped detection in GWAS; or 3) rather than affecting the binding sequence directly, genetic variation involving miRNA regulation may contribute to differential miRNA expression levels. The present analyses focused on 1 kb around the miRNA sequence, and although this cutoff may have led to the overlooking of associated variants, this conservative approach lowered the risk of false-positive findings.

None of the variants identified within 1 kb of the candidate miRNAs showed a testwide significant association, but 25 variants had nominally significant  $P$  values. Of those, genotyping of a replication cohort confirmed association with rs14921515, which, however, remained ambiguous in 2 additional cohorts. This might be explained in part by the low frequency of rs14921515, which is largely restricted to the European population. However, we found additional lines of evidence for a potential role of miR-149 and nsCL/P. First, we reanalyzed miRNA expression data from murine facial tissues at embryonic days E10.5 to E14.5 and found the mouse homologue mmu-miR-149 to be expressed in the frontonasal and maxillary processes as well as in the developing palate (accessed via FaceBase on 06/01/2017; Brinkley et al. 2016; Ding et al. 2016). Second, we also observed expression of miR-149 in the data set GSE20880 (data not shown) and found that the genomic region around miR-149 shows strong regulatory activity in human embryonic craniofacial tissues from early Carnegie stages (13–15) and is classified as a craniofacial super-enhancer region (Wilderman et al. 2018).

The identification of the respective downstream targets was investigated 1) computationally, using in silico target prediction, and 2) by manipulating miRNA levels in a cellular system, followed by expression analysis. Our in silico miR-149 target prediction revealed a large number of potential target genes. Some may represent false-positive results as consequence of the still imperfect performance of the computational algorithms. Alternatively, this abundance may indicate that miR-149 regulates many genes. After overexpression of miR-149 in hNCC in vitro, differential expression was observed for several hundreds of genes. Most of these DE genes (88.0%) showed downregulation, which is consistent with the hypothesis that miRNAs silence gene expression due to a direct interaction between the miRNA molecules and their target RNA. It has been frequently observed that the host gene, in which the miRNA is located, is among the downstream targets (Dill et al. 2012). In support of this observation, the host gene of miR-149, *GPC1*, was among the predicted target genes, but the observed decrease in *GPC1* expression after miR-149 overexpression was not significant. We also analyzed the expression pattern of *GPC1* and all genes located in the topologically associating domain around rs14921515 in hNCC and murine expression data and found only low expression for all genes

except *GPC1*, further supporting its role as direct candidate gene (Appendix Fig. S4).

The differential gene expression analyses revealed several interesting downstream genes, with *TAGLN* and *ANKRD1* being most significant. The gene *TAGLN* encodes transgelin, a mesenchymal marker that is involved in endothelial-mesenchymal transition processes and acts as a central regulator of ventral body wall closure during embryonic development (Aldeiri et al. 2018). An increase in *TAGLN* promoter activity has been associated with the differentiation of NCC into smooth muscle cells (Jain et al. 1998). For *ANKRD1*, research has demonstrated a central function during heart development, and mutations in the gene are associated with a range of cardiac disorders (Mikhailov and Torrado 2008). While the involvement of these 2 genes supports the hypothesis that miR-149 is implicated in developmental processes, no conclusive evidence for their involvement in craniofacial development or diseases has yet been obtained. Nonetheless, the 471 down-regulated genes included 17 previously associated with craniofacial disease in genetic and/or functional studies. These include *NOG*, *BMP4*, and *GREM1* (Nakatomi et al. 2020). Notably, 2 additional members of the FGF family, *FGF1* and *FGF23*, have also been observed as DE genes upon miR-149 overexpression in our data set. Previous research has demonstrated that FGFs and their receptors (FGFR) are involved in general craniofacial development and orofacial clefting (Stanier and Pauws 2012) and are modulators of hNCC development (Zhang et al. 2018). Finally, we also identified *TGFB2* among our DE genes, which is noteworthy as it was recently revealed as a target of miR-193a-3p during craniofacial development in mice (Fu et al. 2021).

The present study had some inherent limitations. First, our expression analyses were restricted to cells generated from 4 independent donors and therefore require replication in larger samples. This small sample size also prevented us from performing systematic analyses of correlations between miRNA and mRNA expression levels. Also, as miR-149 was not sequenced in the donors, additional (rare) variants that might have affected miR-149 expression or cell culture results remained undetected. Second, the low frequency of the rs149251515 risk allele precluded the analysis of the potential allele-dependent effects of rs149251515 in the donors. In the future, such issues relating to donor numbers will be resolved via collaborative efforts such as HipSci (Streeter et al. 2017). However, for more systematic analyses using hiPSC-derived hNCC (e.g., expression quantitative trait locus), provision of the time and resources required will remain a considerable challenge. Finally, our differentiation protocol, including the scratch assay, represent only a small fraction of hNCC-derived facial morphogenesis, which opens up further questions on how miR-149 contributes to facial development in vivo.

In summary, the present study identified a set of miRNAs that are expressed in hNCC and that represent molecular candidates for disease mechanisms involving hNCC. Our data suggest a role of miR-149-3p in nsCL/P etiology, which may be mediated through modification of regulatory mechanisms in hNCC. Future research is warranted to determine the exact

mechanism through which rs149251515 and miR-149 interact and whether additional types of variation (such as rare variants) confer nsCL/P risk.

### Author Contributions

L.G. Stüssel, contributed to conception, design, data acquisition, analysis, and interpretation, drafted the manuscript; R. Hollstein, contributed to conception, design, data acquisition, analysis, and interpretation, drafted and critically revised the manuscript; M. Laugsch, contributed to data acquisition, analysis, and interpretation, critically revised the manuscript; L.M. Hochfeld, contributed to data analysis and interpretation, critically revised the manuscript; J. Welzenbach, J. Schröder, F. Thieme, N. Ichorst, L. Weinhold, T. Hess, J. Gehlen, S. Heilmann-Heimbach, M. Knapp, contributed to data analysis, critically revised the manuscript; R. Olmos Romero, E. Mangold, A. Rada-Iglesias, contributed to data acquisition, critically revised the manuscript; A. Mostowska, C.P. Schaaf, contributed to data acquisition and analysis, critically revised the manuscript; K.U. Ludwig, contributed to conception, design, data analysis, and interpretation, drafted and critically revised the manuscript. All authors gave final approval and agree to be accountable for all aspects of the work.

### Acknowledgments

We thank the NGS Core Facility of the University Hospital of Bonn, which is part of the West German Genome Center, for support in sequencing. Further, we thank Susanne Theiß for support with microscope imaging.

### Declaration of Conflicting Interests


The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the German Research Council (Deutsche Forschungsgemeinschaft DFG) through funding provided to K.U.L. (LU-1944/3-1). L.G.S. received support from the BONFOR program of the Medical Faculty Bonn (SciMed program). R.H. received a travel grant from the BFB (Bonner Forum Biomedizin).

### ORCID iDs

R. Hollstein  <https://orcid.org/0000-0003-0571-9690>

L.M. Hochfeld  <https://orcid.org/0000-0002-2808-0693>

### URLs

BBMap: <https://github.com/BioInfoTools/BBMap/blob/master/sh/bbduk.sh>

ConsensusPathDB: <http://cpdb.molgen.mpg.de/>

Ensembl: <https://www.ensembl.org/index.html>

FaceBase: [www.facebase.org](http://www.facebase.org)

FastQC: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>

Locus Zoom: <http://locuszoom.org/>

miRWalk 3.0: <http://mirwalk.umm.uni-heidelberg.de/>

PubMed: <https://www.ncbi.nlm.nih.gov/pubmed/>  
 UCSC Genome Browser: <https://genome.ucsc.edu/>

## References

- Adameyko I, Fried K. 2016. The nervous system orchestrates and integrates craniofacial development: a review. *Front Physiol.* 7:49.
- Aldeiri B, Roostalu U, Albertini A, Behnsen J, Wong J, Morabito A, Cossu G. 2018. Abrogation of TGF-beta signalling in TAGLN expressing cells recapitulates Pentagony of Cantrell in the mouse. *Sci Rep.* 8(1):3658.
- Betancur P, Bronner-Fraser M, Sauka-Spengler T. 2010. Assembling neural crest regulatory circuits into a gene regulatory network. *Annu Rev Cell Dev Biol.* 26:581–603.
- Brinkley JF, Fisher S, Harris MP, Holmes G, Hooper JE, Jabs EW, Jones KL, Kesselman C, Klein OD, Maas RL, et al. 2016. The FaceBase Consortium: a comprehensive resource for craniofacial researchers. *Development.* 143(14):2677–2688.
- Chamorro-Jorganes A, Araldi E, Rotllan N, Cirera-Salinas D, Suarez Y. 2014. Autoregulation of alipican-1 by intronic microRNA-149 fine tunes the angiogenic response to FGF2 in human endothelial cells. *J Cell Sci.* 127(pt 6):1169–1178.
- Dill H, Linder B, Fehr A, Fischer U. 2012. Intronic miR-26b controls neuronal differentiation by repressing its host transcript, *ctdps2*. *Genes Dev.* 26(1):25–30.
- Ding H, Hooper JE, Batzel P, Eames BF, Postlethwait JH, Artinger KB, Clouthier DE. 2016. MicroRNA profiling during craniofacial development: potential roles for *Mir23b* and *Mir133b*. *Front Physiol.* 7:281.
- Ding X, Park SI, McCauley LK, Wang CY. 2013. Signaling between transforming growth factor beta (TGF-β) and transcription factor *SNAIL2* represses expression of microRNA *miR-203* to promote epithelial-mesenchymal transition and tumor metastasis. *J Biol Chem.* 288(15):10241–10253.
- Eberhart JK, He X, Swartz ME, Yan Y-L, Song H, Boling TC, Kunerth AK, Walker MB, Kimmel CB, Postlethwait JH. 2008. MicroRNA *Mir140* modulates *Pdgfr* signaling during palatogenesis. *Nat Genet.* 40(3):290–298.
- Fu C, Lou S, Zhu G, Fan L, Yu X, Zhu W, Ma L, Wang L, Pan Y. 2021. Identification of new miRNA-mRNA networks in the development of non-syndromic cleft lip with or without cleft palate. *Front Cell Dev Biol.* 9:631057.
- Gomez-Salazar M, Gonzalez-Galofre ZN, Casamitjana J, Crisan M, James AW, Péault B. 2020. Five decades later, are mesenchymal stem cells still relevant? *Front Bioeng Biotechnol.* 8:148.
- Grosen D, Chevrier C, Skytthe A, Bille C, Mølsted K, Sivertsen Å, Murray JC, Christensen K. 2010. A cohort study of recurrence patterns among more than 54000 relatives of oral cleft cases in Denmark: Support for the multifactorial threshold model of inheritance. *J Med Genet.* 47(3):162–168.
- Hochfeld LM, Anhalt T, Reinbold CS, Herrera-Rivero M, Fricker N, Nothen MM, Heilmann-Heimbach S. 2017. Expression profiling and bioinformatic analyses suggest new target genes and pathways for human hair follicle related microRNAs. *BMC Dermatol.* 17(1):3. doi:10.1186/s12895-017-0054-9
- Jain MK, Layne MD, Watanabe M, Chin MT, Feinberg MW, Sibinga NE, Hsieh CM, Yet SF, Stemple DL, Lee ME. 1998. In vitro system for differentiating pluripotent neural crest cells into smooth muscle cells. *J Biol Chem.* 273(11):5993–5996.
- Kumari P, Singh SK, Raman R. 2018. A novel non-coding RNA within an intron of *CDH2* and association of its SNP with non-syndromic cleft lip and palate. *Gene.* 658:123–128.
- Laugsch M, Bartusel M, Rehim R, Alirzayeva H, Karaolidou A, Crispatzu G, Zentis P, Nikolic M, Bleckwehl T, Kolovos P, et al. 2019. Modeling the pathological long-range regulatory effects of human structural variation with patient-specific hiPSCs. *Cell Stem Cell.* 24(5):736–752.e12.
- Leslie EJ, Carlson JC, Shaffer JR, Feingold E, Wehby G, Laurie CA, Jain D, Laurie CC, Doheny KF, McHenry T, et al. 2016. A multi-ethnic genome-wide association study identifies novel loci for non-syndromic cleft lip with or without cleft palate on 2p24.2, 17q23 and 19q13. *Hum Mol Genet.* 25(13):2862–2872. doi: 10.1093/hmg/ddw104.
- Li L, Meng T, Jia Z, Zhu G, Shi B. 2010. Single nucleotide polymorphism associated with nonsyndromic cleft palate influences the processing of *miR-140*. *Am J Med Genet Part A.* 152(4):856–862.
- Ludwig KU, Böhmer AC, Bowes J, Nikolic M, Ishorst N, Wyatt N, Hammond NL, Gözl L, Thieme F, Barth S, et al. 2017. Imputation of orofacial clefting data identifies novel risk loci and sheds light on the genetic background of cleft lip ± cleft palate and cleft palate only. *Hum Mol Genet.* 26(4):829–842.
- Mikhailov AT, Torrado M. 2008. The enigmatic role of the ankyrin repeat domain 1 gene in heart development and disease. *Int J Dev Biol.* 52(7):811–821.
- Mossey PA, Modell B. 2012. Epidemiology of oral clefts 2012: an international perspective. *Front Oral Biol.* 16:1–18.
- Mostowska A, Gaczowska A, Zukowski K, Ludwig KU, Hozyasz KK, Wojcicki P, Mangold E, Bohmer AC, Heilmann-Heimbach S, Knapp M, et al. 2018. Common variants in *DLG1* locus are associated with non-syndromic cleft lip with or without cleft palate. *Clin Genet.* 93(4):784–793.
- Nakatomi M, Ludwig KU, Knapp M, Kist R, Lisgo S, Ohshima H, Mangold E, Peters H. 2020. *Mx1* deficiency interacts with hypoxia and induces a morphogenetic regulation during mouse lip development. *Development.* 147(21):dev189175.
- Prescott SL, Srinivasan R, Marchetto MC, Grishina I, Narvaiza I, Selli L, Gage FH, Swigut T, Wysocka J. 2015. Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest. *Cell.* 163(1):68–83.
- Rada-Iglesias A, Bajpai R, Prescott S, Brugmann SA, Swigut T, Wysocka J. 2012. Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. *Cell Stem Cell.* 11(5):633–648.
- Rada-Iglesias A, Prescott SL, Wysocka J. 2013. Human genetic variation within neural crest enhancers: molecular and phenotypic implications. *Phil Trans R Soc B.* 368:20120360. doi: 10.1098/rstb.2012.0360.
- Schoen C, Aschrafi A, Thonissen M, Poelmans G, Von den Hoff JW, Carels CEL. 2017. MicroRNAs in palatogenesis and cleft palate. *Front Physiol.* 8:165.
- Sekine K, Tsuzuki S, Yasui R, Kobayashi T, Ikeda K, Hamada Y, Kanai E, Camp JG, Treutlein B, Ueno Y, et al. 2020. Robust detection of undifferentiated iPSC among differentiated cells. *Sci Rep.* 10(1):10293.
- Shen W, Liu J, Zhao G, Fan M, Song G, Zhang Y, Weng Z, Zhang Y. 2017. Repression of Toll-like receptor-4 by microRNA-149-3p is associated with smoking-related COPD. *Int J Chron Obstruct Pulmon Dis.* 2:705–715. doi: 10.2147/COPD.S128031.
- Stanier P, Pauws E. 2012. Development of the lip and palate: FGF signalling. *Cleft Lip Palate Epidemiol Aetiol Treat.* 16:71–80.
- Sticht C, De La Torre C, Parveen A, Gretz N. 2018. miRWalk: an online resource for prediction of microRNA binding sites. *PLoS One.* 13(10):e0206239.
- Streeter J, Harrison PW, Faulconbridge A, The HipSci Consortium; Flicek P, Parkinson H, Clarke L. 2017. The human-induced pluripotent stem cell initiative-data resources for cellular genetics. *Nucleic Acids Res.* 45(D1):D691–D697.
- Thieme F, Ludwig KU. 2017. The role of noncoding genetic variation in isolated orofacial clefts. *J Dent Res.* 96(11):1238–1247.
- Trainor PA, Tam PP. 1995. Cranial paraxial mesoderm and neural crest cells of the mouse embryo: co-distribution in the craniofacial mesenchyme but distinct segregation in branchial arches. *Development.* 121(8):2569–2582.
- Wang J, Bai Y, Li H, Greene SB, Klysiak E, Yu W, Schwartz RJ, Williams TJ, Martin JF. 2013. MicroRNA-17-92, a direct *Ap-2α* transcriptional target, modulates T-box factor activity in orofacial clefting. *PLoS Genet.* 9(9):e1003785.
- White JD, Indencleef K, Naqvi S, Eller RJ, Hoskens H, Roosenboom J, Lee MK, Li J, Mohammed J, Richmond S, et al. 2021. Insights into the genetic architecture of the human face. *Nat Genet.* 53(1):45–53.
- Wilderman A, VanOudenhove J, Kron J, Noonan JP, Cotney J. 2018. High-resolution epigenomic atlas of human embryonic craniofacial development. *Cell Rep.* 23(5):1581–1597.
- Yao J, Wu X. 2019. Upregulation of *miR-149-3p* suppresses spinal chordoma malignancy by targeting *Smad3*. *Oncol Targets Ther.* 12:9987–9997.
- Zhang X-T, Wang G, Li Y, Chuai M, Lee KKH, Yang X. 2018. Role of FGF signalling in neural crest cell migration during early chick embryo development. *Zygote.* 26(6):457–464.

### **3. Danksagung**

Ich bedanke mich bei meiner Doktormutter Frau Dr. rer. nat. Kerstin Ludwig für die Übertragung dieses Forschungsthemas und für ihre Unterstützung, Geduld und sachkundige Kritik, mit der sie mir stets zur Seite stand.

Weiterhin gilt mein Dank Frau Dr. rer. nat. Ronja Hollstein und Frau Dr. rer. nat. Magdalena Laugsch, ohne die diese Veröffentlichung nicht möglich gewesen wäre.

Bei allen MitarbeiterInnen der Humangenetik Bonn, mit denen ich bei der Erstellung dieser Arbeit zusammengearbeitet habe und die mich unterstützt haben, möchte ich mich ebenfalls bedanken.

Für ihre nicht endende moralische Unterstützung danke ich meinem Mann Niklas Freier und meinen Freundinnen Barbara Kalthoff, Leonie Kalthoff, Luisa Wiedemann und Isabelle Winterhagen.

Ein besonderer Dank gilt meinen Eltern für ihren Beistand und ihre Ermutigung in allen Lebensphasen. Ihnen ist diese Arbeit gewidmet.