

Aus dem Departement für Diagnostische Labormedizin der  
Universität Tübingen

Institut für medizinische Genetik und angewandte Genomik

**Erzeugung eines Modells zur Quantifizierung der  
Promotoraktivität des KPNA3-Promotors**

**Inaugural-Dissertation  
zur Erlangung des Doktorgrades  
der Medizin**

**der Medizinischen Fakultät  
der Eberhard Karls Universität  
zu Tübingen**

**vorgelegt von**

**Wanner, Manuel Pascal**

**2021**

Dekan: Professor Dr. B. Pichler

1. Berichterstatter: Professor Dr. O. Rieß

2. Berichterstatter: Privatdozentin Dr. A. Nolte-Karayel

Tag der Disputation: 18.01.2021

## Inhaltsverzeichnis

1. Einleitung.....	1
1.1 Ätiologie .....	1
1.2 Symptomatik .....	3
1.3 Epidemiologie.....	5
1.4 Diagnostik .....	5
1.5 Behandlung .....	6
1.6 Rolle des KPNA3-Transportproteins .....	7
1.6.1 Nukleärer Import allgemein.....	7
1.6.2 Nukleärer Import von Ataxin-3 und Bedeutung von KPNA3 für die SCA3 .....	9
1.7 Zielsetzung.....	10
2. Material und Methoden.....	12
2.1 Material .....	12
2.1.1 Puffer/Lösungen/Medien.....	12
2.1.2 Kits.....	13
2.1.3 Enzyme .....	14
2.1.4 Antikörper .....	15
2.1.5 Geräte.....	15
2.1.6 Primer .....	16
2.1.7 DNA und Konstrukte .....	17
2.2 Methoden .....	18
2.2.1 Plasmid Extraktion .....	18
2.2.2 Sequenzierung.....	19
2.2.3 Generierung chemisch-kompetenter Zellen .....	20
2.2.4 Nested-PCR.....	20
2.2.5 Klonierung.....	21
2.2.6 Transfektion der Konstrukte für Western-Blot.....	26
2.2.7 Western Blot .....	27
2.2.8 Fast-Forward Transfektion für Luciferase-Assay .....	29
2.2.9 Luciferase-Assay .....	30
2.2.11 Identifizierung von Transkriptionsfaktoren .....	31
3. Ergebnisse .....	33

3.1 Generierung eines Konstrukts zur Untersuchung des KPNA3-Promotors .....	35
3.1.1 Klonierung des Abschnitts B .....	37
3.1.2 Klonierung des Abschnitts A .....	40
3.1.3 Fertigstellung des Konstrukts .....	43
3.2 Zusätzliche Konstrukte mit Varianten des Promotors .....	45
3.2.1 5'-Kürzeres Konstrukt .....	47
3.2.2 5'-kürzeres Konstrukt mit modifiziertem SNP .....	49
3.2.3 Konstrukte ohne 5' UTR des Exon 1 .....	52
3.3 Expression durch den KPNA3-Promotor .....	57
3.3.1 Expression im Western-Blot .....	57
3.3.2 Expression im Luciferase-Assay .....	58
3.4 Identifizierung relevanter Transkriptionsfaktoren .....	62
4. Diskussion .....	71
5. Zusammenfassung .....	81
6. Literaturverzeichnis .....	82
7. Erklärungen zum Eigenanteil .....	92
8. Anhang .....	93
8.1. Sequenzen der Konstrukte .....	93

## Abbildungsverzeichnis

Abbildung 1 – Vereinfachtes Schema des Nukleären Imports mittels NLS und Karyopherinen .....	8
Abbildung 2 – Schema des KPNA3-Gens am 5' Ende .....	34
Abbildung 3 – Nested-PCR mit einer doppelten Amplifikation .....	36
Abbildung 4 – Testverdau des Konstrukts nach Einklonierung des Abschnitts B .....	38
Abbildung 5 – Vorgehen zur Einklonierung des Abschnitt B .....	39
Abbildung 6 – Testverdau nach Klonierung des Abschnitts A in das Konstrukt aus 3.1.1 .....	41
Abbildung 7 – Vorgehen zur Einklonierung des Abschnitt A .....	42
Abbildung 8 – Testverdau für Länge des Abschnitt B im neuen Konstrukt .....	44
Abbildung 9 – SNP rs11618521 im Ausgangskonstrukt .....	45
Abbildung 10 – Übersicht der generierten Konstrukte .....	46
Abbildung 11 – Testverdau der angezogenen Klone der Klonierung für 5'-kürzeres Konstrukt mit EcoRI und NheI .....	47
Abbildung 12 – Vorgehen bei der Klonierung für 5'-kürzeres Konstrukt .....	48
Abbildung 13 – Sequenzierungsergebnisse des Konstrukts pGL3-KPNA3-Clone I und des 5'- kürzeren Konstrukts .....	49
Abbildung 14 – Sequenzierungsergebnisse der 5'-kürzeren Konstrukte mit den Varianten des SNP rs11618521 .....	50
Abbildung 15 – Vorgehen bei der Klonierung für 5'-kürzeres Konstrukt mit verändertem SNP rs11618521 .....	51
Abbildung 16 – Vorgehen bei der Klonierung für 3'-kürzeres Konstrukte ohne den 5'-untranslatierten Bereich des KPNA3-Gens .....	53
Abbildung 17 – Testverdau für 3'-kürzere Konstrukte ohne den 5'-untranslatierten Bereich des KPNA3-Gens mit PstI und XhoI .....	54
Abbildung 18 – Sequenzierungsergebnis der beiden Varianten des Konstrukts ohne den 5'-untranslatierten Bereich des KPNA3-Gens .....	55
Abbildung 19 – Ergebnisse der Expressionsanalyse mittels Western-Blot .....	58
Abbildung 20 – Unterschiede bei Vorgehen mit und ohne Mediumswechsel ...	60
Abbildung 21 – Analyse eines Versuchsansatzes in verschiedenen Geräten ..	61

Abbildung 22 – Relative Expression der Konstrukte im Luciferase-Assay .....	62
Abbildung 23 – Genome Browser mit ENCODE V2-Datensatz .....	65
Abbildung 24 – Ergebnis der Analyse mit dem PROMO Online-Tool.....	66
Abbildung 25 – Ergebnisse der Suche in der STRING-Datenbank .....	69

## **Tabellenverzeichnis**

Tabelle 1 – Übersicht über die Subtypen der Spinozerebellären Ataxie Typ 3 (Riess <i>et al.</i> 2008).....	4
Tabelle 2 – Übersicht zu allen genutzten Puffern, Lösungen und Medien.....	12
Tabelle 3 – Übersicht zu den in den Experimenten verwendeten Kits.....	13
Tabelle 4 – Übersicht zu den in den Experimenten verwendeten Enzymen.....	14
Tabelle 5 – Übersicht zu den in den Experimenten verwendeten Antikörpern .	15
Tabelle 6 – Übersicht zu den für die Experimente verwendeten Geräten .....	15
Tabelle 7 – Bezeichnung und Sequenz der in den Experimenten verwendeten Primer .....	16
Tabelle 8 – Zusammensetzung der Reaktionsansätze für die Sequenzierungen .....	19
Tabelle 9 – PCR-Programm der Sequenzierungen .....	19
Tabelle 10 – Zusammensetzung der Ansätze der Klonierungs-PCR .....	21
Tabelle 11 – Touchdown-PCR-Programm der Klonierungs-PCR.....	21
Tabelle 12 – Zusammensetzung der Ansätze für den präparativen Verdau.....	22
Tabelle 13 – Zusammensetzung der Ansätze für die Ligationen.....	24
Tabelle 14 – Zusammensetzung der Ansätze für Verdau zur Überprüfung der Klonierungsprodukte .....	26
Tabelle 15 – Rezept für die Bis-Tris-Polyacrylamid-Gele .....	28
Tabelle 16 – Übersicht über die erzeugten Konstrukte mit ihren Eigenschaften .....	56

## Abkürzungsverzeichnis

5' UTR	5' untranslated region, 5' nicht-translatierter Bereich
ADCA	Autosomal dominante zerebelläre Ataxie
DNA	Desoxyribonukleinsäure
GAPDH	Glycerinaldehyd-3-phosphat-Dehydrogenase
KPNA3	Karyopherin alpha 3
KPNB1	Karyopherin beta 1
MCS	Multiple Cloning Site
MJD	Machado-Joseph-Krankheit
mRNA	Messenger Ribonukleinsäure
MRT	Magnetresonanztomographie
NLS	Nuclear Localisation Signal
NPC	Nuclear Pore Complex
PCR	Polymerasekettenreaktion
SCA3	Spinozerebelläre Ataxie Typ 3
siRNA	Small interfering Ribonukleinsäure
SNP	Single Nucleotide Polymorphism (Einzelnukleotid-Polymorphismus)

### 1. Einleitung

Die Spinozerebelläre Ataxie Typ3 (SCA3) oder auch Machado-Joseph-Krankheit (MJD) ist eine erbliche neurodegenerative Erkrankung, die zu den autosomal dominanten zerebellären Ataxien gezählt wird. Die Ursache der Erkrankung ist ein verlängerter CAG-Repeat Bereich im *ATXN3*-Gen, wodurch im darin kodierten Protein Ataxin-3 ein Polyglutaminbereich expandiert wird. Im Folgenden werden die Hintergründe dieser Erkrankung näher beleuchtet und die Symptomatik, Diagnostik und aktuelle Behandlungsoptionen aufgezeigt. Zuletzt soll auf die Rolle des KPNA3-Transportproteins in Zusammenhang mit der SCA3 eingegangen und die Zielsetzung dieser Arbeit, in der die Promotorregion des KPNA3-Transportproteins im Mittelpunkt steht, erläutert werden.

### 1.1 Ätiologie

Die Ursache der Spinozerebellären Ataxie Typ 3 ist eine Mutation im langen Arm des Chromosom 14 (14q24.3-q32) (Takiyama *et al.* 1993) im *ATXN3*-Gen, das für das Protein Ataxin-3 kodiert. In diesem Gen ist ein Bereich mit CAG-Triplet Wiederholungen enthalten, der bei SCA3 Patienten verlängert ist (Kawaguchi *et al.* 1994). Bei Gesunden sind etwa 12 bis 44 CAG-Triplet-Wiederholungen vorhanden, bei SCA3-Patienten ist dieser Bereich auf 52 bis 86 Triplet-Wiederholungen erweitert (Sun *et al.* 2016). In seltenen Fällen, in denen die CAG-Wiederholungen zwischen diesen Bereichen lagen wurden sowohl Betroffene mit SCA3 typischen Symptomen beobachtet (Gu *et al.* 2004), aber auch symptomfreie Fälle sind beschrieben (Maciel *et al.* 2001). Die CAG-Wiederholungen kodieren für einen Polyglutamin-Bereich im durch das *ATXN3*-Gen kodierten Protein Ataxin-3.

Die Länge der CAG-Triplet Wiederholungen im *ATXN3*-Gen determiniert jedoch nicht nur, ob die Krankheit ausbricht oder nicht, sondern hat auch Einfluss auf das Alter bei Krankheitsbeginn. Hierbei ist das Erkrankungsalter umso niedriger desto mehr der Triplet-Bereich verlängert ist (Dürr *et al.* 1996). Das mittlere Erkrankungsalter liegt hierbei bei 36 Jahren, wobei auch Fälle mit Krankheitsbeginn ab 4 Jahren und mit bis zu 70 Jahren beschrieben sind (Sun *et al.* 2016). Weiterhin korreliert die Schwere der Symptome mit der Anzahl der

CAG-Wiederholungen, die Patienten sind umso schwerer betroffen, desto mehr CAG-Wiederholungen sie im *ATXN3*-Gen aufweisen (Netravathi *et al.* 2009; Silva-Fernandes *et al.* 2010). Bei präsymptomatischen Betroffenen ist dennoch keine eindeutige Prognose zum Krankheitsverlauf möglich (Schöls *et al.* 1996).

Physiologisch besitzt das Ataxin-3 deubiquitinierende Aktivität, wobei angenommen wird, dass es Polyubiquitinketten kürzt (Mao *et al.* 2005). Die physiologischen Substrate des Ataxin-3 sind bisher nicht hinreichend geklärt, wodurch die genaue Funktion des Proteins noch unklar ist. Bekannt ist, dass das Protein p53 ein Substrat von Ataxin-3 ist. p53 gilt vor allem als Tumorsuppressor und spielt eine Rolle bei Entzündungsreaktion (Uehara und Tanaka 2018). Ataxin-3 bindet an p53 und vermindert durch Deubiquitinierung und Stabilisierung dessen Abbau (Liu *et al.* 2016). Außerdem agiert Ataxin-3 als Transporter bei der Translokation von Proteinen zu Proteasomen der Zelle (Doss-Pepe *et al.* 2003). Insbesondere bei der Translokation fehlgefalteter Proteine des Endoplasmatischen Retikulums in das Zytosol der Zelle ist Ataxin-3 beteiligt (Wang *et al.* 2006). Weiterhin ist eine Rolle bei der Regulation der Zellantwort bei oxidativem Stress beschrieben (Araujo *et al.* 2011).

Durch Verlängerung des Polyglutaminbereichs im Ataxin-3 bei der SCA3 kommt es dagegen zu pathologischen Vorgängen, die im Folgenden näher erläutert werden. Vor allem die intranukleäre Lokalisation und Aggregation von expandiertem Ataxin-3 (Paulson *et al.* 1997; Schmidt *et al.* 1998) spielt eine zentrale Rolle bei der Pathogenese (Bichelmeier *et al.* 2007; Sowa *et al.* 2018). Hierbei wird zunächst das Ataxin-3 mit verlängertem Polyglutaminbereich durch proteolytische Enzyme gespalten, wodurch sogenannte „toxische Fragmente“ entstehen. Diese initiieren die Aggregation und führen durch Konformationsänderung des nicht-gespaltenen Ataxin-3 auch zu dessen Aggregation (Haacke *et al.* 2006). Bei der Spaltung von Ataxin-3 sind Caspasen (Berke *et al.* 2004; Jung *et al.* 2009) und Calpaine beteiligt. Insbesondere das Enzym Calpain-2 zeigte relevanten Einfluss auf die Entstehung toxischer Fragmente durch Spaltung von Ataxin-3 *in vitro* und *in vivo*, wobei expandiertes Ataxin-3 besonders sensitiv für die Spaltung durch Calpain-2 ist (Hübener *et al.* 2013). Das intranukleäre, aggregierte Ataxin-3 verursacht dann eine

Dysregulation der Transkription durch Beeinflussung der Funktion von Cajal-Körpern im Zellkern, die im Prozess des Splicings von RNA involviert sind (Sun *et al.* 2007), und durch Erhöhung der Expression pro-apoptotische Mediatoren (Chou *et al.* 2008). Weiterhin wurde gezeigt, dass durch expandiertes Ataxin-3 die Genexpression von Entzündungsmediatoren gesteigert wird und Proteine, die unter anderem das Überleben der Zelle regulierten, herunterreguliert werden (Evert *et al.* 2003). Außerdem ist die Fähigkeit des expandierten Ataxin-3, die Zellregulation bei oxidativem Stress anzupassen, reduziert. (Araujo *et al.* 2011). Durch das expandierte Ataxin-3 kommt es auch zu Störungen im Proteinstoffwechsel der Zelle, insbesondere beim Abbau von fehlgefalteten Proteinen des endoplasmatischen Retikulums, da die oben beschriebene physiologische Funktion des Ataxin-3 gestört ist (Li *et al.* 2015). Weiterhin haben insbesondere die durch die Calpain-vermittelte Spaltung entstandenen toxischen Fragmente auch Einfluss auf die Funktion der Mitochondrien. Die Funktion der Mitochondrien wird durch sie gestört und der mitochondriale Apoptoseweges wird initiiert (Chou *et al.* 2006; Harmuth *et al.* 2018).

Die dargestellten Vorgänge führen insgesamt zu einer Degeneration der betroffenen Neuronen. Davon sind mehrere zentralnervöse Regionen betroffen, unter anderem das Zerebellum, der Hirnstamm und die Basalganglien, aber auch Hirnnerven und das Rückenmark können betroffen sein. Diese Degenerationen werden im Verlauf auch makroskopisch sichtbar und durch die reduzierte Hirnmasse messbar (Rüb *et al.* 2008) und führen zur komplexen Symptomatik der betroffenen Patienten.

### 1.2 Symptomatik

Die Symptomatik der SCA3 ist sehr variabel. Im Allgemeinen können die typischen Symptome einer autosomal dominanten zerebellären Ataxie vorkommen. Hierzu gehören neben Gang- und Standataxie, Ataxie der Extremitäten und Einschränkungen von Sprache und der Okulomotorik auch Symptome wie extrapyramidale Bewegungsstörungen, Spastiken und peripherer Neuropathie (Schöls *et al.* 2004).

Ein mögliches Frühsymptom der SCA3 stellt das Sehen von Doppelbildern aufgrund der beginnenden Ophthalmoplegie dar (Globas *et al.* 2008). Im Verlauf tritt dann weitere Kleinhirnsymptomatik auf. Hierzu zählt die Ataxie der Extremitäten, was sich dann als Stand- und Gangataxie, Dysmetrie der Hände und Probleme beim Schreiben, sowie Sprach- und Schluckstörungen äußern kann. Neben dem oben genannten Wahrnehmen von Doppelbildern kann sich die gestörte Okulomotorik beziehungsweise Ophthalmoplegie auch durch einen beobachtbaren Nystagmus oder Pseudoexophthalmus äußern. Zusätzlich oder isoliert können jedoch auch Symptome hinzukommen, die nicht primär als Kleinhirnsymptomatik einzuordnen sind, und die sich je nach Verlaufsform der Erkrankung zeigen. Hierzu zählen das Auftreten eines Parkinsonismus, eines Restless-Leg-Syndroms, einer Dystonie oder einer axonalen Neuropathie, die durch ein vermindertes Vibrationsempfinden, Atrophie von Muskulatur oder Ausfall von Reflexen auffallen kann. Bei einigen Verlaufsformen können auch Pyramidenbahnzeichen, wie Spastiken und (im Gegensatz zu Verlaufsformen mit Neuropathie) gesteigerte Reflexe, bei Patienten vorzufinden sein. Häufig klagen Patienten auch über Muskelkrämpfe und Dysästhesien. Ein relativ spezifisches Symptom der SCA3 scheint die verminderte Temperaturdiskrimination in den Extremitäten, sowie am Körperstamm oder im Gesichtsbereich zu sein. (Dürr *et al.* 1996; Schöls *et al.* 1996; Schöls *et al.* 2004; Globas *et al.* 2008).

Wie sich durch diese Darstellung der Symptome zeigt, ist das klinische Bild der SCA3 sehr variabel. Dennoch wird die SCA3 klinisch in vier Subtypen, abhängig von vorherrschender Symptomatik und typischem Alter bei Krankheitsbeginn, eingeteilt (Riess *et al.* 2008). Die Subtypen sind in Tabelle 1 dargestellt.

*Tabelle 1 – Übersicht über die Subtypen der Spinozerebellären Ataxie Typ 3 (Riess *et al.* 2008)*

Subtyp	Symptomatik	Typisches Erkrankungsalter
Typ 1	Extrapyramidale (dystonische) und pyramidale Defizite	10-30 Jahre
Typ 2	Pyramidale und zerebelläre Defizite	20-50 Jahre
Typ 3	Zerebelläre Defizite und periphere Neuropathie	40-75 Jahre
Typ 4	Neuropathie und Parkinsonismus	variabel

### 1.3 Epidemiologie

Weltweit sind etwa 0,3 bis 2 pro 100.000 Einwohnern von autosomal-dominant vererbten spinozerebellären Ataxien betroffen (van de Warrenburg *et al.* 2002). Die Spinozerebelläre Ataxie Typ 3 ist mit einem weltweiten Anteil von ca. 21% die häufigste Form (Schöls *et al.* 2004). Bei Subgruppen portugiesischer und brasilianischer Herkunft macht die SCA3 sogar über 60% der spinozerebellären Ataxien aus (Silveira *et al.* 2002). In Deutschland ist davon auszugehen, dass etwa 42% der Patienten mit autosomal-dominant vererbten spinozerebellären Ataxien an SCA3 leiden (Schöls *et al.* 1997).

### 1.4 Diagnostik

Die Diagnostik der SCA3 umfasst mehrere Dimensionen. Diese sind in den Empfehlungen der S1-Leitlinie zu Ataxien des Erwachsenenalters, herausgegeben von der Deutschen Gesellschaft für Neurologie, aufgeführt (Klockgether *et al.* 2018) und sollen laut dieser in Kombination betrachtet werden. Hierzu gehört die Anamnese, die körperliche Untersuchung, bildgebende und andere apparative Verfahren, sowie die molekulargenetischen Untersuchungen. Bei der Anamnese sollte insbesondere auf das Erheben der Familienanamnese aufgrund des Erbgangs Wert gelegt werden, auch zur Abgrenzung von anderen Ataxien. In der körperlichen Untersuchung sollten die speziellen neurologischen Symptome der spinozerebellären Ataxien und insbesondere der SCA3 systematisch untersucht werden. Außerdem sollte eine MRT-Bildgebung des Gehirns erfolgen, unter anderem um fokale Kleinhirnprozesse auszuschließen. In der Diagnostik der SCA3 kann auch eine Neurographie Hinweise geben, da Neuropathien auftreten können. Die bisher genannten Untersuchungen lassen jedoch keine sichere Diagnose der SCA3 zu, sondern können höchstens Hinweise darauf geben. Letztlich muss die Diagnose mittels molekulargenetischer Untersuchung gestellt werden. Hier ist die Mutation im *ATXN3*-Gen mit vermehrten CAG-Triplet-Wiederholungen beziehungsweise ein verlängerter Polyglutaminbereich im Ataxin-3-Protein wegweisend für die Spinozerebelläre Ataxie Typ 3 (Klockgether *et al.* 2018).

### 1.5 Behandlung

Zum jetzigen Stand ist keine kausale Therapie der SCA3 möglich. Die Leitlinien für Ataxien im Erwachsenenalter der Deutschen Gesellschaft für Neurologie empfehlen eine symptomatische Behandlung; etwa bei Parkinsonismus den Einsatz von L-Dopa oder bei Dystonie Anticholinergika (Klockgether *et al.* 2018). In den Leitlinien wird außerdem regelmäßige Physiotherapie empfohlen, da diese bei Ataxien generell sehr wirksam ist. Ein physiotherapeutisches intensives Koordinationstraining kann insbesondere bei zerebellären Ataxien motorische Fähigkeiten verbessern und auch ataktische Symptome reduzieren (Ilg *et al.* 2009). Die Patienten profitieren auch im Langzeitverlauf, vor allem wenn sie selbständig Übungen durchführen, und erfahren letztlich eine Verbesserung in ihren Alltagsaktivitäten (Ilg *et al.* 2010).

Ein spezifischer Therapieversuch der SCA3 mit Riluzol ist laut Leitlinien prinzipiell möglich. Es gibt Hinweise darauf, dass Riluzol den Krankheitsprogress zerebellärer Ataxien positiv beeinflusst (Romano *et al.* 2015). Jedoch wurde im Tiermodell speziell für die SCA3 beobachtet, dass Riluzol hier keine Verbesserung der Symptomatik bewirkte und sogar eher eine Verschlechterung verursachte (Schmidt *et al.* 2016). Im Rahmen von Studien wurden weitere Medikamente angewandt. Valproat scheint bei der Spinozerebellären Ataxie Typ 3 einen positiven Effekt zu besitzen (Lei *et al.* 2016), dieser kann vermutlich darauf zurückgeführt werden, dass der Nukleäre Transport von Ataxin-3 vermindert wird (Wang *et al.* 2018). Ein weiterer Therapieansatz mit Lithium konnte in einer Phase 2 Studie zwar keinen signifikanten Einfluss auf den allgemeinen Progress der SCA3 zeigen, jedoch wurden positive Effekte auf den Krankheitsverlauf etwa in Hinsicht auf das Sprechen und die Koordination gefunden (Saute *et al.* 2014). Therapieversuche mit Vareniclin oder Cotrimoxazol erbrachten bisher keine ausreichende Evidenz zu ihrer Wirksamkeit (Zesiewicz *et al.* 2018). In experimentellen Ansätzen werden weiterhin Substanzen eingesetzt, wie der NGF (Nerve growth factor), der die Apoptose der Neurone im Zerebellum verhindern soll (Tan *et al.* 2015), oder auch IGF-1 (Insulin-like growth factor 1), der einen stabilisierenden Effekt auf den Krankheitsprogress bei ADCAs zeigte (Sanz-Gallego *et al.* 2014). Diese Ansätze haben aber bisher keine

klinische Relevanz. Aufgrund eines bislang fehlenden wirksamen Therapieansatzes für die SCA3 und dem Mangel an aussagekräftigen Studien zu Therapien, die präklinisch vielversprechende Ergebnisse zeigten, werden vermehrt randomisierte, placebokontrollierte klinische Studien gefordert. Hierbei werden multizentrische Studien, die etablierte Scores zur Evaluation und Vergleichbarkeit der Ergebnisse nutzen, empfohlen, sowie das Einschließen auch von präsymptomatischen SCA3-Patienten gefordert, um größere Fortschritte in der Suche nach Behandlungsoptionen zu erreichen (Saute *et al.* 2015; Saute und Jardim 2018; Da Silva *et al.* 2019).

### 1.6 Rolle des KPNA3-Transportproteins

Ataxin-3 ist ein vorwiegend im Zytosol der Zelle lokalisiertes Protein, das wie oben bereits beschrieben dort als deubiquitinierendes Protein agiert und beim Abbau fehlgefalteter Proteine eine Rolle spielt. Physiologisch kommt Ataxin-3 aber auch im Zellkern vor (Reina *et al.* 2010). Dort hat es vermutlich genregulatorische Funktion, insbesondere bei der Reaktion der Zelle auf Stress (Reina *et al.* 2012). Bei der SCA3 bildet das verlängerte Ataxin-3 im Zellkern hingegen Aggregate, die letztlich maßgeblich für die Pathogenese verantwortlich sind (Bichelmeier *et al.* 2007). Es ist also relevant zu verstehen, wie Ataxin-3 in den Zellkern gelangt. Hierzu wird im Folgenden der Nukleäre Import allgemein beleuchtet und anschließend auf den nukleären Import von Ataxin-3 eingegangen. Zuletzt wird die wesentliche Rolle des KPNA3-Transportprotein, dessen Promotorregion und damit dessen Expression im Mittelpunkt dieser Arbeit steht, beim Import des Ataxin-3 und somit bei der Pathogenese der SCA3 aufgezeigt.

#### 1.6.1 Nukleärer Import allgemein

Es gibt mehrere Möglichkeiten, wie Proteine in einer Zelle vom Zytoplasma in den Zellkern importiert werden können. Im Folgenden wird der für das Ataxin-3 und KPNA3-Transportprotein relevante Mechanismus mittels Nuclear Localisation Signal (NLS) und Karyopherinen näher beleuchtet.

Die Membran eines Zellkerns enthält Proteinkomplexe, die passive Diffusion von Ionen und kleineren Proteinen zwischen den beiden Kompartimenten,

Zytoplasma und Karyoplasma, zulassen. Diese werden als Nuclear Pore Complexes (NPCs) bezeichnet. Die Schwelle für die passive Diffusion liegt bei etwa 60kDa, wobei die Diffusionsgeschwindigkeit mit zunehmender Proteingröße sinkt (Allen *et al.* 2000) Größere Proteine benötigen Transportproteine, die sogenannten Karyopherine (teilweise auch als Importine bzw. Exportine bezeichnet) um diese NPCs zu nutzen. Die Karyopherine werden in Untergruppen kategorisiert, wobei vor allem die Gruppen alpha ( $\alpha$ -Karyopherine) und beta ( $\beta$ -Karyopherine) im Folgenden relevant sind. Die zu importierenden Proteine enthalten eine spezifische Aminosäuresequenz, die als Nuclear Localisation Signal (NLS) bezeichnet wird. An diese können die Karyopherine binden und das jeweilige Protein über den NPC in den Zellkern transportieren. Hierbei kann entweder ein  $\beta$ -Karyopherin direkt am NLS binden und den Import des Proteins bewirken, oder es bindet zunächst ein  $\alpha$ -Karyopherin am NLS, an dem wiederum ein  $\beta$ -Karyopherin binden kann, um den Import zu ermöglichen. Der Import erfolgt hierbei entlang eines Ran-Gradienten und die Bindungen dieses Komplexes werden im Zellkern gelöst (Lange *et al.* 2007). Ein vereinfachtes Schema dieses Vorgangs ist in Abbildung 1 dargestellt.

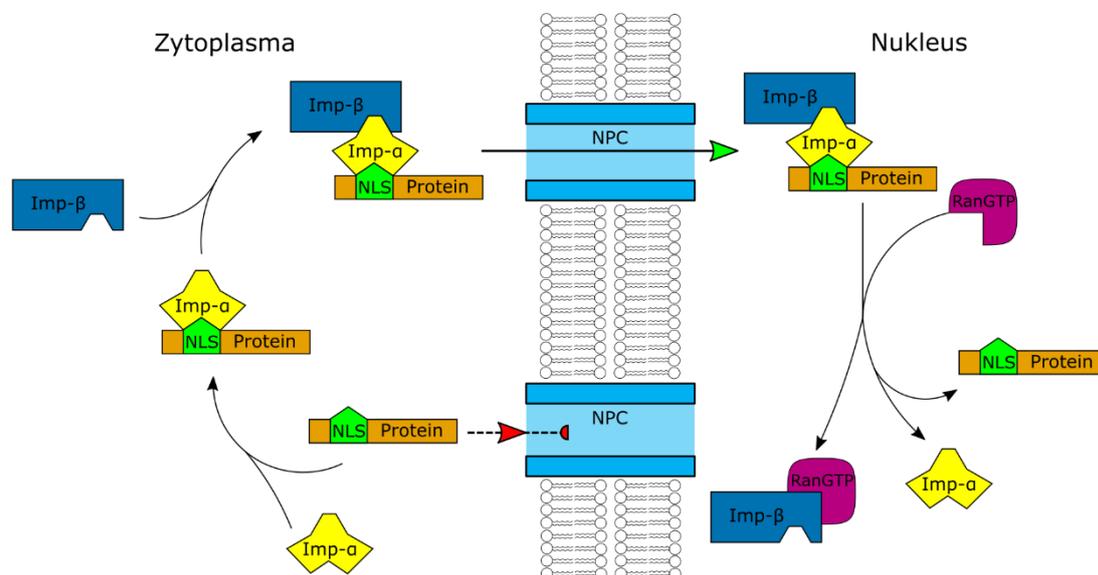


Abbildung 1 – Vereinfachtes Schema des Nukleären Imports mittels NLS und Karyopherinen – Das Protein kann aufgrund seiner Größe nicht allein mittels Diffusion über den Nuclear Pore Complex (NPC) in den Zellkern gelangen. Durch Bindung von Karyopherin  $\alpha$  (Imp- $\alpha$ ) an das Nuklear Import Signal (NLS) des Proteins und Bindung von Karyopherin- $\beta$  (Imp- $\beta$ ) an das Karyopherin- $\alpha$  kann das dabei entstandene Trimer über den NPC in den Zellkern gelangen. Dort werden die Bindungen durch RanGTP wieder gelöst. (In Anlehnung an: Pemberton und Paschal 2005)

### 1.6.2 Nukleärer Import von Ataxin-3 und Bedeutung von KPNA3 für die SCA3

Ataxin-3 wird mittels aktivem Transport in und aus dem Zellkern transportiert (Macedo-Ribeiro *et al.* 2009). Es enthält in seiner Aminosäuresequenz einen Bereich, der einem NLS entspricht, sowie auch sogenannte Nuclear Export Signals (NES), die jedoch für den Transport aus dem Zellkern hinaus Relevanz haben (Albrecht *et al.* 2004; Antony *et al.* 2009). Es wurde gezeigt, dass insbesondere das  $\alpha$ -Karyopherin KPNA3 (Karyopherin alpha 3 oder alternativ Importin- $\alpha$ 4) die Lokalisation des Ataxin-3 innerhalb der Zelle kontrolliert. KPNA3 bindet an das NLS des Ataxin-3 und bildet anschließend durch Bindung von KPNB1 an KPNA3 ein Trimer. Dieses wird durch den Nuclear Pore Complex in den Zellkern importiert. Fehlt KPNA3 beziehungsweise ist es vermindert wurde *in vitro* und *in vivo* weniger Ataxin-3 im Zellkern nachgewiesen und auch die pathologische Aggregation war vermindert (Sowa *et al.* 2018). Die nukleäre Lokalisation von Ataxin-3 ist jedoch wesentlicher Bestandteil dafür, dass die SCA3 symptomatisch wird (Bichelmeier *et al.* 2007). Wie bereits erwähnt spielt auch die nukleäre Aggregation von expandiertem Ataxin-3 eine Rolle in der Pathogenese. Insgesamt scheint KPNA3 also ein Schlüsselprotein bei der Pathogenese der Spinozerebellären Ataxie Typ 3 zu sein. Wenn weniger KPNA3 in der Zelle vorhanden ist, hat dies somit vermutlich positiven Effekt auf die SCA3, da bei fehlendem oder vermindertem KPNA3 weniger Ataxin-3 in den Zellkern gelangt und pathologisch aggregiert. Im Mausmodell konnte bereits gezeigt werden, dass das Fehlen von KPNA3 auch die Symptomatik SCA3 signifikant verbessert (Sowa *et al.* 2018). Das KPNA3-Transportprotein ist somit eine potenzielle Zielstruktur, um durch dessen Reduktion in der Zelle in den pathophysiologischen Vorgang der Translokation und Aggregation von verlängertem Ataxin-3 im Zellkern einzugreifen und dadurch die Symptomatik und eventuell den Krankheitsverlauf der SCA3 positiv zu beeinflussen.

### 1.7 Zielsetzung

KPNA3 ist wie oben beschrieben ein Schlüsselprotein der Pathogenese der SCA3. Folglich ist es sinnvoll, Beeinflussungsmöglichkeiten dieses Proteins zu untersuchen. Da die vorhandene Menge des KPNA3-Proteins die Menge an intrazellulärem und aggregiertem Ataxin-3 beeinflusst, ist ein Ansatz, die vorhandene Menge von KPNA3 zu reduzieren. Proteine werden durch Transkription von DNA in mRNA im Zellkern und anschließender Translation der mRNA in eine Aminosäuresequenz im Zytoplasma gebildet. Die Menge der gebildeten mRNA korreliert hierbei mit der Menge an Protein, die am Ende gebildet wird (Csardi *et al.* 2015). Auf DNA-Ebene wird hierbei die Transkription der mRNA durch Transkriptionsfaktoren, die an spezifische Promotorregionen des jeweiligen Gens binden, reguliert (Mitchell und Tjian 1989). Wenn also die Bindung von Transkriptionsfaktoren an die Promotorregion des *KPNA3*-Gens beeinflusst wird, wird folglich auch die translatierte Menge an KPNA3-Protein beeinflusst, was letztendlich den o.g. Einfluss auf die Menge an Ataxin-3 im Nukleus hat.

Ziel dieser Arbeit ist nun, ein Modell zu erzeugen, mit dem die Expression durch den KPNA3-Promotor untersucht werden kann. Hierfür soll ein Plasmid-Konstrukt erzeugt werden, das die Promotorregion des *KPNA3*-Gens enthält. Hierbei sollen Konstrukte in mehreren Varianten mit verschiedener Länge des enthaltenen Promotors erzeugt werden, um auf relevante (Teil-)Bereiche des Promotors rückzuschließen. In Vorarbeiten unserer Arbeitsgruppe sind weiterhin Polymorphismen im *KPNA3*-Promotor aufgefallen, die einen Einfluss auf das Erkrankungsalter der SCA3 hatten und somit im Verdacht standen, die Promotoraktivität zu beeinflussen. Hierbei handelte es sich um die Einzelnukleotid-Polymorphismen (SNPs), die die Identifikationsnummern rs9535343, rs12868235 und rs11618521 tragen (vergeben durch die dbSNP Datenbank der NCBI (Sherry *et al.* 2001)). Die beiden SNPs rs9535343 und rs12868235 zeigten einen Zusammenhang zum Erkrankungsalter, sind jedoch beide deutlich (über 2500bp in 5' Richtung) vom Transkriptionsstart des *KPNA3*-Gens entfernt (Köbsch 2020). Der SNP rs11618521 zeigte ebenfalls einen signifikanten Zusammenhang mit dem Erkrankungsalter der SCA3 und ist im

Gegensatz zu den beiden anderen SNPs deutlich näher zum Transkriptionsstart des Gens gelegen (ca. 1100bp 5' davon) (Dieterle 2019). Deshalb sollen auch die Varianten dieses Polymorphismus in den Konstrukten abgebildet werden. Die verschiedenen Konstrukte sollen anschließend hinsichtlich ihres Expressionsniveaus untersucht und verglichen werden und der Einfluss der Länge des Promotors und des jeweiligen Polymorphismus auf die Expression analysiert werden. Weiterhin sollen dann Transkriptionsfaktoren identifiziert werden, die die Expressionsunterschiede erklären können. Am Ende soll ein fertiges Modell stehen, an dem zukünftige *in vitro* Wirkstofftestungen vorgenommen werden können, bei denen der Einfluss auf die Expression des KPNA3-Proteins durch Wirkstoffe untersucht werden kann. Hierdurch könnten potenzielle neue Therapeutika für die SCA3 gefunden werden.

2. Material und Methoden2.1 Material2.1.1 Puffer/Lösungen/Medien

Alle genannten Chemikalien, Medien und Lösungen wurden in größtmöglicher Reinheit von Sigma-Aldrich Chemie GmbH (Steinheim), Carl Roth GmbH & Co. KG (Karlsruhe), Lonza Group AG (Basel), Fresenius Kabi Deutschland GmbH (Bad Homburg), Advansta Inc. (Menlo Park, USA) und Becton Dickinson GmbH (Heidelberg) bezogen.

*Tabelle 2 – Übersicht zu allen genutzten Puffern, Lösungen und Medien*

<b>Bezeichnung</b>	<b>Zusammensetzung/Herstellerbezeichnung</b>
Agarose	SeaKem® LE Agarose
Ampicillin (1000x)	Ampicillin 100mg/ml
Bis-Tris Puffer (3,5x)	1,25M BisTris
DMEM	Gibco™ DMEM, Fisher Scientific GmbH, Schwerte 10% FBS 1% A/A
ECL-Puffer	1ml WesternBright Peroxide Chemiluminescent peroxide solution, Advansta Inc., San Jose (USA) 1ml WesternBright ECL Luminol/enhancer solution, Advansta Inc., San Jose (USA) 1ml TBS-Puffer
Ladepuffer (Agarosegel-elektrophorese, 6x)	10mM Tris HCl (pH 7.6) 0,03% Bromphenolblau 60% Glycerol 60mM EDTA
LB-Medium	25g/l BD Difco™ LB Broth, Miller, Fisher Scientific GmbH, Schwerte
LDS Puffer (4x)	1M Tris-Puffer pH=8.5 0,3M Lithiumdodecylsulfat 2mM EDTA 43% Glycerol 0,025% Phenolrot
MES-Puffer (20x)	1M Morpholinoethansulfonsäure 1M Tris 96,3mM SDS 20,5mM EDTA
Opti-MEM	Gibco™ Opti-Mem™, Fisher Scientific GmbH, Schwerte

SOC-Medium	20g/l BD Bacto™ Tryptone 5g/l BD Bacto™ Yeast Extract 10mM NaCl 2,5mM KCl 20mM Mg <sup>2+</sup> , Filter-sterilisiert 20mM Glucose, Filter-Sterilisiert
Steriles Wasser	Ampuwa®, Fresenius Kabi Deutschland GmbH, Bad Homburg
TAE-Puffer	40mM Tris-Azetat 1mM EDTA
TBE-Puffer	90mM Tris-Borat 1mM EDTA
TBS (10x)	0,3M Tris 1,5M NaCl
TBST (10x)	1M Tris 5M NaCl 1% Tween 20
TFB1	100mM RbCl 50mM MnCl <sub>2</sub> 30mM Kaliumacetat 10mM CaCl <sub>2</sub> 15% Glycerol pH 5.8
TFB2	10 mM MOPS 10 mM RbCl 75 mM CaCl <sub>2</sub> 15% Glycerol pH 6.8
Transferpuffer (Western-Blot, 20x)	0,5M Bicin 0,5M Bis-Tris 0,02M EDTA

### 2.1.2 Kits

*Tabelle 3 – Übersicht zu den in den Experimenten verwendeten Kits*

<b>Verwendung</b>	<b>Bezeichnung</b>	<b>Hersteller</b>
Plasmid Extraktion	QIAprep Spin Miniprep Kit	QIAGEN GmbH, Hilden
Aufreinigen von DNA aus Agarosegelen	QIAquick Gel Extraction Kit	QIAGEN GmbH, Hilden
Aufreinigen von PCR-Produkten	QIAquick PCR Purification Kit	QIAGEN GmbH, Hilden
Sequenzierung	GenomeLab DTCS Quick Start Kit	Beckman Coulter GmbH, Sinsheim
Transfektion	Attractene Transfection Reagent	QIAGEN GmbH, Hilden
Bradford-Assay	Bio-Rad Protein Assay Kit II	Bio-Rad Laboratories GmbH, München

Luciferase Assay	Dual-Glo Luciferase Assay System	Promega GmbH, Mannheim
(Nested)-PCR	Expand Long Template PCR System	Roche Diagnostics GmbH, Mannheim

### 2.1.3 Enzyme

*Tabelle 4 – Übersicht zu den in den Experimenten verwendeten Enzymen*

<b>Enzym</b>	<b>Artikelnummer</b>	<b>Hersteller</b>
CIP (Alkaline Phosphatase, Calf Intestinal)	M0290S	New England Biolabs GmbH, Frankfurt
EcoRI-HF	R3101S	New England Biolabs GmbH, Frankfurt
NdeI	R0111S	New England Biolabs GmbH, Frankfurt
NheI	R0131S	New England Biolabs GmbH, Frankfurt
peqGold pwo-Polymerase	01-5010	PEQLAB Biotechnologie GmbH, Erlangen
PfifI	R0595S	New England Biolabs GmbH, Frankfurt
Psyl	ER1331	Fisher Scientific GmbH, Schwerte
T4 DNA Ligase	M0202S	New England Biolabs GmbH, Frankfurt
T4 DNA Ligase	EL0011	Fisher Scientific GmbH, Schwerte
XhoI	R0146S	New England Biolabs GmbH, Frankfurt

Für die Reaktionsansätze wurden die vom Hersteller für das jeweilige Enzym empfohlenen Reaktionspuffer genutzt. Hierbei kamen für die Enzyme von New England Biolabs GmbH (Frankfurt) die herstellereigenen Puffer NEBuffer™ 2.1, NEBuffer™ CutSmart® und T4 DNA Ligase Reaction zum Einsatz. Für Enzyme von Fisher Scientific GmbH (Schwerte) wurden die herstellereigenen Puffer T4 DNA Ligase Buffer (10X) und Buffer B genutzt. Für die peqGold pwo-Polymerase von PEQLAB Biotechnologie GmbH (Erlangen) wurde der im Set enthaltene 10x Reaktionspuffer-komplett eingesetzt.

### 2.1.4 Antikörper

*Tabelle 5 – Übersicht zu den in den Experimenten verwendeten Antikörpern*

<b>Antikörperbezeichnung</b>	<b>Artikelnummer</b>	<b>Hersteller</b>
Anti-Luciferase	L0159	Sigma-Aldrich Chemie GmbH, Steinheim
Goat anti-Rabbit IgG H&L (HRP)	ab6721	Abcam plc., Cambridge (UK)
Goat anti-Rabbit IgG H&L (IRDye® 800CW)	ab216773	Abcam plc., Cambridge (UK)
Anti-GAPDH antibody [mAbcam 9484] - Loading Control	ab9484	Abcam plc., Cambridge (UK)
Donkey anti-Mouse IgG H&L (IRDye® 800CW)	ab216774	Abcam plc., Cambridge (UK)

### 2.1.5 Geräte

*Tabelle 6 – Übersicht zu den für die Experimente verwendeten Geräten*

<b>Verwendung</b>	<b>Gerätebezeichnung</b>
PCR-Cycler	Alphamatrix Biotech G-Storm Bio-Rad T100 Thermal Cycler Bio-Rad C1000 Touch Thermal Cycler
Vakuumtrockner	Eppendorf Concentrator 5301
Photometer	Eppendorf Biophotometer ( $\mu$ -Küvette G1,0)
Zentrifugen	Eppendorf Centrifuge 5417C Eppendorf Centrifuge 5810R Eppendorf Centrifuge 5415R
Sequenz Analyse	Beckman Coulter CEQ 8000
Bradford-Assay	Biotek Synergy HT
Western Blot	Bio-Rad Mini-PROTEAN® Tetra Handcast Systems Bio-Rad Mini-PROTEAN® Tetra Vertical Electrophoresis Cell Hoefer™ TE22 Mini Tank Blotting Unit
Detektion	Li-Cor Odyssey Fc
Messung Lumineszenz	Berthold Technologies MicroLumat Plus LB96V PerkinElmer EnVision multimode plate reader Biotek Synergy HT

### 2.1.6 Primer

Die Primer wurden von der metabion international AG (Planegg) bezogen.

*Tabelle 7 – Bezeichnung und Sequenz der in den Experimenten verwendeten Primer*

<b>Bezeichnung</b>	<b>Sequenz von 5' nach 3'</b>
S058_Rvprimer3	CTAGCAAATAGGCTGTCCC
S059_Glprimer2	CTTTATGTTTTTGCGTCTTCCA
S073_KPNA3Pr-Xho-R	CCTCCTCGAGGGCTGCGCGCGGCTCCGG
S074_KPNA3Pr-Apa-F	TATGCTTGTGTTGCCTACGG
S075_KPNA3Pr-Eco-R	CTTAAAGAACCTGAGGCCCA
S077_KPNA3Pr-Sac-R	GGGAGAGAAGAGCACGTTCT
S078_KPNA3Pr-2600-F	AGGGTGCCACATTTGTGAAT
S079_KPNA3Pr-2100-F	CCTGACCTCAGGTGATCCTC
S080_KPNA3Pr-1600-F	TGCATGTTTTTGTGGGTTG
S081_KPNA3Pr-1100-F	CCCACCTGCTCCAGAAAGTA
S082_KPNA3Pr-600-F	ACTGCCAGAGTGAAGGTTGC
S088_KPNA3_-3746-F	GTGGACTGCTTAGGAGGTGT
S093_KPNA3_+058-R	GGCCCTTGTCTTGAAGCTC
S096_KPNA3_+168-R	ACAAGCCGACGAGAACCAG
S097_KPNA3Pr-Nhe-F2	GTGAGCTAGCCTGTTATGCTTGTGTTGCCT
S098_KPNA3Pr-Eco-F	GTCCAGTATGACCAGAGGATCT
S159_KPNA3Pr2-Nhe-F	GTGAGCTAGCGGTGTTGCAAAGTGGAGACC
S160_KPNA3Pr3-Nhe-F	GTGAGCTAGCCACTCACAAGACGAATGCC
S782_KPNA3Pr-2400-R	CAGAACTTGTGTGTTTGCCA
S785_KPNA3Pr-1300-F	CCAAGGCCCTTTCGCTTTAA
S786_KPNA3Pr-1189-F	CAAGCAAATGAAGCCCACCT
S787_KPNA3Pr-850-R	GGAGGGCAGATCACAAGGT
S788_KPNA3PrEx1Xho-R	CCTCCTCGAGGGCGGTCGTCAACTACAGTTC
S789_KPNA3Pr-550-R	ACTTGTTGAACTGGCGCTTT
S942_KPNA3Pr-2177-R	CATGGTGGTGCATGCCTGTAGTCA
S943_KPNA3Pr-619-R	TATCAGGAGGAAGAGGGCGGG
S944_KPNA3Pr-376-F	GCAGTAAAAGTCGCCAGGTCCGG
S946_KPNA3Pr-1245-R	ACTGAGGTCAAAGTTGGTAGGACT
S983_KPNA3Pr-Nde-F	CAAGTAGCTGAGATTACAGGTATG
S984_KPNA3Pr-Asp-R	CCAGCCTGGCGACAGAGG

### 2.1.7 DNA und Konstrukte

Als Ausgangsvektor wurde der pGL3-Basic Vektor (Artikelnummer E175A) und als Kontrollvektor für das Luciferase-Assay der Vektor pRL-CMV (Artikelnummer E2261) von der Promega GmbH (Mannheim) verwendet.

Die genomische DNA stammte von einer gesunden Kontrollperson nach deren Einwilligung.

Ein ca. 420bp großes DNA-Stück für die Klonierung wurde als gBlocks Gene Fragment von Integrated DNA Technologies Inc. (Skokie, USA) bezogen.

Folgende Konstrukte wurden im Verlauf selbst erzeugt:

- pGL3-basic construct
- pGL3-KPNA3-clone I
- pGL3-KPNA3-abEcoG
- pGL3-KPNA3-abEcoA
- pGL3-KPNA3-noUTR-G
- pGL3-KPNA3-noUTR-A

### 2.2 Methoden

#### 2.2.1 Plasmid Extraktion

Zur Plasmidgewinnung wurden die entsprechenden Stämme in LB-Medium mit Zusatz von Ampicillin in einer Endkonzentration von 100µg/ml angezogen. Inkubiert wurde über Nacht bei 37°C auf dem Schüttler bei 220rpm.

Die Extraktion der Plasmide erfolgte mit Hilfe des QIAprep Spin Miniprep Kits und dem zugehörigen Protokoll „Plasmid DNA Purification using the QIAprep Spin Miniprep Kit and a Microcentrifuge“ nach Angaben des Herstellers QIAGEN GmbH (Hilden) wie folgt. Zunächst wurde die Übernachtskultur für 10min bei ca. 450 g abzentrifugiert und der Überstand abgeschüttet. Die Bakterien wurden dann in 250µl Puffer P1 resuspendiert und in ein 1,5ml Reaktionsgefäß gegeben. Es erfolgte darauf die Lyse der Bakterien durch Zugabe von 250µl Puffer P2. Durch 350µl Puffer N3 wurde anschließend die Lysereaktion gestoppt und die Zellbestandteile für 10 min bei 17800 g abzentrifugiert. 800µl des Überstandes der die DNA enthielt wurden in die im Kit enthaltenen Spin Columns gegeben und diese anschließend für 1 min zentrifugiert, wobei die DNA in der enthaltenen Membran verbleibt. Nach Entfernen des Durchflusses wurde 500µl Puffer PB zugegeben und erneut für 1 min zentrifugiert. Der Durchfluss wurde wieder entfernt und 750µl Puffer PE zugegeben und für 1 min zentrifugiert. Hierdurch wurden zurückgebliebene Verunreinigungen ausgewaschen. Anschließend wurde der Durchfluss erneut entfernt und der restliche Puffer durch Zentrifugation entfernt. Nun wurde die Spin Column in ein frisches 1,5ml Reaktionsgefäß gegeben und anschließend die gebundene DNA mit 50µl Puffer EB eluiert. Nach Zugabe des Puffers wurde für eine Minute das Reaktionsgefäß bei Raumtemperatur inkubiert und dann für 1min zentrifugiert, um die DNA letztlich in das 1,5ml Reaktionsgefäß zu transferieren.

Die DNA-Konzentration der erhaltenen Plasmid-Lösungen wurde am Photometer (Eppendorf BioPhotometer) mit der Voreinstellung „dsDNA“ bei 260nm gemessen. Zur Bestimmung des Leerwerts kam der zur Lösung der Plasmide genutzte Puffer EB aus dem o.g. Kit zum Einsatz.

### 2.2.2 Sequenzierung

Die Sequenzierung erfolgte nach der Methode nach Sanger (Sanger *et al.* 1977). Die Reaktionsansätze wurden wie in Tabelle 8 aufgeführt gewählt und in 0,2ml PCR-Reaktionsgefäßen mit einem Zielvolumen von 10µl angesetzt.

*Tabelle 8 – Zusammensetzung der Reaktionsansätze für die Sequenzierungen*

H <sub>2</sub> O	0-5µl (gemäß Zielvolumen)
DNA-Template	260-520ng
Primer (10pmol/µl)	0,5µl
DTCS Quick Start Master Mix*	4µl

\*aus dem GenomeLab™ DTCS Quick Start Kit von Beckman Coulter GmbH (Sinsheim).

Die DNA wurde jeweils vor Zugabe der restlichen Lösungen bei 96°C für 1min vordenaturiert, anschließend wurden die kompletten Ansätze mit dem in Tabelle 9 aufgeführten PCR-Programm in den PCR-Cycler gegeben.

*Tabelle 9 – PCR-Programm der Sequenzierungen*

30x	20sec	96°C
	20sec	50°C
	4min	60°C
1x	Halten	10°C

Das Sequenzierprodukt wurde nach dem Protokoll „GenomeLab™ Dye Terminator Cycle Sequencing with Quick Start Kit“ von Beckman Coulter GmbH (Sinsheim) wie folgt aufgereinigt. Zunächst wurde in einem 1,5ml Reaktionsgefäß 5µl Stop-Solution (bestehend aus 2µl 3M Na-Azetat, 2µl 100mM Na<sub>2</sub>-EDTA und 1µl 20mg/ml Glykogen) vorgelegt, die die vorangegangene PCR-Reaktion beenden sollte. Das Sequenzierprodukt wurde anschließend zugegeben, das 0,2ml PCR-Reaktionsgefäß nochmals mit 10µl Wasser gespült und diese 10µl auch in das 1,5ml Reaktionsgefäß gegeben, um möglichst alle bei der PCR entstandene DNA zu überführen. Um das PCR-Produkt aufzureinigen wurde anschließend 60µl eiskaltes 95% Ethanol zugegeben und bei 4°C und 16100 g für 30 min zentrifugiert. Nach Entfernen des Überstandes wurde das entstandene Pellet nochmals zweimal mit 200µl eiskaltem 70% Ethanol gewaschen mit jeweils anschließender Zentrifugation für 15min bei 16100 g und Entfernen des

Überstandes. Anschließend wurde das Pellet für 10min vakuumgetrocknet, um verbliebenes Ethanol zu entfernen, und danach in 40µl Sample Loading Solution (enthalten im Quick Start Kit, s.o.) resuspendiert. Die Proben wurden vor der Analyse für 30-60min bei 4°C gelagert, um sich vollständig in der Sample Loading Solution zu lösen. Die Analyse der Proben erfolgte dann im Beckman Coulter CEQ 8000 nach Angaben des Herstellers.

### 2.2.3 Generierung chemisch-kompetenter Zellen

Zur Generierung chemisch-kompetenter Zellen wurden zunächst in 20ml LB-Medium eine Übernachtskultur von DH5alpha angezogen. Von der Kultur wurden 2ml zu 200ml vorgewärmtem LB-Medium gegeben und dieses bei 37°C auf dem Schüttler inkubiert bis eine optische Dichte von 0,45 bei 595nm erreicht war (als Leerwert diente LB-Medium). Dann wurde die Lösung für 5min auf Eis inkubiert und anschließend in 50ml Falcon-Tubes bei 1500g und 4°C für 5min abzentrifugiert. Bei allen folgenden Schritten wurden die Tubes auf Eis gelagert. Der Überstand wurde verworfen und jeweils 10ml 4°C-kalter TFB1-Puffer zugegeben. Nach 90min Inkubation auf Eis wurde erneut für 5min bei 1500g und 4°C zentrifugiert, der Überstand wiederum verworfen und jeweils 4ml eiskalter TFB2-Puffer zugegeben. Die Lösungen wurden für 15min auf Eis inkubiert und anschließend in 500µl Aliquots in flüssigem Stickstoff schockgefroren. Die chemisch-kompetenten Zellen wurden bis zur Transformation bei -80°C gelagert.

### 2.2.4 Nested-PCR

Zur Amplifikation von großen DNA-Bereichen aus genomischer DNA kam eine Nested-PCR zum Einsatz. Hierbei wurden zunächst Primer gewählt, die einen größeren Bereich als den letztlich gewünschten Zielbereich einschlossen, jedoch sehr spezifische und starke Bindungsstellen in der DNA besaßen, um nur den gewählten Bereich ohne unerwünschte Nebenprodukte zu amplifizieren. Das hierbei entstandene Produkt wurde als Ausgangspunkt für Klonierungs-PCRs genutzt. Hierdurch wurde sichergestellt, dass bei nachfolgenden Klonierungs-PCRs die Klonierungs-Primer nicht an potenziell zusätzliche Bindungsstellen in der genomischen DNA binden können und unerwünschte Produkte generiert werden. Zur Nested-PCR aus genomischer DNA wurde nach dem Protokoll

„Expand Long Template PCR System“ mit gleichnamigem Kit der Firma Roche Diagnostics GmbH (Mannheim) vorgegangen.

2.2.5 Klonierung

*A Klonierungs-PCR*

Die Generierung von Inserts für die Klonierung erfolgte mittels PCR mit der pwo-Polymerase und zugehörigem Reaktionspuffer von PEQLAB Biotechnologie GmbH (Erlangen). Die PCRs wurden wie in Tabelle 10 aufgeführt angesetzt.

*Tabelle 10 – Zusammensetzung der Ansätze der Klonierungs-PCR*

H <sub>2</sub> O	Bis zu Zielvolumen von 20µl
10 x Reaktionspuffer komplett (mit 20 mM MgSO <sub>4</sub> )	2 µl
2mM dNTPs	2 µl
Template DNA	10-25ng
Forward-Primer (10 pmol/µl)	0,6 µl
Reverse-Primer (10 pmol/µl)	0,6 µl
peqGOLD Pwo-DNA-Polymerase (1 U/µl)	0,8 µl

Die Ansätze wurden in den PCR-Cycler gegeben und die Inserts mit dem in Tabelle 11 dargestellten PCR-Programm als Touchdown-PCR amplifiziert.

*Tabelle 11 – Touchdown-PCR-Programm der Klonierungs-PCR*

1 x	2 min	94 °C
10 x	15 s	94 °C
	30 s	65 °C -1°C pro Zyklus
	1 min	68 °C
25 x	15 s	94 °C
	30 s	55 °C
	1 min	68 °C
1 x	7 min	68 °C
1 x	Halten	10°C

Anschließend wurden die PCR-Produkte mit dem QIAquick PCR Purification Kit und zugehörigem „Quick Start Protocol“ wie folgt aufgereinigt: Zunächst wurde

zum PCR-Produkt Puffer PB in 5-fachem Volumen des Produktes zugegeben und gemischt. Die Lösung wurde dann komplett in eine QIAquick Column des Kits gegeben und für 1min bei 17800 g zentrifugiert, wobei die DNA in der enthaltenen Membran zurückblieb. Der Durchfluss wurde anschließend verworfen. Um unerwünschte Rückstände zu entfernen, wurde 750µl Puffer PE zugegeben und wieder für 1min zentrifugiert. Der Durchfluss wurde dann erneut verworfen und zum Entfernen des restlichen Puffers nochmals für eine Minute zentrifugiert. Die QIAquick Column wurde anschließend in ein frisches 1,5ml Reaktionsgefäß gegeben. Zuletzt wurde die DNA durch Zugabe von 30µl Wasser auf die Membran der Column eluiert, eine Minute bei Raumtemperatur inkubiert und anschließend die DNA-Lösung durch Zentrifugation für 1 min in das Reaktionsgefäß überführt.

*B Präparativer Verdau*

Präparative Verdauungen erfolgten in 0,2ml PCR-Reaktionsgefäßen. Es wurde jeweils mit zwei verschiedenen Enzymen verdaut, um Klonierungsstellen zu erzeugen, die sich voneinander unterscheiden. Insert und Vektor wurden jeweils parallel verdaut. Die Ansätze wurden wie in Tabelle 12 aufgeführt gewählt:

*Tabelle 12 – Zusammensetzung der Ansätze für den präparativen Verdau*

<u>Insertverdau</u>	<u>Vektorverdau</u>
26µl aufgereinigtes PCR-Produkt	1-1,5µg Vektor-DNA
3µl 10x Reaktionspuffer	3µl 10x Reaktionspuffer
0,5µl Enzym 1	0,5µl Enzym 1
0,5µl Enzym 2	0,5µl Enzym 2
	Wasser bis zum Zielvolumen von 30µl

Die Ansätze wurden jeweils für 3h bei 37°C inkubiert und anschließend je nach Herstellerangaben des Enzyms für 20 min bei 65°C bzw. 80°C Hitzeinaktiviert.

Zum Einsatz kamen Reaktionspuffer NEBuffer™ 2.1 und NEBuffer™ CutSmart® von New England Biolabs GmbH (Frankfurt) und Buffer B von Fisher Scientific GmbH (Schwerte) je nach Herstellerangaben der gewählten Enzyme.

### *C Präparative Gele*

Die präparativen Verdauungen wurden auf präparativen Agarosegelen aufgetrennt. Als Puffer wurde TAE-Puffer genutzt. Die Agarosekonzentration der Gele betrug je nach aufzutrennenden Fragmenten zwischen 0,8% und 2%. Es wurden 50ml-Gele mit Ethidiumbromid als Färbemittel genutzt.

Vor Laden der Gele wurde den Verdauungen 6x Ladepuffer zugegeben, sodass dieser im zu ladenden Produkt in einfacher Konzentration vorlag, also auf ein Sechstel verdünnt.

Nach Laden der Gele wurden diese je nach aufzutrennenden Fragmentgrößen zwischen 30-60min bei 90-110V laufen gelassen.

Die entsprechenden gewünschten Banden wurden dann am UV-Tisch unter Sicht herauspräpariert und die Gelfragmente in 2ml Reaktionsgefäße gegeben.

Die Aufreinigung der DNA aus dem Gel erfolgte mit Hilfe des QIAquick Gel Extraction Kit und zugehörigem Quick Start Protocol wie folgt: Zunächst wurde das Gewicht der präparierten Gelfragmente bestimmt, woraus sich auf das Volumen der Fragmente schließen ließ (100mg~100µl). Dann wurde Puffer QG in dreifachem Volumen des Gels hinzugegeben und die Lösungen für 10min bei 50°C inkubiert, um das Gel zu verflüssigen und im Puffer zu lösen. Anschließend wurde der gesamte Ansatz in QIAquick Spin Columns gegeben und für 1min bei 17800 g zentrifugiert, wobei die DNA in der enthaltenen Membran zurückblieb. Der Durchfluss wurde verworfen und 750µl Puffer PE zum Waschen zugegeben, 4min bei Raumtemperatur inkubiert und für 1min zentrifugiert. Der Durchfluss wurde dann erneut verworfen und zum Entfernen des restlichen Puffers nochmals für eine Minute zentrifugiert. Dann wurde das Säulchen in ein frisches 1,5ml Reaktionsgefäß gegeben. Auf die Membran des Säulchens wurde zum Eluieren der DNA zuletzt 30µl Wasser gegeben, 4min lang bei Raumtemperatur inkubiert und dann für eine Minute zentrifugiert, um die DNA-Lösung in das Reaktionsgefäß zu überführen. Die Lösungen wurden bis zur Ligation bei -20°C gelagert.

### D Ligation

Zur Abschätzung der DNA-Mengen für die Ligationen wurden je 2µl der aufgereinigten DNA-Lösungen auf TBE-Agarosegelen aufgetragen. Es wurden Gele mit Agarosekonzentrationen zwischen 1%-2% genutzt. Die Laufzeiten betragen zwischen 30-50min bei 90-120V. Zur Färbung wurde den Gelen Ethidiumbromid zugegeben.

Die Intensität von Inserts und Vektoren wurde verglichen und dann zur Abschätzung für die einzusetzenden Mengen für die Ligation folgende Formeln angewendet:

- $\mu l_{Vektor} = \mu l_{Insert} \times f$  bzw.  $\mu l_{Insert} = \frac{\mu l_{Vektor}}{f}$
- wobei  $f = \frac{Länge_{Vektor}}{Länge_{Insert}} \times \frac{Intensität_{Insert}}{Intensität_{Vektor}} \times \frac{Volumen_{Vektor}}{Volumen_{Insert}} \times \frac{1}{Überschuss}$
- Volumen entspricht dem aufgetragenen Volumen auf dem Agarosegel; als Überschuss wurde 3 bis 5 gewählt.

Die Ansätze wurden dann mit den berechneten Mengen nach dem in Tabelle 13 aufgeführten Schema gewählt.

Tabelle 13 – Zusammensetzung der Ansätze für die Ligationen

H <sub>2</sub> O	Bis zu Zielvolumen von 10µl
10x T4 Ligase Puffer*	1µl
Insert (je nach Berechnung)	0,5µl-7,5µl
Vektor (je nach Berechnung)	0,5µl-7,5µl
T4 Ligase*	1µl

\*von New England Biolabs GmbH (Frankfurt) oder Fisher Scientific GmbH (Schwerte)

Die Ligationsansätze wurden entweder für 2h bei 22°C oder bei 16°C über Nacht inkubiert und anschließend bis zur Transformation bei 4°C gelagert.

### E Transformation

Transformationen erfolgten mittels Heat-Shock-Transformation. Die Ligationsprodukte wurden vor der Transformation für 10-20min bei 65°C inkubiert zur Hitzeinaktivierung der Ligase und anschließend auf Eis abgekühlt.

Transformiert wurden DH5alpha chemisch kompetente Zellen. Zunächst wurden diese schonend auf Eis aufgetaut.

Dann wurden 100-150µl der kompetenten Zellen in vorgekühlte 2ml Reaktionsgefäße gegeben. Pro Ansatz wurden 5µl des jeweiligen Ligationsansatzes zugegeben und vorsichtig durch Rühren mit der Pipettenspitze und leichtes Klopfen gegen das Reaktionsgefäß durchmischt, um die enthaltenen Zellen möglichst zu schonen und nicht durch zu große Manipulation zu zerstören. Die Ansätze wurden dann für 30min auf Eis inkubiert, um eine Annäherung der Ligationsprodukte und der Zellen zu erreichen.

Anschließend erfolgte der Hitzeschock bei 42°C für 45s im Wasserbad, der die Membran der Zellen für die erzeugten Plasmide durchgängig machen sollte. Direkt danach wurden die Ansätze nochmals für 2min auf Eis inkubiert, um die Membranen wieder in ihren Ausgangszustand zu überführen. Nach Zugabe von 900µl LB- oder SOC-Medium wurden die Ansätze für 1-2h bei 37°C auf dem Schüttler (220rpm) inkubiert, wodurch sich die Zellen vom Hitzeschock erholen sollten.

Dann wurden die Ansätze für 5min bei 600g abzentrifugiert, 800µl des Überstandes entfernt und die Zellen im übrigen Medium resuspendiert.

Die kompletten Ansätze wurden dann auf LB-Agarplatten, die mit Ampicillin in einer Konzentration von 100µg/ml versetzt waren, ausplattiert und die Platten über Nacht bei 37°C inkubiert. Der Ampicillinzusatz diente der Selektion, da Zellen nur bei erfolgreicher Transformation eine Ampicillinresistenz (kodiert auf dem transformierten Plasmid) besaßen und somit wachsen konnten.

Aus den auf den Platten gewachsenen Kolonien wurden Übernachtskulturen angelegt und die Plasmide gewonnen (siehe unter 2.2.1 Plasmid Extraktion).

### *F Überprüfen der Klonierungsprodukte*

Zur Überprüfung der Klonierungsergebnisse wurden die erhaltenen Plasmide mit den für die Klonierung genutzten Enzymen verdaut, um die Restriktionsschnittstellen und die Größe des Inserts zu überprüfen.

Die Verdauungen wurden wie folgt angesetzt:

*Tabelle 14 – Zusammensetzung der Ansätze für Verdauung zur Überprüfung der Klonierungsprodukte*

500ng DNA
1µl 10x Reaktionspuffer
0,5µl Enzym 1
0,5µl Enzym 2
Wasser bis zum Zielvolumen von 10µl

Der Verdauung erfolgte für 3h bei 37°C mit anschließender Hitzeinaktivierung bei (je nach Enzym-Spezifikationen) 65°C-80°C für 20min.

Anschließend wurden die Verdauung mit 2µl 6x Ladepuffer versehen und auf TBE-Agarosegelen aufgetrennt und die Banden analysiert

Anschließend wurden die Bereiche von Interesse der Plasmide, die im Testverdau positive Resultate brachten, mittels Sanger-Sequenzierung sequenziert, um die gewünschte Klonierung zu bestätigen und eventuelle ungewünschte Mutationen auszuschließen. Vorgegangen wurde hierbei wie unter 2.2.2 Sequenzierung.

### 2.2.6 Transfektion der Konstrukte für Western-Blot

Die Konstrukte für die Western Blots wurden in HEK293T-Zellen transfiziert. Hierfür wurden aus bestehenden 10ml Kulturflaschen mit HEK293T-Zellen die Zellen auf 6-well Zellkulturplatten wie folgt ausplattiert: Zunächst wurde das Medium aus den Kulturflaschen entfernt. Anschließend wurden 10ml vorgewärmter PBS-Puffer zugegeben und darauf geachtet, dass der gesamte Zellrasen bedeckt war. Der PBS Puffer wurde wiederum entfernt und die Zellen dann mit 1ml Trypsinlösung trypsinisiert, um die Zellverbindungen zu lösen und die Zellen frei in Suspension bringen zu können. Nach 2min Inkubationszeit bei 37°C wurden die Zellen nach Zugabe von 9ml DMEM-Medium im Gesamtvolumen von 10ml resuspendiert. 9ml des Mediums mit gelösten Zellen wurden entnommen und die Zellen unter dem Mikroskop zur Abschätzung der Konzentration gezählt. Dann wurden pro Well einer 6-Well-Platte  $3 \times 10^5$  Zellen ausplattiert und die Platten anschließend für 24h bei 37°C inkubiert, um ein Wachstum der Zellen zu ermöglichen.

Am Folgetag wurden zur Transfektion pro Well einer 6-Well-Platte 1,2µg Plasmid-DNA mit Opti-MEM zum Zielvolumen von 100µl gemischt. Nach Zugabe von 4,5µl Attractene-Transfektionsreagenz wurden die Ansätze für 10min bei Raumtemperatur inkubiert und anschließend der gesamte Ansatz in die entsprechenden Wells mit den ausplattierten Zellen gegeben. Die Platten wurden dann 72h bei 37°C inkubiert. Anschließend wurden die Analysen mittels Western Blot wie unter 2.2.7 aufgeführt durchgeführt.

Um Positionseffekte bei der Transfektion und Inkubation der Ansätze zu vermeiden, wurden die Positionen der Konstrukte in den Wells der Kulturplatten jeweils zwischen den Experimenten variiert.

### 2.2.7 Western Blot

Für die Analyse der Expression der Luciferase wurden Western Blots durchgeführt. Zunächst wurden die Zellen aus den Kulturen in 6-Well Platten, die aus der unter 2.2.6 beschriebenen Transfektion stammen, gesammelt. Hierfür wurde das Kulturmedium entfernt und die Zellen in 1ml kaltem DPBS (pro Well) resuspendiert. Die gelösten Zellen wurden dann jeweils in 1,5ml Reaktionsgefäße gegeben und diese für 5min bei 300 G zentrifugiert. Der Überstand wurde anschließend entfernt, die entstandenen Pellets in PBS+1%Triton-X+cOmplete-Puffer resuspendiert und die Zellen mittels Ultraschalls lysiert.

Der Proteingehalt der Lösungen wurde mittels Bradford-Analyse unter Zuhilfenahme des Bio-Rad Protein Assay Kits II bestimmt. Hierfür wurden je 2µl des Zellhomogenats mit 798µl H<sub>2</sub>O verdünnt und 200µl des im Kit enthaltenen Bradfordreagenz zugegeben. Jeweils 100µl der Lösungen wurden in eine 96-Well Platte gegeben und bei diesen die OD<sub>595</sub> (Absorption bei Wellenlänge 595nm) bestimmt. Die Eichkurve zur Konzentrationsbestimmung wurde parallel mit BSA (Bovines Serumalbumin, ebenfalls im Kit vorhanden) in den Verdünnungen 1µg/ml, 2µg/ml, 5µg/ml, 10µg/ml und 15µg/ml erzeugt. Aus den bestimmten OD<sub>595</sub> wurden mithilfe der Eichkurve die Proteinkonzentrationen errechnet.

Die Ansätze für den Western Blot wurden wie folgt vorbereitet. Pro Ansatz wurden 15µg Protein aus den Zellhomogenaten mit Wasser auf ein Zielvolumen von 15µl gemischt. Zu den Lösungen wurde 2,3µl 1M DTT und 5,75µl 4x LDS gegeben.

Um die Proben aufzutrennen wurden 8%-Bis-Tris-Polyacrylamid-Trenngele mit 5%-Bis-Tris-Polyacrylamid-Sammelgelen hergestellt, das Vorgehen dazu ist in Tabelle 15 dargestellt. Die Proben wurden darin mit MES-Laufpuffer aufgetrennt. Als Größenstandard kam Precision Plus Protein™ Dual Color Standards von Bio-Rad Laboratories GmbH (München) zum Einsatz.

*Tabelle 15 – Rezept für die Bis-Tris-Polyacrylamid-Gele*

	8% Trenngel	5% Sammelgel
3,5x Bis-Tris Puffer	1,42ml	0,5ml
H <sub>2</sub> O	2,24ml	0,95ml
30% Acrylamid	1,34ml	0,3ml
10% APS	25µl	17,5µl
TEMED	7µl	3,35µl

Geblottet wurde mittels Tankblot auf eine Amersham Protran Premium 0.2 NC Membran von GE Healthcare Europe GmbH (Freiburg) mit Transferpuffer+15% Methanol für 90min bei 4°C und 80V.

Die Membran wurde anschließend mit 10ml 5% BSA-Lösung (in TBS-Puffer) für 1h bei Raumtemperatur geblockt.

Nach Waschen der Membran mit TBST-Puffer wurde der primäre Antikörper gegen die Luciferase in einer Verdünnung von 1:2000 in 5ml TBS-Puffer zur Membran gegeben und über Nacht bei 4°C inkubiert. Nach Entfernen des primären Antikörpers und Waschen mit TBST wurde ein sekundärer Antikörper, entweder HRP-konjugiert oder direkt fluoreszierend konjugiert, jeweils in einer Verdünnung von 1:10 000 in 5ml TBS zugegeben und für 1h bei Raumtemperatur inkubiert. Anschließend wurde mit TBST, bei HRP-konjugiertem sekundären Antikörper zusätzlich mit Wasser und TBS, gewaschen. Bei Antikörper mit direkter Fluoreszenz wurde die Membran nach dem Waschen direkt detektiert,

bei HRP-konjugiertem Antikörper wurde nach Zugabe des Peroxidase-Substrats (ECL-Puffer) für 2min inkubiert und anschließend die Membran detektiert

Als Ladekontrolle wurde Glycerinaldehyd-3-phosphat-Dehydrogenase (GAPDH) mit einem Primärantikörper in Verdünnung 1:1000 und sekundärem Antikörper in Verdünnung 1:10 000 (jeweils in 5ml TBS) markiert. Es wurde dazu folgendermaßen vorgegangen. Nach Detektion der Luciferase wurde die Membran mit TBST gewaschen und anschließend für 2h bei Raumtemperatur mit dem Primärantikörper für GAPDH inkubiert. Nach der Inkubation wurde die Membran wieder mit TBST gewaschen, mit dem entsprechenden Sekundärantikörper für 1h bei Raumtemperatur inkubiert und nach erneutem Waschen mit TBST detektiert.

### 2.2.8 Fast-Forward Transfektion für Luciferase-Assay

Zellen die später mittels Luciferase-Assay analysiert werden sollten, wurden mit Fast-Forward Transfektion wie folgt transfiziert und auf weißen 96-Well Zellkulturplatten (Nunc™ F96 MicroWell™ Polystyrolplatte, weiß, #136101, Fisher Scientific GmbH, Schwerte) ausplattiert: Die Zellen wurden zunächst analog zu dem Vorgehen, das unter 2.2.6 „Transfektion der Konstrukte für Western-Blot“ beschrieben ist, aus den Kulturflaschen in Medium resuspendiert. Nach Zählen der Zellen wurde eine Zellsuspension für die entsprechende Anzahl an Versuchsansätzen vorbereitet, deren Zellgehalt 40 000 Zellen/100µl betrug. Hiervon wurden je 100µl in entsprechende Wells der 96-Well Platte gegeben (je drei Wells für jedes Konstrukt).

Für die Transfektion wurde für jedes zu testende Konstrukt ein Mastermix angesetzt. In diesem wurde je zu beladendes Well 0,2µg DNA des jeweiligen Konstrukts mit 0,005µg pRL-CMV-DNA (für späteren Standard) gemischt und mit Opti-MEM auf ein Gesamtvolumen von 50µl (je Well) gebracht. Nach Zugabe von 0,75µl Attractene (je Well) wurde für 10min bei Raumtemperatur inkubiert. Anschließend wurden je 50µl des Transfektionsansatzes tröpfchenweise in die entsprechenden Wells gegeben und die Platten anschließend für 48h bei 37°C inkubiert. Anschließend wurde die Luciferase-Aktivität der verschiedenen Konstrukte mittels Luciferase-Assay wie unter 2.2.9 beschrieben bestimmt.

Um Positionseffekte bei der Transfektion und Inkubation der Ansätze zu vermeiden, wurde, wie schon bei der Transfektion für die Western Blots, die Positionen der Konstrukte in den Wells der Kulturplatten jeweils zwischen den Experimenten variiert.

### 2.2.9 Luciferase-Assay

Die Luciferase-Aktivität der transfizierten Zellen wurde mit Hilfe des Dual-Glo® Luciferase Assay Systems von Promega GmbH (Mannheim) bestimmt. Die transfizierten Zellen stammten hierfür aus den unter 2.2.8 beschriebenen Kulturen auf 96-Well Zellkulturplatten, die mittels Fast-Forward Transfektion transfiziert wurden. Zunächst wurde nach Anweisung des Herstellers der gesamte Dual-Glo® Luciferase Puffer und das gesamte Dual-Glo® Luciferase Substrat des Kits gemischt, um Dual-Glo® Luciferase Reagenz zu erhalten. Außerdem wurde der Dual-Glo® Stop & Glo® Puffer und das Dual-Glo® Stop & Glo® Substrat nach Herstellerangaben (100:1) gemischt, um Dual-Glo® Stop & Glo® Reagenz zu erhalten. Das Dual-Glo® Luciferase Reagenz wurde aliquotiert und bis zur Verwendung bei -80°C gelagert, das Dual-Glo® Stop & Glo® Reagenz wurde für die Versuche jeweils frisch angesetzt.

Vor Beginn wurden die vorbereiteten Reagenzien, sowie die Kulturplatte mit den behandelten Zellen auf Raumtemperatur gebracht und 100µl Medium aus den Wells entfernt, um ein Zielvolumen von 50µl Medium für das Assay zu erhalten. Dann wurde den Wells mit den zu analysierenden Zellen jeweils 50µl Dual-Glo® Luciferase Reagenz zugegeben (entspricht dem Volumen des Mediums im Well) und durch Pipettieren gemischt. Nach 10min Inkubation bei Raumtemperatur auf dem Plattenschüttler wurde die Lichtemission durch die Firefly-Luciferase im Luminometer gemessen. Anschließend wurde den Wells jeweils 50µl Dual-Glo® Stop & Glo® Reagenz zugegeben, gemischt und nach 10min Inkubation bei Raumtemperatur auf dem Plattenschüttler die Lichtemission der Renilla-Luciferase im Luminometer gemessen. Die Emission wurde in relativen Lichteinheiten (RLU) gemessen. Um einen Vergleich zu ermöglichen wurden pro Well die Werte der Firefly-Luciferase auf die Werte der Renilla-Luciferase normalisiert, um potenzielle Unterschiede bei der Transfektionseffektivität auszugleichen. Die Berechnung erfolgte folgendermaßen:

$$\text{Firefly\_Luciferase\_Aktivität}_{\text{normalisiert}} = \frac{RLU_{\text{Firefly\_Luciferase}}}{RLU_{\text{Renilla\_Luciferase}}}$$

Nachdem die normalisierte Firefly-Luciferase-Aktivität für jedes Well einzeln berechnet wurde, wurde aus den Ergebnissen der jeweils drei Wells, die mit dem selben Konstrukt transfiziert worden waren, der Mittelwert errechnet. Die Mittelwerte wurden dann zum Vergleich zwischen den einzelnen Konstrukten genutzt.

### 2.2.10 Statistische Auswertung

Die erhobenen Daten der Quantifizierungen der Western-Blots und der Luciferase-Assays wurden mittels des Shapiro-Wilk-Test auf Normalverteilung überprüft. Zur statistischen Auswertung der Daten kam dann bei gegebener Normalverteilung und nach Annahme unabhängiger Stichproben mit ungleichen Varianzen der Welch-T-Test zum Einsatz. Mit diesem Test wurden die erhobenen Daten aus Western-Blots beziehungsweise Luciferase-Assays paarweise (je von zwei Konstrukten) verglichen. Als Signifikanzniveau wurde  $\alpha=0,05$  gewählt. Ergebnisse mit einem p-Wert von  $p \leq 0,05$  wurden somit als signifikant gewertet.

### 2.2.11 Identifizierung von Transkriptionsfaktoren

Zur Identifizierung von Transkriptionsfaktoren wurden Online-Datenbanken und Online-Tools genutzt. Es wurde der UCSC Genome Browser eingesetzt (verfügbar unter [www.genome.ucsc.edu](http://www.genome.ucsc.edu)), mit dem Regionen des menschlichen Genoms mit ihrer Sequenz und zusätzlichen Datensätzen, beispielsweise zu kodierten Genen, Polymorphismen oder auch zu Bindungsstellen von Transkriptionsfaktoren, dargestellt werden können (Kent *et al.* 2002). Als Datensätze für die Transkriptionsfaktoren im UCSC Genome Browser kamen Datensätze der Enzyklopädie der DNA Elemente „ENCODE“ zum Einsatz (Encode Project Consortium 2011; Rosenbloom *et al.* 2013). Weiterhin wurde das Online-Tool „PROMO“ der Algorithmics and Genetics Group der Universität Politècnica de Catalunya genutzt, mit dem Sequenzen spezifisch auf Bindungsstellen von Transkriptionsfaktoren untersucht werden können (Messeguer *et al.* 2002; Farre *et al.* 2003). Dieses Tool ist über die Webadresse [algggen.lsi.upc.es](http://algggen.lsi.upc.es) abrufbar. Mithilfe des Human Protein Atlas (erreichbar unter

www.proteinatlas.org), mit dem unter anderem die Expression einzelner Transkriptionsfaktoren in verschiedenen Geweben nachvollzogen werden kann (Uhlen *et al.* 2015), wurde überprüft, in welchen Geweben die gefundenen Transkriptionsfaktoren exprimiert werden. Um Interaktionen zwischen Transkriptionsfaktoren zu finden, wurde die STRING-Datenbank (abrufbar über [www.string-db.org](http://www.string-db.org)) genutzt. Diese bietet die Möglichkeit spezifische Proteine auf bekannte und vorausgesagte Protein-Protein Interaktionen zu untersuchen (Szklarczyk *et al.* 2019).

### 3. Ergebnisse

In dieser Arbeit sollte ein Modell erzeugt werden, mit dem die Promotorregion des *KPNA3*-Gens untersucht und dessen Promotoraktivität quantifiziert werden kann. Hierfür sollten verschiedene Varianten der Promotorregion mit sich unterscheidenden Polymorphismen als Konstrukte generiert werden, um einerseits relevante Promotorbereiche zu identifizieren, andererseits die beobachtete Assoziation des Polymorphismus auf den Erkrankungsbeginn (vgl. 1.7) der SCA3 weiter zu untersuchen. Am Ende soll das Modell eine Quantifizierung der Promotoraktivität erlauben, die im weiteren Verlauf auch Testung von Wirkstoffen erlaubt um die Aktivität zu beeinflussen.

Hierfür wurde zunächst ein Bereich von 3500bp in einen Vektor inkloniert. Dieser enthielt den Bereich von -3085bp bis +415bp ausgehend vom Transkriptionsstart des *KPNA3*-Gens. Hierbei entsprachen die ersten 415bp des Gens, die mitenthalten waren, der untranslatierten Region des Exon 1 des *KPNA3*-Gens. Das Exon 1 erstreckt sich über die ersten 484bp am 5' Ende des *KPNA3*-Gens und war somit nicht vollständig im Konstrukt vorzufinden, sondern nur dessen untranslatierter Bereich. Das Gen selbst besteht aus 17 Exons und 16 Introns. Am 5' und 3' Ende befinden sich jeweils nicht translatierte Bereiche (UTRs), wobei der 5' UTR wie bereits erwähnt einen Großteil des Exon 1 ausmacht (vgl. Daten des UCSC Genome Browsers (Kent *et al.* 2002)). Der Aufbau des Gens im Bereich des 5' UTR ist in Abbildung 2 schematisch dargestellt. Der 5' untranslatierte Bereich wurde miteinkloniert, da es durchaus denkbar war, dass dieser, wenn er keine Funktion im später translatierten Genprodukt besitzt, eventuell Einfluss auf die Expression und Promotoraktivität besitzt.

Ausgehend vom Ausgangskonstrukt wurden Konstrukte, die einen kleineren Bereich abbildeten, erzeugt. Diese sollten Aufschluss geben, ob überhaupt ein so großer Promotorbereich relevant ist. Diese enthielten 2394bp und umfassten den Bereich von -1981bp bis +415bp ausgehend vom Beginn des *KPNA3*-Gens, wiederum inklusive des untranslatierten Bereichs des Exon 1. In diesen Konstrukten sollten wiederum beide Versionen des Polymorphismus rs11618521, der in Assoziation zum Erkrankungsalter der SCA3 auffiel (Dieterle

2019), abgebildet werden. Dieser befindet sich an Position -1094bp ausgehend vom Transkriptionsstart des *KPNA3*-Gens. Zuletzt wurde aus diesen Konstrukten noch Konstrukte erzeugt, die den 5' UTR des *KPNA3*-Gens nicht enthielten, um dessen Rolle bei der Expression zu klären. Diese beinhalteten folglich 1981bp des Promotorbereichs und umfassten den Bereich von -1981bp bis zum Transkriptionsstart des *KPNA3*-Gens, ohne Anteile des 5' UTR.

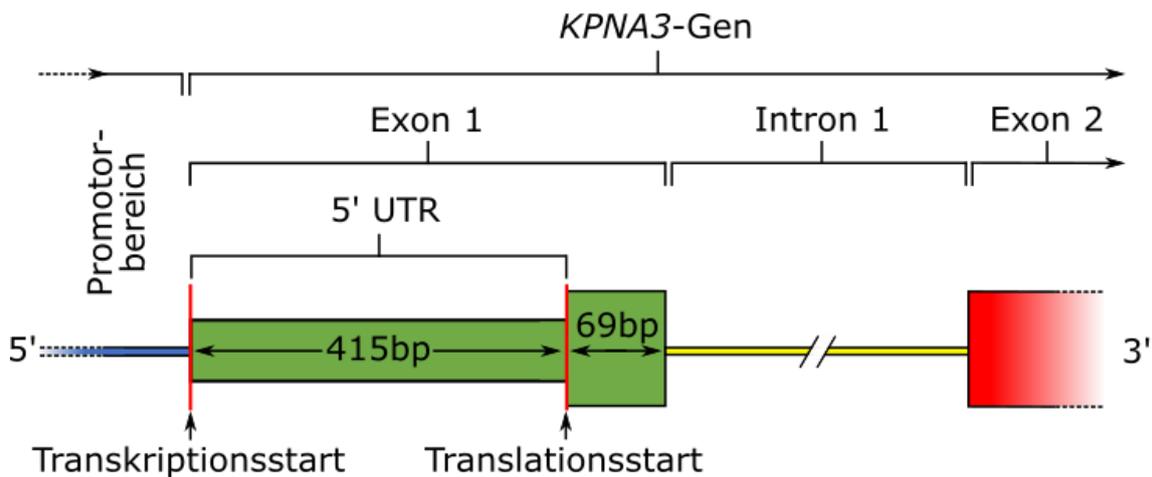


Abbildung 2 – Schema des *KPNA3*-Gens am 5' Ende – Dargestellt ist schematisch das 5' Ende des *KPNA3*-Gens. Das erste Exon besitzt 5' einen Bereich, der nicht translatiert wird (5' UTR). Dieser ist 415bp lang. Der restliche Abschnitt von 69bp Länge des Exon 1, der 3' davon liegt, wird translatiert. Vor dem Exon 1 (5') liegt der Promotorbereich, 3' schließen sich dann Introns und Exons im Wechsel und in verschiedener Länge an das erste Exon an (exemplarisch bis Exon 2 dargestellt). Insgesamt besteht das *KPNA3*-Gen aus 17 Exons und 16 Introns. In den erzeugten Konstrukten sind Varianten des Promotorbereichs, sowie teilweise der 5' UTR mit seinen 415 Basenpaaren (also nicht das komplette Exon 1) enthalten, um deren Relevanz für die Expression des nachgeschalteten Gens zu untersuchen. Das Schema beruht auf Daten des UCSC Genome Browsers (Kent et al. 2002).

Im nächsten Schritt wurde die Expression der verschiedenen Konstrukte untersucht und quantifiziert, um den Einfluss von Größe des gewählten Promotorbereichs, der Variante des Polymorphismus und des untranslatierten Bereichs am 5'-Ende des *KPNA3*-Gens zu untersuchen.

Zuletzt wurden mit Hilfe von Literaturdaten beziehungsweise Datenbanken Transkriptionsfaktoren identifiziert, die in den für die Konstrukte gewählten Promotorbereichen binden, um eine Erklärung für die Unterschiede, die die Konstrukte in der Expression zeigten, zu bieten.

Das einzelne Vorgehen ist in den folgenden Abschnitten dargestellt.

#### 3.1 Generierung eines Konstrukts zur Untersuchung des KPNA3-Promotors

Zunächst wurden 3500bp der KPNA3-Promotorregion (Position -3085 bis +415, inkl. des 5' UTR des Exon 1) in den pGL3-basic Vektor von Promega GmbH (Mannheim) inkloniert. Dieser enthielt ein Gen, das für eine Firefly-Luciferase kodierte und zur späteren Expressionsanalyse diente. Die Promotorregion wurde hierbei vor dieses Gen geschaltet, wodurch die Expression des *Luciferase*-Gens durch den Promotor reguliert werden sollte.

Hierfür wurde zuerst der KPNA3-Promotorbereich aus genomischer DNA einer gesunden Kontrollperson mittels Nested-PCR vervielfältigt, um anschließend die Promotorregion in zwei Abschnitten in den pGL3-basic Vektor einzufügen. Als Primerpaar für den ersten PCR-Schritt wurden S088\_KPNA3\_-3746-F und S093\_KPNA3\_+058-R genutzt. Die PCR-Reaktion wurde hierbei wie unter 2.2.4 beschrieben mit dem Kit „Expand Long Template PCR System“ der Firma Roche Diagnostics GmbH (Mannheim) nach dem vom Hersteller empfohlenen Vorgehen durchgeführt.

Das entstandene Produkt enthielt den gewünschten Promotorbereich mit 3500bp und an 5'- und 3'-Ende jeweils einige Basenpaare als Überschuss und somit eine Gesamtgröße von ca. 4000bp. Im KPNA3-Promotor ist eine EcoRI-Schnittstelle an Position -1720bp enthalten. Diese wurde im Weiteren genutzt, um den Promotor in zwei Schritten und mit dadurch kleineren Inserts (im Folgenden Abschnitt A und Abschnitt B genannt) in den Vektor einzufügen. In Abbildung 3 ist das gesamte Vorgehen zur Erstellung der beiden Inserts schematisch dargestellt.

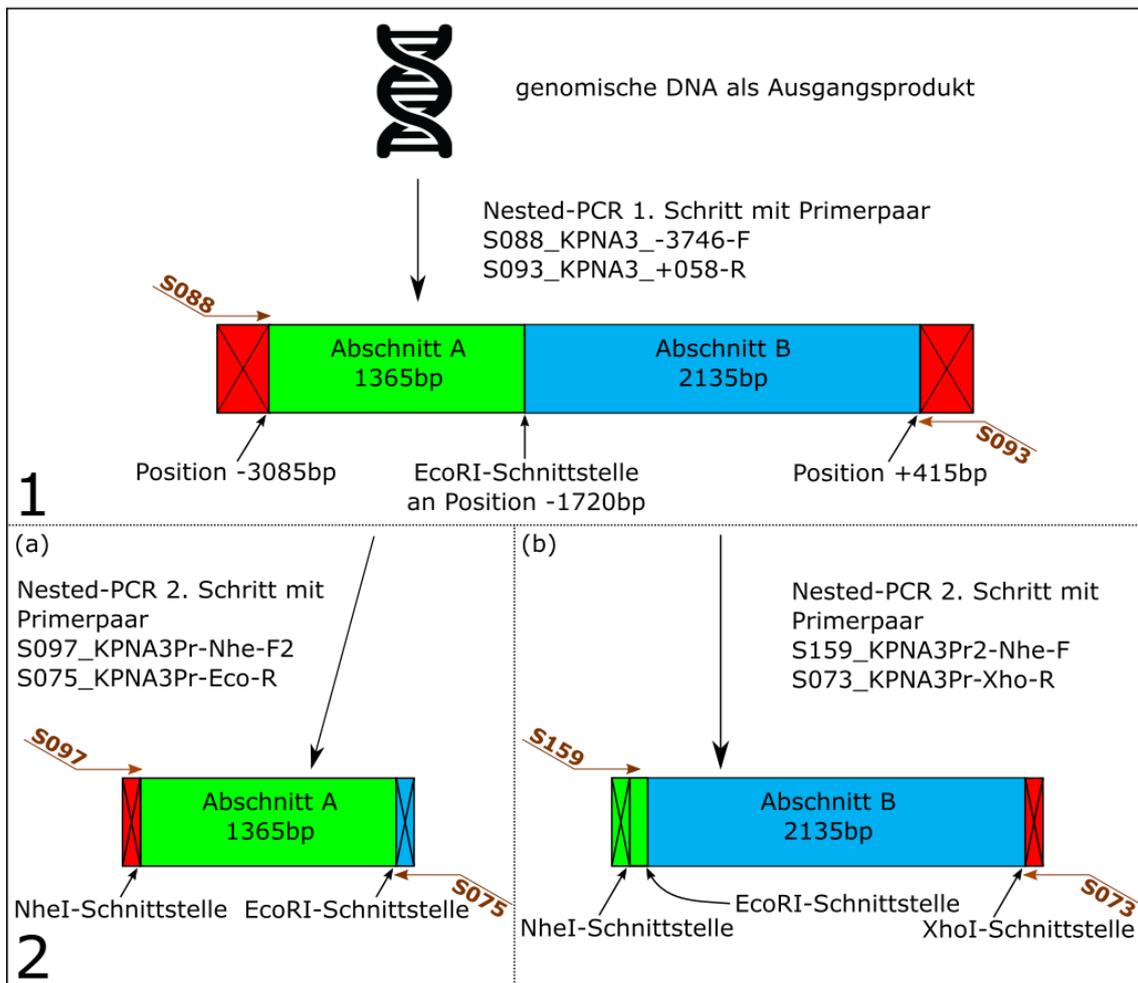


Abbildung 3 – Nested-PCR mit einer doppelten Amplifikation – 1) Im 1. Schritt wurde aus genomischer DNA zunächst ein größerer Bereich, der den gewünschten 3500bp Promotorbereich des KPNA3-Promotors (Position -3085bp bis +415bp) enthielt, mittels PCR amplifiziert. An 5'- und 3'-Ende war jeweils ein Basenüberhang (rot markiert) vorhanden, sodass das Produkt eine Gesamtgröße von ca. 4000bp aufwies. Das Produkt enthält eine Schnittstelle für das Restriktionsenzym EcoRI an Position -1720bp, welches im Folgenden als Grenze der beiden zur Klonierung genutzten Abschnitte dient. Als Primerpaar kamen S088\_KPNA3\_-3746-F und S093\_KPNA3\_+058-R zum Einsatz. – 2) Mit dem Produkt aus dem 1. Schritt als Ausgangs-DNA wurden dann im 2. Schritt die Abschnitte A und B getrennt voneinander amplifiziert, wobei die EcoRI-Schnittstelle als Abgrenzung zwischen den Abschnitten diente. (a) Um Abschnitt A mittels PCR zu amplifizieren wurden die Primer S097\_KPNA3Pr-Nhe-F2 und S075\_KPNA3Pr-Eco-R genutzt. Hierbei wurde durch Primer S097\_KPNA3Pr-Nhe-F2 direkt vor (5'-Ende) Abschnitt A eine Schnittstelle für das Enzym NheI für die spätere Klonierung eingefügt. Im PCR-Produkt ist außerdem die EcoRI-Schnittstelle enthalten. (b) Für die Amplifikation des Abschnitt B kamen die Primer S159\_KPNA3Pr2-Nhe-F und S073\_KPNA3Pr-Xho-R zum Einsatz. S159\_KPNA3Pr2-Nhe-F fügte hierbei vor (5') der EcoRI-Schnittstelle (also noch in einem zum Abschnitt A zugehörigen Bereich) noch eine Schnittstelle für NheI, da diese für die Einklonierung in die Multiple Cloning Site des Vektors nötig war. S073\_KPNA3Pr-Xho-R fügte am Ende (3') des Produkts eine XhoI-Schnittstelle ein, ebenfalls um die Einklonierung in den Vektor zu ermöglichen. Die Bestandteile der jeweiligen PCR-Produkte unter A und B, die jeweils bei der Klonierung nicht mit einkloniert wurden, sind durch Durchstreichen markiert.

#### 3.1.1 Klonierung des Abschnitts B

Aus dem Produkt der Nested-PCR wurde zunächst der Abschnitt B, also der Bereich von der EcoRI-Schnittstelle an Position -1720bp bis zum Ende des 5' UTR des *KPNA3*-Gens an Position +415bp, mittels PCR amplifiziert und inkloniert. Hierzu wurden Primer genutzt die vor (5'-Ende) der Schnittstelle von EcoRI banden und dort im PCR-Produkt eine NheI-Schnittstelle einfügten, beziehungsweise hinter (3'-Ende) dem Ende der Promotorregion banden und dort eine XhoI-Schnittstelle einfügten. Diese beiden Enzyme hatten Schnittstellen in der Multiple-Cloning-Site des pGL3-basic Vektors und konnten somit als Klonierungsenzyme genutzt werden. Das schematische Vorgehen ist in Abbildung 5 dargestellt. Die Primer, die für die PCR genutzt wurden, waren S159\_KPNA3Pr2-Nhe-F, der die NheI-Schnittstelle am 5'Ende einfügte und S073\_KPNA3Pr-Xho-R, der die XhoI-Schnittstelle am 3'Ende des 5' UTR einfügte. Die Klonierung erfolgte nach dem unter 2.2.5 A beschriebenen Vorgehen zur Klonierungs-PCR. Das PCR-Produkt wurde anschließend mit den Restriktionsenzymen NheI und XhoI wie unter 2.2.5 B-D beschrieben in den Vektor eingefügt.

Im Endprodukt ist der Abschnitt B im Vektor enthalten, sowie ein kleiner Bereich des Abschnitts A, bis zur EcoRI-Schnittstelle. Das Klonierungsprodukt wurde zur Amplifikation in DH5 $\alpha$  chemisch kompetente Zellen transformiert wie unter 2.2.5 E beschrieben.

Das Produkt der Klonierung enthielt den Abschnitt B (von Position -1720bp bis +415bp) und einige Basenpaare des Abschnitts A 5' der EcoRI-Schnittstelle. Die erfolgreiche Klonierung wurde mittels Testverdaus mit den Klonierungsenzymen NheI und XhoI überprüft. Hierbei sollte die entstandene Bande die Größe des inklonierten Insert besitzen. Außerdem wurde parallel mit NheI und EcoRI verdaut, um das Vorhandensein der EcoRI-Schnittstelle zu verifizieren. Vorgegangen wurde jeweils wie unter 2.2.5 F aufgeführt. In beiden Ansätzen konnte das erwartete Produkt mit der entsprechenden Größe dargestellt werden. Wie in Abbildung 4 zu sehen, kamen die Banden wie erwartet zwischen 3000bp und 2000bp (NheI und XhoI als Enzyme) beziehungsweise zwischen 300bp und 200bp (NheI und EcoRI als Enzyme) zu liegen. Die Einklonierung des Inserts

schien somit zunächst erfolgreich. In der Sequenzierung des Produkts zeigte sich jedoch, dass am 3' Ende des Abschnitts B ein Bereich von ca. 200bp der Promotorregion fehlten. Der fehlende Bereich erstreckte sich somit von Position +190bp bis +397bp, entsprach also einem Bereich am 3' Ende des 5' UTR, wobei ein Bereich von Position +398bp bis +415bp noch enthalten war. Die Sequenzierung erfolgte hierbei nach dem unter 2.2.2 beschriebenen Vorgehen. Das entstandene Konstrukt enthielt somit den Promotorbereich von -1981bp bis +415bp ausgehend vom Transkriptionsstart des *KPNA3*-Gens, wobei der Bereich von +190bp bis +397bp nicht enthalten war.

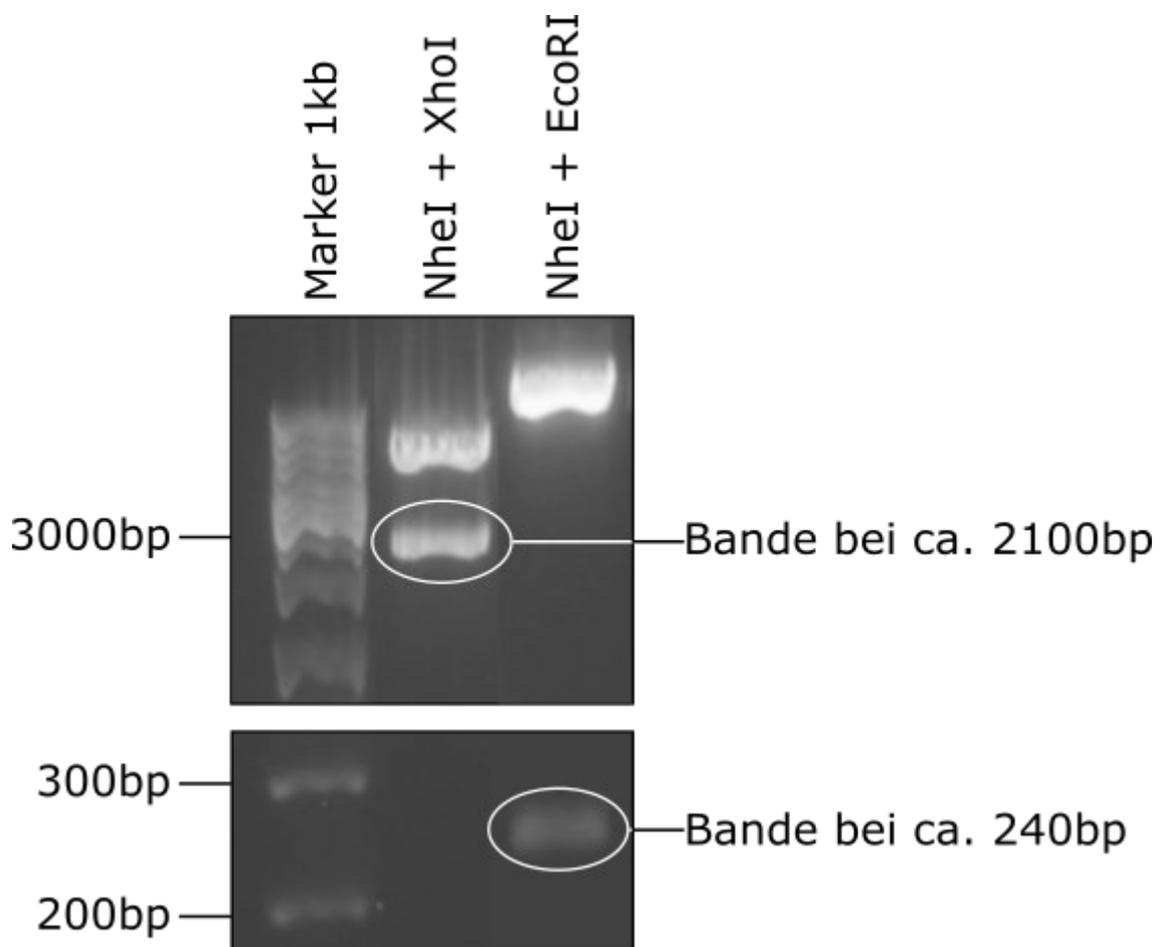


Abbildung 4 – Testverdau des Konstrukts nach Einklonierung des Abschnitts B – Das Klonierungsprodukt, bei dem der Abschnitt B einkloniert wurde, wurde mit den Restriktionsenzymen *NheI* und *XhoI* beziehungsweise *NheI* und *EcoRI* verdaut und auf einem Agarosegel aufgetrennt. Die ausgeschnittenen Bereiche kamen dabei wie erwartet zwischen 3000bp und 2000bp (erwartete Länge ca. 2100bp) beziehungsweise zwischen 300bp und 200bp (erwartete Länge von ca. 240bp) zu liegen. Die erfolgreiche Klonierung des Abschnitts B in den Vektor zeigte sich somit zunächst im Testverdau bestätigt. In der nachfolgend durchgeführten Sequenzierung zeigte sich jedoch, dass ein ca. 200bp langer Bereich von Position +190bp bis +397bp fehlte, also ein Bereich der sich am 3' Ende des 5' UTR befindet (hier nicht dargestellt).

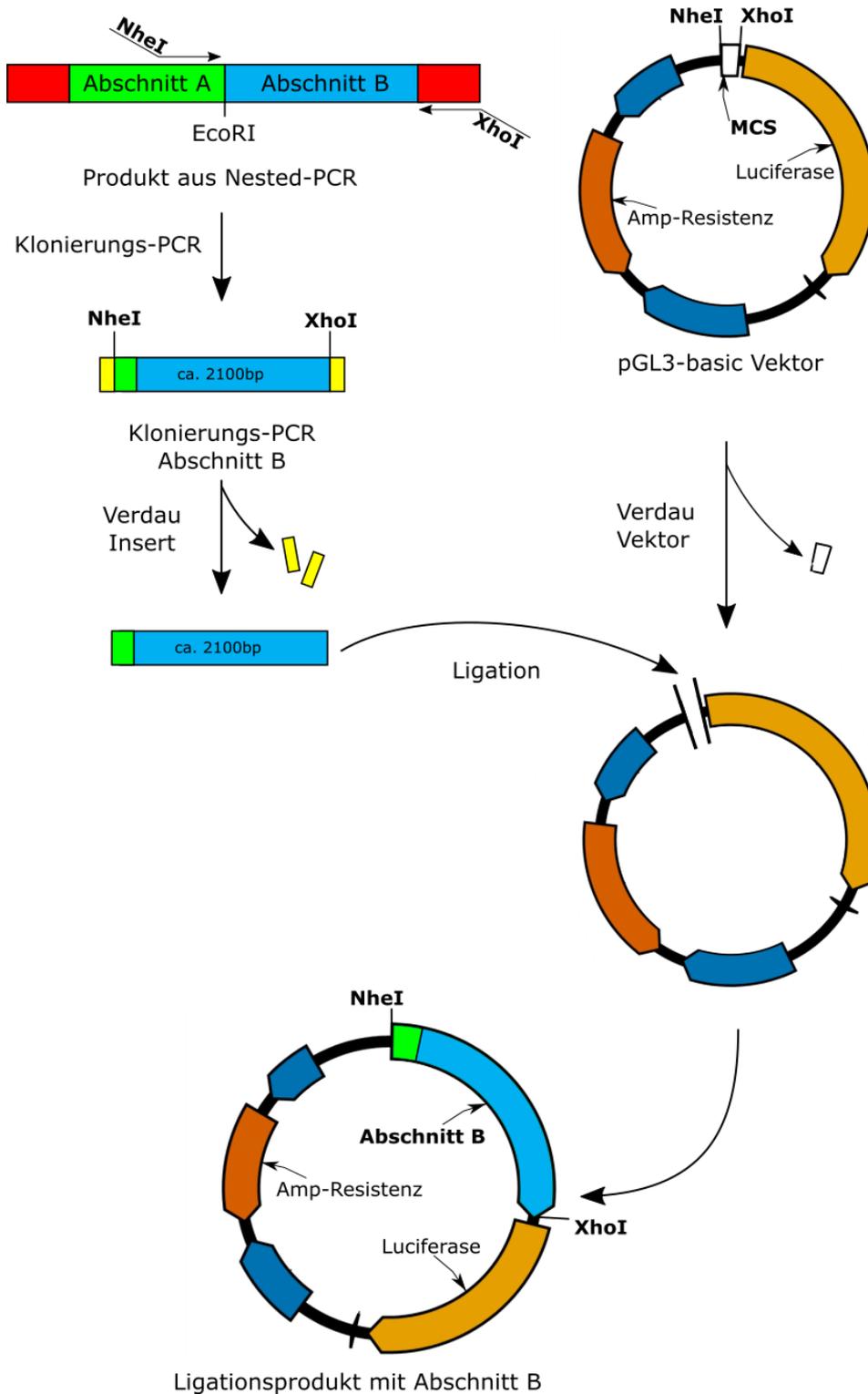


Abbildung 5 – Vorgehen zur Einklonierung des Abschnitt B – Das PCR-Produkt aus der Nested-PCR wurde mit den Primern S159\_KPNA3Pr2-Nhe-F und S073\_KPNA3Pr-Xho-R, die den Abschnitt B umfassen und zusätzlich Schnittstellen für *XhoI* und *NheI* einfügen, amplifiziert (Bindungsstellen mit Pfeilen dargestellt) und anschließend mit den Enzymen *XhoI* und *NheI* verdaut. Der Vektor enthält in der Multiple-Cloning-Site (MCS) Schnittstellen für *XhoI* und *EcoRI* und wurde parallel zum Insert auch mit eben diesen Enzymen verdaut. Das entstandene Insert wurde anschließend in den verdauten Vektor einkloniert. Im Endprodukt war nun der Abschnitt B (hellblau), sowie ein kleiner Bereich des Abschnitts A (grün) bis zur *EcoRI*-Schnittstelle enthalten.

#### 3.1.2 Klonierung des Abschnitts A

Vor Korrektur des Bereichs am 3' Ende des unter 3.1.1. erzeugten Konstrukts, dem ein Abschnitt des Promotorbereich von Position +190bp bis +397bp fehlte, wurde zunächst der Abschnitt A in dieses Konstrukt inkloniert. Hierfür wurde der Abschnitt A des Produkts der Nested-PCR mit Primern amplifiziert, die vor (5') Beginn des Abschnitts binden und dort eine NheI-Schnittstelle einfügen, beziehungsweise hinter (3') der EcoRI-Schnittstelle binden. Das Vorgehen ist in Abbildung 7 schematisch dargestellt. Als Primer kamen hierzu S097\_KPNA3Pr-Nhe-F2, der vor der Position -3085bp, also vor Beginn des Abschnitt A, eine NheI-Schnittstelle einfügte, und S075\_KPNA3Pr-Eco-R, der nach (3') der EcoRI-Schnittstelle band, zum Einsatz. Die PCR-Amplifikation wurde nach dem Protokoll zur Klonierungs-PCR (siehe 2.2.5 A) durchgeführt. Das PCR-Produkt enthielt die eingefügte NheI-Schnittstelle und die EcoRI-Schnittstelle des Promotorbereichs, beides Schnittstellen, die bereits im in 3.1.1 erstellten Konstrukt enthalten waren. Somit konnten die Restriktionsenzyme EcoRI und NheI zur Einklonierung des in der PCR entstandenen Inserts genutzt werden. Der enthaltene Teilbereich des Abschnitts A im Konstrukt aus 3.1.1 wurde hierbei ausgeschnitten und durch den in der Klonierungs-PCR gewonnenen kompletten Abschnitt A als Insert ersetzt. Das Klonierungsprodukt wurde in DH5 $\alpha$  chemisch kompetente Zellen transformiert zur Amplifikation. Die Teilschritte zur Klonierung erfolgten hierbei nach dem unter 2.2.5 B-D dargestellten Vorgehen.

Durch Testverdau nach dem unter 2.2.5 F aufgeführten Vorgehen mit den Klonierungsenzymen NheI und EcoRI wurde die erfolgreiche Klonierung überprüft. Es zeigte sich, dass der bisher durch diese Enzyme ausgeschnittene Bereich von 240bp im Ausgangskonstrukt (vergleiche Abbildung 4) durch ein Fragment mit einer Größe von ca. 1400bp ersetzt wurde (siehe Abbildung 6). Die Einklonierung des Abschnitts A war somit erfolgreich. Die Bereiche um die Übergangspunkte der beiden Abschnitte zueinander und zum Vektor wurden mittels Sequenzierung untersucht und zeigten bis auf das aus 3.1.1 bereits bekannte fehlende Fragment von ca. 200bp am 3'-Ende des Abschnitts B die erwartete Sequenz. Die Sequenzierung erfolgte hierbei nach dem unter 2.2.2 beschriebenen Vorgehen. Das Konstrukt wurde pGL3-basic construct benannt.

Es enthielt den Promotorbereich von Position -3085bp bis +415bp. Es fehlte aber darin ein ca. 200bp langer Abschnitt im Bereich des 3' Endes, der Position +190bp bis +397bp entspricht und somit einen Bereich im 5' UTR des *KPNA3*-Gens darstellt.

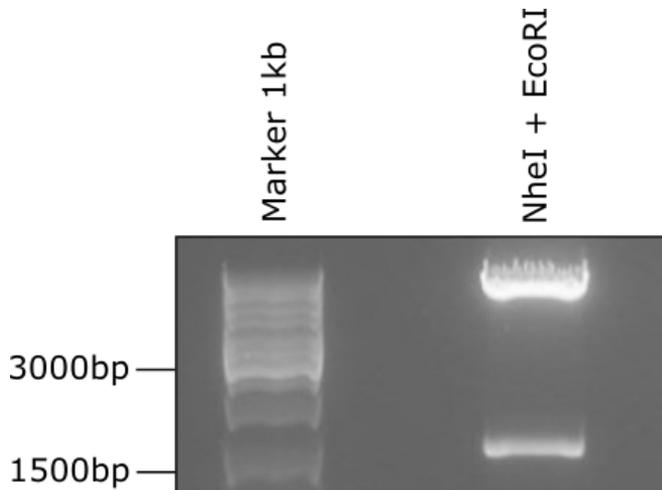


Abbildung 6 – Testverdau nach Klonierung des Abschnitts A in das Konstrukt aus 3.1.1 – Die erfolgreiche Klonierung des Abschnitts A in das in 3.1.1 entstandene Konstrukt wurde mittels Testverdau mit den Klonierungsenzymen *NheI* und *EcoRI* überprüft. Es entstand eine Bande, die etwa bei 1500bp zu liegen kam, was mit der erwarteten Größe von ca. 1400bp des Abschnitts A, der bei diesem Verdau ausgeschnitten wurde, vereinbar ist. Die erfolgreiche Einklonierung des Insert konnte hiermit bestätigt werden, da die bisher durch Verdau mit *NheI* und *EcoRI* entstandene Bande von 240bp (siehe Abbildung 4) durch die größere Bande von ca. 1400bp ersetzt wurde.

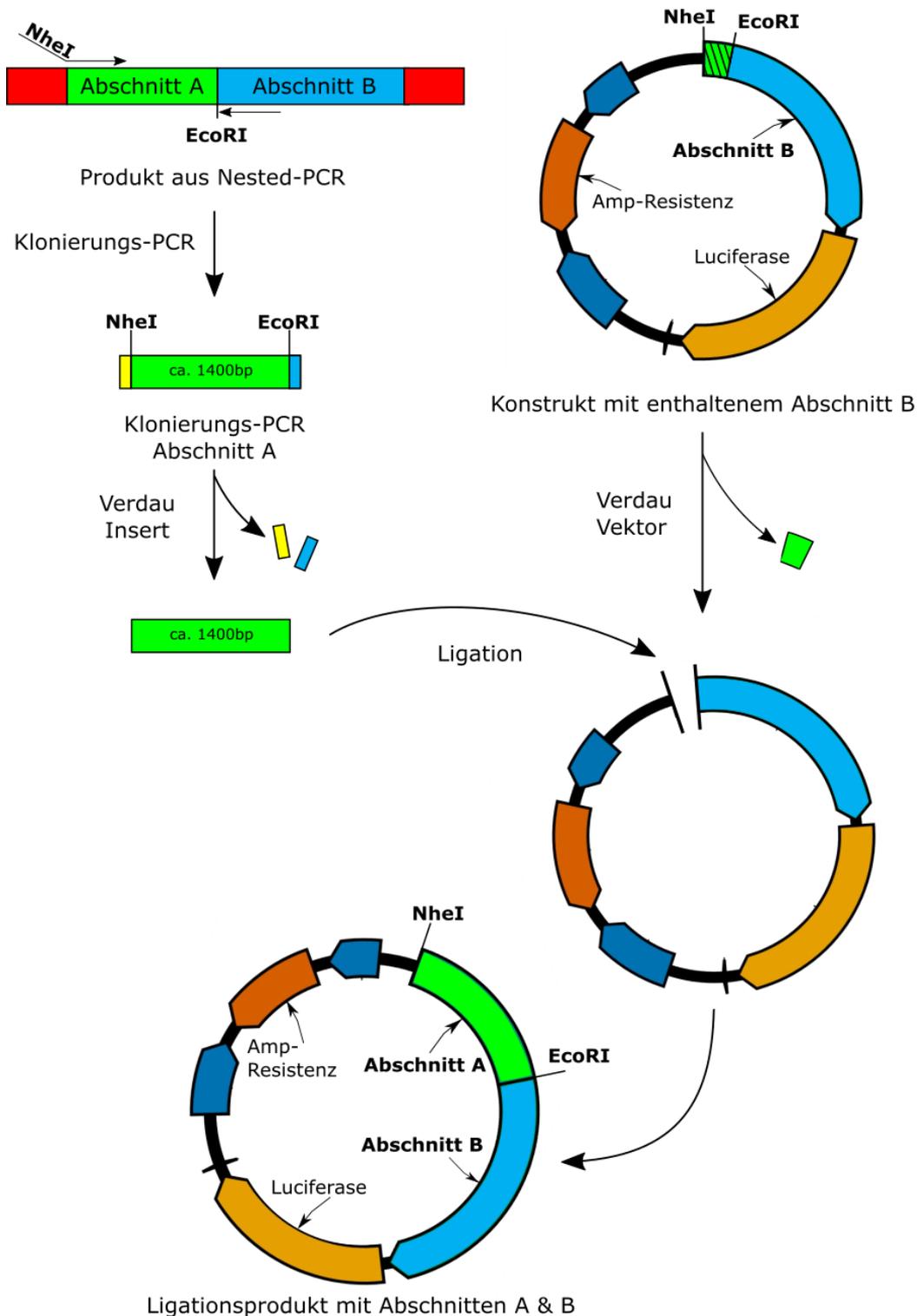


Abbildung 7 – Vorgehen zur Einklonierung des Abschnitt A – Das PCR-Produkt aus der Nested-PCR wurde mit Primern, die den Abschnitt A umfassen und zusätzlich vor (5') den Abschnitt A eine Schnittstelle für *NheI* einfügen, amplifiziert (Bindungsstellen mit Pfeilen dargestellt) und anschließend mit den Enzymen *EcoRI* und *NheI* verdaut. Parallel wurde der Vektor, der bereits Abschnitt B enthält, mit denselben Enzymen verdaut und anschließend das Insert inkloniert, das hierbei das zuvor enthaltene Fragment des Abschnitts A ersetzte. Im Endprodukt sind somit nun Abschnitt A und B enthalten. Es wird mit pGL3-basic construct bezeichnet.

### 3.1.3 Fertigstellung des Konstrukts

Um den fehlenden Bereich am 3'-Ende (Position +190bp bis +397bp) des Abschnitts B zu erhalten, wurde der Bereich des Abschnitts direkt aus genomischer DNA durch PCR amplifiziert, ohne vorhergehende Nested-PCR. Mittels Primer wurde (analog zu 3.1.1) am Ende des Abschnitts B eine XhoI-Schnittstelle eingefügt. Die verwendeten Primer waren S160\_KPNA3Pr3-Nhe-F, der 5' der im Promotor enthaltenen EcoRI-Schnittstelle band, und S073\_KPNA3Pr-Xho-R, der eine XhoI-Schnittstelle hinter (3') Position +415bp einfügte. Die PCR wurde hierbei nach dem unter 2.2.5 A aufgeführten Schema angesetzt. Durch das gewählte Primerpaar konnten zur Klonierung im weiteren Verlauf die Restriktionsenzyme EcoRI und XhoI genutzt werden.

Um zu verifizieren, dass das PCR-Produkt den bisher fehlenden Bereich des Abschnitts B enthielt, wurde ein Testverdau des PCR-Produkts durchgeführt. Im bisher fehlenden Bereich war eine Schnittstelle des Restriktionsenzym SacII an Position +192bp enthalten, weshalb sich dieses Enzym zum Test auf erfolgreiche Einklonierung des nun kompletten Abschnitts B eignete, da das Enzym nur bei Vorhandensein des bisher fehlenden Bereichs schneiden konnte. Im Testverdau des PCR-Produkts zeigte sich, dass das PCR-Produkt geschnitten wurde und Fragmente der erwarteten Größe entstanden. Im PCR-Produkt war somit der bisher fehlende Bereich des Abschnitts B enthalten.

Die Einklonierung des nun kompletten Abschnitts B aus der PCR erfolgte nach dem in 2.2.5 B-E dargestellten Vorgehen mit den Enzymen EcoRI und XhoI. Das Klonierungsprodukt wurde zur Amplifikation wie unter 2.25 E beschrieben in DH5 $\alpha$  chemisch kompetente Zellen transformiert. Ein Testverdau wurde nach 2.2.5 F zur Überprüfung der erfolgreichen Klonierung durchgeführt. Als Restriktionsenzyme kamen EcoRI und XhoI zum Einsatz. Es wurden hierbei sechs der Klone, die nach Transformation Wachstum auf den Agarplatten zeigten, mit dem in 3.1.2 generierten pGL3-basic construct verglichen. Wie in Abbildung 8 zu sehen, konnte bei allen getesteten Klonen das im Vergleich zum Konstrukt pGL3-basic construct längere Insert nachgewiesen werden.

Von den im Testverdau getesteten Klonen wurde das Konstrukt eines der Klone im Verlauf in beide Leserichtungen durchsequenziert, um die Sequenz insgesamt zu verifizieren. Die hierbei bestätigte Sequenz kann im Anhang eingesehen werden. Mit diesem Klon und dem darin enthaltenen Konstrukt wurde im Verlauf weitergearbeitet, das Konstrukt wurde mit pGL3-KPNA3-clone I bezeichnet.

Im Konstrukt pGL3-KPNA3-clone I waren der Promotorbereich des *KPNA3*-Gens von Position -3085bp bis +415bp inklusive des nicht-translatierten Bereichs des Exon 1 des *KPNA3*-Gens (im weiteren Verlauf auch als 5' untranslatierte Bereich oder 5' UTR bezeichnet), sowie der Polymorphismus rs11618521 an Position -1094bp in der Variante mit Guanin enthalten. Im hier generierten Konstrukt sollten die nötigen Bindestellen für relevante Transkriptionsfaktoren auf genetischer Ebene enthalten sein, da ein großer Bereich der Promotorregion, 3500bp inklusive des 5' UTR im Exon 1 des *KPNA3*-Gens, abgebildet ist. Epigenetische Modifikationen, wie beispielsweise Methylierungen, und 3D-Struktur der DNA (das Konstrukt ist ein Plasmid) waren in diesem Modell jedoch nicht abgebildet.

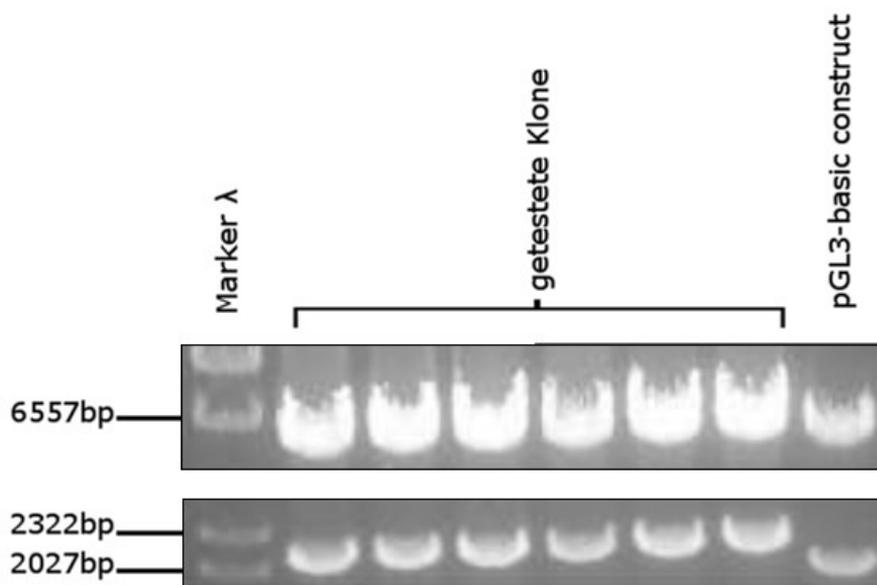


Abbildung 8 – Testverdau für Länge des Abschnitt B im neuen Konstrukt – Die Konstrukte der Klone, die nun den neuen Abschnitt B enthalten, wurden mit den Restriktionsenzymen *EcoRI* und *XhoI* verdaut. Parallel dazu wurde als Kontrolle auch das Konstrukt pGL3-basic construct mit diesen Enzymen verdaut. Die neuen Konstrukte enthalten alle das längere Insert (ca. 2100bp) im Vergleich zum bisherigen Konstrukt pGL3-basic construct.

### 3.2 Zusätzliche Konstrukte mit Varianten des Promotors

Das entstandene Ausgangskonstrukt pGL3-KPNA3-clone I enthält ca. 3500bp der Promotorregion 5' des *KPNA3*-Gens (inklusive des 5' nicht translatierten Bereichs des *KPNA3*-Gens in dessen Exon 1). Aus diesem wurden im Anschluss kürzere Konstrukte erzeugt, die etwa 2400bp 5' des *KPNA3*-Gens enthalten, um das Konstrukt zu verkleinern und später zu klären, ob überhaupt die Größe von 3500bp der Promotorregion zur Expression notwendig ist. In diesem Konstrukt ist der Polymorphismus (Single Nucleotide Polymorphism, SNP) rs11618521 an Position -1094bp enthalten, bei dem beide Versionen in der Bevölkerung häufig vorkommen (in Europa ca. 60% Guanin vs. 40% Adenin) (Sherry *et al.* 2001). Im Ausgangskonstrukt ist er in der Version mit Guanin enthalten (vgl. Abbildung 9). Dieser zeigte, wie bereits erwähnt, in Analysen in unserer Arbeitsgruppe einen signifikanten Zusammenhang mit dem Erkrankungsalter der SCA3 (Dieterle 2019). Es besteht deshalb der Verdacht, dass der SNP Auswirkungen auf die Promotoraktivität hat. Aus den genannten Gründen wurde deshalb jeweils ein Konstrukt mit der jeweiligen Variante des SNPs erzeugt. Aus diesen beiden Konstrukten wurde anschließend jeweils ein Konstrukt, welches den 5' untranslatierten Bereich des *KPNA3*-Gens nicht enthält, erzeugt, um zu untersuchen, ob dieser Einfluss auf die Expression des Gens hat, nachdem er im translatierten Genprodukt aufgrund der fehlenden Translation keine Funktion innehat. Eine Übersicht zum Aufbau der Konstrukte ist in Abbildung 10 dargestellt.

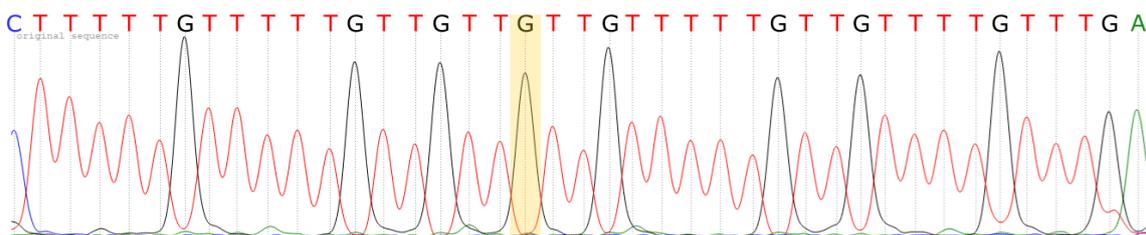


Abbildung 9 – SNP rs11618521 im Ausgangskonstrukt – Sequenzierungsergebnis vom Konstrukt pGL3-KPNA3-clone I im Bereich, in dem der SNP rs11618521 enthalten ist. Gelb markiert ist der SNP rs11618521 an Position -1094bp. Im Ausgangskonstrukt ist die Version mit Guanin enthalten.

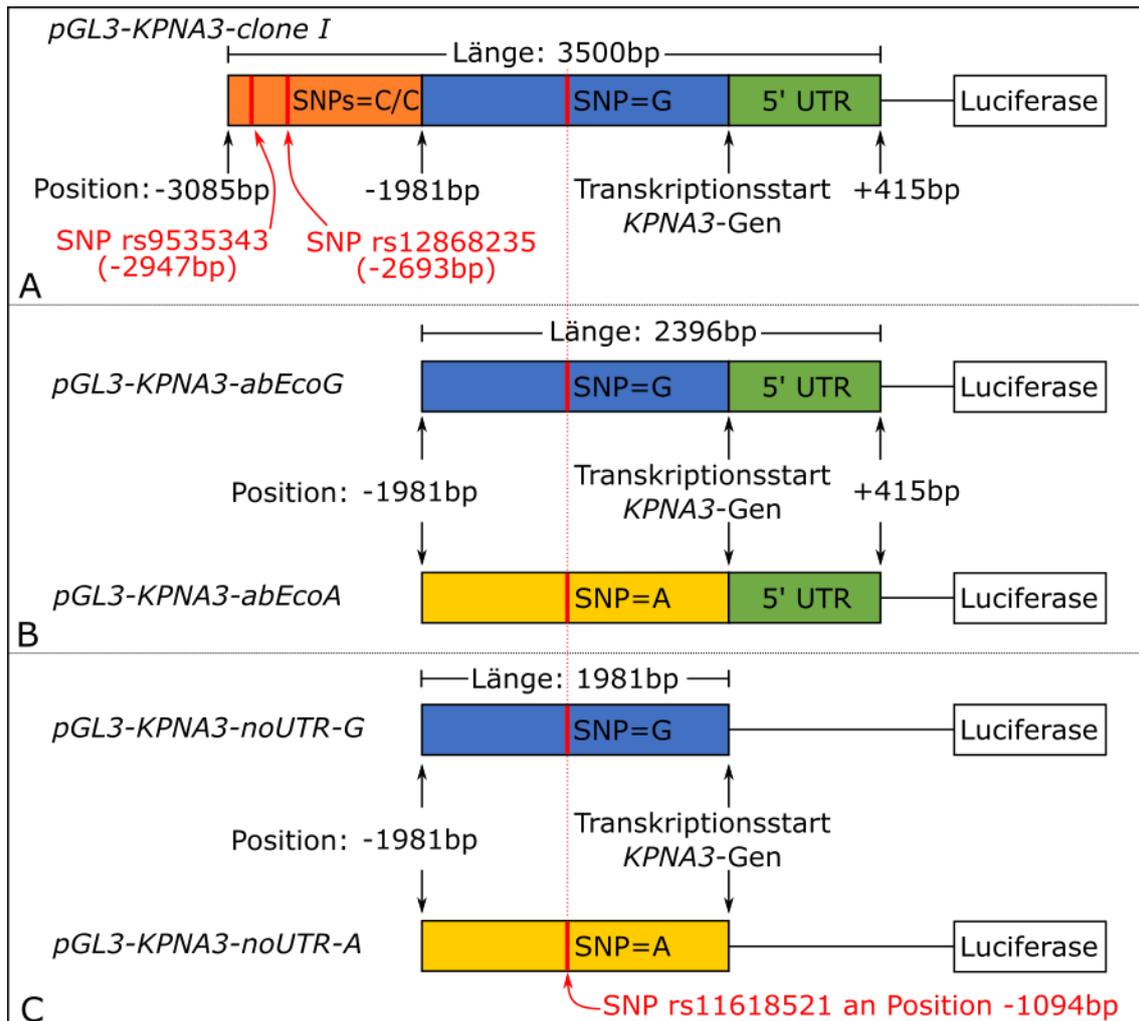


Abbildung 10 – Übersicht der generierten Konstrukte – A) Das Konstrukt *pGL3-KPNA3-clone I* enthielt 3500bp des Promotorbereichs des *KPNA3*-Gens inklusive des 5' nicht-translatierten Bereichs im Exon 1 (5' UTR). Es umfasste den Bereich von -3085bp bis +415bp ausgehend vom Transkriptionsstart. Die beiden SNPs rs9535343 und rs12868235 sind nur in diesem Konstrukt enthalten und liegen in der Version mit Cytosin vor. Der SNP rs11618521 liegt in der Version mit Guanin vor. – B) Die 5'-kürzeren Konstrukte enthielten 2394bp des Promotorbereichs, beginnend bei Position -1981 und endend an Position +415bp und somit beinhalteten sie auch den 5' UTR. Es lag jeweils eine Version des SNPs rs11618521 mit Guanin (*pGL3-KPNA3-abEcoG*) und eine mit Adenin (*pGL3-KPNA3-abEcoA*) vor. – C) Die Konstrukte ohne den 5' untranslatierten Bereich des *KPNA3*-Gens beinhalteten 1981bp des Promotorbereichs, beginnend an Position -1981bp und endend am Transkriptionsstart des *KPNA3*-Gens. Es lag auch jeweils eine Version des SNPs rs11618521 mit Guanin (*pGL3-KPNA3-noUTR-G*) und eine mit Adenin (*pGL3-KPNA3-noUTR-A*) vor. Es ist jeweils die Gesamtlänge der Konstrukte, relevante Positionen in Bezug auf den Transkriptionsstart des *KPNA3*-Gens und die ungefähre Position des SNP rs11618521 (entspricht Position -1094bp) durch eine rote Markierung dargestellt.

### 3.2.1 5'-Kürzeres Konstrukt

Zunächst wurde ein 5'-kürzeres Konstrukt aus dem bereits vorhandenen pGL3-KPNA3-clone I generiert, da der im Ausgangskonstrukt enthaltenen Bereich mit 3500bp recht groß gewählt wurde und mit den 5'-kürzeren Konstrukten untersucht werden konnte, welche Promotor-Regionen relevant sind. Hierfür wurde eine Klonierungs-PCR mit dem Ausgangskonstrukt als Template mit den Primern S159\_KPNA3Pr2-Nhe-F und S075\_KPNA3Pr-Eco-R angesetzt. Das Produkt entsprach einer Länge von 264bp und umfasste den Bereich von der EcoRI-Schnittstelle ca. 250 Basen Richtung 5'-Ende, wobei dort eine NheI Schnittstelle am PCR-Produkt eingefügt wurde. PCR-Produkt und Vektor wurden mit EcoRI und NheI verdaut und das PCR-Produkt anstatt des ca.1100bp langen Stücks aus dem pGL3-KPNA3-clone I inkliert (vgl. Abbildung 12). Es entstand ein ca. 7200bp langes Konstrukt, das 5' ca. 1000bp kürzer war als das Ausgangskonstrukt, und somit ca. 2500bp der Promotorregion enthielt, inklusive des im Exon 1 des *KPNA3*-Gens enthaltenen 5' UTR. Der enthaltene Bereich erstreckte sich von Position -1981bp bis Position +415bp

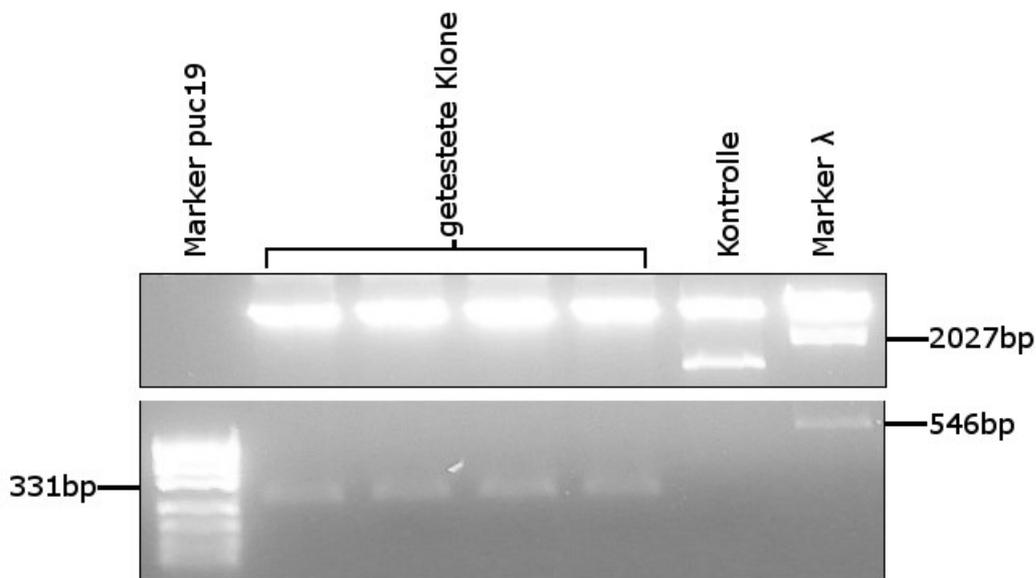


Abbildung 11 – Testverdau der angezogenen Klone der Klonierung für 5'-kürzeres Konstrukt mit *EcoRI* und *NheI* – In allen hier getesteten Klonen war das kürzere Insert mit ca. 250bp enthalten, das im Bereich der Markerbande von 331bp zu liegen kommt. Als Kontrolle diente das Ausgangskonstrukt pGL3-KPNA3-clone I, das sich im Testverdau mit einem längeren Insert mit ca.1100bp darstellt, das unterhalb der Markerbande von 2027bp zu liegen kommt. Für den Testverdau kamen die Klonierungsenzyme *EcoRI* und *NheI* zum Einsatz.

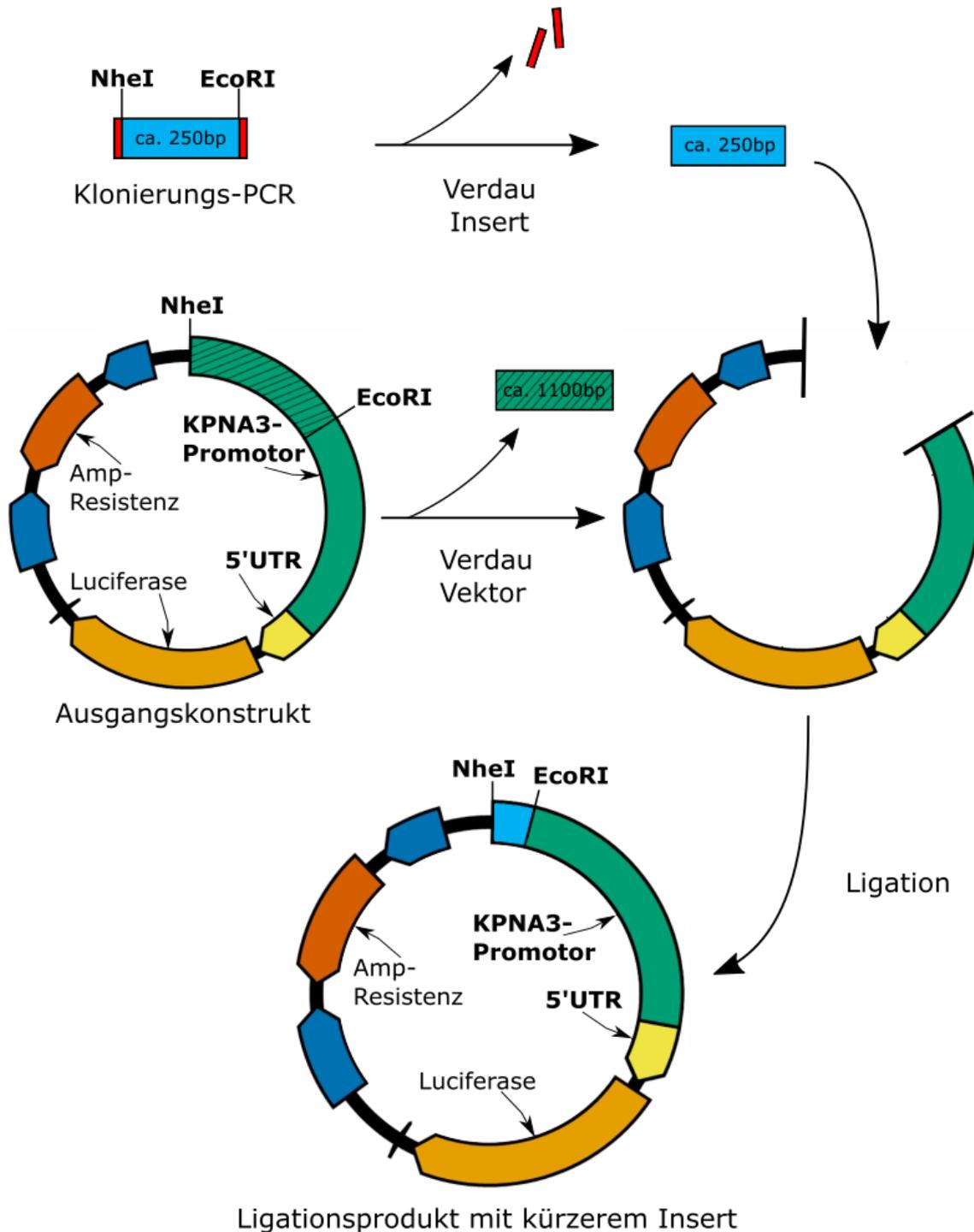


Abbildung 12 – Vorgehen bei der Klonierung für 5'-kürzeres Konstrukt – Als Primer für die Klonierungs-PCR des Inserts kamen S159\_KPNA3Pr2-Nhe-F und S075\_KPNA3Pr-Eco-R zum Einsatz. Das Ausgangskonstrukt pGL3-KPNA3- und das in der Klonierungs-PCR gewonnene Fragment wurden jeweils mit den Restriktionsenzymen *NheI* und *EcoRI* verdaut. Das ca. 1100bp Fragment des Ausgangskonstrukts wurde durch das ca. 250bp Fragment aus der Klonierungs-PCR ersetzt. Das Ligationsprodukt pGL3-KPNA3-abEcoG ist somit am 5'-Ende des Promotors verkürzt.

Das Klonierungsprodukt wurde zur Amplifikation in DH5 $\alpha$  chemisch kompetente Zellen transformiert. Die erfolgreiche Klonierung wurde durch einen Testverdau des Konstrukts mit den Klonierungsenzymen EcoRI und NheI (Abbildung 11) und mit Sequenzierung mit dem Primer S058\_RVprimer3 (Abbildung 13) bestätigt. Im Vergleich mit dem Ausgangskonstrukt war das kürzere Insert enthalten. Der SNP rs11618521 entspricht dem Ausgangskonstrukt (Guanin), da dieser Bereich nicht verändert wurde. Das Konstrukt wurde pGL3-KPNA3-abEcoG bezeichnet, aufgrund der Nähe der Eco-R1 Schnittstelle zum Anfang des Konstrukts und der SNP-Variante mit Guanin.

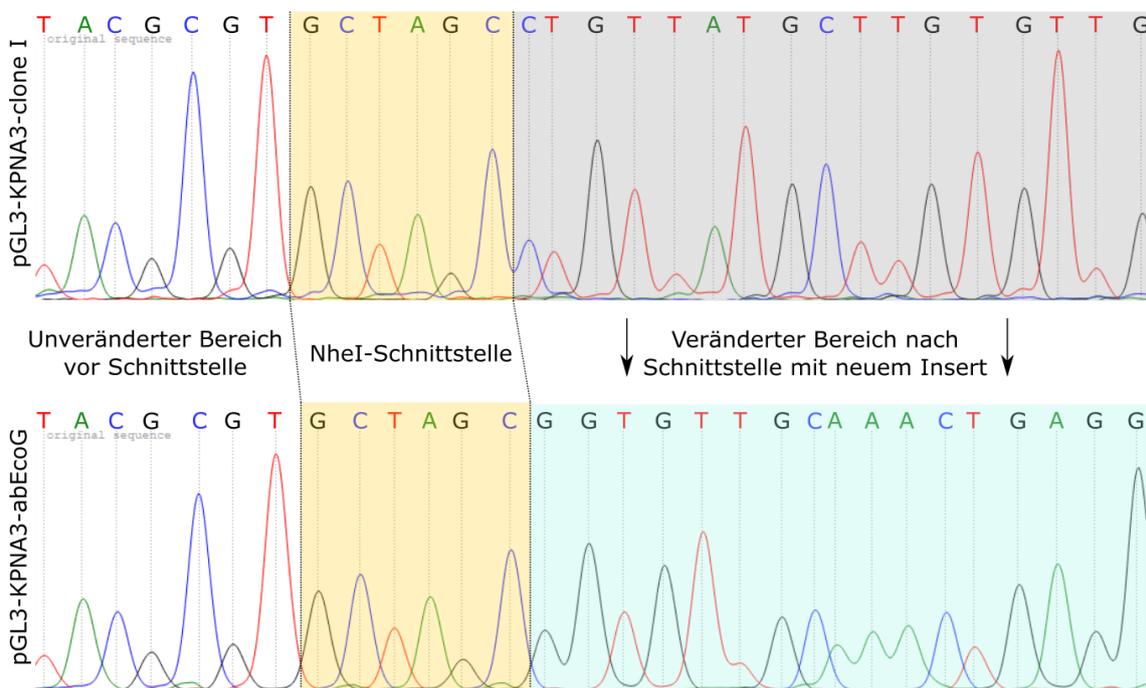


Abbildung 13 – Sequenzierungsergebnisse des Konstrukts pGL3-KPNA3-Clone I und des 5'- kürzeren Konstrukts – In Gelb ist jeweils die NheI-Schnittstelle markiert. Oben ist die Sequenz des Konstrukts pGL3-KPNA3-clone I dargestellt. In Grau ist der Bereich nach der NheI-Schnittstelle markiert. Das Konstrukt pGL3-KPNA3-clone I enthält nach der NheI-Schnittstelle den Promotorbereich beginnend an Position -3085bp. Unten ist zum Vergleich die Sequenz des 5'-Kürzeren Konstrukts pGL3-KPNA3-abEcoG zu sehen. In Türkis ist der Bereich nach der NheI-Schnittstelle markiert, der sich in der erwarteten Form von der Sequenz des Konstrukts pGL3-KPNA3-clone I unterscheidet und somit die erfolgreiche Einklonierung des neuen Inserts bestätigt. Das neue Konstrukt enthält den Promotorbereich ab Position -1981. Als Sequenzierungsprimer kam jeweils S058\_RVprimer3 zum Einsatz

### 3.2.2 5'-kürzeres Konstrukt mit modifiziertem SNP

Um den Polymorphismus rs11618521, der im 5'-kürzeren Konstrukt enthalten ist, in beiden Varianten in Konstrukten enthalten zu haben, wurde ein 420bp langes DNA-Fragment von Integrated DNA Technologies Inc. (Skokie, USA) bezogen,

das den SNP in der zweiten Variante enthält (Adenin statt Guanin) und mit den Restriktionsenzymen NdeI und PstI einklonierbar war. Die restliche Sequenz entsprach dem Ausgangskonstrukt. Dieses DNA-Stück wurde mittels Klonierungs-PCR mit den Primern S983\_KPNA3Pr-Nde-F und S984\_KPNA3Pr-Asp-R vervielfältigt. Anschließend wurden das PCR-Produkt und der Vektor (pGL3-KPNA3-abEcoG) mit NdeI und PstI in parallelen Ansätzen verdaut und das neue Insert in den Vektor einkloniert (Abbildung 15). Das Klonierungsprodukt wurde zur Amplifikation in DH5 $\alpha$  chemisch kompetente Zellen transformiert.

Die erfolgreiche Klonierung wurde mittels Testverdau mit den Klonierungsenzymen NdeI und PstI und Sequenzierung mit den Primern S785\_KPNA3Pr-1300-F beziehungsweise S080\_KPNA3Pr-1600-F (Abbildung 14) bestätigt. Das neue Konstrukt wurde mit pGL3-KPNA3-abEcoA bezeichnet (SNP-Variante mit Adenin). Es enthielt den selben Bereich des Promotors, wie das Konstrukt pGL3-KPNA3-abEcoG, von dem es sich bis auf die Variante des SNP rs11618521 nicht unterschied.

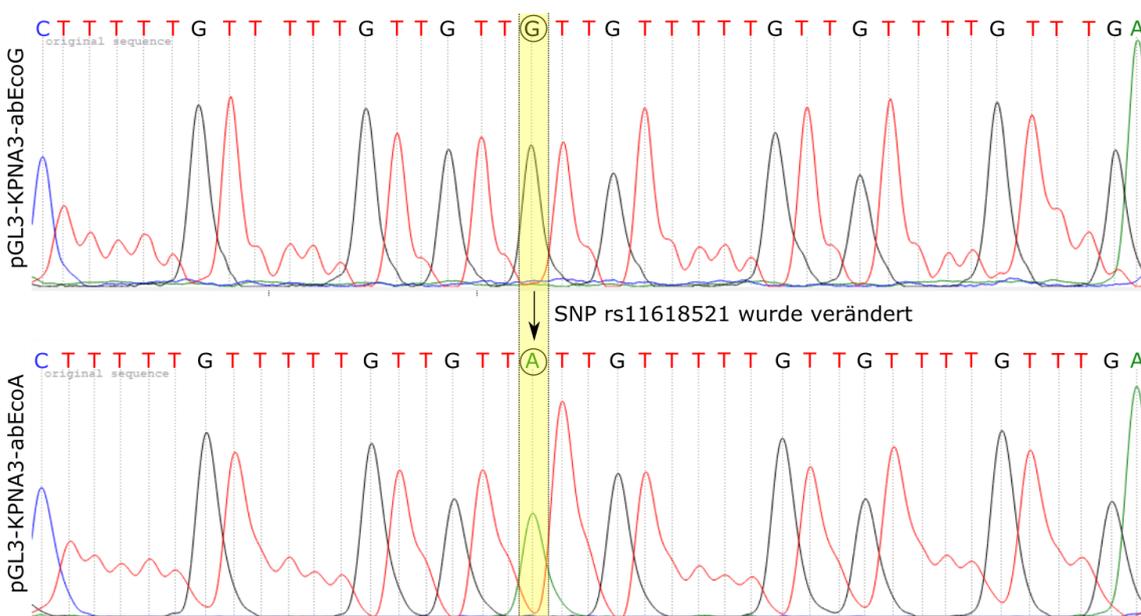


Abbildung 14 – Sequenzierungsergebnisse der 5'-kürzeren Konstrukte mit den Varianten des SNP rs11618521 – Gelb markiert ist jeweils der SNP rs11618521 an Position -1094bp. Im Ausgangskonstrukt pGL3-KPNA3-abEcoG (oben) war der Polymorphismus rs11618521 mit Adenin als Base enthalten (Sequenzierung mit Primer S785\_KPNA3Pr-1300-F). Im neuen Konstrukt pGL3-KPNA3-abEcoA (unten) ist er in der zweiten Variante mit Adenin enthalten (Sequenzierung mit Primer S080\_KPNA3Pr-1600-F). Die erfolgreiche Einklonierung des Inserts konnte hiermit bestätigt werden.

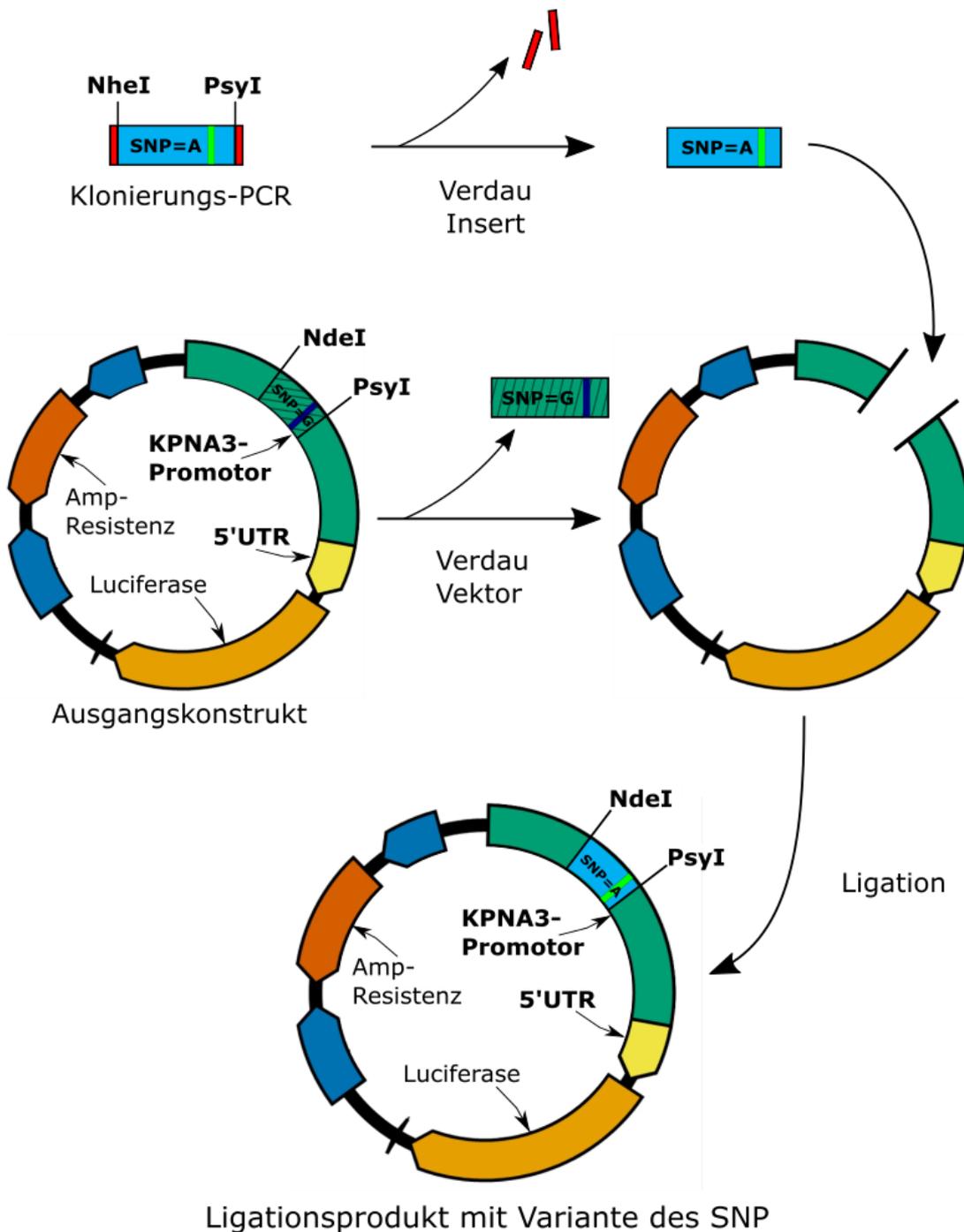


Abbildung 15 – Vorgehen bei der Klonierung für 5'-kürzeres Konstrukt mit verändertem SNP rs11618521 – Als Primer für die Klonierungs-PCR des Inserts kamen S983\_KPNA3Pr-Nde-F und S984\_KPNA3Pr-Asp-R zum Einsatz. Das PCR-Produkt mit der Variante des SNP rs11618521 mit Adenin und der Vektor pGL3-KPNA3-abEcoG wurden jeweils mit NdeI und PstI verdaut und anschließend ligiert. Der Bereich im Ausgangskonstrukt enthielt im ausgeschnittenen Bereich den SNP rs11618521 in der Version mit Guanin (grünes Fragment, SNP-Position dunkelblau markiert), der durch die Einklonierung des Inserts mit der Version mit Adenin ersetzt wurde (hellblaues Fragment, SNP-Position hellgrün markiert). Das Ligationsprodukt enthält somit den modifiziertem SNP rs11618521 in der Version mit Adenin, ansonsten sind die Basenabfolgen zwischen den ausgetauschten Fragmenten identisch.

#### 3.2.3 Konstrukte ohne 5' UTR des Exon 1

Aus den beiden 5'-kürzeren Konstrukten wurden Konstrukte generiert, bei denen der 5' UTR des Exon 1 des *KPNA3*-Gens nicht enthalten ist. Hierfür wurde eine Klonierungs-PCR durchgeführt, die den Bereich von der *Pst*I Schnittstelle der 5'-kürzeren Konstrukte bis zum Beginn des 5' UTR abdeckte und dort eine *Xho*I-Schnittstelle einfügte. Als Primer kamen hierfür S786\_KPNA3Pr-1189-F und S788\_KPNA3PrEx1Xho-R zum Einsatz. Letzterer fügte eine *Xho*I-Schnittstelle im Bereich des Transkriptionsstarts des *KPNA3*-Gens ein. Die beiden 5'-kürzeren Konstrukte, die jeweils eine Variante des SNP rs11618521 enthalten, und das PCR-Produkt wurden mit den Restriktionsenzymen *Pst*I und *Xho*I verdaut und das 1400bp lange Fragment der Ausgangskonstrukte mit dem ca. 1000bp langen Insert ohne 5' UTR der PCR ausgetauscht (Abbildung 16). Somit war nun für jede der zwei Varianten des SNP ein Ansatz erfolgt, dessen Produkt den 5' UTR nicht mehr enthielt.

Die Klonierungsprodukte wurden zur Amplifikation jeweils in DH5 $\alpha$  chemisch kompetente Zellen transformiert. Die erfolgreiche Klonierung wurde mittels Testverdau und Sequenzierung bestätigt. Im Testverdau waren von den getesteten Klonen der Transformation für die Variante des SNP rs11618521 mit Guanin drei Klone positiv für das neue Insert, für die Variante mit Adenin ein Klon (Abbildung 17). Von diesen wurde zur Bestätigung der erfolgreichen Klonierung jeweils ein Klon sequenziert und die SNP Variante dort auch bestätigt (Abbildung 18). Als Primer kam hierzu jeweils S785\_KPNA3Pr-1300-F zum Einsatz. Es wurden somit für beide Varianten des SNPs je ein Konstrukt generiert, das den 5' UTR nicht enthielt. Diese deckten somit den Bereich von Position -1981bp bis zum Transkriptionsstart des *KPNA3*-Gens ab.

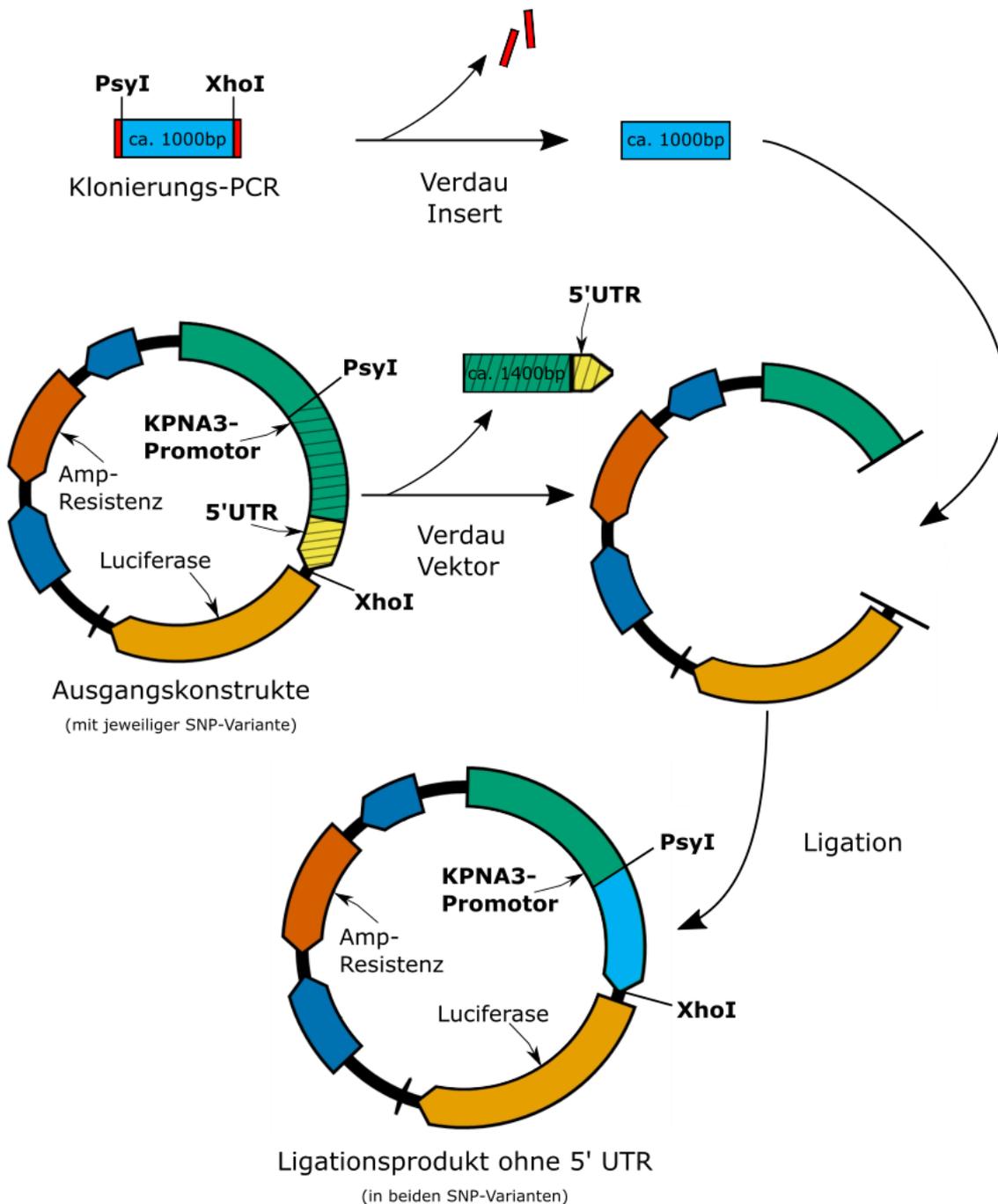


Abbildung 16 – Vorgehen bei der Klonierung für 3'-kürzere Konstrukte ohne den 5'-untranslatierten Bereich des KPNA3-Gens – Es wurden für beide Varianten des SNP rs11618521 parallel Ansätze für 3'-kürzere Konstrukte mit der jeweiligen Variante durchgeführt. Als Primer für die Klonierungs-PCR des Inserts kamen S786\_KPNA3Pr-1189-F und S788\_KPNA3PrEx1Xho-R zum Einsatz. Als Klonierungsenzyme dienten *PstI* und *XhoI*. Im Endprodukt war jeweils der 5' UTR nicht mehr enthalten, das Klonierungsprodukt endete folglich an Position 1, die dem Transkriptionsstart des KPNA3-Gens entspricht.

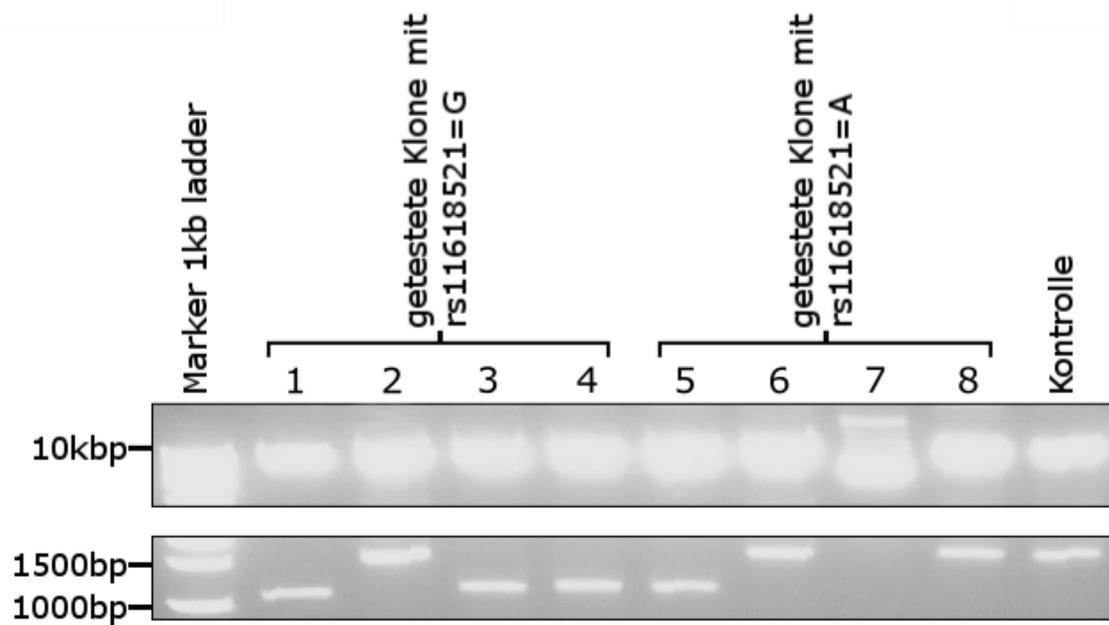


Abbildung 17 – Testverdau für 3'-kürzere Konstrukte ohne den 5'-untranslatierten Bereich des KPNA3-Gens mit PstI und XhoI – Für die Version des SNP rs11618521 mit Guanin enthielten drei der vier getesteten Klone (Nummer 1, 3 und 4) die erwartete Insertbande von ca. 1000bp Länge und für die Version des SNP rs11618521 mit Adenin war es ein Klon der vier getesteten, der die erwartete Insertbande von ca. 1000bp enthielt (Nummer 5). Als Kontrolle diente das Konstrukt pGL3-KPNA3-abEcoG, bei dem das ausgeschnittene Fragment bei ca.1400bp zu liegen kommt. Die Klone mit den Nummern 2, 6 und 8 enthielten ein Insert, das in seiner Länge der Insertbande der Kontrolle entsprach. Hier war die Klonierung nicht erfolgreich, vermutlich kam es entweder beim Ausschneiden der Insertbanden im präparativen Verdau zu Verunreinigungen, wodurch bei der Ligation wieder das Ausgangskonstrukt entstand, oder im präparativen Verdau wurde das Ausgangskonstrukt nicht vollständig verdaut, wodurch bei der anschließenden Ligation und Transfektion wiederum das Ausgangskonstrukt mitgehalten war. Beim Klon Nummer 7 waren keine klaren Bande abzugrenzen, vermutlich hatte hier der Verdau nicht funktioniert, was beispielsweise auf Verunreinigungen beim Ansatz zurückzuführen sein könnte. Als Verdauungsenzyme kamen jeweils die Klonierungsenzyme PstI und XhoI zum Einsatz.

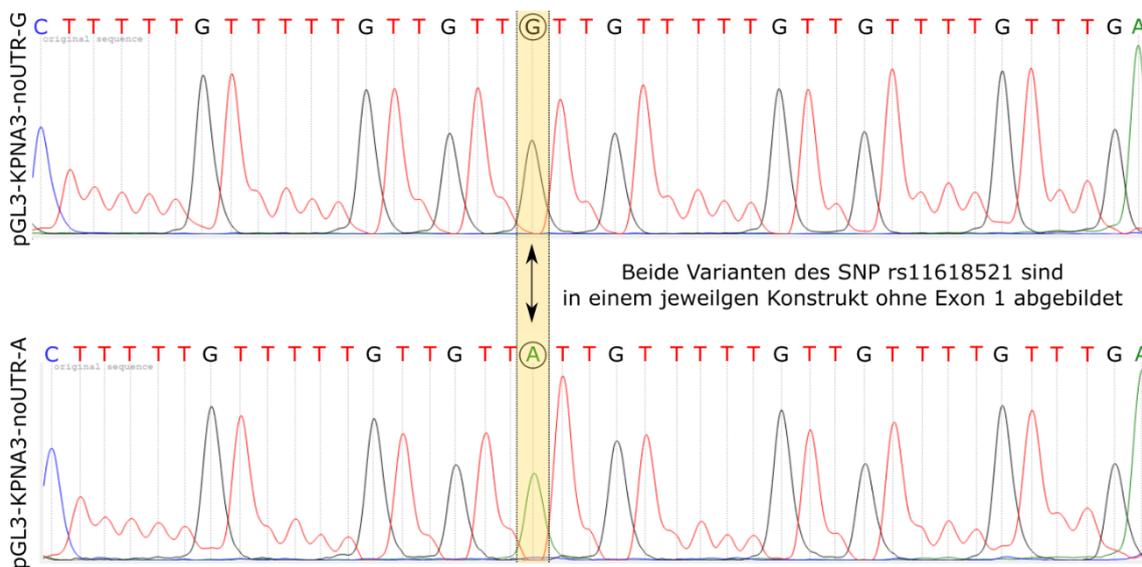


Abbildung 18 – Sequenzierungsergebnis der beiden Varianten des Konstrukts ohne den 5'-untranslatierten Bereich des *KPNA3*-Gens – Oben ist die Sequenzierung des Konstrukts pGL3-KPNA3-noUTR-G dargestellt, unten die Sequenzierung des Konstrukts pGL3-KPNA3-noUTR-A. Markiert in Gelb ist der SNP rs11618521, der beim Konstrukt pGL3-KPNA3-noUTR-G der Version mit Guanin entspricht und beim Konstrukt pGL3-KPNA3-noUTR-A der Version mit Adenin. Ansonsten unterscheiden sich beide Konstrukte nicht in ihrer Sequenz. Sequenziert wurde jeweils mit dem Primer S785\_KPNA3Pr-1300-F.

Insgesamt wurden somit nun fünf Konstrukte erzeugt. Das zunächst erzeugte Ausgangskonstrukt, mit pGL3-KPNA3-clone I bezeichnet, enthielt ca. 3500bp der Promotorregion des *KPNA3*-Gens inklusive dessen untranslatiertem Bereich am 5' Ende des Exon 1. Des Weiteren wurden zwei Konstrukte erzeugt, die 5' kürzer waren und ca. 2500bp der Promotorregion enthielten, ebenfalls inklusive des 5' UTR. Von diesen enthielt eines den SNP rs11618521 in der Version mit Guanin, welches mit pGL3-KPNA3-abEcoG bezeichnet wurde, und das andere den SNP in der Version mit Adenin, welches mit pGL3-KPNA3-abEcoA bezeichnet wurde. Letztlich wurden noch die beiden Konstrukte erzeugt, die den 5' UTR des *KPNA3*-Gens nicht enthielten, und somit etwa 2000bp der Promotorregion beinhalteten. Hiervon enthielt auch eines den SNP rs11618521 in der Version mit Guanin, bezeichnet als pGL3-KPNA3-noUTR-G, und das andere den SNP in der Version mit Adenin, bezeichnet mit pGL3-KPNA3-noUTR-A.

Zur Übersicht aller erzeugten Konstrukte sei nochmals auf Abbildung 10 verwiesen. Weiterhin sind die Konstrukte nochmals in Tabelle 16 mit ihren Bezeichnungen, Eigenschaften und den jeweiligen SNP-Varianten des SNP rs11618521 aufgelistet. Das Konstrukt pGL3-basic construct, welches als Ergebnis aus der Klonierung unter 3.1.2 entstand, wird hier nicht mehr aufgeführt, da es aufgrund des fehlerhaften Abschnitts am 3' Ende im Bereich des enthaltenen 5' UTRs nicht für die weiteren Experimente genutzt wurde.

*Tabelle 16 – Übersicht über die erzeugten Konstrukte mit ihren Eigenschaften*

Bezeichnung [pGL3-KPNA3...]	Länge	Pos. 5'	Pos. 3'	5' UTR	SNP Status
-clone I	3500bp	-3085bp	+415bp	+	G
-abEcoG	2396bp	-1981bp	+415bp	+	G
-abEcoA	2396bp	-1981bp	+415bp	+	A
-noUTR-G	1981bp	-1981bp	-1bp	-	G
-noUTR-A	1981bp	-1981bp	-1bp	-	A

*Hinweise: Alle Konstrukte beginnen bei der Bezeichnung mit pGL3-KPNA3 gefolgt von einer individuellen Endung. Der Transkriptionsstart entspricht der Position +1bp, alle Positionen sind darauf bezogen angegeben. Das Vorhandensein bzw. Fehlen des 5' UTR ist mit einem Plus- bzw. Minuszeichen dargestellt. Im SNP-Status ist jeweils die enthaltene Base des SNPs rs11618521 angegeben (A-Adenin, G-Guanin). Die beiden SNPs rs9535343 und rs12868235 sind nicht aufgeführt, da sie nur im Konstrukt pGL3-KPNA3-clone I vorkommen (beide in der Version mit Cytosin).*

### 3.3 Expression durch den KPNA3-Promotor

#### 3.3.1 Expression im Western-Blot

Zur Überprüfung der Expression durch den KPNA3-Promotor wurden zunächst die fünf Konstrukte mit den verschiedenen Versionen des Promotors in HEK293T-Zellen transfiziert und die Expression der nachgeschalteten Firefly-Luciferase mittels Western-Blot überprüft.

Hierbei zeigte sich, dass bei allen Konstrukten die Firefly-Luciferase exprimiert wird, wodurch davon ausgegangen werden kann, dass in allen Konstrukten für das nachgeschaltete Gen relevante Teile des Promotors enthalten sind. Als negativ Kontrolle wurden Zellen entweder mit dem Konstrukt pRL-CMV, das im nachfolgenden Luciferase-Assay als Standard dienen soll, oder mit dem pGL3-basic Vektor (ohne Promotorregion) transfiziert (Abbildung 19A).

Im Wester-Blot war erkennbar, dass die Konstrukte die Luciferase verschieden stark exprimieren. Hierbei war ein Trend erkennbar, dass das Konstrukt pGL3-KPNA3-abEcoA die stärkste Expression zeigte. Im Vergleich mit dem bis auf den SNP rs11618521 identischen Konstrukt pGL3-KPNA3-abEcoG war ein Unterschied zu erkennen, der jedoch nicht die statistische Signifikanz erreichte ( $p=0,15$ ). Bei den Konstrukten, bei denen der 5' UTR fehlte, war kein solcher Effekt zu sehen (Abbildung 19B).

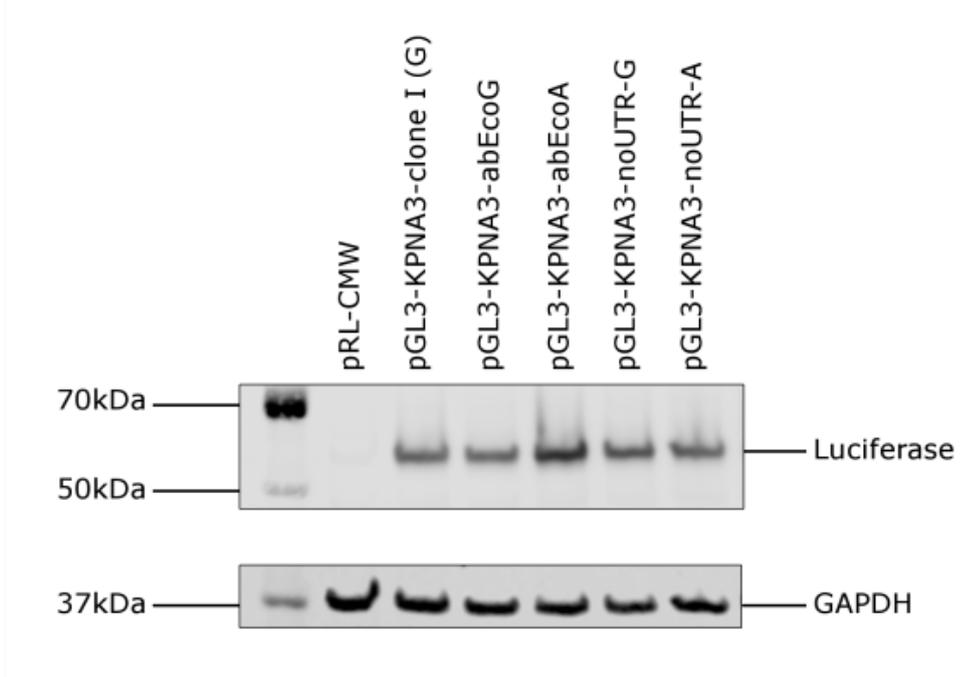
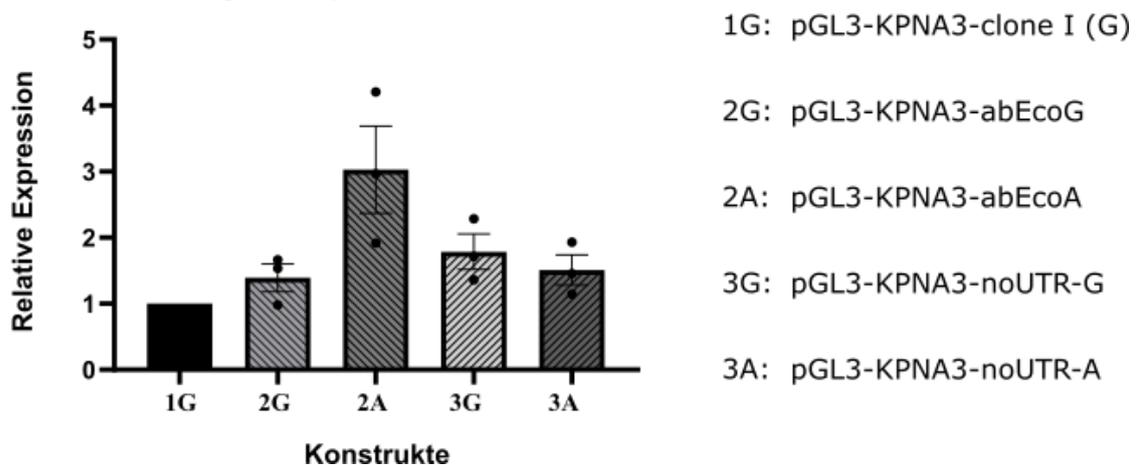
**A Expression der Konstrukte im Western Blot****B Quantifizierung der Expression im Western Blot**

Abbildung 19 – Ergebnisse der Expressionsanalyse mittels Western-Blot – A) Expression der Luciferase im Western-Blot – Alle erzeugten Konstrukte zeigten eine Expression der Firefly-Luciferase. Die Negativkontrolle pRL-CMV zeigt keine Expression der Firefly-Luciferase. Als Ladekontrolle diente GAPDH. – B) Quantifizierung der Expression im Western-Blot– Die Expression der Luciferase durch die Konstrukte wurde mit der Expression des Ausgangskonstrukts pGL3-KPNA3-clone I (1G) als Standard quantifiziert, dieses enthält den SNP rs11618521 in der Version mit Guanin (G). Dessen Expression wurde der Wert 1 zugeordnet und die Expression der anderen Konstrukte hierzu ins Verhältnis gesetzt. Es ist ein Trend zu erkennen, dass sich das Konstrukt pGL3-KPNA3-abEcoA (2A) von den anderen unterscheidet und stärker die Firefly-Luciferase exprimiert. Ein statistisch signifikanter Unterschied ist jedoch nicht vorhanden. Der SNP rs11618521 scheint bei den Konstrukten, die 5'-kürzer als das Ausgangskonstrukt sind (2G und 2A) einen Einfluss auf die Expression zu haben. Fehlt zusätzlich der 5' UTR (3G und 3A) ist kein solcher Effekt erkennbar. – Angegeben sind Messpunkte, Mittelwert und Standardfehler; n=3.

**3.3.2 Expression im Luciferase-Assay**

Um die Erkenntnisse aus den Western-Blots weiter zu untersuchen, wurde die Expression der Konstrukte mittels Luciferase-Assays gemessen. Dieses hat den

Vorteil, dass im Vergleich zum Western-Blot, wo die Quantifizierung der Expression erst in einem zweiten Schritt anhand des erzeugten Bildes erfolgt, im Luciferase-Assay direkt quantitative Werte gemessen werden (nämlich die gemessene Lumineszenz), die auf die Expression der Luciferase schließen lassen. Da das zu testende Konstrukt und das Konstrukt, das als Standard dient (hier pRL-CMV), parallel transfiziert werden, können auch potenzielle Messfehler durch Unterschiede bei der Transfektionseffektivität in den einzelnen Ansätzen minimiert werden. Weiterhin ist es mit dem Luciferase-Assay möglich mehrere Versuchsreihen mit einer hohen Anzahl an Proben (in diesem Modell theoretisch bis zu 96) in kürzeren Zeiteinheiten als im Western-Blot durchzuführen und durch wesentlich weniger Versuchsschritte sind weniger potenzielle Störfaktoren und Fehlerquellen vorhanden.

Zunächst wurden mit dem Ausgangskonstrukt verschiedene Assay-Bedingungen getestet. Hierbei wurde festgestellt, dass eine Transfektion mittels Fast-Forward Transfektion (siehe unter 2.2.8) möglich ist. Außerdem wurde getestet, ob ein Austausch des Zellmediums 24h nach Transfektion einen Einfluss auf das Messergebnis hat. Es zeigte sich, dass die Messungen dadurch nicht bedeutend verändert werden, weshalb auch wegen des Fehlerpotentials durch den Mediumsaustausch bei den folgenden Versuchen darauf verzichtet wurde. Eine solche Messung ist beispielhaft in Abbildung 20 dargestellt. Weiterhin wurde die Zellzahl pro Well getestet, die sich am besten zur Messung eignet, und letztlich auf 40 000 Zellen für jedes Well festgelegt. Außerdem wurden zunächst Messungen in allen drei im Labor vorhanden Geräten, die die Möglichkeit zur Messung von Lumineszenz bieten, durchgeführt, um Vorteile bei der Handhabung und Unterschiede in der Datenausgabe zu erkennen. Hierbei wurde der Berthold Technologies MicroLumat Plus LB96V als am geeignetsten identifiziert, da dieser eine einfache Bedienung, die schnellste Messung im Vergleich und dabei gleichbleibende Messwerte mit großem dynamischem Bereich bot. In Abbildung 21 sind beispielhaft Messdaten eines einzelnen Versuchsansatzes in den drei Geräten graphisch dargestellt. Aus den genannten Erkenntnissen hat sich das definitive Vorgehen zu Expressionsanalyse der verschiedenen Konstrukte ergeben, das auch unter 2.2.8 und 2.2.9 beschrieben

ist. Um Effekte, die durch die Position der Proben in der Platte während der Inkubation (z.B. durch Temperaturunterschiede) entstehen konnten, zu minimieren, wurden die Positionen der verschiedenen Konstrukte jeweils von Experiment zu Experiment variiert und die umgebenden Wells der Platten jeweils mit Zellmedium befüllt. Die rohen Messwerte der einzelnen Konstrukte wurden jeweils auf die von ihnen gezeigte Lumineszenz normiert und im Vergleich zur gewählten Referenz dann daraus die relative Expression ermittelt. Das genaue Vorgehen zur Berechnung ist auch unter 2.2.8 aufgeführt.

### Unterschiede bei Vorgehen mit und ohne Mediumswechsel

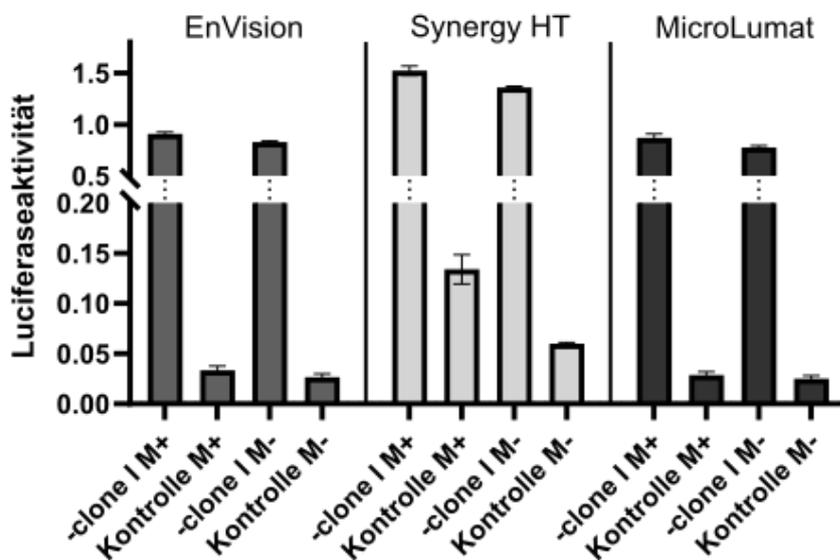


Abbildung 20 – Unterschiede bei Vorgehen mit und ohne Mediumswechsel – Dargestellt ist exemplarisch eine Messung der Lumineszenz bei Versuchsansätzen, bei denen nach 24 Stunden nach Transfektion das Zellmedium durch frisches ersetzt wurde (jeweils mit „M+“ gekennzeichnet). Diesen gegenüber gestellt ist jeweils die Messung der Lumineszenz von Ansätzen, die die kompletten 48 Stunden nach Transfektion im selben Medium ohne dessen Austausch inkubiert wurden (gekennzeichnet durch „M-“). Für diesen Versuch kam das Konstrukt pGL3-KPNA3-clone I zum Einsatz, der Übersichtlichkeit wegen mit „-clone I“ abgekürzt. Als Kontrolle diente der pGL3-Basic Vektor ohne inklonierten Promotorbereich. Es wurde jeweils in den drei verfügbaren Geräten (EnVision, Synergy HT und MicroLumat) analysiert, hierbei wurde kein relevanter Unterschied zwischen den Ansätzen mit und ohne Mediumswechsel beobachtet. In den weiteren Versuchen wurde deshalb auf den Mediumswechsel verzichtet, auch um dies als Fehlerquelle zu vermeiden. Für jedes Konstrukt und Vorgehen bezüglich des Mediumswechsels wurden je drei Wells in der Zellkulturplatte transfiziert. Hier dargestellt ist jeweils der Mittelwert und die Standardabweichung der normalisierten Luciferaseaktivität (Berechnung siehe unter 2.2.9) dieser drei Wells pro Konstrukt.

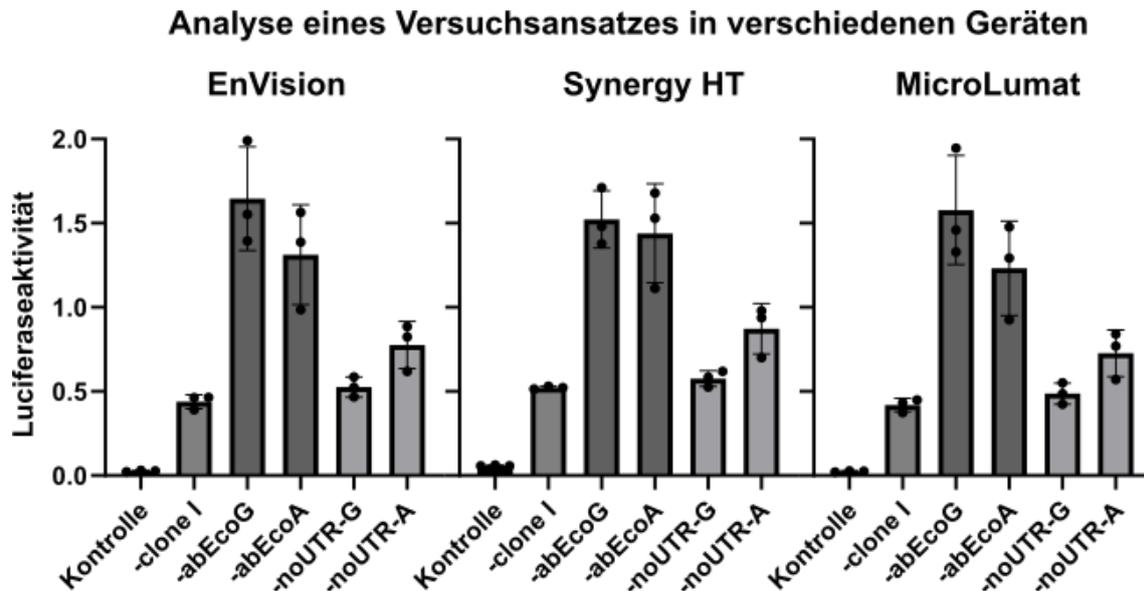


Abbildung 21 – Analyse eines Versuchsansatzes in verschiedenen Geräten – Exemplarisch ist die Analyse eines Versuchsansatzes mit den fünf erzeugten Konstrukten in den drei vorhandenen Geräten im Labor zur Lumineszenzmessung (EnVision, Synergy HT, MicroLumat) dargestellt. Für jedes Konstrukt wurden je drei Wells in der Zellkulturplatte transfiziert. Hier dargestellt ist jeweils der Mittelwert und die Standardabweichung der normalisierten Luciferaseaktivität (Berechnung siehe unter 2.2.9), sowie die Datenpunkte dieser drei Wells pro Konstrukt. Als Kontrolle diente der pGL3-Basic Vektor ohne inklonierten Promotorbereich. Es konnte hierbei kein bedeutender Unterschied der Daten (-qualität) zwischen den einzelnen Analysegeräten beobachtet werden. Aufgrund der einfachen Handhabung, der konstanten Messergebnisse mit großem dynamischem Bereich und der hohen Geschwindigkeit der Messungen wurde der Berthold Technologies MicroLumat Plus LB96V für die endgültigen Versuche genutzt. Aus Gründen der Übersichtlichkeit wurde nur die Endung der jeweiligen Konstrukte angegeben, sie haben alle denselben Stamm mit pGL3-KPNA3[...].

In den durchgeführten Luciferase-Assays diente, wie schon in den Western-Blots, das Ausgangskonstrukt pGL3-KPNA3-cloneI als Referenz für die Expressionsstärke. Wie sich im Western-Blot bereits andeutete, war zu erkennen, dass sich das Konstrukt pGL3-KPNA3-abEcoA signifikant vom Konstrukt pGL3-KPNA3-abEcoG unterschied ( $p=0,0056$ ;  $n=5$ ), das bis auf den SNP rs11618521 (Base Adenin statt Guanin) identisch ist. Der SNP hat also in diesen Konstrukten deutliche Auswirkungen auf die Expression des nachgeschalteten (*Luciferase*-) Gens. Hier war die Expression um über 70% höher, wenn der SNP rs11618521 die Base Adenin enthielt. Außerdem zeigte sich, dass die beiden Konstrukte pGL3-KPNA3-abEcoA und pGL3-KPNA3-noUTR-A, die sich nur im Vorhandensein (pGL3-KPNA3-abEcoA) beziehungsweise der Abwesenheit (pGL3-KPNA3-noUTR-A) des 5' UTRs unterschieden, eine signifikant unterschiedliche Expression besaßen ( $p=0,0245$ ,  $n=5$ ). Das Konstrukt pGL3-KPNA3-abEcoA wies hierbei eine fast 60% stärkere Expression auf. Bei den Konstrukten, bei denen der 5' UTR fehlte, war wie bereits

im Western-Blot vermutet, auch weiterhin kein Unterschied zwischen den Konstrukten mit den zwei SNP rs11618521 Varianten feststellbar (Abbildung 22).

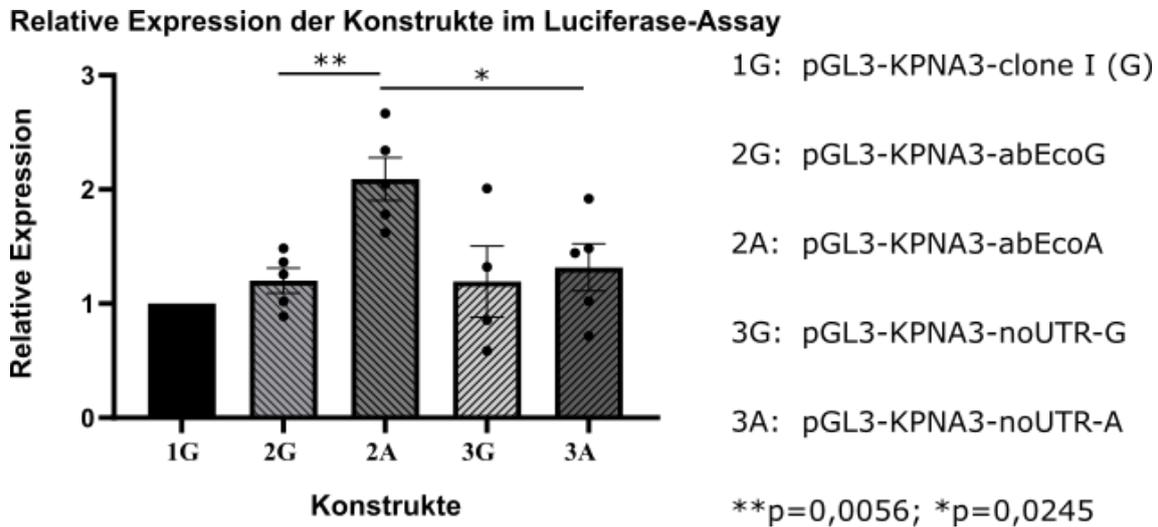


Abbildung 22 – Relative Expression der Konstrukte im Luciferase-Assay – Für die Expression diente das Ausgangskonstrukt pGL3-KPNA3-clone I (1G) als Referenz, dieses enthält den SNP rs11618521 in der Version mit Guanin (G). Dessen Expression wurde der Wert 1 zugeordnet und die Expression der anderen Konstrukte hierzu ins Verhältnis gesetzt. Zwischen den Konstrukten, die sich nur im SNP rs11618521 unterscheiden, war bei denen ohne den 5' UTR (3G und 3A) kein Unterschied erkennbar. Bei denen, die den 5' UTR enthalten (2G und 2A), war ein signifikanter Unterschied erkennbar, wobei das Konstrukt pGL3-KPNA3-abEcoA (2A), bei dem der SNP rs11618521 Adenin enthält, die Luciferase über 70% mehr exprimiert als das Konstrukt pGL3-KPNA3-abEcoG (2G), bei dem der SNP Guanin enthält. Die beiden Konstrukte pGL3-KPNA3-abEcoA (2A) und pGL3-KPNA3-noUTR-A (3A), die sich nur im Vorhandensein (2A) beziehungsweise der Abwesenheit (3A) des 5' UTRs unterscheiden, zeigen eine signifikant unterschiedliche Expression, wobei pGL3-KPNA3-abEcoA (5' UTR vorhanden) eine fast 60% stärkere Expression besitzt. –Angegeben sind Messpunkte, Mittelwert und Standardfehler; \*\*p=0,0056, \*p=0,0245 (Welch-t-Test); n=5.

### 3.4 Identifizierung relevanter Transkriptionsfaktoren

Die Ergebnisse aus den Versuchen unter 3.3 legen nahe, dass der SNP rs11618521 maßgeblich Einfluss auf die Expression durch die Promotorregion hatte, je nachdem in welcher Variante er enthalten war. Es bestand die Möglichkeit, dass es spezifische Transkriptionsfaktoren gab, die nur banden, wenn der SNP in der Variante mit Adenin enthalten war, und die somit für eine verstärkte Expression sorgten. Durch Recherche in Online-Datenbanken und mithilfe von Online-Tools wurden nun potenziell in Frage kommende Transkriptionsfaktoren ermittelt, die als Erklärung für die Expressionsunterschiede in Frage kamen.

Bei der Suche nach Transkriptionsfaktoren, die in der Region des SNPs rs11618521 binden, kam zunächst der UCSC Genome Browser zum Einsatz (Erreichbarkeit und Referenzen zu diesem und allen nachfolgend genannten Tools und Datensätzen sind unter 2.2.11 einzusehen). Als Datensätze für die Transkriptionsfaktoren kamen Datensätze der Enzyklopädie der DNA Elemente „ENCODE“ zum Einsatz. Bei Betrachtung des UCSC Genome Browser mit dem ENCODE-Datensatz „Transcription Factor ChIP-seq from ENCODE (V2)“ (Abbildung 23) zeigte sich, dass die meistens untersuchten Transkriptionsfaktoren erst im Promotorbereich 3' des SNPs rs11618521 in Richtung des *KPNA3*-Gens banden. Ein großer Anteil der dort ermittelten Transkriptionsfaktoren besitzt laut dieser Datenbank auch Bindestellen im Bereich des 5' UTR. Der Transkriptionsfaktor HA-E2F1 (entspricht E2F Transkriptionsfaktor 1/Retinoblastom-Binding Protein 3) kam in dieser Datenbank als einziger relevanter Mediator der Expression bezogen auf die Region um den SNP rs11618521 in Frage, da dieser in der Region um den SNP herum zu binden schien. Da in dem genannten Datensatz der SNP rs11618521 in der Version mit Guanin enthalten war, erklärte die Bindestelle jedoch noch nicht, weshalb die Version mit Adenin eine stärkere Expression zeigte, und ob es Unterschiede in den Bindestellen zwischen den beiden SNP-Varianten gibt.

Um dies weiter zu untersuchen, wurde das Online-Tool „PROMO“ genutzt, welches die Untersuchung spezifischer Sequenzen auf Transkriptionsfaktor-Bindestellen erlaubt. Zur Voreinstellung wurde gewählt, dass nur auf humane Transkriptionsfaktoren und Bindungsstellen im humanen Genom getestet werden sollte. Es wurde eine Sequenz gewählt, die jeweils 100bp des Promotorbereichs vor und nach der Position des SNP rs11618521 enthielt, und die Sequenz jeweils einmal in der Version des SNP mit Guanin und einmal in der Version mit Adenin auf Transkriptionsfaktorbindestellen untersucht. Es zeigte sich hierbei, dass bei der Version des SNPs mit Adenin fünf zusätzlich Transkriptionsfaktoren Bindestellen zeigten, die bei der Variante mit Guanin nicht vorhanden waren. Die fünf zusätzlichen Transkriptionsfaktoren, die nur in der Version des SNPs mit Adenin Bindungsstellen besaßen, waren HOXD9, HOXD10, GR-beta, C/EBPalpha und HNF-1C (Abbildung 24). In der SNP-Variante mit Guanin

zeigten die Transkriptionsfaktoren FOXP3 und C/EBPbeta an der Position des SNPs eine zusätzliche Bindestelle. Jedoch hatten diese im Bereich um den SNP und unabhängig von diesem bei beiden Varianten mehrere Bindestellen, waren also bei beiden Varianten nur verschieden oft vertreten (somit mehr bei Guanin-Variante). Die übrigen ermittelten Transkriptionsfaktoren unterschieden sich zwischen beiden Varianten nicht. Das Ergebnis des relevanten Abschnitts, der den SNP enthält, ist in Abbildung 24 dargestellt. Der über die ENCODE-Datenbank ermittelte Transkriptionsfaktor E2F1 zeigte in dieser Analyse auch eine Bindestelle, allerdings nicht unmittelbar im Bereich des SNP rs11618521. Dessen Bindestelle befand sich beim Online-Tool „PROMO“ etwa 350bp 5' des SNPs und somit im Bereich von Position -1440bp. Der SNP scheint nach dieser Datenbank keinen Einfluss auf die Bindung von E2F1 zu haben.

Zur Klärung, ob die identifizierten Transkriptionsfaktoren eine Relevanz auch in Bezug auf die SCA3 haben können, wurde mithilfe des Human Protein Atlas untersucht, ob die gefundenen Transkriptionsfaktoren im Gehirn exprimiert werden. Hierbei zeigte sich, dass sowohl C/EBPalpha als auch C/EBPbeta im Gehirn exprimiert werden und dort auch im Kleinhirn. Die Transkriptionsfaktoren GR-beta und FOXP3 werden ebenfalls im Gehirn exprimiert, wiederum insbesondere auch im Kleinhirn. Die Transkriptionsfaktoren HOXD9, HOXD10 und HNF-1C werden nach dieser Datenbank nicht im Gehirn exprimiert, haben deswegen vermutlich weniger Relevanz in Bezug auf die SCA3. Der Transkriptionsfaktor E2F1, der bei Durchsicht des ENCODE-Datensatzes auffiel, wird laut dem Human Protein Atlas ebenfalls im Gehirn exprimiert, da er jedoch scheinbar unabhängig vom SNP bindet wurde er im Folgenden nicht weiter untersucht.

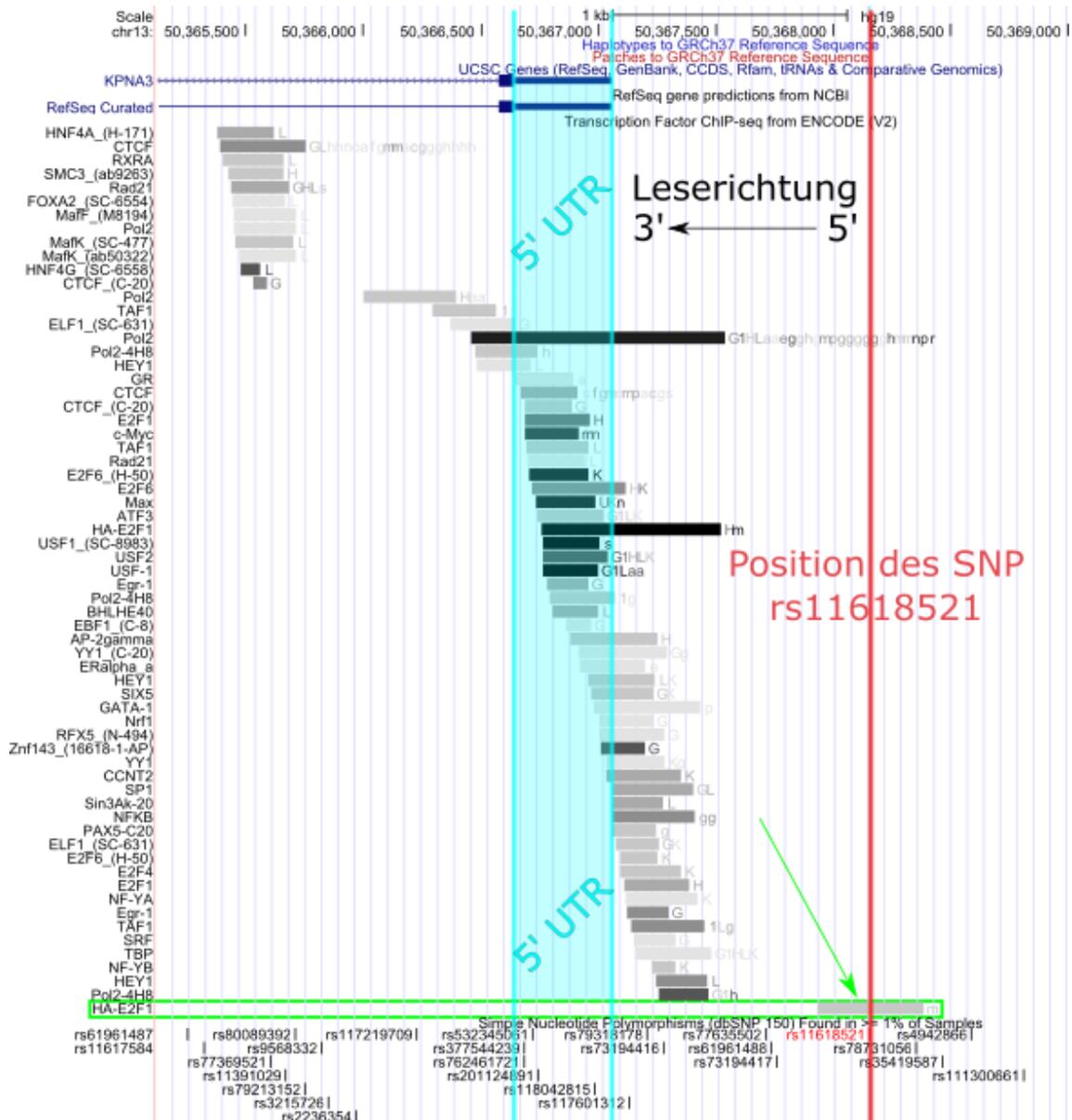
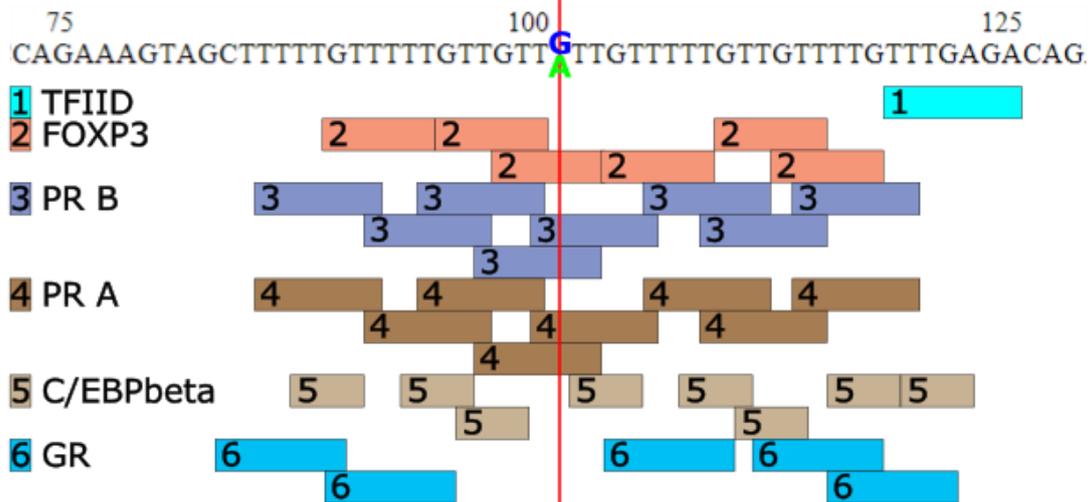
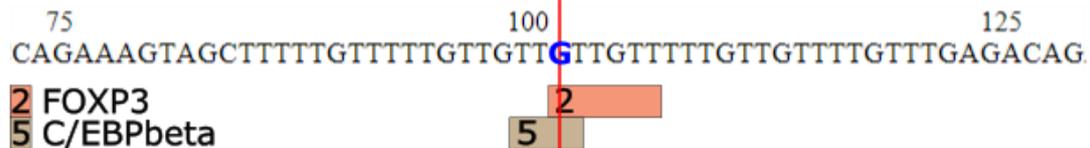


Abbildung 23 – Genome Browser mit ENCODE V2-Datensatz – In der Region, in der sich der SNP11618521 befindet (rot markiert), scheint der Transkriptionsfaktor HA-E2F1, der hier sehr zentral zu liegen kommt, (grün markiert und grüner Pfeil) einen Bindungsbereich zu haben. Die anderen hier dargestellten Transkriptionsfaktoren binden erst nach dem SNP in 3' Richtung hin zum KPNA3-Gen. Ein großer Anteil davon scheint auch im Bereich des 5' UTR (hellblau hinterlegt) zu binden. Bearbeiteter Screenshot von UCSC Genome Browser, <https://genome.ucsc.edu>, Assembly: Human Feb. 2009 (GRCh37/hg19), Region: chr13:50,365,120-50,370,000, mit Datensatz „Transcription Factor CHIP-seq from ENCODE (V2)“ (Kent et al. 2002).

## Bindestellen unabhängig vom Status des SNP rs11618521:



## Zusätzliche Bindestellen bei SNP rs11618521 mit Guanin:



## Zusätzliche Bindestellen bei SNP rs11618521 mit Adenin:

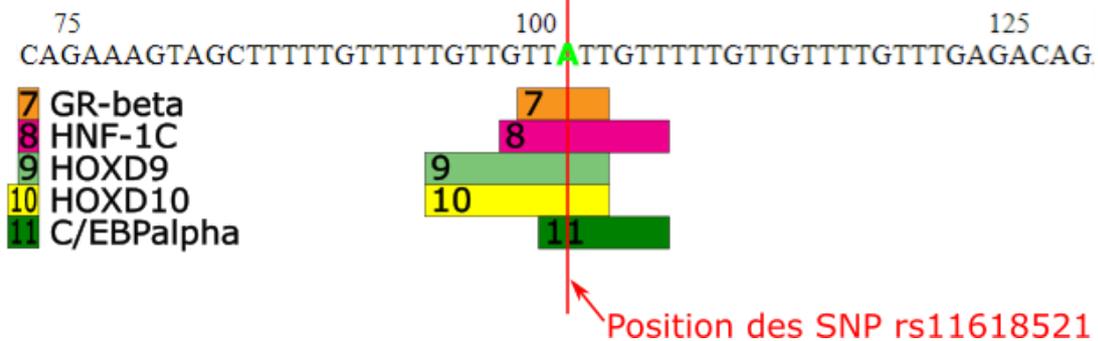


Abbildung 24 – Ergebnis der Analyse mit dem PROMO Online-Tool – Eine Sequenz den Bereichs 100bp vor und nach dem SNP rs11618521 beinhaltend wurde mit dem PROMO Online-Tool auf Transkriptionsfaktor-Bindestellen untersucht (Messeguer et al. 2002; Farre et al. 2003) und die beiden Varianten des SNPs dahingehend verglichen. Hier ist ein jeweils ein Ausschnitt um den Bereich des SNPs dargestellt, dessen Position ist in rot markiert. Oben sind die Transkriptionsfaktoren dargestellt, die unabhängig von der jeweiligen SNP-Variante eine Bindestelle in diesem Bereich besitzen (TFIID, FOXP3, PR A, PR B, C/EBPbeta, GR). Darunter jeweils sind jeweils die Transkriptionsfaktoren dargestellt, die je nach Variante des SNPs mit Guanin (Mitte) beziehungsweise Adenin (unten) binden. Im Vergleich der beiden Varianten zeigt sich, dass in der Variante mit Guanin zusätzlich zwei Transkriptionsfaktoren unmittelbar an Position des SNPs Bindungsstellen besitzen, nämlich FOXP3 und C/EBPbeta. Diese besitzen allerdings auch im Bereich um den SNP herum und unabhängig von diesem mehrere Bindungsstellen (siehe oben). Bei der SNP-Variante mit Adenin binden diese beiden an der Position des SNPs nicht, jedoch binden fünf Transkriptionsfaktoren zusätzlich, die bei der Variante mit Guanin keine Bindestelle in diesem Bereich besitzen. Diese sind GR-beta, HNF-1C, HOXD9, HOXD10 und C/EBPalpha. Der Transkriptionsfaktor E2F1, der im ENCODE Datensatz aufgefallen war, zeigte in der Analyse mit diesem Tool auch eine Bindestelle. Allerdings befand sich diese etwa 350bp 5' des SNPs und somit ist nicht von einer Anhängigkeit seiner Bindung vom SNP auszugehen (hier nicht dargestellt). Angepasste Darstellung anhand der graphischen Auswertung des PROMO Online-Tools (Messeguer et al. 2002; Farre et al. 2003), abrufbar über [gen.lsi.upc.es](http://gen.lsi.upc.es).

Unsere Experimente zeigten, dass eine signifikante Mehrexpression bei der SNP-Variante des rs11618521 mit Adenin nur auftrat, wenn auch der 5' UTR ebenso im Konstrukt vorhanden war. Deshalb wurde im Folgenden auch der 5' UTR auf Transkriptionsfaktor-Bindestellen untersucht. Somit sollten Transkriptionsfaktoren identifiziert werden, die als potenzielle Co-Faktoren der Transkriptionsfaktoren, die bei der SNP-Variante mit Adenin binden, in Frage kommen und somit die Expressionsunterschiede erklären könnten. In der Analyse im der UCSC Genome Browser mit dem ENCODE-Datensatz zeigte sich bereits, wie in Abbildung 23 dargestellt, dass im Bereich des 5' UTR einige Transkriptionsfaktoren eine Bindestelle besitzen. Auch die anderen ENCODE-Datensätze, die im UCSC Genome Browser hinterlegt sind, wurden dahingehend begutachtet. Zusätzlich wurde auch mit dem Online-Tool „PROMO“ der gesamte 5' UTR mit denselben Voreinstellungen wie bereits der Bereich um den SNP rs11618521 untersucht. In diesem wurden 37 Transkriptionsfaktoren ermittelt, die eine potenzielle Bindestelle im 5' UTR besitzen. In Hinsicht auf oben genannte Fragestellung wurde nachgehend recherchiert, ob es bekannte Interaktionen gibt zwischen den Transkriptionsfaktoren, die nur bei der SNP-Variante mit Adenin binden, und den ermittelten Transkriptionsfaktoren, die im 5' UTR Bindestellen besitzen (sowohl laut ENCODE-Datenbanken als auch im PROMO Tool). Da die Expression letztlich Relevanz für die SCA3 haben soll, wurde zusätzlich der Fokus auf die im Gehirn exprimierten Faktoren gelegt. Es wurden somit Interaktion für GR-beta und C/EBPalpha untersucht.

Hierfür wurde die STRING-Datenbank genutzt, mit der Proteine auf bekannte und vorausgesagte Protein-Protein Interaktionen untersucht werden können. In dieser Datenbank wurden Interaktion von C/EBPalpha und GR-beta mit weiteren Proteinen gesucht, wobei in dieser Datenbank nur Daten allgemein für den Glukokortikoid-Rezeptor (dort als NR3C1 bezeichnet) und nicht für die einzelnen Isoformen (also spezifisch GR-beta) hinterlegt sind. Die dort gewonnenen Daten wurden dann mit den Transkriptionsfaktoren abgeglichen, die laut PROMO Online-Tool und den ENCODE-Datenbanken im Bereich des 5' UTR Bindestellen besitzen. Es zeigte sich dann in Zusammenschau, dass C/EBPalpha (in der Datenbank als CEBPA bezeichnet) Interaktionen aufweist mit den

Transkriptionsfaktoren Jun und EP300, die im 5' UTR Bindestellen besitzen. Für GR-beta beziehungsweise genauer für NR3C1 wurden Interaktionen mit dem Transkriptionsfaktor NCOR1 (besitzt Bindestelle im 5' UTR), sowie ebenfalls mit EP300 und Jun gefunden. Die Auswertung der STRING Datenbankergebnisse sind nochmals in Abbildung 25 graphisch dargestellt. (Anmerkung: Es sind mehrere ENCODE-Datensätze im UCSC Genome Browser hinterlegt, die sich auch in den aufgeführten Transkriptionsfaktoren unterscheiden. Deshalb sind hier genannte Transkriptionsfaktoren nicht in der Abbildung 23 zu finden). Protein-Interaktionen müssen zwar nicht funktionelle Interaktion bedeuten, könnten es aber. Synergistische Effekte der genannten Transkriptionsfaktoren, sind also durchaus denkbar und könnten die beobachteten Expressionsunterschiede erklären.

Insgesamt wurden somit die Transkriptionsfaktoren C/EBPalpha, C/EBPbeta, FOXP3 und GR-beta als möglicherweise relevant identifiziert. Sie binden jeweils spezifisch nur bei der Variante des SNP rs11618521 mit Adenin (C/EBPalpha und GR-beta) beziehungsweise zeigen bei der Variante mit Guanin im Vergleich zu Adenin mehr Bindungsstellen (C/EBPbeta und FOXP3) und weisen zugleich auch eine Expression im Gehirn auf. Somit kommen diese zur Erklärung der Expressionsunterschiede zwischen den Konstrukten pGL3-KPNA3-abEcoG und pGL3-KPNA3-abEcoA in Zusammenschau mit der Assoziation, die der SNP rs11618521 zum Erkrankungsalter der SCA3 zeigte (siehe unter 1.7), in Frage.

Wie aufgezeigt, sind insbesondere für den Transkriptionsfaktor C/EBPalpha Interaktionen mit Transkriptionsfaktoren, die im Bereich des 5' UTR des *KPNA3*-Gens binden, beschrieben. Wenn man synergistische Effekte annimmt, bietet dies eine Erklärung, weshalb die Mehrexpression bei der SNP-Variante mit Adenin nur bei zusätzlichem Vorhandensein des 5' UTR zu beobachten war. Da die Interaktions-Analyse nicht spezifisch für die Isoform GR-beta durchführbar war, sondern nur allgemein für den Glukokortikoid-Rezeptor, ist hier eine analoge Aussage nur mit Vorsicht zu tätigen.

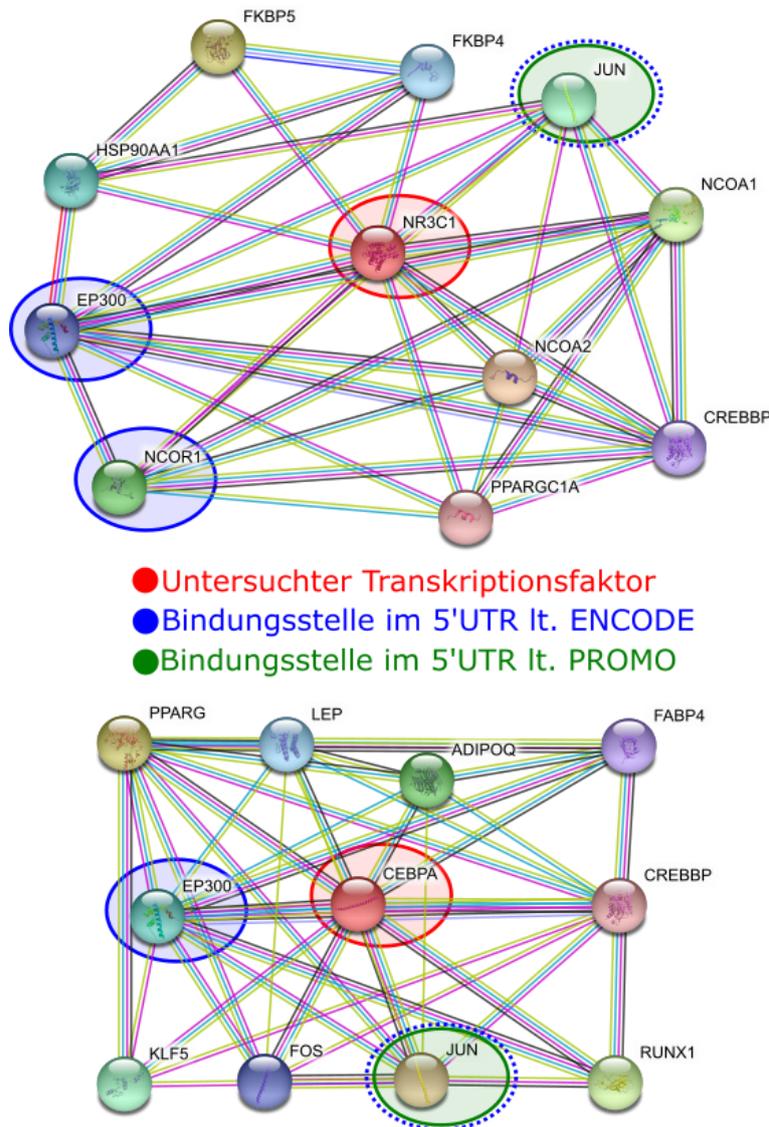


Abbildung 25 – Ergebnisse der Suche in der STRING-Datenbank – Die beiden Transkriptionsfaktoren GR-beta und C/EBPalpha, die nur bei der SNP rs11618521 Variante mit Adenin binden, wurden auf Protein-Protein Interaktionen mithilfe der STRING-Datenbank untersucht. Oben: In der STRING-Datenbank sind die Isoformen des Glukokortikoid-Rezeptors, zu denen auch GR-beta gehört, nicht einzeln aufgeschlüsselt, sondern nur allgemein in Form des NR3C1 (andere Bezeichnung für den Glukokortikoid-Rezeptor) zu finden. NR3C1 zeigt jedoch Interaktionen mit den Transkriptionsfaktoren EP300, NCOR1 und Jun, die laut PROMO Online-Tool bzw. der ENCODE-Datenbank Bindungsstellen im Bereich des 5' UTR des KPNA3-Gens besitzen. Unten: Für C/EBPalpha (in dieser Datenbank mit CEBPA bezeichnet) wurde ebenfalls nach Protein-Protein Interaktionen gesucht. Hier wurden Interaktionen mit den Transkriptionsfaktoren EP300 und Jun gefunden, die jeweils eine Bindestelle laut PROMO Online-Tool bzw. der ENCODE-Datenbank im Bereich des 5' UTR des KPNA3-Gens besitzen. Diese Interaktionen bieten eine mögliche Erklärung, weshalb die Expression bei den Konstrukten mit der Variante des SNPs rs11618521 mit Adenin in Kombination mit dem Vorhandensein des 5' UTR signifikant erhöht ist im Vergleich zu den Konstrukten mit dem SNP mit Adenin, denen jedoch der 5' UTR fehlt. – In rot ist jeweils das Protein markiert, von dem die Suche ausging. Die Transkriptionsfaktoren, die laut den ENCODE-Datenbanken im Bereich des 5' UTR des KPNA3-Gens binden, sind blau markiert (Anmerkung: Es sind mehrere ENCODE-Datensätze im UCSC Genome Browser hinterlegt, die sich auch in den aufgeführten Transkriptionsfaktoren unterscheiden. Deshalb sind hier genannte Transkriptionsfaktoren nicht in der Abbildung 23 zu finden). Der Transkriptionsfaktor, der laut dem PROMO Online-Tool im Bereich des 5' UTR des KPNA3-Gens bindet, ist grün markiert. Der Transkriptionsfaktor Jun ist zusätzlich blau gestrichelt markiert, da laut ENCODE-Datenbank der Transkriptionsfaktor JunD eine Bindestelle im 5'UTR des KPNA3-Gens besitzt, der wohl zur Jun-Familie gehört, aber nicht explizit in dieser Analyse mit STRING auftaucht. – Graphische Auswertung einer Suche in der STRING-Datenbank (Szklarczyk et al. 2019), abrufbar unter [www.string-db.org](http://www.string-db.org).

Zusammenfassend wurden in dieser Arbeit fünf Konstrukte erzeugt, die jeweils einer Variante der Promotorregion des *KPNA3*-Gens enthielten. Sie unterschieden sich in ihrer Gesamtlänge, der Variante des SNPs rs11618521 und dem Vorhandensein beziehungsweise dem Fehlen des 5' UTR des *KPNA3*-Gens. Ihre Sequenzen wurden mittels Sequenzierung bestätigt. Die Expression der Konstrukte wurde nachfolgend mit Western-Blots und Luciferase-Assays verglichen. Hier zeigte sich eine signifikant höhere Expression durch das Konstrukt mit dem SNP rs11618521 in der Version mit Adenin, bei dem zusätzlich der 5' UTR enthalten war, im Vergleich zum Konstrukt, das den SNP in der Version mit Guanin, aber sonst die identische Sequenz (inklusive 5' UTR) enthielt. Wenn der 5' UTR nicht enthalten war, war ein solcher Effekt zwischen den SNP-Varianten nicht festzustellen. Als mögliche Erklärung dieser beobachteten Unterschiede wurden Transkriptionsfaktoren identifiziert, deren Bindung Abhängigkeit von der SNP-Variante zeigte. Es wurden hierbei Transkriptionsfaktoren gefunden, die nur bei der SNP-Variante mit Adenin banden. Nachfolgend wurde nach Interaktionen dieser Transkriptionsfaktoren mit anderen Faktoren, die im Bereich des 5' UTR binden, gesucht. Hierbei zeigten sich vor allem C/EBPalpha mit seinen potenziellen Interaktionspartnern EP300 und Jun als vielversprechende Kandidaten zur Erklärung der beobachteten Expressionsunterschiede.

#### 4. Diskussion

Ziel dieser Arbeit war die Untersuchung der Promotorregion des KPNA3-Transportproteins. Zunächst wurde die Promotorregion des *KPNA3*-Gens in einen Vektor, der einen für die Firefly-Luciferase kodierenden Bereich enthält, inkloniert. Die Expression des *Luciferase*-Gens hing somit von der Aktivität des Promotors ab. Das Vorgehen mittels präparativen Verdaus, Ligation mit T4-Ligasen und Transformation in *Escherichia coli* Bakterien stellte hierbei ein seit Jahrzehnten bewährtes Verfahren dar (Bolivar und Backman 1979). Das erhaltene Konstrukt wurde mit pGL3-KPNA3-clone I bezeichnet und enthielt etwa 3500bp der Promotorregion, inklusive des nicht-translatierten Bereichs am 5' Ende des ersten Exons des *KPNA3*-Gens. Der Polymorphismus (SNP) rs11618521 lag in der Version mit Guanin vor. Dieses Konstrukt wurde als Ausgangspunkt für die weiteren Experimente genutzt. Aus diesem Konstrukt wurden durch Klonierung weitere Konstrukte erzeugt, die am 5'-Ende des Promotors etwa 1000bp kürzer waren. Wir wählten diesen Schritt, da wir davon ausgingen, dass auch eine Promotorregion kürzerer Länge zur Expression ausreicht. Bei Durchsicht des ENCODE-Datensatzes (Rosenbloom *et al.* 2013) im UCSC Genome Browser unter <https://genome.ucsc.edu> (Kent *et al.* 2002) wurde diese Annahme bestärkt, da die dort identifizierten Transkriptionsfaktoren in einem Bereich der nicht mehr als 2000bp vom Transkriptionsstart des *KPNA3*-Gens entfernt ist binden (vgl. Abbildung 23). Die kürzeren Konstrukte enthielten noch ca. 2500bp der Promotorregion, inklusive des oben genannten Exons 1. Es wurden zwei dieser Konstrukte erzeugt, sodass der enthaltene SNP rs11618521 in seinen beiden Varianten vorlag, und diese mit pGL3-KPNA3-abEcoG (SNP rs11618521 enthielt Guanin) beziehungsweise pGL3-KPNA3-abEcoA (SNP rs11618521 enthielt Adenin). Aus diesen Konstrukten wurden wiederum zwei weitere Konstrukte erzeugt, in denen der nicht-translatierte Bereich am 5' Ende des *KPNA3*-Gens (5' UTR) nicht mehr enthalten war, die aber jeweils den SNP rs11618521 in einer der beiden Varianten enthielten. Durch diese Konstrukte sollte überprüft werden, ob dieser nicht-translatierte Bereich eventuell Einfluss auf die Expression hat. Solche Phänomene durch 5' UTRs sind bereits für die Regulation der Translation beschrieben (van der Velden und Thomas 1999). So besitzt etwa der 5' UTR des

Gens, das für das Protein bcl-2 kodiert, regulatorische Funktion für dessen Expression durch den vorgeschalteten Promotor (Young und Korsmeyer 1993). Somit war die Annahme, dass der 5' UTR des *KPNA3*-Gens ebenfalls Einfluss auf dessen Expression hat, durchaus sinnvoll zu überprüfen. Die Konstrukte ohne 5' UTR wurden mit pGL3-KPNA3-noUTR-G beziehungsweise pGL3-KPNA3-noUTR-A bezeichnet, abhängig von der Variante des SNP. Alle erzeugten Konstrukte wurden mittels Sequenzierung untersucht. Hierbei wurde sichergestellt, dass diese sich nur in den gewünschten Bereichen am 5'-Ende, durch Vorhandensein oder Fehlen des 5' UTR, beziehungsweise in der jeweiligen Variante des Polymorphismus unterschieden, die restliche Sequenz bei der Klonierung jedoch unverändert blieb und sich zwischen den einzelnen Konstrukten nicht unterschied. Es kann also davon ausgegangen werden, dass sich die Unterschiede, die sich in der Expression der Konstrukte zeigten, ausschließlich durch die ausgewählten Variationen der Sequenz zustande kamen. Das genutzte Verfahren zur Sequenzierung, das auf der Sanger-Sequenzierung basiert, kann als sehr zuverlässig mit niedriger Fehlerrate bezeichnet werden (Metzker 2005), weshalb davon ausgegangen werden kann, dass die Ergebnisse der Sequenzierungen korrekt sind.

Die Aktivität der jeweiligen Variante des Promotors der verschiedenen Konstrukte wurde mittels Western-Blots und Luciferase-Assays untersucht. Im Western-Blot zeigte sich zunächst kein signifikanter Unterschied zwischen den Konstrukten, jedoch ein Trend, dass das Konstrukt pGL3-KPNA3-abEcoA mehr Expression aufweist. Im Luciferase-Assay zeigte sich, dass sich das Konstrukt, das den SNP rs11618521 in der Variante Adenin enthielt, signifikant ( $p=0,0056$ ) von der Variante des Konstrukts mit dem SNP mit Guanin unterschied und die Expression bei dem Konstrukt, das im SNP Adenin enthielt, um etwa 70% höher war. Bei den Konstrukten ohne den 5' UTR des Exon 1 wurde kein Unterschied zwischen den beiden SNP-Varianten gefunden. Weiterhin zeigte sich bei den Konstrukten, die jeweils den SNP mit Adenin enthielten, eine um ca. 60% signifikant ( $p=0,0245$ ) höhere Expression, wenn der 5' UTR im Konstrukt enthalten war. Dies weist darauf hin, dass der 5' UTR im Exon 1 des *KPNA3*-Gens tatsächlich regulatorische Eigenschaften besitzt, diese sich aber vornehmlich bei der SNP-

Variante mit Adenin auswirkt. Dass ein 5' UTR regulatorische Aktivitäten zeigen kann ist wie oben geschildert bereits beschrieben. Als mögliche Erklärung für solche Effekte nannten beispielsweise Bugaut und Balasubramanian sogenannte „5' UTR RNA G-quadrupels“, bei denen es sich um Sekundärstrukturen der mRNA handelt, die sich in Guanin-reichen mRNA-Strängen ausbilden können. Diese haben laut ihren Untersuchungen Auswirkungen auf die Expression gewisserer Proteine durch Beeinflussung der Translation und können die Expression verstärken oder abschwächen (Bugaut und Balasubramanian 2012). Tatsächlich besitzt der 5' UTR des *KPNA3*-Gens einen hohen GC-Anteil von 73%, weshalb hier durchaus vorstellbar ist, dass dieses Phänomen auftreten kann. Jedoch scheint das *KPNA3*-Protein noch nicht auf solche Effekte untersucht worden zu sein, weshalb offenbleibt, ob die im Luciferase-Assay beobachteten Expressions-Unterschiede auf 5' UTR RNA G-quadrupels zurückzuführen sind.

Das Ausgangskonstrukt in dieser Arbeit enthielt ca. 3500bp Promotorbereich des *KPNA3*-Gens inklusive des 5' UTR des Gens. Es beinhaltet somit einen Bereich bis ca. 3000bp 5' des Transkriptionsstart des *KPNA3*-Gens. Es diente vor allem als Ausgangs- und Referenzwert für die Untersuchung der anderen Konstrukte. Es zeigte keinen signifikanten Expressionsunterschied zu den Konstrukten, die den SNP rs11618521 in derselben Variante mit Guanin enthielten und 5' ca. 1000bp kürzer waren, beziehungsweise zusätzlich den 5' UTR nicht enthielten. Die zusätzlichen 1000bp am 5'-Ende haben somit wahrscheinlich keinen Einfluss auf die Expression des nachgeschalteten Gens. In der Literatur wird keine maximale Länge eines Promotors ausgehend vom Transkriptionsstart angegeben, es gibt jedoch Hinweise, dass vor allem die ersten 1000-2000bp vor dem Transkriptionsstart eines Gens relevant sein könnten (Pedersen *et al.* 1999). Dies deckt sich in Bezug auf das *KPNA3*-Gen mit den Daten der ENCODE-Datenbank im UCSC Genome Browser (Encode Project Consortium 2011; Rosenbloom *et al.* 2013). Hier binden alle identifizierten Transkriptionsfaktoren in einem Bereich, der sich weniger als 2000bp 5' des Transkriptionsstart befindet (vgl. hierzu auch Abbildung 23 unter 3.4). Dies unterstützt die These, dass die zusätzlichen ca. 1000bp, die das Ausgangskonstrukt im Vergleich zu den

anderen Konstrukten am 5' Ende besitzt, keinen Einfluss auf die Genexpression haben. Die beiden SNPs rs9535343 und rs12868235, die in vorherigen Experimenten eine Assoziation zum Erkrankungsalter der SCA3 zeigten (Köbsch 2020), befinden sich ebenfalls in diesem ca. 1000bp langem Bereich, was ihren Einfluss auf die Expression des nachgeschalteten Gens unwahrscheinlicher macht. Jedoch wurde ihr Einfluss auf die Expression hier nicht gezielt untersucht und eine zuverlässige Aussage ist somit nicht möglich.

In den hier durchgeführten Experimenten wurde der SNP rs11618521 im Promotor des *KPNA3*-Gens als maßgeblicher Regulator der Expressionsstärke identifiziert. Der SNP rs11618521 ist bisher nicht in Zusammenhang mit veränderter *KPNA3*-Expression untersucht worden. Jedoch ist für andere SNPs beschrieben, dass sie maßgeblich die Expression beeinflussen können. So konnten etwa Wang *et al.*, ebenfalls durch Untersuchungen mit Luciferase-Assays, nachweisen, dass im Promotor der Metalloproteinase MMP8 gewisse SNPs die Expression um das bis zu Dreifache verändern können (Wang *et al.* 2004). Dies bestärkt die Annahme, dass tatsächlich der SNP der entscheidende Faktor bei den beobachteten Unterschieden ist.

In dieser Arbeit wurden potenzielle Transkriptionsfaktoren, die die Expressionsunterschiede durch den SNP rs11618521 erklären könnten, identifiziert. Dass SNPs prinzipiell die Bindung von Transkriptionsfaktoren beeinflussen können, wurde bereits für den Transkriptionsfaktor OCT-1 im Zusammenhang mit SNPs im Promotor des Tumornekrosefaktors alpha nachgewiesen (Hohjoh und Tokunaga 2001).

Als relevante Transkriptionsfaktoren in Bezug auf die SCA3 wurden C/EBPalpha, C/EBPbeta, FOXP3 und GR-beta identifiziert. C/EBPalpha ist bisher vor allem in Bezug auf Stoffwechselforgänge im Fettgewebe, wo es die Expression des Leptins beeinflusst (Miller *et al.* 1996), und in Zusammenhang mit dem Energiestoffwechsel (Wang *et al.* 1995) beschrieben. Auch in Zusammenhang mit der Hämatopoese ist C/EBPalpha aufgefallen, unter anderem in Bezug auf die Akute myeloische Leukämie. Hierbei wurde unter anderem das Molekül ICCB280 identifiziert, das die Aktivität von C/EBPalpha induziert (Takei und

Kobayashi 2019). Sollte sich C/EBPalpha als entscheidender Transkriptionsfaktor bei der SNP rs11618521 Variante mit Adenin erweisen, wäre es umgekehrt denkbar bei SCA3-Patienten mit dieser SNP-Variante die Aktivität von C/EBPalpha medikamentös zu reduzieren, was zu weniger KPNA3-Expression und somit weniger Ataxin-3 im Zellkern führen würde.

C/EBPbeta ist ein in vielen Geweben vorkommender Transkriptionsfaktor, der eine Rolle bei Apoptose, Differenzierung und Seneszenz von Zellen zu spielen scheint. Durch seine Funktion wird es bereits als potenzielle Zielstruktur für epitheliale Karzinome diskutiert (Zahnow 2009). Ebenso soll C/EBPbeta bei der Expression von pro-inflammatorischen Zytokinen involviert sein. Seine Aktivität kann außerdem durch Quercetin vermindert werden (Galleggiante *et al.* 2019).

Bei GR-beta handelt es sich um eine Isoform des Glukokortikoid-Rezeptors, die in vielen Geweben im Körper, und wie bereits unter 3.4 vermerkt auch im Gehirn, exprimiert wird. GR-beta wird im Verhältnis zu der alpha-Isoform des Glukokortikoid-Rezeptors deutlich weniger exprimiert und besitzt nur bei Vorhandensein der alpha-Isoform regulatorische Aktivität. Diese Aktivität besteht darin, die Hormon-induzierte Aktivität des GR-alphas zu hemmen (Oakley *et al.* 1996). GR-beta wird deshalb als möglicher Faktor bei der Glukokortikoid-Resistenz bei Patienten diskutiert (Lewis-Tuffin und Cidlowski 2006). Wenn man diese Erkenntnisse beachtet, scheint es unwahrscheinlich, dass die hier beobachtete gesteigerte Expression bei der SNP rs11618521 Variante mit Adenin durch den Transkriptionsfaktor GR-beta erklärbar ist, da dieser laut Literatur eher einen negativen Effekt auf die Expression besitzt und zusätzlich, wie unter 3.4 erwähnt, auch unabhängig vom SNP einige Bindestelle im Bereich des Promotors besitzt.

FOXP3 besitzt je nach Gen, für das es als Transkriptionsfaktor fungiert, regulatorische Funktionen, die mehr oder weniger Expression des jeweiligen Gens bewirken (Zheng *et al.* 2007). FOXP3 zeigte bisher vor allem als Faktor bei der Differenzierung von regulatorischen T-Zellen Relevanz (Hori *et al.* 2003). Weiterhin ist er als Faktor bei der Entstehung und in Hinsicht auf die Prognose von verschiedenen Krebsformen beschrieben (Martin *et al.* 2010). Auch in Bezug

auf die Entstehung von Autoimmun-Erkrankungen soll FOXP3 eine Rolle spielen (Sakaguchi *et al.* 2006). Da FOXP3 in der Literatur vornehmlich in Zusammenhang mit der Expression in T-Zellen auftaucht, ist eine Relevanz für die SCA3, die sich vornehmlich in Kleinhirnzellen abspielt, eher unwahrscheinlich. Außerdem zeigt FOXP3, analog zu GR-beta, auch unabhängig vom SNP rs11618521 einige Bindestelle im Bereich des Promotors (siehe unter 3.4) und erklärt somit letztlich nicht ausreichend die Expressionsunterschiede zwischen den beiden SNP-Varianten.

Auffällig war in dieser Arbeit, dass sich die signifikante Mehrexpression bei der SNP-Variante des rs11618521 mit Adenin nur zeigte, wenn zusätzlich der 5' UTR in den Konstrukten enthalten war. Die Mehrexpression bei Vorhandensein des 5' UTR könnte durch synergistische Effekte erklärbar sein, die die Transkriptionsfaktoren GR-beta und C/EBPalpha, die nur in der SNP-Variante mit Adenin eine Bindungsstelle besitzen, mit Transkriptionsfaktoren, die im Bereich des 5' UTR binden, ausbilden. Da GR-beta, wie oben erwähnt, eher einen negativen Effekt auf die Expression besitzt, sind vor allem die Interaktionspartner von C/EBPalpha von Interesse. Für C/EBPalpha wurden die Proteine Jun und EP300 als Interaktionspartner identifiziert, die jeweils eine Bindestelle im 5' UTR des *KPNA3*-Gens besitzen. Bei EP300 handelt es sich nicht um einen Transkriptionsfaktor im eigentlichen Sinne, sondern um die Histon Acetyltransferase p300 (Ogryzko *et al.* 1996). Dennoch hat EP300 regulatorisch Funktion als Co-Aktivator von Transkriptionsfaktoren (Hussain und Habener 1999; Teufel *et al.* 2007). Jun ist ein Bestandteil des Transkriptionsfaktors AP-1 und spielt eine Rolle bei der Regulation des Zellzyklus (Wisdom *et al.* 1999). Jun wird als ein Faktor bei der Entstehung verschiedener Tumoren beschrieben (Szabo *et al.* 1996; Eferl *et al.* 2003). Weiterhin ist Jun auch bei der Regeneration beschädigter Axone von Nervenzellen beteiligt (Mahar und Cavalli 2018). Die gefundene Protein-Protein-Interaktion ist zwar nicht gleichbedeutend mit einer funktionellen Interaktion, jedoch wurde EP300 bereits als Co-Aktivator von C/EBPalpha beschrieben (Erickson *et al.* 2001). Ebenso wurde auch Jun als Co-Faktor von C/EBPalpha identifiziert und soll weiterhin die Fähigkeit besitzen, in Kombination mit C/EBPalpha die Promotoraktivität zu steigern (Cai *et al.* 2008;

Hong *et al.* 2011). In Zusammenschau aller Erkenntnisse kommt somit der Transkriptionsfaktor C/EBPalpha, in Interaktion mit EP300 oder Jun als potenzielle Co-Faktoren, zur Erklärung der ermittelten Expressionsunterschiede zwischen den SNP-Varianten bei Vorhandensein des 5' UTR als vielversprechendste Erklärung in Betracht.

Um zu verifizieren, ob die identifizierten Transkriptionsfaktoren tatsächlich relevant im Bereich des SNPs binden und ob in diesem Fall die beiden Varianten einen Einfluss auf die Bindung der jeweiligen Transkriptionsfaktoren haben, würde es sich empfehlen diese Region und die oben genannten Transkriptionsfaktoren beispielsweise mittels ChIP-seq zu untersuchen. ChIP-seq ermöglicht die Protein-DNA-Interaktionen der Faktoren mit der jeweiligen Variante der DNA-Region zu untersuchen (Furey 2012) und somit den Einfluss des SNPs auf die Bindung der Transkriptionsfaktoren. Sollte sich hier ein vielversprechender Kandidat herauskristallisieren, wäre im Folgenden eine erneute Expressionsanalyse, beispielsweise wieder mit einem Luciferase-Assay und den in dieser Arbeit erzeugten Konstrukten, mit verschiedenen Ansätzen in Hinsicht auf die Transkriptionsfaktor-Verfügbarkeit denkbar. Hierfür können beispielweise Versuchsansätze mit siRNAs genutzt werden. Hierbei handelt es sich um kurze RNA-Stücke, die sich an eine spezifische mRNA anlagern und deren Abbau initiieren (Chu und Rana 2007). Die siRNA kann gezielt genutzt werden, um spezifische mRNA abzubauen und somit die Expression von gewünschten Genen abzuschalten (Lagana *et al.* 2015). Durch Abschalten bestimmter Gene für Transkriptionsfaktoren, beispielsweise für C/EBPalpha, der sich in dieser Arbeit als vielversprechend zeigte, und nachfolgende Expressionsanalyse, könnte untersucht werden, ob dann immer noch Expressionsunterschiede zwischen den SNP rs1161851 Varianten und den 5'UTR-Varianten nachweisbar sind. Dadurch kann dann rückgeschlossen werden, welcher Transkriptionsfaktor tatsächlich maßgeblich für die bisher gefundenen Expressionsunterschiede war.

Möglicherweise wäre es sinnvoll, die Expression der Konstrukte, denen der 5' UTR des *KPNA3*-Gens fehlt, erneut beziehungsweise mit anderen Voraussetzungen zu untersuchen, um zu sehen, ob hier doch der SNP

rs11618521 einen signifikanten Faktor für die Expression darstellt. Während der Experimente konnte bisher kein signifikanter Unterschied zwischen den beiden SNP-Varianten bei den Konstrukten ohne den 5' UTR gefunden werden. Jedoch haben die Ergebnisse bei diesen beiden Konstrukten deutlich höhere Schwankungen gezeigt im Vergleich zu den anderen Konstrukten. Eine erneute Testung mit einer höheren Anzahl an Wiederholungen könnte überprüfen, ob auch zwischen den Konstrukten ohne enthaltenen 5' UTR ein signifikanter Unterschied in der Expression durch den SNP nachweisbar ist.

Die Experimente in dieser Arbeit wurden ausschließlich in HEK-293T Zellen durchgeführt. Diese Zelllinie leitet sich von embryonalen Nierenzellen ab, besitzt jedoch ein neuronales Expressionsprofil (Shaw *et al.* 2002). Da der *KPNA3*-Promotor in Hinblick auf die Relevanz für die *SCA3* untersucht werden sollte, ist die Ähnlichkeit der Zellen zu Neuronen durchaus vorteilhaft. Bei dieser Zellreihe ist eine Dynamik ihres Genoms beschrieben (Lin *et al.* 2014), weshalb natürlich weiter zu klären ist, inwiefern die aus den Zellversuchen gewonnen Erkenntnisse auf tatsächliche Zellen des Gehirns übertragbar sind in Hinblick auf Aussagen zur *SCA3*. Außerdem wurden Transkriptionsfaktoren identifiziert, die zwar nicht im Gehirn exprimiert werden, jedoch potenziell in den HEK-293T-Zellen. Deshalb ist es im Verlauf sinnvoll, die gefundenen Ergebnisse durch Wiederholung der Experimente auch in anderen Zelllinien zu verifizieren, um auszuschließen, dass es sich nur um einen spezifischen Effekt dieser Zelllinie handelt. Weiterhin könnten auch die hier fraglichen Transkriptionsfaktoren mittels Ansätzen mit siRNA wie oben beschrieben untersucht werden, um auszuschließen, dass sich die beobachteten Effekte doch nur auf diese, nur in HEK-293T Zellen exprimierten, Transkriptionsfaktoren zurückführen lassen.

Die Untersuchung mittels Wester-Blot brachte in der gewählten Anzahl an Wiederholung keine signifikanten Ergebnisse. Sie wurde hauptsächlich durchgeführt, um die prinzipielle Expression von Luciferase in den erzeugten Konstrukten nachzuweisen, und wurde dann zur eigentlichen Messung der Expressionsstärke durch das Luciferase-Assay abgelöst. Dieser Weg wurde gewählt, da für die Testung der Expression durch einen Promotor das Luciferase-Assay ein etabliertes und einfach durchzuführendes Verfahren ist, das effizient

und sensitiv ist und sich insbesondere zur Untersuchung von Promotoren eignet (Xu *et al.* 2013). Zwar wiesen die Ergebnisse der Western-Blots in dieser Arbeit in die Richtung der Ergebnisse der Luciferase-Assays, jedoch könnte eine Wiederholung der Western-Blots, eventuell auch mit anderen Zelllinien, zur Verifizierung der Luciferase-Assays dienen, oder auch Unterschiede zwischen den Konstrukten ohne 5' UTR aufzeigen.

Es sollte weiterhin abgeklärt werden, ob der SNP rs11618521 auch *in vivo* einen denselben Effekt auf die Expression des *KPNA3*-Gens hat, etwa durch Analyse der Menge des *KPNA3*-Proteins im Gehirn bei den beiden SNP Varianten. Auch wäre es interessant abzuklären, ob diese SNP-Varianten einen Effekt auf Alter bei Krankheitsbeginn, Schwere der Symptomatik und Krankheitsverlauf bei *SCA3*-Patienten haben. Da eine niedrigere Menge von *KPNA3* zu weniger intranukleärem Ataxin-3 und Aggregation führt (Sowa *et al.* 2018), sollten theoretisch bei Patienten mit der Variante des SNP rs11618521 mit Guanin weniger Ataxin-3 und Aggregate intranukleär nachweisbar sein und diese weniger stark im Hinblick auf Symptomatik und Krankheitsverlauf betroffen sein, als Patientin mit der Variante mit Adenin.

Abgesehen von diesen Punkten kann jedoch nun auf Grundlage dieser Arbeit weiter mit dem erzeugten Modell gearbeitet werden. Bisher wurde der *KPNA3*-Promotor so in der Literatur noch nicht untersucht und das Modell kann Ansatz für weitere Untersuchungen bieten. Es empfiehlt sich zu untersuchen, ob die Expression in diesem Modell durch Substanzen, im Speziellen beispielsweise bereits zugelassene Medikamente, beeinflusst werden kann. Da das Vorhandensein des *KPNA3*-Transportproteins bei der *SCA3* eine Rolle spielt und auch bei anderen Erkrankungen Einfluss zu haben scheint, könnten hiermit mögliche neue Therapieoptionen entdeckt werden, wenn diese Substanzen in der Lage sind, die Expression des *KPNA3* Gens zu verringern. Das verminderte Vorhandensein des *KPNA3*-Transportproteins korreliert mit der Schwere der Erkrankung der *SCA3*-Patienten. Substanzen, die die Expression durch den Promotor im Modell reduzieren und somit die Menge an synthetisiertem *KPNA3*-Transportprotein *in vivo* potenziell reduzieren, stellen eine bisher nicht genutzte Therapieoption dar.

Das KPNA3 Transportprotein scheint auch in anderen Krankheiten eine Rolle zu spielen. So scheinen gewisse Variationen des KPNA3-Proteins die Entstehung einiger psychischer Erkrankungen, wie Schizophrenie, Depression und Alkoholabhängigkeit, zu begünstigen (Morris *et al.* 2012). Weiterhin scheint in der B-Zell chronisch lymphatischen Leukämie (B-CLL) das KPNA3-Protein herunterreguliert zu sein (Mertens *et al.* 2002) und auch bei Influenza spielt das KPNA3-Protein eine Rolle, wobei es bei der Virusreplikation ein wichtiger Faktor zu sein scheint (Gabriel *et al.* 2011). Die Untersuchung mittels des in dieser Arbeit erzeugten Modells auf Beeinflussung der KPNA3-Menge auf transkriptioneller Ebene könnte Substanzen identifizieren, die die Expression von KPNA3 verändern und somit neue Ansätze zum Verständnis oder sogar zu neuen Therapien auch bei diesen und anderen Erkrankungen liefern.

Zusammengefasst ist eine nähere Untersuchung des Promotors, insbesondere in Hinsicht auf die relevanten Transkriptionsfaktoren sinnvoll, um den Effekt den der SNP rs11618521 auf die Expression besser zu verstehen. Unabhängig davon ermöglicht das in dieser Arbeit erzeugte Modell die nähere Untersuchung des KPNA3-Promotors und dessen Beeinflussbarkeit, einerseits spezifisch um neue Therapietargets für die SCA3 zu identifizieren und andererseits auch als potentielle Struktur zur Untersuchung und Therapiefindung in weiteren Erkrankungen, in denen das KPNA3-Protein eine Rolle spielt.

### 5. Zusammenfassung

Die Spinozerebelläre Ataxie Typ 3 (SCA3) ist eine neurodegenerative Erkrankung, die durch die vermehrte Wiederholung eines CAG-Basentriplets im *ATXN3*-Gen verursacht wird. Hierdurch wird das darin kodierte Protein Ataxin-3 fehlerhaft gebildet, wodurch das Protein im Zellkern akkumuliert und Aggregate bildet. Dies beeinflusst verschiedene Stoffwechselfvorgänge in der Zelle. Letztendlich kommt es zum Untergang von Neuronen vor allem im Bereich des Kleinhirns, mit einer Vielzahl von teils ausgeprägten Symptomen bei betroffenen Patienten. Die Lokalisation des Ataxin-3 im Zellkern ist ein wesentlicher Faktor für die Krankheitsmanifestation. Das KPNA3-Transportprotein spielt beim Transport des Ataxin-3 in den Zellkern eine entscheidende Rolle. Bei Verminderung oder Fehlen von KPNA3 in der Zelle wurde *in vitro* und *in vivo* eine Verringerung der Menge an Ataxin-3 im Zellkern und auch weniger Aggregation desselben im Zellkern beobachtet. KPNA3 wird somit als ein Schlüsselprotein bei der Pathogenese der SCA3 gesehen. In dieser Arbeit wurde ein Modell erzeugt, mit dem die KPNA3 Expression durch den KPNA3-Promotor untersucht werden kann. Es wurden Konstrukte des Promotorbereichs generiert, die sich in Gesamtlänge, Variante des enthaltenen SNP rs11618521 und Vorhandensein des 5' UTR des *KPNA3*-Gens unterschieden. Bei der Expressionsanalyse der Konstrukte mittels Western-Blots und Luciferase-Assays zeigte sich, dass der SNP rs11618521 einen signifikanten Einfluss auf die Expression durch den Promotor hat, sofern der 5' UTR des *KPNA3*-Gens ebenfalls im Konstrukt enthalten ist. Es wurden Transkriptionsfaktoren identifiziert, die eine Erklärung für die beobachteten Expressionsunterschiede liefern könnten. Der Transkriptionsfaktor C/EBPalpha stellte sich, zusammen mit den potenziellen Co-Faktoren Jun und EP300, als vielversprechend in dieser Hinsicht heraus. Insgesamt kann nun mit diesem Modell weitergearbeitet werden, um etwa Substanzen zu identifizieren, die die Menge an KPNA3-Protein in der Zelle über eine verminderte Expression reduzieren, als neue Therapiemöglichkeiten der SCA3. Auch bei anderen Erkrankungen, bei denen KPNA3 in der Pathogenese als Faktor diskutiert wird, kann dieses Modell zum weiteren Krankheitsverständnis eingesetzt werden und zur Suche nach neuen Therapiemöglichkeiten dienen.

6. Literaturverzeichnis

- Albrecht, M., Golatta, M., Wullner, U. und Lengauer, T. (2004) 'Structural and functional analysis of ataxin-2 and ataxin-3', *Eur J Biochem*, 271(15), 3155-70.
- Allen, T.D., Cronshaw, J.M., Bagley, S., Kiseleva, E. und Goldberg, M.W. (2000) 'The nuclear pore complex: mediator of translocation between nucleus and cytoplasm', *J Cell Sci*, 113 ( Pt 10), 1651-9.
- Antony, P.M., Mantele, S., Mollenkopf, P., Boy, J., Kehlenbach, R.H., Riess, O. und Schmidt, T. (2009) 'Identification and functional dissection of localization signals within ataxin-3', *Neurobiol Dis*, 36(2), 280-92.
- Araujo, J., Breuer, P., Dieringer, S., Krauss, S., Dorn, S., Zimmermann, K., Pfeifer, A., Klockgether, T., Wuellner, U. und Evert, B.O. (2011) 'FOXO4-dependent upregulation of superoxide dismutase-2 in response to oxidative stress is impaired in spinocerebellar ataxia type 3', *Hum Mol Genet*, 20(15), 2928-41.
- Berke, S.J., Schmied, F.A., Brunt, E.R., Ellerby, L.M. und Paulson, H.L. (2004) 'Caspase-mediated proteolysis of the polyglutamine disease protein ataxin-3', *J Neurochem*, 89(4), 908-18.
- Bichelmeier, U., Schmidt, T., Hübener, J., Boy, J., Rüttiger, L., Häbig, K., Poths, S., Bonin, M., Knipper, M., Schmidt, W.J., Wilbertz, J., Wolburg, H., Laccone, F. und Riess, O. (2007) 'Nuclear localization of ataxin-3 is required for the manifestation of symptoms in SCA3: in vivo evidence', *J Neurosci*, 27(28), 7418-28.
- Bolivar, F. und Backman, K. (1979) 'Plasmids of Escherichia coli as cloning vectors', *Methods Enzymol*, 68, 245-67.
- Bugaut, A. und Balasubramanian, S. (2012) '5'-UTR RNA G-quadruplexes: translation regulation and targeting', *Nucleic Acids Res*, 40(11), 4727-41.
- Cai, D.H., Wang, D., Keefer, J., Yeaman, C., Hensley, K. und Friedman, A.D. (2008) 'C/EBP alpha:AP-1 leucine zipper heterodimers bind novel DNA elements, activate the PU.1 promoter and direct monocyte lineage commitment more potently than C/EBP alpha homodimers or AP-1', *Oncogene*, 27(19), 2772-9.
- Chou, A.H., Yeh, T.H., Kuo, Y.L., Kao, Y.C., Jou, M.J., Hsu, C.Y., Tsai, S.R., Kakizuka, A. und Wang, H.L. (2006) 'Polyglutamine-expanded ataxin-3 activates mitochondrial apoptotic pathway by upregulating Bax and downregulating Bcl-xL', *Neurobiol Dis*, 21(2), 333-45.
- Chou, A.H., Yeh, T.H., Ouyang, P., Chen, Y.L., Chen, S.Y. und Wang, H.L. (2008) 'Polyglutamine-expanded ataxin-3 causes cerebellar dysfunction of SCA3 transgenic mice by inducing transcriptional dysregulation', *Neurobiol Dis*, 31(1), 89-101.

- Chu, C.Y. und Rana, T.M. (2007) 'Small RNAs: regulators and guardians of the genome', *J Cell Physiol*, 213(2), 412-9.
- Csardi, G., Franks, A., Choi, D.S., Airoidi, E.M. und Drummond, D.A. (2015) 'Accounting for experimental noise reveals that mRNA levels, amplified by post-transcriptional processes, largely determine steady-state protein levels in yeast', *PLoS Genet*, 11(5), e1005206.
- Da Silva, J.D., Teixeira-Castro, A. und Maciel, P. (2019) 'From Pathogenesis to Novel Therapeutics for Spinocerebellar Ataxia Type 3: Evading Potholes on the Way to Translation', *Neurotherapeutics*, 16(4), 1009-1031.
- Dieterle, M.C. (2019) 'Untersuchung des Einflusses von Polymorphismen als modifizierender Faktor der Spinozerebellären Ataxie Typ 3', unveröffentlichte Dissertation, Eberhard Karls Universität Tübingen.
- Doss-Pepe, E.W., Stenroos, E.S., Johnson, W.G. und Madura, K. (2003) 'Ataxin-3 interactions with rad23 and valosin-containing protein and its associations with ubiquitin chains and the proteasome are consistent with a role in ubiquitin-mediated proteolysis', *Mol Cell Biol*, 23(18), 6469-83.
- Dürr, A., Stevanin, G., Cancel, G., Duyckaerts, C., Abbas, N., Didierjean, O., Chneiweiss, H., Benomar, A., Lyon-Caen, O., Julien, J., Serdaru, M., Penet, C., Agid, Y. und Brice, A. (1996) 'Spinocerebellar ataxia 3 and Machado-Joseph disease: clinical, molecular, and neuropathological features', *Ann Neurol*, 39(4), 490-9.
- Eferl, R., Ricci, R., Kenner, L., Zenz, R., David, J.P., Rath, M. und Wagner, E.F. (2003) 'Liver tumor development. c-Jun antagonizes the proapoptotic activity of p53', *Cell*, 112(2), 181-92.
- Encode Project Consortium (2011) 'A user's guide to the encyclopedia of DNA elements (ENCODE)', *PLoS Biol*, 9(4), e1001046.
- Erickson, R.L., Hemati, N., Ross, S.E. und MacDougald, O.A. (2001) 'p300 coactivates the adipogenic transcription factor CCAAT/enhancer-binding protein alpha', *J Biol Chem*, 276(19), 16348-55.
- Evert, B.O., Vogt, I.R., Vieira-Saecker, A.M., Ozimek, L., de Vos, R.A., Brunt, E.R., Klockgether, T. und Wullner, U. (2003) 'Gene expression profiling in ataxin-3 expressing cell lines reveals distinct effects of normal and mutant ataxin-3', *J Neuropathol Exp Neurol*, 62(10), 1006-18.
- Farre, D., Roset, R., Huerta, M., Adsuara, J.E., Rosello, L., Alba, M.M. und Messeguer, X. (2003) 'Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN', *Nucleic Acids Res*, 31(13), 3651-3.
- Furey, T.S. (2012) 'ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions', *Nat Rev Genet*, 13(12), 840-52.

- Gabriel, G., Klingel, K., Otte, A., Thiele, S., Hudjetz, B., Arman-Kalcek, G., Sauter, M., Schmidt, T., Rother, F., Baumgarte, S., Keiner, B., Hartmann, E., Bader, M., Brownlee, G.G., Fodor, E. und Klenk, H.D. (2011) 'Differential use of importin-alpha isoforms governs cell tropism and host adaptation of influenza virus', *Nat Commun*, 2, 156.
- Galleggiante, V., De Santis, S., Liso, M., Verna, G., Sommella, E., Mastronardi, M., Campiglia, P., Chieppa, M. und Serino, G. (2019) 'Quercetin-Induced miR-369-3p Suppresses Chronic Inflammatory Response Targeting C/EBP-beta', *Mol Nutr Food Res*, 2019/07/25, e1801390.
- Globas, C., du Montcel, S.T., Baliko, L., Boesch, S., Depondt, C., DiDonato, S., Durr, A., Filla, A., Klockgether, T., Mariotti, C., Meleggh, B., Rakowicz, M., Ribai, P., Rola, R., Schmitz-Hubsch, T., Szymanski, S., Timmann, D., Van de Warrenburg, B.P., Bauer, P. und Schols, L. (2008) 'Early symptoms in spinocerebellar ataxia type 1, 2, 3, and 6', *Mov Disord*, 23(15), 2232-8.
- Gu, W., Ma, H., Wang, K., Jin, M., Zhou, Y., Liu, X., Wang, G. und Shen, Y. (2004) 'The shortest expanded allele of the MJD1 gene in a Chinese MJD kindred with autonomic dysfunction', *Eur Neurol*, 52(2), 107-11.
- Haacke, A., Broadley, S.A., Boteva, R., Tzvetkov, N., Hartl, F.U. und Breuer, P. (2006) 'Proteolytic cleavage of polyglutamine-expanded ataxin-3 is critical for aggregation and sequestration of non-expanded ataxin-3', *Hum Mol Genet*, 15(4), 555-68.
- Harmuth, T., Prell-Schicker, C., Weber, J.J., Gellerich, F., Funke, C., Driessen, S., Magg, J.C.D., Kriebehl, G., Wolburg, H., Hayer, S.N., Hauser, S., Kruger, R., Schols, L., Riess, O. und Hubener-Schmid, J. (2018) 'Mitochondrial Morphology, Function and Homeostasis Are Impaired by Expression of an N-terminal Calpain Cleavage Fragment of Ataxin-3', *Front Mol Neurosci*, 11, 368.
- Hohjoh, H. und Tokunaga, K. (2001) 'Allele-specific binding of the ubiquitous transcription factor OCT-1 to the functional single nucleotide polymorphism (SNP) sites in the tumor necrosis factor-alpha gene (TNFA) promoter', *Genes Immun*, 2(2), 105-9.
- Hong, S., Skaist, A.M., Wheelan, S.J. und Friedman, A.D. (2011) 'AP-1 protein induction during monopoiesis favors C/EBP: AP-1 heterodimers over C/EBP homodimerization and stimulates FosB transcription', *J Leukoc Biol*, 90(4), 643-51.
- Hori, S., Nomura, T. und Sakaguchi, S. (2003) 'Control of regulatory T cell development by the transcription factor Foxp3', *Science*, 299(5609), 1057-61.

- Hübener, J., Weber, J.J., Richter, C., Honold, L., Weiss, A., Murad, F., Breuer, P., Wullner, U., Bellstedt, P., Paquet-Durand, F., Takano, J., Saido, T.C., Riess, O. und Nguyen, H.P. (2013) 'Calpain-mediated ataxin-3 cleavage in the molecular pathogenesis of spinocerebellar ataxia type 3 (SCA3)', *Hum Mol Genet*, 22(3), 508-18.
- Hussain, M.A. und Habener, J.F. (1999) 'Glucagon gene transcription activation mediated by synergistic interactions of pax-6 and cdx-2 with the p300 co-activator', *J Biol Chem*, 274(41), 28950-7.
- Ilg, W., Brotz, D., Burkard, S., Giese, M.A., Schols, L. und Synofzik, M. (2010) 'Long-term effects of coordinative training in degenerative cerebellar disease', *Mov Disord*, 25(13), 2239-46.
- Ilg, W., Synofzik, M., Brotz, D., Burkard, S., Giese, M.A. und Schols, L. (2009) 'Intensive coordinative training improves motor performance in degenerative cerebellar disease', *Neurology*, 73(22), 1823-30.
- Jung, J., Xu, K., Lessing, D. und Bonini, N.M. (2009) 'Preventing Ataxin-3 protein cleavage mitigates degeneration in a Drosophila model of SCA3', *Hum Mol Genet*, 18(24), 4843-52.
- Kawaguchi, Y., Okamoto, T., Taniwaki, M., Aizawa, M., Inoue, M., Katayama, S., Kawakami, H., Nakamura, S., Nishimura, M., Akiguchi, I. *et al.* (1994) 'CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1', *Nat Genet*, 8(3), 221-8.
- Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. und Haussler, D. (2002) 'The human genome browser at UCSC', *Genome Res*, 12(6), 996-1006.
- Klockgether T. *et al.* (2018) 'Ataxien des Erwachsenenalters, S1-Leitlinie', in: Deutsche Gesellschaft für Neurologie (Hrsg.), Leitlinien für Diagnostik und Therapie in der Neurologie. Online: [www.dgn.org/leitlinien](http://www.dgn.org/leitlinien) (abgerufen am 16.08.2018).
- Köbsch, A. (2020) 'Untersuchung des Einflusses von Polymorphismen auf die Expression von Proteinen des nukleozytoplasmatischen Transports als modifizierende Faktoren der Spinozerebellären Ataxie Typ 3', unveröffentlichte Dissertation, Eberhard Karls Universität Tübingen.
- Lagana, A., Veneziano, D., Russo, F., Pulvirenti, A., Giugno, R., Croce, C.M. und Ferro, A. (2015) 'Computational design of artificial RNA molecules for gene regulation', *Methods Mol Biol*, 1269, 393-412.
- Lange, A., Mills, R.E., Lange, C.J., Stewart, M., Devine, S.E. und Corbett, A.H. (2007) 'Classical nuclear localization signals: definition, function, and interaction with importin alpha', *J Biol Chem*, 282(8), 5101-5.
- Lei, L.F., Yang, G.P., Wang, J.L., Chuang, D.M., Song, W.H., Tang, B.S. und Jiang, H. (2016) 'Safety and efficacy of valproic acid treatment in SCA3/MJD patients', *Parkinsonism Relat Disord*, 26, 55-61.

- Lewis-Tuffin, L.J. und Cidlowski, J.A. (2006) 'The physiology of human glucocorticoid receptor beta (hGRbeta) and glucocorticoid resistance', *Ann N Y Acad Sci*, 1069, 1-9.
- Li, X., Liu, H., Fischhaber, P.L. und Tang, T.S. (2015) 'Toward therapeutic targets for SCA3: Insight into the role of Machado-Joseph disease protein ataxin-3 in misfolded proteins clearance', *Prog Neurobiol*, 132, 34-58.
- Lin, Y.C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., Chen, J., Speleman, F., Lambrechts, D., Van de Peer, Y., Tavernier, J. und Callewaert, N. (2014) 'Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations', *Nat Commun*, 5, 4767.
- Liu, H., Li, X., Ning, G., Zhu, S., Ma, X., Liu, X., Liu, C., Huang, M., Schmitt, I., Wullner, U., Niu, Y., Guo, C., Wang, Q. und Tang, T.S. (2016) 'The Machado-Joseph Disease Deubiquitinase Ataxin-3 Regulates the Stability and Apoptotic Function of p53', *PLoS Biol*, 14(11), e2000733.
- Macedo-Ribeiro, S., Cortes, L., Maciel, P. und Carvalho, A.L. (2009) 'Nucleocytoplasmic shuttling activity of ataxin-3', *PLoS One*, 4(6), e5834.
- Maciel, P., Costa, M.C., Ferro, A., Rousseau, M., Santos, C.S., Gaspar, C., Barros, J., Rouleau, G.A., Coutinho, P. und Sequeiros, J. (2001) 'Improvement in the molecular diagnosis of Machado-Joseph disease', *Arch Neurol*, 58(11), 1821-7.
- Mahar, M. und Cavalli, V. (2018) 'Intrinsic mechanisms of neuronal axon regeneration', *Nat Rev Neurosci*, 19(6), 323-337.
- Mao, Y., Senic-Matuglia, F., Di Fiore, P.P., Polo, S., Hodsdon, M.E. und De Camilli, P. (2005) 'Deubiquitinating function of ataxin-3: insights from the solution structure of the Josephin domain', *Proc Natl Acad Sci U S A*, 102(36), 12700-5.
- Martin, F., Ladoire, S., Mignot, G., Apetoh, L. und Ghiringhelli, F. (2010) 'Human FOXP3 and cancer', *Oncogene*, 29(29), 4121-9.
- Mertens, D., Wolf, S., Schroeter, P., Schaffner, C., Dohner, H., Stilgenbauer, S. und Lichter, P. (2002) 'Down-regulation of candidate tumor suppressor genes within chromosome band 13q14.3 is independent of the DNA methylation pattern in B-cell chronic lymphocytic leukemia', *Blood*, 99(11), 4116-21.
- Messeguer, X., Escudero, R., Farre, D., Nunez, O., Martinez, J. und Alba, M.M. (2002) 'PROMO: detection of known transcription regulatory elements using species-tailored searches', *Bioinformatics*, 18(2), 333-4.
- Metzker, M.L. (2005) 'Emerging technologies in DNA sequencing', *Genome Res*, 15(12), 1767-76.

- Miller, S.G., De Vos, P., Guerre-Millo, M., Wong, K., Hermann, T., Staels, B., Briggs, M.R. und Auwerx, J. (1996) 'The adipocyte specific transcription factor C/EBPalpha modulates human ob gene expression', *Proc Natl Acad Sci U S A*, 93(11), 5507-11.
- Mitchell, P.J. und Tjian, R. (1989) 'Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins', *Science*, 245(4916), 371-8.
- Morris, C.P., Baune, B.T., Domschke, K., Arolt, V., Swagell, C.D., Hughes, I.P., Lawford, B.R., Mc, D.Y.R. und Voisey, J. (2012) 'KPNA3 variation is associated with schizophrenia, major depression, opiate dependence and alcohol dependence', *Dis Markers*, 33(4), 163-70.
- Netravathi, M., Pal, P.K., Purushottam, M., Thennarasu, K., Mukherjee, M. und Jain, S. (2009) 'Spinocerebellar ataxias types 1, 2 and 3: age adjusted clinical severity of disease at presentation correlates with size of CAG repeat lengths', *J Neurol Sci*, 277(1-2), 83-6.
- Oakley, R.H., Sar, M. und Cidlowski, J.A. (1996) 'The human glucocorticoid receptor beta isoform. Expression, biochemical properties, and putative function', *J Biol Chem*, 271(16), 9550-9.
- Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H. und Nakatani, Y. (1996) 'The transcriptional coactivators p300 and CBP are histone acetyltransferases', *Cell*, 87(5), 953-9.
- Paulson, H.L., Perez, M.K., Trotter, Y., Trojanowski, J.Q., Subramony, S.H., Das, S.S., Vig, P., Mandel, J.L., Fischbeck, K.H. und Pittman, R.N. (1997) 'Intranuclear inclusions of expanded polyglutamine protein in spinocerebellar ataxia type 3', *Neuron*, 19(2), 333-44.
- Pedersen, A.G., Baldi, P., Chauvin, Y. und Brunak, S. (1999) 'The biology of eukaryotic promoter prediction--a review', *Comput Chem*, 23(3-4), 191-207.
- Pemberton, L.F. und Paschal, B.M. (2005) 'Mechanisms of receptor-mediated nuclear import and nuclear export', *Traffic*, 6(3), 187-98
- Reina, C.P., Nabet, B.Y., Young, P.D. und Pittman, R.N. (2012) 'Basal and stress-induced Hsp70 are modulated by ataxin-3', *Cell Stress Chaperones*, 17(6), 729-42.
- Reina, C.P., Zhong, X. und Pittman, R.N. (2010) 'Proteotoxic stress increases nuclear localization of ataxin-3', *Hum Mol Genet*, 19(2), 235-49.
- Riess, O., Rub, U., Pastore, A., Bauer, P. und Schols, L. (2008) 'SCA3: neurological features, pathogenesis and animal models', *Cerebellum*, 7(2), 125-37.

- Romano, S., Coarelli, G., Marcotulli, C., Leonardi, L., Piccolo, F., Spadaro, M., Frontali, M., Ferraldeschi, M., Vulpiani, M.C., Ponzelli, F., Salvetti, M., Orzi, F., Petrucci, A., Vanacore, N., Casali, C. und Ristori, G. (2015) 'Riluzole in patients with hereditary cerebellar ataxia: a randomised, double-blind, placebo-controlled trial', *Lancet Neurol*, 14(10), 985-91.
- Rosenbloom, K.R., Sloan, C.A., Malladi, V.S., Dreszer, T.R., Learned, K., Kirkup, V.M., Wong, M.C., Maddren, M., Fang, R., Heitner, S.G., Lee, B.T., Barber, G.P., Harte, R.A., Diekhans, M., Long, J.C., Wilder, S.P., Zweig, A.S., Karolchik, D., Kuhn, R.M., Haussler, D. und Kent, W.J. (2013) 'ENCODE data in the UCSC Genome Browser: year 5 update', *Nucleic Acids Res*, 41(Database issue), D56-63.
- Rüb, U., Brunt, E.R. und Deller, T. (2008) 'New insights into the pathoanatomy of spinocerebellar ataxia type 3 (Machado-Joseph disease)', *Curr Opin Neurol*, 21(2), 111-6.
- Sakaguchi, S., Ono, M., Setoguchi, R., Yagi, H., Hori, S., Fehervari, Z., Shimizu, J., Takahashi, T. und Nomura, T. (2006) 'Foxp3+ CD25+ CD4+ natural regulatory T cells in dominant self-tolerance and autoimmune disease', *Immunol Rev*, 212, 8-27.
- Sanger, F., Nicklen, S. und Coulson, A.R. (1977) 'DNA sequencing with chain-terminating inhibitors', *Proc Natl Acad Sci U S A*, 74(12), 5463-7.
- Sanz-Gallego, I., Rodriguez-de-Rivera, F.J., Pulido, I., Torres-Aleman, I. und Arpa, J. (2014) 'IGF-1 in autosomal dominant cerebellar ataxia - open-label trial', *Cerebellum Ataxias*, 1, 13.
- Saute, J.A., de Castilhos, R.M., Monte, T.L., Schumacher-Schuh, A.F., Donis, K.C., D'Avila, R., Souza, G.N., Russo, A.D., Furtado, G.V., Gheno, T.C., de Souza, D.O., Portela, L.V., Saraiva-Pereira, M.L., Camey, S.A., Torman, V.B., de Mello Rieder, C.R. und Jardim, L.B. (2014) 'A randomized, phase 2 clinical trial of lithium carbonate in Machado-Joseph disease', *Mov Disord*, 29(4), 568-73.
- Saute, J.A., Rieder, C.R., Castilhos, R.M., Monte, T.L., Schumacher-Schuh, A.F., Donis, K.C., D'Avila, R., Souza, G.N., Russo, A.D., Furtado, G.V., Gheno, T.C., Souza, D.O., Saraiva-Pereira, M.L., Portela, L.V., Camey, S., Torman, V.B. und Jardim, L.B. (2015) 'Planning future clinical trials in Machado Joseph disease: Lessons from a phase 2 trial', *J Neurol Sci*, 358(1-2), 72-6.
- Saute, J.A.M. und Jardim, L.B. (2018) 'Planning Future Clinical Trials for Machado-Joseph Disease', *Adv Exp Med Biol*, 1049, 321-348.
- Schmidt, J., Schmidt, T., Golla, M., Lehmann, L., Weber, J.J., Hubener-Schmid, J. und Riess, O. (2016) 'In vivo assessment of riluzole as a potential therapeutic drug for spinocerebellar ataxia type 3', *J Neurochem*, 138(1), 150-62.

- Schmidt, T., Landwehrmeyer, G.B., Schmitt, I., Trottier, Y., Auburger, G., Laccone, F., Klockgether, T., Volpel, M., Epplen, J.T., Schols, L. und Riess, O. (1998) 'An isoform of ataxin-3 accumulates in the nucleus of neuronal cells in affected brain regions of SCA3 patients', *Brain Pathol*, 8(4), 669-79.
- Schöls, L., Amoiridis, G., Büttner, T., Przuntek, H., Epplen, J.T. und Riess, O. (1997) 'Autosomal dominant cerebellar ataxia: phenotypic differences in genetically defined subtypes?', *Ann Neurol*, 42(6), 924-32.
- Schöls, L., Amoiridis, G., Epplen, J.T., Langkafel, M., Przuntek, H. und Riess, O. (1996) 'Relations between genotype and phenotype in German patients with the Machado-Joseph disease mutation', *J Neurol Neurosurg Psychiatry*, 61(5), 466-70.
- Schöls, L., Bauer, P., Schmidt, T., Schulte, T. und Riess, O. (2004) 'Autosomal dominant cerebellar ataxias: clinical features, genetics, and pathogenesis', *Lancet Neurol*, 3(5), 291-304.
- Shaw, G., Morse, S., Ararat, M. und Graham, F.L. (2002) 'Preferential transformation of human neuronal cells by human adenoviruses and the origin of HEK 293 cells', *FASEB J*, 16(8), 869-71.
- Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M. und Sirotkin, K. (2001) 'dbSNP: the NCBI database of genetic variation', *Nucleic Acids Res*, 29(1), 308-11.
- Silva-Fernandes, A., Costa Mdo, C., Duarte-Silva, S., Oliveira, P., Botelho, C.M., Martins, L., Mariz, J.A., Ferreira, T., Ribeiro, F., Correia-Neves, M., Costa, C. und Maciel, P. (2010) 'Motor uncoordination and neuropathology in a transgenic mouse model of Machado-Joseph disease lacking intranuclear inclusions and ataxin-3 cleavage products', *Neurobiol Dis*, 40(1), 163-76.
- Silveira, I., Miranda, C., Guimaraes, L., Moreira, M.C., Alonso, I., Mendonca, P., Ferro, A., Pinto-Basto, J., Coelho, J., Ferreirinha, F., Poirier, J., Parreira, E., Vale, J., Januario, C., Barbot, C., Tuna, A., Barros, J., Koide, R., Tsuji, S., Holmes, S.E., Margolis, R.L., Jardim, L., Pandolfo, M., Coutinho, P. und Sequeiros, J. (2002) 'Trinucleotide repeats in 202 families with ataxia: a small expanded (CAG)<sub>n</sub> allele at the SCA17 locus', *Arch Neurol*, 59(4), 623-9.
- Sowa, A.S., Martin, E., Martins, I.M., Schmidt, J., Depping, R., Weber, J.J., Rother, F., Hartmann, E., Bader, M., Riess, O., Tricoire, H. und Schmidt, T. (2018) 'Karyopherin alpha-3 is a key protein in the pathogenesis of spinocerebellar ataxia type 3 controlling the nuclear localization of ataxin-3', *Proc Natl Acad Sci U S A*, 115(11), E2624-E2633.
- Sun, J., Xu, H., Negi, S., Subramony, S.H. und Hebert, M.D. (2007) 'Differential effects of polyglutamine proteins on nuclear organization and artificial reporter splicing', *J Neurosci Res*, 85(11), 2306-17.

- Sun, Y.M., Lu, C. und Wu, Z.Y. (2016) 'Spinocerebellar ataxia: relationship between phenotype and genotype - a review', *Clin Genet*, 90(4), 305-14.
- Szabo, E., Riffe, M.E., Steinberg, S.M., Birrer, M.J. und Linnoila, R.I. (1996) 'Altered cJUN expression: an early event in human lung carcinogenesis', *Cancer Res*, 56(2), 305-15.
- Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., Jensen, L.J. und Mering, C.V. (2019) 'STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets', *Nucleic Acids Res*, 47(D1), D607-D613.
- Takei, H. und Kobayashi, S.S. (2019) 'Targeting transcription factors in acute myeloid leukemia', *Int J Hematol*, 109(1), 28-34.
- Takiyama, Y., Nishizawa, M., Tanaka, H., Kawashima, S., Sakamoto, H., Karube, Y., Shimazaki, H., Soutome, M., Endo, K., Ohta, S. *et al.* (1993) 'The gene for Machado-Joseph disease maps to human chromosome 14q', *Nat Genet*, 4(3), 300-4.
- Tan, S., Wang, R.H., Niu, H.X., Shi, C.H., Mao, C.Y., Zhang, R., Song, B., Sun, S.L., Liu, X.J., Hou, H.M., Liu, Y.T., Gao, Y., Fang, H., Kong, X.D. und Xu, Y.M. (2015) 'Nerve growth factor for the treatment of spinocerebellar ataxia type 3: an open-label study', *Chin Med J (Engl)*, 128(3), 291-4.
- Teufel, D.P., Freund, S.M., Bycroft, M. und Fersht, A.R. (2007) 'Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53', *Proc Natl Acad Sci U S A*, 104(17), 7009-14.
- Uehara, I. und Tanaka, N. (2018) 'Role of p53 in the Regulation of the Inflammatory Tumor Microenvironment and Tumor Suppression', *Cancers (Basel)*, 10(7).
- Uhlen, M., Fagerberg, L., Hallstrom, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigartyo, C.A., Odeberg, J., Djureinovic, D., Takanen, J.O., Hober, S., Alm, T., Edqvist, P.H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J.M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J. und Ponten, F. (2015) 'Proteomics. Tissue-based map of the human proteome', *Science*, 347(6220), 1260419.
- van de Warrenburg, B.P., Sinke, R.J., Verschuuren-Bemelmans, C.C., Scheffer, H., Brunt, E.R., Ippel, P.F., Maat-Kievit, J.A., Dooijes, D., Notermans, N.C., Lindhout, D., Knoers, N.V. und Kremer, H.P. (2002) 'Spinocerebellar ataxias in the Netherlands: prevalence and age at onset variance analysis', *Neurology*, 58(5), 702-8.

- van der Velden, A.W. und Thomas, A.A. (1999) 'The role of the 5' untranslated region of an mRNA in translation regulation during development', *Int J Biochem Cell Biol*, 31(1), 87-106.
- Wang, H., Parry, S., Macones, G., Sammel, M.D., Ferrand, P.E., Kuivaniemi, H., Tromp, G., Halder, I., Shriver, M.D., Romero, R. und Strauss, J.F., 3rd (2004) 'Functionally significant SNP MMP8 promoter haplotypes and preterm premature rupture of membranes (PPROM)', *Hum Mol Genet*, 13(21), 2659-69.
- Wang, N.D., Finegold, M.J., Bradley, A., Ou, C.N., Abdelsayed, S.V., Wilde, M.D., Taylor, L.R., Wilson, D.R. und Darlington, G.J. (1995) 'Impaired energy homeostasis in C/EBP alpha knockout mice', *Science*, 269(5227), 1108-12.
- Wang, Q., Li, L. und Ye, Y. (2006) 'Regulation of retrotranslocation by p97-associated deubiquitinating enzyme ataxin-3', *J Cell Biol*, 174(7), 963-71.
- Wang, Z.J., Hanet, A., Weishaupl, D., Martins, I.M., Sowa, A.S., Riess, O. und Schmidt, T. (2018) 'Divalproex sodium modulates nuclear localization of ataxin-3 and prevents cellular toxicity caused by expanded ataxin-3', *CNS Neurosci Ther*, 24(5), 404-411.
- Wisdom, R., Johnson, R.S. und Moore, C. (1999) 'c-Jun regulates cell cycle progression and apoptosis by distinct mechanisms', *EMBO J*, 18(1), 188-97.
- Xu, Y.Z., Kanagaratham, C., Jancik, S. und Radzioch, D. (2013) 'Promoter deletion analysis using a dual-luciferase reporter system', *Methods Mol Biol*, 977, 79-93.
- Young, R.L. und Korsmeyer, S.J. (1993) 'A negative regulatory element in the bcl-2 5'-untranslated region inhibits expression from an upstream promoter', *Mol Cell Biol*, 13(6), 3686-97.
- Zahnow, C.A. (2009) 'CCAAT/enhancer-binding protein beta: its role in breast cancer and associations with receptor tyrosine kinases', *Expert Rev Mol Med*, 11, e12.
- Zesiewicz, T.A., Wilmot, G., Kuo, S.H., Perlman, S., Greenstein, P.E., Ying, S.H., Ashizawa, T., Subramony, S.H., Schmahmann, J.D., Figueroa, K.P., Mizusawa, H., Schols, L., Shaw, J.D., Dubinsky, R.M., Armstrong, M.J., Gronseth, G.S. und Sullivan, K.L. (2018) 'Comprehensive systematic review summary: Treatment of cerebellar motor dysfunction and ataxia: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology', *Neurology*, 90(10), 464-471.
- Zheng, Y., Josefowicz, S.Z., Kas, A., Chu, T.T., Gavin, M.A. und Rudensky, A.Y. (2007) 'Genome-wide analysis of Foxp3 target genes in developing and mature regulatory T cells', *Nature*, 445(7130), 936-40.

## 7. Erklärungen zum Eigenanteil

Die Arbeit wurde im Universitätsklinikum Tübingen im Institut für medizinische Genetik und angewandte Genomik unter Betreuung von Professor Dr. Olaf Rieß (Ärztlicher Direktor des Instituts für medizinische Genetik und angewandte Genomik) durchgeführt.

Die Konzeption der experimentellen Arbeit erfolgte durch Dr. Thorsten Schmidt (Arbeitsgruppenleiter SCA3, Instituts für medizinische Genetik und angewandte Genomik).

Die Versuche wurden von mir eigenständig durchgeführt, mit Ausnahmen von den Versuchen, die zu den Ergebnissen im Ergebnisteil unter 3.1.1, 3.1.2 und 3.1.3 führten. Diese Versuche wurden durch Anna Katharina Kuderna (3.1.1 und 3.1.2) und Anna Kohl (3.1.3) im Rahmen einer Bachelorarbeit (Kuderna) beziehungsweise eines Praktikums (Kohl) durchgeführt.

Die statistische Auswertung erfolgte, nach Anleitung durch Dr. Thorsten Schmidt, eigenständig durch mich.

Ich versichere, das Manuskript selbständig verfasst zu haben und keine weiteren als die von mir angegebenen Quellen verwendet zu haben.

---

Ort, Datum

Unterschrift

## 8. Anhang

## 8.1. Sequenzen der Konstrukte

Im Folgenden ist die Sequenz des Ausgangskonstrukts pGL3-KPNA3-clone I mit Strang (5' nach 3') und Gegenstrang (3'-5') dargestellt. Es sind darin jeweils Beginn und Ende der Bereiche markiert, die von den verschiedenen Konstrukten umfasst werden.

- Legende:
- █ Restriktionsenzym-Schnittstellen
  - █ Position in Bezug auf Transkriptionsstart des *KPNA3*-Gens
  - █ Exklusiver Bereich von pGL3-KPNA3-clone I
  - █ Bereich der in allen Konstrukten enthalten ist
  - █ 5'-UTR des Exon 1
  - █ SNPs rs9535343, rs12868235 und rs11618521
  - █ Luciferase-Gen
  - █ Ampicillin-Resistenz (auf Gegenstrang lokalisiert)

NheI └┘ Position -3085 / Beginn: pGL3-KPNA3-clone I

```

5'- GGTACCGAGC TCTTACGCGT GCTAGCCTGT TATGCTTGTG TTGCCTACGG CATACTGCTA
3'- CCATGGCTCG AGAATGCGCA CGATCGGACA ATACGAACAC AACGGATGCC GTATGACGAT

61 CCTGACTGAC TGCTGGCATG ACTCCCCATG CTTGCGGAAG GTGGAGAGTG AGTCAGAGGA
GGACTGACTG ACGACCGTAC TGAGGGGTAC GGAACCCTTC CACCTCTCAC TCAGTCTCCT

121 TTTGATTGAC CCCTGGGCTG AGGGCCCTTC CTTCTGTCT CCCCAGAGAC CCCCTAGTCC
AAACTAACTG GGGACCCGAC TCCCGGGAAG GAAGGACAGA GGGGGCTCTG GGGGATCAGG

181 TTCATACATG TGCCCGATGA ATACAATGCT TGGTATACAG TAGACACTCA CAATATGTGT
AAGTATGTAC ACGGGCTACT TATGTTACGA ACCATATGTC ATCTGTGAGT GTTATACACA

241 GGAAAATAAA TCCATTTAAA AGAAGAGCTG TCCAGAGAAG TTTCAGGCCA AGGATGATGA
CCTTTTATTT AGGTAAATTT TCTTCTCGAC AGGTCTCTTC AAAGTCCGGT TCCTACTACT

301 TTAGAGTGAG GCAAAAAGAGA AGGAGATAGT TTCTAACAAT GAGAAGCACC AAATTTGAGT
AATCTCACTC CGTTTTCTCT TCCTCTATCA AAGATTGTTA CTCTTCGTGG TTTAAACTCA

361 CCCTGGAGGT GGTCATTTCC CAGAGATTAA GGTCTCCAAG GTCTGGTGCA GGCTTGCACT
GGGACCTCCA CCAGTAAAGG GTCTCTAATT CCAGAGGTTC CAGACCACGT CCGAACGTGA

421 TTTTACCCTG TGAGTGGTAA AAAACTTGGT GAAGTTAAGG ATGTAGAAAAG ACTGTGGTGT
AAAATGGTGA ACTCACCATT TTTTGAACCA CTCAATTCC TACATCTTTC TGACACCACA

481 CAGGGTGCCA CATTTGTGAA TGCTGAAGTT ACCAGGGGTA CTGGCAAAGA ATAAAAATGA
GTCCACGGT GTAAACACTT ACGACTTCAA TGGTCCCAT GACCGTTTCT TATTTTTACT

```

541 GTAAACTAGA TCCAGGCAAA AGGGCAGCCC TGCTTTGGAC TGGGTGTGAA GAAAGGTGGC  
 CATTTGATCT AGGTCCGTTT TCCCGTCGGG ACGAAACCTG ACCCACACTT CTTTCCACCG

601 ATAAAGTAAC CTGAAAAGGG AATACCTACT AAATTATAAT AGAGACTGAG CAACAGTCTG  
 TATTTTCATTG GACTTTTCCC TTATGGATGA TTTAATATTA TCTCTGACTC GTTGTGACAG

661 AAAATTTCTC TTCCATTTAG CCAAGGAGAA TGGCAAACAC ACAAGTTCTG TTTCCCTTAAA  
 TTTTAAAGAG AAGGTAAATC GGTTCCTCTT ACCGTTTGTG TGTTCAAGAC AAAGGAATTT

721 ACCCAGAGTT GAGCTAATGT ATTCTTTCTT TTTTCTTTTT TTTGAGACAG AGTCTCATTC  
 TGGGTCTCAA CTCGATTACA TAAGAAAGAA AAAAGAAAAA AAACCTGTGC TCAGAGTAAG

781 TGTTACCCAG GCTGGAGTGC AGTGGCGCGA TCTCAACTCA CTGCAACCTC CGCCTCCAG  
 ACAATGGGTC CGACCTCACG TCACCGCGCT AGAGTTGAGT GACGTTGGAG GCGGAGGGTC

841 GTTCAAGCGA TTCTCCTGCC TCAGCTTCTT GAGCACCTGT GACTACAGGC ATGCACCACC  
 CAAGTTCGCT AAGAGGACGG AGTCGAAGGA CTCGTGGACA CTGATGTCCG TACGTGGTGG

901 ATGCCTGCCT AATTTTTGTA TTTTTGTAGA GAAAGGGTTT CACTATGTGG GCCATGGTTG  
 TACGGACGGA TTAAAAACAT AAAAACATCT CTTTCCCAA GTGATACACC CGGTACCAAC

961 GCCAGACTGG TCTCAAACCT CTGACCTCAG GTGATCCTCC CACCTCAGCC TCCCAAAGTG  
 CGGTCTGACC AGAGTTTGAG GACTGGAGTC CACTAGGAGG GTGGAGTCGG AGGGTTTCAC

1021 CTGGGATTAC AGGCATGAGC CACTGTGCC TGCCTAATGT ATTCCTTCTA AATTCATAGC  
 GACCCTAATG TCCGTACTCG GTGACACGGG ACGGATTACA TAAGGAAGAT TTAAGTATCG

Beginn: pGL3-KPNA3-abEcoG/pGL3-KPNA3-abEcoA

und pGL3-KPNA3-noUTR-G/pGL3-KPNA3-noUTR-A an **Position -1981**

1081 TGGTGTGTGT GTGTGTGTTT CCCTTCAAT GAGTGGTTAG ACTAAGATCA GGTGTGCAA  
 ACCACACACA CACACACAAA GGAAGTTTA CTCACCAATC TGATTCTAGT CCACAACGTT

1141 ACTGAGGACC TGATATTCTG AAGTGACTGA AATGCTTAAG TCCATTAACA AAATGTGAAT  
 TGACTCCTGG ACTATAAGAC TTCACTGACT TTACGAATTC AGGTAATTGT TTTACTACTA

1201 TAACATTTAA AAATAAAGAG AAAAAGAATC TACCTTTTTG GTTTTCCTTG AAAAGTGGAA  
 ATTGTAAATT TTTATTTCTC TTTTCTTAG ATGGAAAAAC CAAAAGGAAC TTTTCACCTT

1261 TCTTCTGGCA ACACTGGGCC ACAGTCCAGT ATGACCAGAG GATCTGAAGC CTGTGTTTAT  
 AGAAGACCGT TGTGACCCGG TGTGAGTCA TACTGGTCTC CTAGACTTCG GACACAAATA

1321 CAATAGGCCA CACCCACCA CTCACAAGAC GAATGCCCTT GGACATAGCT TGGCCACCAA  
 GTTATCCGGT GTGGGGTGGT GAGTGTTCTG CTTACGGGGA CCTGTATCGA ACCGGTGGTT

**EcoRI** 7

1381 AGGCTTCTGA ATTCCTGACT TCTGGGCCTC AGGTTCTTTA AGGCCCTTTC GCTTTTTTTT  
 TCCGAAGACT TAAGGACTGA AGACCCGGAG TCCAAGAAAT TCCGGGAAAG CGAAAAAAA

1441 TTTTTTTTTT TTCTGTATCT CCTCATATAT TTATTCTTTT TCTGTGATCA GTGTCCTTTG  
 AAAAAAAA AAGACATAGA GGAGTATATA AATAAGAAAA AGACACTAGT CACAGGAAAC

1501 CATGTTTTTG TTGGGTGTT TGAATTTTTT ATTTTTATTT TTTGGGGATG GAGTCTCACT  
 GTACAAAAAC AACCAACAA ACTTAAAAA TAAAAATAA AAACCCCTAC CTCAGAGTGA

1561 CTGTCGCCTA GGCTGGAGTG CAGTGGCAGG ATCTCAGCTC ACTGCAACCT CGCCTCCCA  
 GACAGCGGAT CCGACCTCAC GTCACCGTGC TAGAGTCGAG TGACGTTGGA GCGGAGGGT

1621 GGTTCAAGCA ATTCTCCTGC CTCAGGCTCC CAAGTAGCTG AGATTACAGG TATGCGC **CAT**  
 CCAAGTTCGT TAAGAGGACG GAGTCCGAGG GTTCATCGAC TCTAATGTCC ATACGCG **GTA**

└ NdeI

1681 **ATG**CCCCGTCT AATTTTTGTA TTTTTTTTAG TAGAGGTGGG GTTTCACCAT GTTGGCCAGG  
**TAC**GGGCAGA TTAAAAACAT AAAAAAATC ATCTCCACCC CAAAGTGGTA CAACCGGTCC

1741 CTGGTCTCGA CCTCCTGACC TTGTGATCCA CCCACCTCGG CCTCCCAAAG TGCTGGGATT  
 GACCAGAGCT GGAGGACTGG AACACTAGGT GGGTGGAGCC GGAGGGTTTC ACGACCCTAA

1801 ACAGGCATGA GCCACCGCGC CCAGCCAAGG CCCTTTCGCT TTAAAGATTG TCCTTTTTGA  
 TGTCCGTACT CGGTGGCGCG GGTCGTTTC GGGAAAGCGA AATTTCTAAC AGGAAAAACT

1861 TCTCTTCATG GTATAGTCTT ACCAACTTTT GACCTCAGTT CAATTATAAG TACTACAACA  
 AGAGAAGTAC CATATCAGGA TGGTTGAAAA CTGGAGTCAA GTTAATATTC ATGATGTTGT

1921 TTACATACCT GTCATGTCT ACCATAGATT CTGAAAATTA ACATACAAGC AAATGAAGCC  
 AATGTATGGA CGAGTACAGA TGGTATCTAA GACTTTTAAT TGTATGTTTC TTTACTTCGG

SNP rs11618521<sub>T</sub> (entspricht  $\frac{A}{T}$  bei Konstrukten mit Adenin)

1981 CACCTGCTCC AGAAAGTAGC TTTTGTGTTT TGTGTTG **G**TT GTTTTTGTTG TTTTGTGTTGA  
 GTGGACGAGG TCTTTCATCG AAAAAACAAA ACAACAAC **A**AAAAACAAC AAACAAAAC

└ PstI

2041 **GACAGAGTCT** CCCTCTGTGC CCAGGCTGGA GTGCAGTGGC GCCATCTCGG CTCGCTGCAA  
**CTGTCTCAGA** GGGAGACAGC GGTCCGACCT CACGTCACCG CGGTAGAGCC GAGCGACGTT

2101 CCTCCGCCTA CCGGGTTCAA GCGATTCTCC TGCCTCAGCC TCCAGGCATG CGCCACCAGC  
 GGAGGCGGAT GGCCCAAGTT CGCTAAGAGG ACGGAGTCGG AGGTCCGTAC GCGGTGGTGC

2161 CCCAGCTAAT TTTTGTATAT TTAGTAAAGA AGGGGTTTCA CCATGCGGGC CATGGTTGGC  
 GGGTCGATTA AAAACATATA AATCATTCTT TCCCAAAGT GGTACGCCCC GTACCAACCG

2221 CAGACTGGTC TCGGTCTCTT GACCTTGTGA TCTGCCCTCC TCGGCCTCCC AAAGTGCTGG  
 GTCTGACCAG AGCCAGAGAA CTGGAACACT AGACGGGAGG AGCCGGAGGG TTTCACGACC

2281 GATTACAGGC GTGAGCCACC ACACCCGGCC GAAAGCAGCT CTTTATCAGT GTAATCCATC  
 CTAATGTCCG CACTCGGTGG TGTGGGCCGG CTTTCGTCGA GAAATAGTCA CATTAGGTAG

2341 CAAATTTCCC TCTTTCATTC TTTCGATGCC CCCAGAGGGA GTGGGAGTGC TACACAGGAG  
 GTTTAAAGGG AGAAAGTAAG AAAGCTACGG GGTCTCCCT CACCCTCACG ATGTGTCCCTC

2401 AGTCTGCGGA CCAGGGCTGG GCTGCTGCCT GAAGGCTAAG ATTCTGTGTG GTGGTAGAAC  
 TCAGACGCCT GGTCCCGACC CGACGACGGA CTCCGATTC TAAGACACAC CACCATCTTG

2461 AAGAAGTGTT GGGGGGCATG TAAATACAGG GACTGCCAGA GTGAAGGTTG CAGAACCGGG  
 TTCTTCACAA CCCCCGTAC ATTTATGTCC CTGACGGTCT CACTTCCAAC GTCTTGGCCC

2521 CTTGGAAGGA TGATGAAGAA ACTGGGCGAC ACTAAAGCGC CAGTTCAACA AGTATTGCGT  
 GAACCTTCCT ACTACTTCTT TGACCCGCTG TGATTTCCGG GTCAAGTTGT TCATAACGCA

2581 ATCTACGACG TGGTAGGCAT TGTTTTGACA CTAGAGATTC AGCAGTAATT CAACCAAAGT  
 TAGATGCTGC ACCATCCGTA ACAAACACTGT GATCTCTAAG TCGTCATTAA GTTGGTTTCA

2641 ACCTGCTCTC ATTGAGCTGA ATTTCCGCGC GGCACCTAG GTAACGCATA TAAACACTCG  
 TGGACGAGAG TAACTCGACT TAAAGGCGCG CCCGTGGATC CATTGCGTAT ATTTGTGAGC

2701 TATCACATAC ATACACATGT ACATGTCAGG TAGAAATGCT ATTTATAAAA CTGAAGCAGA  
 ATAGTGTATG TATGTGTACA TGTACAGTCC ATCTTTACGA TAAATATTTT GACTTCGTCT

2761 GCATGGGAAG AGCGAATGAA GGTGGCTATT TGAGAAGACC CCCAAAAAAG CGAGACTGCA  
 CGTACCCTTC TCGCTTACTT CCACCGATAA ACTCTTCTGG GGGTTTTTTC GCTCTGACGT

2821 AAGCACTCGC GGTTGTTCTT TTTCTGATCT TCTAGCTACA GGCAGCGAGG TGCGCTCCCA  
 TTCGTGAGCG CCAACAAGGA AAAGACTAGA AGATCGATGT CCGTCGCTCC ACGCGAGGGT

2881 GGAAGGCCCG CCCTCTTCTT CCTGATAGGG CGCAAGGTGC CAGGGACGCT CCCCCGCTCT  
 CCTTCCGGGC GGGAGAAGGA GGACTATCCC GCGTTCCACG GTCCCTGCGA GGGGGCGAGA

2941 TCCCCGCCCA GAGGCGGCCT TGTGGGGGCG GAGCTTTGTG CCCAGCAACA TGAGCGCGGC  
 AGGGGCGGGT CTCCGCCGGA ACACCCCGC CTCGAAACAC GGTTCGTTGT ACTCGCGCCG

3001 GCGCAAAGAG CCGCGCGCGG AAGGGTCCCG GCATTCTCCT CGTTCTCACT AGGCGTGGAG  
 CGCGTTTCTC GCGCGCGGCC TTCCAGGGC CGTAAGAGGA GCAAGAGTGA TCCGCACCTC

Ende: pGL3-KPNA3-noUTR-G/pGL3-KPNA3-noUTR-A

und Beginn: 5' UTR des Exon 1 an Position 1

3061 AATTACTCCA GCGGATCGT GGCATGATGG GAACTGTAGT TGACGACCGC CGCCCCGCGC  
 TTAATGAGGT CCGCCTAGCA CCGTACTACC CTTGACATCA ACTGCTGGCG GCGGGGCGCG

3121 CTGAGGGGCA GTAAAAGTCG CCAGGTCCGG CTCCATTTCT GGCACAAAAC TTGCAGCACC  
 GACTCCCCGT CATTTTCAGC GGTCCAGGCC GAGGTAAAGA CCGTGTTTTG AACGTCGTGG

3181 GAGGGGTTGT GGAGAGCCCT TGCAGGGGAA GAGGGCAGGG TCATCCCGAG AACCAACGGG  
 CTCCCCAACA CCTCTCGGGA ACGTCCCCTT CTCCCGTCCC AGTAGGGCTC TTGGTTGCC

3241 CACGTATAGC CCGGCGAACG CCCAAGCCGG TCACCGCCCC CGGTCACGTG TCGCCAGCCT  
 GTGCATATCG GGCCGCTTGC GGGTTCGGCC AGTGGCGGGG GCCAGTGCAC AGCGGTGCGA

3301 CCGCGGCCCG GCGCCGCTCT CAGCACCGTT CCCGCCCCAC CCGGCCCGGC AGTCGGCCCC  
 GCGCGCCGGC GCGCGCGAGA GTCGTGGCAA GGGCGGGGTG GGCCGGGCCG TCAGCCGGGC

3361 CGCCTCCCC GCGGCTACTG CCACCTCGCG CTCGGAGGCG TCACAGAACG TGCTCTTCTC  
 GCGGAGGGGG CCGCGATGAC GGTGGAGCGC GAGCCTCCGC AGTGTCTTGC ACGAGAAGAG

3421 TCCCCTCCCC CCTCCCGCTC TCCCCTCCT CCCCTCCCG CTCCAAGATT CGCCGCCGCC  
 AGGGGAGGGG GGAGGGCGAG AGGGGAGGA GGGGAGGGC GAGGTTCTAA GCGGCGGCGG

Ende: pGL3-KPNA3-abEcoG/pGL3-KPNA3-abEcoA

und pGL3-KPNA3-clone I an Position +415

3481 GCCGCCGAG CCGCAGGAGT AGCCGCCGCC GGAGCCGCGC GCAGCCCTCG AGATCTGCGA  
 CGGCGGCGTC GCGTCTCTCA TCGGCGGCGG CCTCGGCGCG CGTCCGGAGC TCTAGACGCT

3541 TCTAAGTAAG CTTGGCATTG CCGTACTGTT GGTAAGCCA CCATGGAAGA CGCCAAAAAC  
 AGATTCATTG GAACCGTAAG GCCATGACAA CCATTTCCGGT GGATACCTTCT GCGGTTTTTG

3601 ATAAAGAAAG GCCCGGCGCC ATTCTATCCG CTGGAAGATG GAACCGCTGG AGAGCAACTG  
 TATTTCTTTC CGGGCGCGG TAAGATAGGC GACCTTCTAC CTTGGCGACC TCTCGTTGAC

Luciferase-Gen

3661 CATAAGGCTA TGAAGAGATA CGCCCTGGTT CCTGGAACAA TTGCTTTTAC AGATGCACAT  
GTATTCCGAT ACTTCTCTAT GCGGGACCAA GGACCTTGTT AACGAAAATG TCTACGTGTA

3721 ATCGAGGTGG ACATCACTTA CGCTGAGTAC TTCGAAATGT CCGTTCGGTT GGCAGAAGCT  
TAGCTCCACC TGTAGTGAAT GCGACTCATG AAGCTTTACA GGCAAGCCAA CCGTCTTCGA

3781 ATGAAACGAT ATGGGCTGAA TACAAATCAC AGAATCGTCG TATGCAGTGA AAACCTCTCTT  
TACTTTGCTA TACCCGACTT ATGTTTAGTG TCTTAGCAGC ATACGTCACT TTTGAGAGAA

3841 CAATTCTTTA TGCCGGTGTG GGGCGCGTTA TTTATCGGAG TTGCAGTTGC GCCCGCGAAC  
GTTAAGAAAT ACGGCCACAA CCCGCGCAAT AAATAGCCTC AACGTCAACG CGGGCGCTTG

3901 GACATTTATA ATGAACGTGA ATTGCTCAAC AGTATGGGCA TTTCGCAGCC TACCGTGGTG  
CTGTAAATAT TACTTGCCTT TAACGAGTTG TCATACCCGT AAAGCGTCGG ATGGCACCAC

3961 TTCGTTTCCA AAAAGGGGTT GCAAAAAATT TTGAACGTGC AAAAAAAGCT CCCAATCATC  
AAGCAAAGGT TTTTCCCAA CGTTTTTTAA AACTTGCACG TTTTTTTTCA GGGTTAGTAG

4021 CAAAAAATTA TTATCATGGA TTCTAAAACG GATTACCAGG GATTTTCAGTC GATGTACACG  
GTTTTTTAAT AATAGTACCT AAGATTTTGC CTAATGGTCC CTAAAGTCAG CTACATGTGC

4081 TTCGTACACAT CTCATCTACC TCCCGGTTTT AATGAATACG ATTTTGTGCC AGAGTCCTTC  
AAGCAGTGTA GAGTAGATGG AGGGCCAAAA TTACTTATGC TAAAACACGG TCTCAGGAAG

4141 GATAGGGACA AGACAATTGC ACTGATCATG AACTCCTCTG GATCTACTGG TCTGCCTAAA  
CTATCCCTGT TCTGTTAACG TGACTAGTAC TTGAGGAGAC CTAGATGACC AGACGGATTT

4201 GGTGTGCGTC TGCCTCATAG AACTGCCTGC GTGAGATTCT CGCATGCCAG AGATCCTATT  
CCACAGCGAG ACGGAGTATC TTGACGGACG CACTCTAAGA GCGTACGGTC TCTAGGATAA

4261 TTTGGCAATC AAATCATTCC GGATACTGCG ATTTTAAGTG TTGTTCCATT CCATCACGGT  
AAACCGTTAG TTTAGTAAGG CCTATGACGC TAAAATTCAC AACAAGGTAA GGTAGTGCCA

4321 TTTGGAATGT TTACTIONACT CGGATATTTG ATATGTGGAT TTCGAGTCGT CTTAATGTAT  
AAACCTTACA AATGATGTGA GCCTATAAAC TATACACCTA AAGCTCAGCA GAATTACATA

4381 AGATTTGAAG AAGAGCTGTT TCTGAGGAGC CTTGAGGATT ACAAGATTCA AAGTGCCTG  
TCTAAACTTC TTCTCGACAA AGACTCCTCG GAAGTCCTAA TGTCTAAGT TTCACGCGAC

4441 CTGGTGCCAA CCCTATTCTC CTTCTTCGCC AAAAGCACTC TGATTGACAA ATACGATTTA  
GACCACGGTT GGGATAAGAG GAAGAAGCGG TTTTCGTGAG ACTAACTGTT TATGCTAAAT

4501 TCTAATTTAC ACGAAATTGC TTCTGGTGGC GCTCCCCTCT CTAAGGAAGT CGGGGAAGCG  
AGATTAAATG TGCTTTAACG AAGACCACCG CGAGGGGAGA GATTCCTTCA GCCCCTTCGC

4561 GTTGCCAAGA GGTTCATCT GCCAGGTATC AGGCAAGGAT ATGGGCTCAC TGAGACTACA  
CAACGGTTCT CCAAGGTAGA CGGTCCATAG TCCGTTTCTA TACCCGAGTG ACTCTGATGT

4621 TCAGCTATTC TGATTACACC CGAGGGGGAT GATAAACCGG GCGCGGTCGG TAAAGTTGTT  
AGTCGATAAG ACTAATGTGG GCTCCCCTA CTATTTGGCC CGCGCCAGCC ATTTCAACAA

4681 CCATTTTTTG AAGCGAAGGT TGTGGATCTG GATACCGGGA AAACGCTGGG CGTTAATCAA  
GGTAAAAAAC TTCGCTTCCA ACACCTAGAC CTATGGCCCT TTTGCGACCC GCAATTAGTT

4741 AGAGGCGAAC TGTGTGTGAG AGGTCCTATG ATTATGTCCG GTTATGTAAA CAATCCGGAA  
TCTCCGCTTG ACACACACTC TCCAGGATAC TAATACAGGC CAATACATTT GTTAGGCCTT

4801 GCGACCAACG CCTTGATTGA CAAGGATGGA TGGCTACATT CTGGAGACAT AGCTTACTGG  
CGCTGGTTGC GGAACATACT GTTCTTACCT ACCGATGTAA GACCTCTGTA TCGAATGACC

4861 GACGAAGACG AACACTTCTT CATCGTTGAC CGCCTGAAGT CTCTGATTAA GTACAAAGGC  
CTGCTTCTGC TTGTGAAGAA GTAGCAACTG GCGGACTTCA GAGACTAATT CATGTTTTCCG

4921 TATCAGGTGG CTCCCGCTGA ATTGGAATCC ATCTTGCTCC AACACCCCAA CATCTTCGAC  
ATAGTCCACC GAGGGCGACT TAACCTTAGG TAGAACGAGG TTGTGGGGTT GTAGAAGCTG

4981 GCAGGTGTCG CAGGTCTTCC CGACGATGAC GCCGGTGAAC TTCCC GCCG C GTTGTGTGTT  
CGTCCACAGC GTCCAGAAGG GCTGCTACTG CGGCCACTTG AAGGGCGGGC GCAACAACAA

5041 TTGGAGCACG GAAAGACGAT GACGGAAAAA GAGATCGTGG ATTACGTCGC CAGTCAAGTA  
AACCTCGTGC CTTTCTGCTA CTGCCTTTTT CTCTAGCACC TAATGCAGCG GTCAGTTCAT

5101 ACAACCGCGA AAAAGTTGCG CGGAGGAGTT GTGTTTGTGG ACGAAGTACC GAAAGGTCTT  
TGTTGGCGCT TTTTCAACGC GCCTCCTCAA CACAAACACC TGCTTCATGG CTTTCCAGAA

5161 ACCGGAAAAAC TCGACGCAAG AAAAAATCAGA GAGATCCTCA TAAAGGCCAA GAAGGGCGGA  
TGGCCTTTTG AGCTGCGTTC TTTTATAGTCT CTCTAGGAGT ATTTCCGGTT CTTCCCGCT

## Luciferase-Gen 7

5221 AAGATCGCCG TGTAATTCTA GAGTCGGGGC GGCCGGCCGC TTCGAGCAGA CATGATAAGA  
TTCTAGCGGC ACATTAAAGAT CTCAGCCCCG CCGGCCGGCG AAGCTCGTCT GTACTATTCT

5281 TACATTGATG AGTTTGGACA AACCACAAC T AGAATGCAGT GAAAAAATG CTTTATTTGT  
ATGTAAC TAC AAACCTGT TTGGTGTGTA TCTTACGTCA CTTTTTTTAC GAAATAAACA

5341 GAAATTTGTG ATGCTATTGC TTTATTTGTA ACCATTATAA GCTGCAATAA ACAAGTTAAC  
CTTTAAACAC TACGATAACG AAATAAACAT TGGTAATATT CGACGTTATT TGTTCAATTG

5401 AACACAATT GCATTCATTT TATGTTTCAG GTTCAGGGGG AGGTGTGGGA GGTTTTTTAA  
TTGTTGTAA CGTAAGTAAA ATACAAAGTC CAAGTCCCCC TCCACACCCT CCAAAAAATT

5461 AGCAAGTAAA ACCTCTACAA ATGTGGTAAA ATCGATAAGG ATCCGTCGAC CGATGCCCTT  
TCGTTTCAATTT TGGAGATGTT TACACCATTT TAGCTATTCC TAGGCAGCTG GCTACGGGAA

5521 GAGAGCCTTC AACCCAGTCA GTCCTTCCG GTGGGCGCGG GGCATGACTA TCGTCGCCGC  
CTCTCGGAAG TTGGGTCAGT CGAGGAAGGC CACCCGCGCC CCGTACTGAT AGCAGCGGGC

5581 ACTTATGACT GTCTTCTTTA TCATGCAACT CGTAGGACAG GTGCCGGCAG CGCTCTTCCG  
TGAATACTGA CAGAAGAAAT AGTACGTTGA GCATCCTGTC CACGGCCGTC GCGAGAAGGC

5641 CTTCTCGCT CACTGACTCG CTGCGCTCGG TCGTTCGGCT GCGGCGAGCG GTATCAGCTC  
GAAGGAGCGA GTGACTGAGC GACGCGAGCC AGCAAGCCGA CGCCGCTCGC CATAGTCGAG

5701 ACTCAAAGGC GGTAATACGG TTATCCACAG AATCAGGGGA TAACGCAGGA AAGAACATGT  
TGAGTTTCCG CCATTATGCC AATAGGTGTC TTAGTCCCCT ATTGCGTCTT TTCTTGTACA

5761 GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG GCGTTTTTCC  
CTCGTTTTCC GGTGTTTTTC CGGTCCTTGG CATTTTTTCCG GCGCAACGAC CGCAAAAAGG

5821 ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG AGGTGGCGAA  
TATCCGAGGC GGGGGGACTG CTCGTAGTGT TTTTAGCTGC GAGTTCAGTC TCCACCGCTT

5881 ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC GTGCGCTCTC  
TGGGCTGTCC TGATATTTCT ATGGTCCGCA AAGGGGGACC TTCGAGGGAG CACGCGAGAG

5941 CTGTTCCGAC CCTGCCGCTT ACCGGATAACC TGTCCGCCTT TCTCCCTTCG GGAAGCGTGG  
GACAAGGCTG GGACGGCGAA TGGCCTATGG ACAGGCGGAA AGAGGGAAGC CCTTCGCACC

6001 CGCTTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCGGT GTAGGTCGTT CGCTCCAAGC  
GCGAAAGAGT ATCGAGTGCG ACATCCATAG AGTCAAGCCA CATCCAGCAA GCGAGGTTCG

6061 TGGGCTGTGT GCACGAACCC CCCGTTTCAGC CCGACCGCTG CGCCTTATCC GGTAACCTATC  
ACCCGACACA CGTGCTTGGG GGGCAAGTCG GGCTGGCGAC GCGGAATAGG CCATTGATAG

6121 GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC ACTGGTAACA  
CAGAACTCAG GTTGGGCCAT TCTGTGCTGA ATAGCGGTGA CCGTCGTCGG TGACCATTGT

6181 GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG TGGCCTAACT  
CCTAATCGTC TCGCTCCATA CATCCGCCAC GATGTCTCAA GAACTTCACC ACCGGATTGA

6241 ACGGCTACAC TAGAAGAACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA GTTACCTTCG  
TGCCGATGTG ATCTTCTTGT CATAAACCAT AGACGCGAGA CGACTTCGGT CAATGGAAGC

6301 GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC CGCTGGTAGC GGTGGTTTTT  
CTTTTTCTCA ACCATCGAGA ACTAGGCCGT TTGTTTGGTG GCGACCATCG CCACCAAAAA

6361 TTGTTTGCAA GCAGCAGATT ACGCGCAGAA AAAAAGGATC TCAAGAAGAT CCTTTGATCT  
AACAAACGTT CGTCGTCTAA TGCGCGTCTT TTTTTCCTAG AGTTCTTCTA GGAAACTAGA

6421 TTTCTACGGG GTCTGACGCT CAGTGGAACG AAAACTCACG TTAAGGGATT TTGGTCATGA  
AAAGATGCCC CAGACTGCGA GTCACCTTGC TTTTGAGTGC AATTCCCTAA AACCAGTACT

6481 GATTATCAAA AAGGATCTTC ACCTAGATCC TTTTAAATTA AAAATGAAGT TTTAAATCAA  
CTAATAGTTT TTCCTAGAAG TGGATCTAGG AAAATTTAAT TTTTACTTCA AAATTTAGTT

6541 TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTACCA ATGCTTAATC AGTGAGGCAC  
AGATTTTCATA TATACTCATT TGAACCAGAC TGTC AATGGT TACGAATTAG TCACTCCGTG  
↳ Ampicillin-Resistenz-Gen ←

6601 CTATCTCAGC GATCTGTCTA TTTCGTTTCAT CCATAGTTGC CTGACTCCCC GTCGTGTAGA  
GATAGAGTCG CTAGACAGAT AAAGCAAGTA GGTATCAACG GACTGAGGGG CAGCACATCT

6661 TAACTACGAT ACGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA CCGCGAGACC  
ATTGATGCTA TGCCCTCCCG AATGGTAGAC CGGGGTACAG ACGTTACTAT GGCGCTCTGG

6721 CACGCTCACC GGCTCCAGAT TTATCAGCAA TAAACCAGCC AGCCGGAAGG GCCGAGCGCA  
GTGCGAGTGG CCGAGGTCTA AATAGTCGTT ATTTGGTTCGG TCGGCCTTCC CGGCTCGCGT

6781 GAAGTGGTCC TGCAACTTTA TCCGCCTCCA TCCAGTCTAT TAATTGTTGC CGGGAAGCTA  
CTTCACCAGG ACGTTGAAAT AGGCGGAGGT AGGTCAGATA ATTAACAACG GCCCTTCGAT

6841 GAGTAAGTAG TTCGCCAGTT AATAGTTTGC GCAACGTTGT TGCCATTGCT ACAGGCATCG  
CTCATTTCATC AAGCGGTCAA TTATCAAACG CGTTGCAACA ACGGTAACGA TGTCCGTAGC

6901 TGGTGTACAG CTCGTCGTTT GGTATGGCTT CATTTCAGCTC CGGTTCCCAA CGATCAAGGC  
ACCACAGTGC GAGCAGCAA CCATACCGAA GTAAGTCGAG GCCAAGGGTT GCTAGTTCCG

6961 GAGTTACATG ATCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT CCTCCGATCG  
CTCAATGTAC TAGGGGGTAC AACACGTTTT TTCGCCAATC GAGGAAGCCA GGAGGCTAGC

7021 TTGTCAGAAG TAAGTTGGCC GCAGTGTTAT CACTCATGGT TATGGCAGCA CTGCATAATT  
AACAGTCTTC ATTCAACCGG CGTCACAATA GTGAGTACCA ATACCGTCGT GACGTATTA

7081 CTCTTACTGT CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC TCAACCAAGT  
GAGAATGACA GTACGGTAGG CATTCTACGA AAAGACACTG ACCACTCATG AGTTGGTTCA

7141 CATTCTGAGA ATAGTGTATG CGGCGACCGA GTTGCTCTTG CCCGGCGTCA ATACGGGATA  
GTAAGACTCT TATCACATAC GCCGCTGGCT CAACGAGAAC GGGCCGCAGT TATGCCCTAT

7201 ATACCGCGCC ACATAGCAGA ACTTTAAAAG TGCTCATCAT TGGAAAACGT TCTTCGGGGC  
TATGGCGCGG TGTATCGTCT TGAAATTTTC ACGAGTAGTA ACCTTTTGCA AGAAGCCCCG

7261 GAAAACCTCTC AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC ACTCGTGCAC  
CTTTTGAGAG TTCCTAGAAT GGCACAACT CTAGGTCAAG CTACATTGGG TGAGCACGTG

7321 CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGCGTTTC TGGGTGAGCA AAAACAGGAA  
GGTTGACTAG AAGTCGTAGA AAATGAAAGT GGTCGCAAAG ACCCACTCGT TTTTGTCTTT

7381 GGCAAAATGC CGCAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA CTCATACTCT  
CCGTTTTACG GCGTTTTTTC CCTTATTCCC GCTGTGCCTT TACAATTAT GAGTAGGAGA

← Ampicillin-Resistenz-Gen ↓

7441 TCCTTTTTCA ATATTATTGA AGCATTTATC AGGGTTATTG TCTCATGAGC GGATACATAT  
AGGAAAAAGT TATAATAACT TCGTAAATAG TCCAATAAC AGAGTACTCG CCTATGTATA

7501 TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCCGCG CACATTTCCC CGAAAAGTGC  
AACTTACATA AATCTTTTTA TTTGTTTATC CCCAAGGCGC GTGTAAAGGG GCTTTTCACG

7561 CACCTGACGC GCCCTGTAGC GGCGCATTAA GCGCGGCGGG TGTGGTGGTT ACGCGCAGCG  
GTGGACTGCG CGGGACATCG CCGCGTAATT CCGCCGCCC ACACCACCAA TGCGCGTTCG

7621 TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGTCTCTTT CGCTTTCTTC CTTTCTTTTC  
ACTGGCGATG TGAACGGTTCG CGGGATCGCG GGCAGGAAA GCGAAAGAAG GGAAGGAAAG

7681 TCGCCACGTT CGCCGGCTTT CCCCCTCAAG CTCTAAATCG GGGGCTCCCT TTAGGGTTCC  
AGCGGTGCAA GCGGCCGAAA GGGGCAGTTC GAGATTTAGC CCCCAGGGA AATCCCAAGG

7741 GATTTAGTGC TTTACGGCAC CTCGACCCCA AAAAATTGA TTAGGGTGAT GGTTACGTA  
CTAAATCAGC AAATGCCGTG GAGCTGGGGT TTTTGAATC AATCCCACTA CCAAGTGCAT

7801 GTGGGCCATC GCCCTGATAG ACGGTTTTTC GCCCTTTGAC GTTGGAGTCC ACGTTCTTTA  
CACCCGGTAG CGGGACTATC TGCCAAAAG CGGGAAACTG CAACCTCAGG TGCAAGAAAT

7861 ATAGTGGACT CTTGTTCCAA ACTGGAACAA CACTCAACCC TATCTCGGTC TATTCTTTTG  
TATCACCTGA GAACAAGGTT TGACCTTGTT GTGAGTTGGG ATAGAGCCAG ATAAGAAAAC

7921 ATTTATAAGG GATTTTGCCG ATTTCCGGCT ATTGTTTAAA AAATGAGCTG ATTTAACAAA  
TAAATATTCC CTAAAACGGC TAAAGCCGGA TAACCAATTT TTTACTCGAC TAAATTGTTT

7981 AATTTAACGC GAATTTTAAC AAAATATTAA CGCTTACAAT TTGCCATTCG CCATTCAGGC  
TTAAATTGCG CTTAAAATTG TTTTATAATT GCGAATGTTA AACGGTAAGC GGTAAGTCCG

8041 TGCGCAACTG TTGGGAAGGG CGATCGGTGC GGGCCTCTTC GCTATTACGC CAGCCCAAGC  
ACGCGTTGAC AACCCCTCCC GCTAGCCACG CCCGGAGAAG CGATAATGCG GTCGGGTTCG

8101 TACCATGATA AGTAAGTAAT ATTAAGGTAC GGGAGGTACT TGGAGCGGCC GCAATAAAAT  
ATGGTACTAT TCATTCATTA TAATCCATG CCCTCCATGA ACCTCGCCGG CGTTATTTTA

8161 ATCTTTATTT TCATTACATC TGTGTGTTGG TTTTTGTGT GAATCGATAG TACTAACATA  
TAGAAATAAA AGTAATGTAG ACACACAACC AAAAAACACA CTAGCTATC ATGATTGTAT

8221 CGCTCTCCAT CAAAACAAAA CGAAACAAAA CAAACTAGCA AAATAGGCTG TCCCCAGTGC  
GCGAGAGGTA GTTTTGTTTT GCTTTGTTTT GTTTGATCGT TTTATCCGAC AGGGGTCACG

8281 AAGTGCAGGT GCCAGAACAT TTCTCTATCG ATA -3`  
TTCACGTCCA CGGTCTTGTA AAGAGATAGC TAT -5`