

Molekulargenetische Analyse der equinen *CRISP* Gene und ihrer Rolle in der Fruchtbarkeit des Hengstes

Von dem Fachbereich Biologie
der Universität Hannover

zur Erlangung des Grades eines
Doktors der Naturwissenschaften

Dr. rer. nat.

genehmigte Dissertation

von

Dipl. Biol. Alexander Giese

geboren am 22.07.1971, in Kassel

2002

Referent: Prof. Dr. Tosso Leeb
Korreferent: Prof. Dr. Achim Gossler
Tag der Promotion: 29.11.2002

Meiner Familie

Zusammenfassung

Die Familie der Cystein-reichen sekretorischen Proteine (CRISP) umfasst drei Mitglieder mit der Bezeichnung AEG1, TPX1 und AEG2, die durch 16 hochkonservierte Cystein-Reste im C-terminalen Bereich der Proteine charakterisiert werden. Die CRISP Proteine werden primär in verschiedenen Teilen des männlichen Genitaltraktes exprimiert. Sie sind vermutlich an Zell-Zell Interaktionen von männlichen Keimzellen mit anderen Zellen während der Spermienreifung oder während der Befruchtung beteiligt. Aufgrund dessen erscheinen die *CRISP* Gene als aussichtsreiche Kandidatengene für die Vererbung von männlichen Fruchtbarkeitsstörungen und als mögliche QTL (quantitative trait loci) für die männliche Fruchtbarkeit. In der vorliegenden Arbeit wurden 227 kb genomischer Pferde-DNA vom equinen Chromosom (ECA) 20q22 mit den Genen *AEG1*, *TPX1* und *AEG2* kloniert und sequenziert. Das equine *AEG1* Gen umfasst acht Exons, die sich über 31 kb erstrecken. Die Analyse der equinen *AEG1* Transkripte zeigte keine Hinweise auf alternative Spleißprodukte, es wurden jedoch drei verschiedene Transkriptionsstartstellen gefunden. Die erste Transkriptionsstartstelle wurde 20 nt stromabwärts einer TATA-Box lokalisiert. RT-PCR Analysen zeigten, dass das *AEG1* in verschiedenen Teilen der Epididymis exprimiert wird. Im Hoden konnte es nur in geringen Mengen nachgewiesen werden. Das equine *TPX1* Gen umfasst 10 Exons, die sich über 18 kb erstrecken, während das *AEG2* Gen aus acht Exons besteht, die sich auf mehr als 24 kb verteilen. Die Expression von *TPX1* im Hoden und von *AEG2* in der Speicheldrüse wurde bestätigt. Mit Hilfe der RT-PCR wurde ein alternativ gespleißtes Transkript von *AEG2* gefunden, in welchem das zweite Exon fehlte. Alle drei *CRISP* Gene liegen beim Pferd auf ECA 20q22 an der selben chromosomalen Lokalisation. Die orthologen Gene liegen beim Menschen auf HSA 6p21 und bei der Maus auf MMU 17. Dies bestätigt die konservierte Syntanie in diesem Bereich zwischen diesen Spezies. Sequenzvergleiche zeigten außerdem, dass das intronlose equine *PGK2* Gen für die Testis-spezifische Phosphoglyceratkinase etwa 39 kb stromabwärts des *AEG1* Gens ebenfalls in dieser Region lokalisiert ist.

Schlagworte: Pferd, Fruchtbarkeit, CRISP

Abstract

The cysteine-rich secretory protein (CRISP) family consists of three members called AEG1, TPX1 and AEG2, which share 16 conserved cysteine residues at their C-termini. The CRISP proteins are primarily expressed in different sections of the male genital tract and are thought to mediate cell-cell interactions of male germ cells with other cells during sperm maturation or during fertilization. Therefore, their genes are of interest as candidate genes for inherited male fertility dysfunctions and as putative quantitative trait loci (QTL) for male fertility traits. In this report, the cloning and DNA sequence of 227 kb of horse genomic DNA from equine chromosome 20q22 containing the equine *AEG1*, *TPX1* and *AEG2* genes are described. The equine *AEG1* gene consists of eight exons spanning 31 kb. Analysis of equine *AEG1* transcripts did not reveal any evidence for alternative splicing, however three different transcription start sites are used. The first transcription start site is located 20 nt downstream of a TATA box motif. RT-PCR analysis demonstrated that *AEG1* is expressed in different parts of the epididymis, whereas it is hardly detectable in testis. The equine *TPX1* gene consists of 10 exons spanning 18 kb while the *AEG2* gene consists of eight exons that are spread over 24 kb. The expression of *TPX1* in testis, and of *AEG2* in the submandibular gland was confirmed. RT-PCR revealed an alternatively spliced transcript for the *AEG2* gene, in which exon 2 is missing. Comparative genome analysis between horse, human and mouse indicates that all three *CRISP* genes are clustered on one chromosomal location, which shows conserved synteny between these species. Finally, sequence comparisons revealed that the intronless equine *PGK2* gene for the testis-specific phosphoglycerate kinase is located approximately 39 kb downstream of the *AEG1* gene.

Keywords: Horse, fertility, CRISP

Inhaltsverzeichnis

1 EINLEITUNG	1
2 LITERATURÜBERSICHT	2
2.1 DIE FAMILIE DER CYSTEIN-REICHEN SEKRETORISCHEN PROTEINE (CRISPs)	2
2.2 DAS CYSTEIN-REICHE SEKRETORISCHE PROTEIN 1 (CRISP1, AEG1)	5
2.3 DAS CYSTEIN-REICHE SEKRETORISCHE PROTEIN 2 (CRISP2, TPX1)	7
2.4 DAS CYSTEIN-REICHE SEKRETORISCHE PROTEIN 3 (CRISP3, AEG2)	8
2.5 FRUCHTBARKEIT BEIM PFERD	9
3 MATERIAL	12
3.1 BAKTERIEN, PLASMIDE, GENBANKEN	12
3.1.1 Bakterien	12
3.1.2 Plasmide	12
3.1.3 Genbanken	12
3.2 ENZYME	13
3.2.1 Enzyme	13
3.2.2 Oligonukleotide	13
3.2.3 Chemikalien und Verbrauchsmaterialien	13
3.3 LABORGERÄTE	14
4 METHODEN	17
4.1 VERMEHRUNG UND ISOLIERUNG VON PLASMID-DNA	17
4.1.1 Bakterienkultur (Ausubel et al., 1990)	17
4.1.2 Isolierung von Plasmid-DNA aus E. coli	18
4.1.3 Isolierung von BAC-DNA aus E. coli	22
4.1.4 Herstellung transformationskompetenter E. coli XL1-Blue (Hanahan, 1983)	23
4.1.5 Transformation kompetenter E. coli	24
4.1.6 Erstellen einer Shotgun Plasmid Genbank mit Hilfe des TOPO Shotgun Subkloning Kits	26
4.2 ENZYMATISCHE MANIPULATION UND ANALYSE VON DNA	27
4.2.1 Auftrennung von kleinen DNA-Fragmenten	27
4.2.2 Pulsfeldgelelektrophorese (Schwartz und Cantor, 1984)	28
4.2.3 Isolierung und Reinigung von DNA-Fragmenten aus Agarosegelen mittels QIAEX II-Kit	29
4.2.4 Isolierung und Reinigung von DNA-Fragmenten aus Lösungen mittels QIAEX II-Kit	29
4.2.5 Enzymatische Manipulation von DNA	30
4.2.6 Alkalischer Transfer von DNA auf Trägermembranen (Reed und Mann, 1985)	33
4.2.7 Hybridisierung und Detektion von DNA-Fragmenten mit ECL direct labelling system	33
4.3 POLYMERASE-KETTENREAKTION (PCR) (SAIKI ET AL., 1988)	34
4.4 DNA-SEQUENZIERUNG	35
4.4.1 Nichtradioaktive Sequenzierung doppelsträngiger DNA	35

4.4.2 Gelelektrophorese von Sequenzprodukten auf einem automatischen DNA-Sequencer	35
4.5 ISOLIERUNG UND ANALYSE VON RNA	36
4.5.1 Isolierung von Gesamt-RNA aus Gewebe	36
4.5.2 Amplifikation von cDNA mit Hilfe des First Choice RLM-RACE Kit.....	36
4.6 BIOINFORMATIK.....	37
4.6.1 Programme (lokal).....	37
4.6.2 Internetseiten	38
5 ERGEBNISSE	40
5.1 CDNA DER EQUINEN <i>CRISP</i> GENE	40
5.1.1 Sequenzanalyse der equinen <i>AEG1</i> cDNA und der abgeleiteten Proteinsequenz	40
5.1.2 Sequenzanalyse der equinen <i>TPX1</i> cDNA und der abgeleiteten Proteinsequenz	43
5.1.3 Sequenzanalyse der equinen <i>AEG2</i> cDNA und der abgeleiteten Proteinsequenz	46
5.2 GENOMSTRUKTUR DER EQUINEN <i>CRISP</i> GENE	49
5.2.1 Sequenzanalyse des <i>AEG1</i> Gens	49
5.2.2 Sequenzanalyse des <i>TPX1</i> Gens.....	52
5.2.3 Sequenzanalyse des <i>AEG2</i> Gens	53
5.3 SEQUENZANALYSE DER FLANKIERENDER REGIONEN	56
5.3.1 Sequenzanalyse des equinen <i>PGK2</i> Gens	56
5.3.2 Sequenzanalyse der flankierenden Regionen der equinen <i>CRISP</i> Gene	57
5.3.3 Analyse der Mikrosatelliten im Bereich der equinen <i>CRISP</i> Gene.....	58
5.4 CHROMOSOMALE LOKALISATION DER EQUINEN <i>CRISP</i> GENE.....	59
6 DISKUSSION	61
6.1 INTERPRETATION DER SEQUENZDATEN.....	61
6.1.1 Diskussion der Sequenzdaten des <i>AEG1</i> Gens.....	61
6.1.2 Diskussion der Sequenzdaten des <i>TPX1</i> Gens	63
6.1.3 Diskussion der Sequenzdaten des <i>AEG2</i> Gens.....	65
6.2 ANALYSE DER CHROMOSOMALEN LOKALISATION DER EQUINEN <i>CRISP</i> GENE.....	66
6.3 ANALYSE DER ZUVERLÄSSIGKEIT DER SEQUENZEN.....	68
6.4 BEDEUTUNG DER ARBEIT FÜR DAS PFERDE-GENOMPROJEKT	69
6.5 AUSBLICK AUF WEITERFÜHRENDE UNTERSUCHUNGEN.....	69
7 LITERATURVERZEICHNIS.....	70
8 ANHANG	76
VERWENDETE OLIGONUKLEOTIDE:.....	76
PCR-Primer zum Durchmustern der equinen BAC-Genbank.....	76
Sequenzprimer.....	76
RACE-Primer	79

Verzeichnis der verwendeten Abkürzungen

aa	Aminosäure (engl.: amino acid)
amp	Ampicillin
ARE	Androgenrezeptor Bindeelement (engl.: androgen response element)
ATP	Adenosintriphosphat
BAC	Künstliches Bakterienchromosom (engl.: bacterial artificial chromosome)
bp	Basenpaar
cDNA	komplementäre DNA
DMF	N,N-Dimethylformamid
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure (engl.: desoxyribonucleic acid)
dNTP	Desoxyribonukleosidtriphosphat
ECA	equines Chromosom
EDTA	Ethylendiamintetraacetat
FISH	Fluoreszenz <i>in situ</i> Hybridisierung
HSA	humanes Chromosom
kana	Kanamycin
kb	Kilobasen
mRNA	Boten-RNA (engl.: messenger RNA)
MMU	murines Chromosom
nt	Nukleotide
OAc	Acetat
PCR	Polymerase-Kettenreaktion (engl.: polymerase chain reaction)
QTL	Quantitativer Genort (engl.: quantitative trait locus)
RACE	Amplifikation von cDNA Enden (engl.: rapid amplification of cDNA ends)
RNA	Ribonukleinsäure (engl.: ribonucleic acid)
RT-PCR	Reverse Transkriptase-Polymerase-Kettenreaktion
SDS	Natriumdodecylsulfat (engl.: sodium dodecylsulfate)
snRNP	kleine Kern-Ribonucleopartikel (engl.: small nuclear ribonucleoprotein particles)
TBE	Tris- Borat-EDTA-Lösung
TE	Tris-EDTA-Lösung
tet	Tetracyclin
Tris	Tris-(hydroxymethyl)-aminomethan
U	Einheit (engl.: unit)
Upm	Umdrehungen pro Minute
X-Gal	5-Chlor-4-Brom-3-indolyl- β -D-Galaktosid

Verzeichnis der Abbildungen

Abbildung 1:	Genomstruktur des <i>Aeg1</i> Gens der Ratte.	6
Abbildung 2:	<i>AEG1</i> cDNA mit Translation.	41
Abbildung 3:	Alignment des equinen AEG1 Proteins mit den Sequenzen der bekannten orthologen AEG1 Proteine.	42
Abbildung 4:	<i>TPX1</i> cDNA mit Translation.	44
Abbildung 5:	Alignment des equinen TPX1 Proteins mit den Sequenzen der bekannten orthologen TPX1 Proteine.	45
Abbildung 6:	<i>AEG2</i> cDNA mit Translation.	47
Abbildung 7:	Alignment des equine AEG2 Proteins mit den Sequenzen der bekannten orthologen AEG2 Proteine.	48
Abbildung 8:	Genomische Organisation des equinen <i>AEG1</i> Locus.	50
Abbildung 9:	Schematische Darstellung des <i>AEG2</i> und <i>TPX1</i> BAC-Contig.	52
Abbildung 10:	Genomische Organisation des equinen <i>TPX1</i> und <i>AEG2</i> Locus.	55
Abbildung 11:	Chromosomale Lokalisation der untersuchten Gene durch FISH Analyse.	60
Abbildung 12:	Darstellung der <i>AEG1</i> Promotorregion von Pferd und Maus.	62
Abbildung 13:	Darstellung der equinen <i>TPX1</i> Promotorregion.	64
Abbildung 14:	Darstellung der equinen <i>AEG2</i> Promotorregion.	66
Abbildung 15:	Vergleichende Genkarte der <i>CRISP</i> Genregion zwischen Mensch, Maus und Pferd.	67

Verzeichnis der Tabellen

Tabelle 1:	Übersicht über die Nomenklatur der CRISP Proteinfamilie	2
Tabelle 2:	Exon/Intron Übergänge des equinen <i>AEG1</i> Gens.	51
Tabelle 3:	Exon/Intron Übergänge des equinen <i>TPX1</i> Gens.	53
Tabelle 4:	Exon/Intron Übergänge des equinen <i>AEG2</i> Gens.	54
Tabelle 5:	Verteilung der Mikrosatelliten im Bereich der equinen <i>CRISP</i> Gene.	58

1 Einleitung

Spermien durchlaufen während ihrer Passage durch den männlichen Genitaltrakt einen Reifungsprozess, der die Bindung verschiedener Seminalplasmaproteine auf der Oberfläche der Spermienmembran umfasst. Eine wichtige Klasse der daran beteiligten Plasmaproteine sind die Cystein-reichen sekretorischen Proteine (CRISPs). Charakterisierend für diese Proteinfamilie sind 16 hochkonservierte Cysteine im C-terminalen Bereich der Proteine, die eine charakteristische Struktur von Disulfidbrücken bilden. Zur Zeit sind bei den Säugern drei Mitglieder der CRISP Familie beschrieben, die als AEG1, TPX1 und AEG2 bezeichnet werden. Die Nomenklatur dieser Proteine und ihrer codierenden Gene ist unübersichtlich, da sie von verschiedenen Forschergruppen beim Mensch und in Nagetieren unterschiedlich benannt wurden. So wird das *AEG1* Gen beim Menschen auch als *CRISP1* oder *AEGL1* bezeichnet, während es bei der Ratte als Gen für das Nebenhodenprotein DE bekannt ist. Das *TPX1* Gen ist auch als *CRISP2* oder *AA1* Gen bekannt, und das *AEG2* Gen wird auch als *CRISP3* oder *SGP28* Gen bezeichnet. Die CRISP Proteine sind vermutlich an der Fusion von Spermium und Eizelle beteiligt. Sie werden in verschiedenen Geweben des männlichen Genitaltraktes unterschiedlich stark exprimiert und in das Seminalplasma sezerniert. Aus dem Seminalplasma assoziieren die CRISP Proteine dann an den Kopf reifer Spermien. Werden die CRISP Proteine durch spezifische Antikörper blockiert, wird eine Fusion von Spermium und Eizelle verhindert, obwohl die Spermien nach der Akrosomenreaktion normal an die Eizelle binden. Des weiteren wurde festgestellt, dass die Menge der auf Hengstspermien gebundenen CRISP Proteine mit der Befruchtungsleistung dieser Hengste korreliert ist. Die CRISP-Menge guter Befruchter ist bis zu 10 mal höher als bei subfertilen Hengsten. Diese Reduktion der messbaren CRISP-Menge könnte durch eine Mutation im Promoterbereich eines *CRISP* Gens oder durch eine Mutation im codierenden Bereich und daraus resultierendem Aminosäureaustausch verursacht werden. Aufgrund dieser Daten sind die *CRISP* Gene vielversprechende Kandidaten für Genorte mit einem Einfluss auf die Fruchtbarkeit von Hengsten.

2 Literaturübersicht

2.1 Die Familie der Cystein-reichen sekretorischen Proteine (CRISPs)

Die Cystein-reichen sekretorischen Proteine (CRISPs) repräsentieren eine evolutionär hoch konservierte Familie, die zuerst im männlichen Genitaltrakt von Nagern (Cameo and Blaquier, 1976; Lea et al., 1978; Kasahara et al., 1989b) und später auch in der Speicheldrüse gefunden wurde (Mizuki und Kasahara, 1992; Haendler et al., 1993). Der Name der CRISP Proteine wurde von einer Gruppe Cystein-Reste in ihrem C-terminalen Bereich abgeleitet. Alle diese Cystein-Reste bilden Disulfidbrücken, wobei die sechs in N-terminaler Richtung lokalisierten Cystein Reste untereinander drei Bindungen eingehen. Dies weist auf die Existenz von zwei diskreten Domänen innerhalb des Proteins hin (Eberspaecher et al., 1995).

Tabelle 1: Übersicht über die Nomenklatur der CRISP Proteinfamilie.

Protein	synonyme Namen	Spezies	Länge	Literatur
CRISP1	ARP, AEGL1	Mensch	249 aa	Krätzschmar et al., 1996
	MEP7, Aeg1	Maus	244 aa	Mizuki und Kasahara, 1992
	D, E	Ratte	246 aa	Brooks et al., 1986
CRISP2	TPX1	Mensch	243 aa	Krätzschmar et al., 1996
	Tpx1	Maus	243 aa	Kasahara et al., 1989a
	Tpx1	Ratte	243 aa	Maeda et al., 1998
	AA1	Meerschweinchen	244 aa	Foster und Gerton, 1996
CRISP3	SGP28	Mensch	245 aa	Krätzschmar et al., 1996; Kjeldsen et al., 1996
	Aeg2	Maus	241 aa	Mizuki und Kasahara, 1992
	AEG2	Pferd	244 aa	Schambony et al., 1998

AEG1 (acidic epididymal glycoprotein 1, CRISP1) wird hauptsächlich in der Epididymis gebildet und repräsentiert 15 % des Proteinanteils des epididymalen Sekrets. Bisherige Veröffentlichungen zum murinen Epididymis AEG1 Protein, das 68 % Sequenzübereinstimmung zum humanen AEG1 aufweist, berichten von einer schwachen Bindung des Proteins an Spermien und weisen außerdem darauf hin, dass das Protein von hellen Zytoplasmazellen der Epididymis aufgenommen wird (Rankin et al., 1992; Vierula et al., 1995; Eberspaecher et al., 1995). Die funktionelle Bedeutung dieses Phänomens ist jedoch noch unbekannt.

Das zum AEG1 Protein orthologe epididymale Protein DE der Ratte wurde bisher am genauesten untersucht (Hinton and Palladino, 1995). Die zu AEG1 orthologen Proteine von Ratte und Maus weisen eine 70 %ige Sequenzübereinstimmung inklusive der 16 Cystein-Reste auf. Beide Proteine wurden hauptsächlich in der Epididymis lokalisiert (Haendler et al., 1993; Brooks et al., 1986; Chares et al., 1988). Das Aeg1 bindet vermutlich über einen spezifischen Rezeptor an den Kopf von Spermien. (Brooks and Tiver, 1983, Brooks, 1987). Die Beteiligung von Aeg1 an der Rosettenbildung von Spermien, so wie der Anwesenheit von komplementären Bindungsstellen auf Oozyten wurde ebenfalls nachgewiesen (Fornes und Burgos, 1994; Rochwerger et al., 1992). Die Immunisierung von Ratten mit Aeg1 führte zur Ausbildung von anti-Aeg1 Antikörpern, was zu einer signifikanten und reversiblen Reduktion der Fruchtbarkeit führte (Perez Martinez et al., 1995). Dies lässt eine Beteiligung von AEG1 an der Fusion von Spermium und Oocyte vermuten. Die genaue Funktion von AEG1 ist jedoch zur Zeit noch ungeklärt. Im Hoden wurden ebenfalls Transkripte von *CRISP* Genen identifiziert. Das *TPX1* (*CRISP2*) Gen, dessen Protein bei der Maus eine Homologie von 47 % zu AEG2 und 55 % zu AEG1 aufweist, wird dort hoch und spezifisch in den haploiden männlichen Keimzellen exprimiert (Kasahara et al., 1989b; Haendler et al., 1993; Mizuki et al., 1992). Es wird vermutet, dass TPX1 am Zell-Zell Kontakt zwischen für die Spermatogenese wichtigen Zellen und Sertoli Zellen beteiligt ist (Maeda et al., 1998). Bei Untersuchungen von TPX1 im Meerschweinchen wurde herausgefunden, dass TPX1 bis zu 6,4 % der akrosomalen Proteine von Spermien repräsentiert. Die Blockierung von TPX1 mit spezifischen Antikörpern zeigte beim Meerschweinchen jedoch keinen Einfluss auf *in vitro* Befruchtungsraten (Hardy et al., 1988). Auch zur genauen Funktion von TPX1 gibt es zur Zeit noch keine Hinweise.

Das dritte Mitglied der CRISP Familie wurde in der Speicheldrüse nachgewiesen. Neben geringen Mengen AEG1 wurde dort eine hohe Expression von AEG2 (CRISP3) gefunden (Mizuki und Kasahara, 1992; Haendler et al., 1993). Das AEG2 Protein weist bei der Maus 77 % Sequenzidentität zu AEG1 auf, und beinhaltet ebenfalls die 16 konservierten Cystein-Reste (Haendler et al., 1993). Durch Kastration und Hormonunterdrückung wurde gezeigt, dass die Expression von *AEG1* und *AEG2* in der Speicheldrüse androgenabhängig ist (Mizuki und Kasahara, 1992; Haendler et al., 1993). Dies wurde durch die unterschiedlich hohe Expression in männlichen und weiblichen Individuen bestätigt. In der Promotorregion des murinen *AEG2* wurden vermutliche Androgenrezeptor Bindeelemente (ARE) identifiziert (Schwidetzky et al., 1995).

Aufgrund der konservierten Struktur der Domänen wird davon ausgegangen, dass *AEG1*, *AEG2* und *TPX1* von einem einzigen Urgen abstammen. Die *CRISP* Gene sind eng gekoppelt auf dem murinen Chromosom 17 beziehungsweise auf dem humanen Chromosom 6 lokalisiert und liegen in enger Nachbarschaft zur genomischen Region des Haupthistokompatibilitäts-Komplex (Kasahara et al., 1989a; Kasahara et al., 1989b; Kasahara et al., 1995; Hayashi et al., 1996).

Es wird außerdem vermutet, dass die CRISP Proteine potenziell eine Rolle in der angeborenen Immunantwort haben könnten. Sie werden in spezifischen humanen und murinen Immunzellen, in Drüsen mit exokrinen Funktionen und auf Mukosaepitheloberflächen exprimiert (Mizuki, et al., 1992; Haendler et al., 1993; Kjeldsen et al., 1996; Pfisterer et al., 1996; Eberspaecher et al., 1995; Krätzschar et al., 1996). Dies weist auf eine mögliche Beteiligung an der unspezifischen Immunantwort hin.

2.2 Das Cystein-reiche sekretorische Protein 1 (CRISP1, AEG1)

Aeg1 wurde zuerst aus der Epididymis der Maus isoliert. Dort durchlaufen die Spermien ihren letzten Reifungsprozess und erlangen ihre Bewegungsfähigkeit und die Fähigkeit zur Befruchtung. Dieser epididymale Reifungsprozess umfasst die Remodulierung der molekularen Oberflächenstruktur von Spermien durch die Anlagerung von epididymalen sekretorischen Proteinen und Lipiden. Gleichzeitig werden bereits existierende Plasmamembran-Komponenten des Spermiums entfernt oder verändert (Yanagimachi, 1994; Cooper, 1992). Das *AEG1* Protein weist eine Struktur mit zwei Domänen auf. Die N-terminale Domäne beinhaltet sechs paarweise verbundene Cysteine und die C-terminale Domäne zehn Cysteine, die einen für die CRISP Familie charakteristischen Abstand aufweisen. Transkripte des *Aeg1* Gens wurden in der Epididymis und in geringer Konzentration auch in der Speicheldrüse von männlichen Mäusen nachgewiesen. Die Expression der *Aeg1* Transkripte in der Speicheldrüse weiblicher Mäuse war weitaus geringer als in den entsprechenden Geweben der männlichen Tiere (Mizuki et al., 1992b; Haendler et al., 1993). Die Proteinsequenz des murinen *Aeg1* zeigt eine Sequenzübereinstimmung von 41 % zum humanen *AEG1* und 70 % zum *Aeg1* der Ratte. Die beiden orthologen Proteine von Mensch und Ratte werden ebenfalls in großen Mengen in der Epididymis exprimiert (Krätzschar et al., 1996; Hayashi et al., 1996; Cameo et al., 1976; Brooks et al., 1986; 1987).

Das *AEG1* Gen wurde bisher von der Maus und der Ratte kloniert (Schwidetzky et al., 1997; Klemme et al., 1999). Das murine *Aeg1* Gen besteht aus acht Exons mit einer Länge von 53 bis 751 bp, die durch sieben Introns unterbrochen werden. In der Promotorregion des murinen *Aeg1* wurde 29 bp stromaufwärts des Transkriptionsstartes eine TATA Box beschrieben. Es wurden 12 AREs identifiziert, deren funktionale Bedeutung experimentell bestätigt wurde. Hiermit wurde eine androgenabhängige Regulation des *Aeg1* Promotors nachgewiesen. (Haendler et al., 1997, 2001). Als weitere interessante Elemente wurden eine Poly(CA) und eine Poly(CT) Wiederholung identifiziert, deren genaue Bedeutung jedoch nicht bekannt ist (Schwidetzky et al., 1997). Im 3'-untranslatierten Bereich des murinen *Aeg1* Gens, der im Exon 8 enthalten ist, wurde ein kanonisches Polyadenylierungssignal lokalisiert.

Anders als im murinen *Aeg1* Gen wurde bei der Ratte ein weiteres 5'-gelegenes Exon gefunden. Somit beinhaltet das *Aeg1* Gen der Ratte eine Genomstruktur mit neun Exons, die durch acht Introns unterbrochen werden. Es wurden zwei Transkriptionsstartstellen identifiziert, die vor dem 1. Exon beziehungsweise im 1. Intron liegen. Dies lässt eine alternative Transkription des *Aeg1* Gens der Ratte vermuten. Durch RT-PCR Experimente wurde bestätigt, dass beide Transkriptionsstartstellen verwendet werden. Das kürzere Produkt entspricht in Länge und Sequenz den *AEG1* Transkripten, die auch in anderen Spezies konserviert sind. Das längere Produkt enthält zusätzliche Sequenzinformationen im 5'-gelegenen Bereich (Klemme et al., 1999; Roberts et al., 2001).

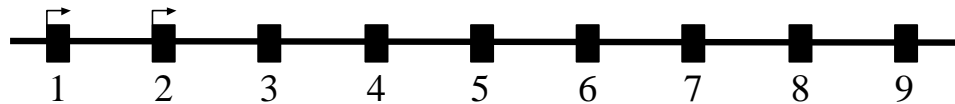


Abbildung 1: Genomstruktur des *Aeg1* Gens der Ratte. Die Exons sind als schwarze Kästen dargestellt. Die horizontalen Pfeile zeigen die Position der ermittelten Transkriptionsstartstellen. Die Exons 2-9 entsprechen den Exons 1-8 des *AEG1* Gens in anderen Spezies.

Die Größe der ermittelten Exons beträgt 70 bp bis 237 bp, wobei der 3'-untranslatierte Bereich im letzten Exon nicht bestimmt wurde. Untersuchungen der regulatorischen Regionen des *AEG1* Gens der Ratte zeigten ebenso wie im Promotor des murinen *Aeg1* Gens eine große Anzahl von AREs.

Das größte Spektrum an verschiedenen alternativen *AEG1* Transkripten wurde bisher beim Menschen gefunden. Hier wurden sowohl Transkripte des normalen AEG1 Proteins als auch Transkripte einer verkürzten Form des Proteins beschrieben. (Krätzschar et al., 1996). Dennoch weisen die verschiedenen humanen *AEG1* Transkripte ähnliche 5'-Enden auf und werden innerhalb eines Exons initiiert. Bei der Maus wurden keine alternativen Transkripte nachgewiesen. Das humane *AEG1* Gen ist auf HSA 6p21.1-p21.2 lokalisiert während das orthologe *Aeg1* Gen bei der Maus auf Chromosom MMU 17 lokalisiert ist (Hayashi et al., 1996; Yoshino et al., 1998).

2.3 Das Cystein-reiche sekretorische Protein 2 (CRISP2, TPX1)

Durch *in situ* Hybridisierung von Teilen des Hodens wurde gezeigt, dass TPX1 in hohem Maße von haploiden männlichen Keimzellen gebildet wird (Haendler et al., 1997). Während der Spermatogenese in Säugetieren treten unterschiedliche an der Spermatogenese beteiligte Zellen in Kontakt mit somatischen Sertoli Zellen (Russel und Petersen, 1985). Diese Assoziation scheint essentiell für den Verlauf der Spermatogenese und für die zyklische Funktion der Sertoli Zellen zu sein (Kierszenbaum, 1994). Es wird angenommen, dass die beiden Zelltypen wichtige Signale austauschen und dieser Austausch möglicherweise durch direkten Zell-Zell Kontakt zustande kommt. Obwohl verschiedenen Zelladhäsionsmoleküle auf der Oberfläche von an der Spermatogenese beteiligten Zellen und Sertoli Zellen nachgewiesen wurden, ist der Mechanismus dieser Interaktion noch immer ungeklärt (Kierszenbaum, 1994). TPX1 wurde im Zusammenhang mit der spezifischen Adhäsion von an der Spermatogenese beteiligten Zellen an Sertoli Zellen untersucht. Hierbei wurde festgestellt, dass diese Adhäsion in Gegenwart eines anti-TPX1 Antikörpers um bis zu 50 % reduziert werden kann (Maeda et al., 1998). Das weist darauf hin, dass TPX1 an der spezifischen Assoziation dieser beiden Zelltypen beteiligt ist. Das *TPX1* Gen wurde bei Mensch und Maus kloniert (Kasahara et al., 1989a; Kasahara et al., 1987; Mizuki et al., 1992a). Die Sequenzinformationen des humanen *TPX1* Gens wurden im Rahmen des humanen Genomprojektes ermittelt, und stehen in der EMBL Nucleotide Sequence Database zur Verfügung (Accession AL121950). Das humane *TPX1* Gen erstreckt sich über 21 kb und umfasst zehn Exons mit einer Länge von 37 bp bis 541 bp. Im Exon 8 des humanen *TPX1* Gens wurde ein Polyadenylierungssignal 18 bp stromaufwärts der Polyadenylierungsstelle lokalisiert.

Die Gesamtlänge des murinen *TPX1* Gens beträgt 35 kb. Es ist in 10 Exons mit einer Länge von 50 bp bis 601 bp und 9 Introns unterteilt (Kasahara et al., 1987; Mizuki et al., 1992a). Der 5'-untranslatierte Bereich des murinen *TPX1* Gens erstreckt sich über vier Exons. Der kodierende Bereich umfasst die Exons 4 bis 10. Das Exon 4 enthält einen Teil des 5'- untranslatierten Bereichs und kodiert für das Signalpeptid sowie den N-terminalen Bereich des reifen TPX1 Proteins. Das Exon 10 enthält den gesamten 3'- untranslatierten Bereich. Zwischen Exon 8 und Exon 9 wurde ein ungewöhnlich langes Intron von 17 kb

Länge beschrieben (Mizuki et al., 1992a). Im Gegensatz zu den *AEG1* und *AEG2* Genen enthält der *TPX1* Promotor keine AREs und beinhaltet keine TATA-Box.

Von *TPX1* wurde außerdem die cDNA-Sequenz beim Meerschweinchen und der Ratte ermittelt (Foster and Gerton, 1996; O'Bryan et al., 1998).

2.4 Das Cystein-reiche sekretorische Protein 3 (CRISP3, AEG2)

Das AEG2 Protein wurde zuerst in der Speicheldrüse von Mensch und Maus gefunden wo es in den Zellen der granularen Tubuli gebildet wird (Mizuki et al., 1992b). In der Speicheldrüse wurde eine geschlechtsspezifische Expression von verschiedenen Genen beschrieben, die einer androgenabhängig kontrollierten Expression unterliegen. Diese androgenabhängige Kontrolle der Expression weist darauf hin, dass die beteiligten Gene ebenfalls im Genitaltrakt exprimiert werden. Diese Annahme wurde für das *AEG2* Gen bestätigt. In der Ratte wird das *Aeg2* Gen in der Epididymis hoch exprimiert, in der Speicheldrüse wurde jedoch eine niedrige Expression beschrieben (Brooks et al., 1987). Interessanterweise ist bei der Maus und im Menschen die *AEG2* Expression in der Speicheldrüse viel höher als in anderen Geweben. *AEG2* Transkripte wurden in vielen verschiedenen Geweben und Zelltypen beschrieben wie zum Beispiel unreifen B-Zellen, humanen neutrophilen Granulozyten, Dickdarm, Bauchspeicheldrüse, Eierstock und der Prostata (Krätzschar et al., 1996). Die murine *Aeg2* cDNA codiert für ein Protein mit 241 aa Länge. Das murine *Aeg2* Protein zeigt eine Sequenzübereinstimmung von 78 % zum murinen *Aeg1* und 55 % zum *Tpx1* Protein (Schwidetzky et al., 1995).

Das *Aeg2* Gen wurde bisher nur in der Maus beschrieben (Schwidetzky et al., 1995). Die Genomstruktur des murinen *Aeg2* Gens weist acht Exons zwischen 55 bp und 740 bp auf. Das erste Exon enthält den gesamten 5'-untranslatierten Bereich. Im zweiten Exon wurde der kodierende Bereich für das 20 aa umfassende Signalpeptid und vier der sechs N-terminalen Cysteine lokalisiert. Die Exons 7 und 8 kodieren den cysteinreichen C-terminalen Bereich des AEG2 Proteins. Der 3'-untranslatierte Bereich sowie das Polyadenylierungssignal sind in Exon 8 enthalten. Der Transkriptionsstart wurde 57 bp stromaufwärts des ATG-Startcodons lokalisiert (Schwidetzky et al., 1995). Der Promotor

beinhaltet eine TATA-Box sowie einige experimentell nachgewiesene AREs die auf eine androgenabhängige Expression hinweisen (Haendler et al., 1997; Haendler et al., 2001). Die cDNA Sequenz des *AEG2* Gens wurde beim Pferd schon beschrieben (Schambony et al., 1998b). Im Unterschied zu Mensch und Maus wird das equine *AEG2* Gen sehr stark in der Ampulle exprimiert und es ist möglich eine große Menge *AEG2* Protein aus dem Seminalplasma des Hengstes zu isolieren. Diese speziesspezifische Expression des *AEG2* Gens führt zu der Annahme, das *AEG2* eine besondere Rolle in der Fortpflanzung von Pferden spielt (Schambony et al., 1998a).

2.5 Fruchtbarkeit beim Pferd

Die Fruchtbarkeit hat eine sehr hohe wirtschaftliche Bedeutung in der Nutztierzucht. Da die Kosten für die Erzeugung eines Fohlens im Vergleich zu anderen Nutztieren sehr hoch sind, ist die Fruchtbarkeit von Pferden ein wichtiges Leistungsmerkmal in der Pferdezucht. Die Fruchtbarkeit von Pferden war jedoch viele Jahre kein bedeutendes Selektionskriterium für Zuchtpferde, da diese Tiere besonders wegen ihrer Arbeitsleistung beziehungsweise Leistung als Sportpferd und dem Exterieur für die Zucht ausgewählt wurden. In früheren Zeiten wurde das Pferd als vielseitiges Arbeitstier benötigt, das vor allem eine hohe Arbeitleistung erbrachte. In der heutigen Zeit werden die Tiere hauptsächlich im Sport oder in der Freizeit eingesetzt. Hier sind Äußerungen des Geschlechtsverhaltens wie Rosse oder Hengstigkeit häufig unerwünscht. Eine Selektion der Tiere auf die Fruchtbarkeit ist schwierig, da es zur Zeit keine direkte Testmöglichkeit für die Fruchtbarkeit gibt. Momentan werden die Trächtigkeits- und Abfohlrate als Hilfsmerkmale für die Beurteilung der Fruchtbarkeit verwendet. Die Fruchtbarkeit wird jedoch durch komplexe Faktoren beeinflusst, die aus einer paternalen, einer maternalen, einer fetalen sowie umweltbedingte Komponenten bestehen.

Aufgrund dieser Umstände weist das Pferd die geringste Fortpflanzungsrate unter den landwirtschaftlichen Nutztieren auf (Merkt und Klug, 1989). Bei Rindern liegt die Trächtigkeitsrate bei 95% und die Abkalberate bei 91,9 % (Christmas, 2001). Bei Schwein und Ziege liegen die Befruchtungsergebnisse zwischen 96 % bis 97 % (Dohms, 2001).

Pferde haben dagegen eine Befruchtungsrate der gedeckten Stuten von 79 % und eine Abfohlrate der gedeckten Stuten bei 70 %. In der breiten Landeszucht liegt in der Bundesrepublik Deutschland die Abfohlrate nur bei 60 % der gedeckten Stuten (Merkt, 1995). Da auch mit schwierigen Problemstuten und subfertilen Hengsten Züchtungsversuche unternommen werden, besteht die Gefahr einer Selektion auf Subfertilität (Merkt, 1995).

Die Verbesserung der Fruchtbarkeit stellt daher einen wichtigen ökonomischen Faktor für die Zuchtbetriebe dar. Des weiteren ermöglicht eine verbesserte Bestandsentwicklung eine erhöhte Selektionsintensität und dadurch eine Steigerung des züchterischen Fortschrittes in der Pferdezucht.

Durch den zunehmenden Einsatz der künstlichen Besamung in der Pferdezucht nimmt die Fortpflanzungsleistung von Hengsten einen immer größeren Stellenwert ein (Dohms et al., 1999). Die künstliche Besamung hat sich zum Erreichen optimaler Befruchtungsergebnisse in Form der Frischsamenübertragung zur zentralen Biotechnologie auf dem Sektor der equinen Reproduktion entwickelt (Sieme, 1996). Durch Einführung der künstlichen Besamung und eine zunehmende Intensivierung des Zuchtmanagements konnte die Fruchtbarkeitsleistung in der Reitpferdezucht in den letzten 20 Jahren deutlich verbessert werden. Die Abfohlraten wurden jedoch im Vergleich von 1821 bis 1830 mit der Abfohlrate von 1961 bis 1970, trotz wesentlicher Fortschritte in der Veterinärmedizin nicht verbessert und lagen bei 54,3 % (Jacobs, 1977). Die Befruchtungsleistung von Besamungshengsten liegt bis zu 5 % höher als die von Hengsten, die im Natursprung decken. Eine absolute Erfassung der Befruchtungsrate von Hengsten erweist sich als schwierig, da nicht nur die Stute und das Management, sondern auch eine Schwankung der Befruchtungsrate von Jahr zu Jahr einen Einfluss auf die Fruchtbarkeit hat.

Die Fertilität des Hengstes ist ein wichtiger Faktor für den Erfolg und die Wirtschaftlichkeit eines Zuchtbetriebes. Die Befruchtungsfähigkeit eines Hengstes kann im Rahmen einer andrologischen Untersuchung annähernd prognostiziert werden (Klug, 2002). Dabei wird eine morphologische Untersuchung der Geschlechtsorgane, eine Beobachtung des Paarungsverhaltens, eine Untersuchung des Samens sowie eine mikrobielle Untersuchung der Genitalsekrete durchgeführt. Um eine genaue Beurteilung der Fruchtbarkeit eines Hengstes zu erlangen, muss diese retrospektiv anhand der tatsächlich erbrachten Befruchtungsleistung bewertet werden, wobei aber alle anderen

einflussnehmenden Faktoren wie Fruchtbarkeit der Stute und Management des praktischen Zuchtbetriebs in die Analyse einfließen müssen. Eine genetische Testmöglichkeit von fruchtbarkeitsbestimmenden Faktoren würde dabei eine wichtige Ergänzung zur Selektion und Züchtung von Pferden darstellen.

3 Material

3.1 Bakterien, Plasmide, Genbanken

3.1.1 Bakterien

Aus der Stammsammlung des Instituts für Tierzucht und Vererbungsforschung, Abteilung Molekulargenetik wurde der Bakterienstamm *E. coli* XL1-Blue mit den genetischen Eigenschaften $\Delta(mcrA)183 \Delta(mcrCB-hsd\ SMR-mrr)173$, *endA1*, *sup E44*, *thi-1 gyrA96*, *tellA1*, *lac* [*F'proAB*, *lacI^qZ* $\Delta M15$, Tn10 (*tet^r*)] (Stratagene, Heidelberg) verwendet.

Des weiteren wurden die *E. coli* Stämme TOP10 One Shot, chemisch kompetent, und TOP10 One Shot, elektrokompent, (Invitrogen, Groningen, Niederlande) mit den genetischen Eigenschaften *F- mcrA* $\Delta(mrr-hsdRMS-mcrBC)$ $\phi 80lacZ\Delta M15$ $\Delta lacX74deoR$ *recA1 araD139* $\Delta(ara-leu)7697$ *galU galK rpsL* (*Str^R*) *endA1 nupG* verwendet.

3.1.2 Plasmide

Es wurden die Plasmide pBeloBac11 (Wang *et al.*, 1997), pGEMTM-4Z (Promega, Heidelberg), pCR4Blunt-Topo (Invitrogen, Groningen, Niederlande) und pDrive Cloning Vector (QIAGEN, Hilden) verwendet.

3.1.3 Genbanken

BAC-Klone wurden von Dr. Gérard Guérin (INRA, Jouy et Josas, Frankreich) zur Verfügung gestellt.

3.2 Enzyme, Oligonukleotide und Chemikalien

3.2.1 Enzyme

Es wurden Enzyme der Firmen Roche Diagnostics GmbH (Mannheim), New England Biolabs (Schwalbach) und Amersham Biosciences (Freiburg) verwendet.

3.2.2 Oligonukleotide

Für die Sequenzierung wurden folgende fluoreszenzmarkierte Standardprimer verwendet, die ebenso wie unmarkierte Oligonukleotide von der Firma MWG-Biotech GmbH (Ebersberg) synthetisiert wurden:

M13universal (-21), 18-mer: 5'-TGT AAA ACG ACG GCC AGT-3'

M13reverse (-29), 18-mer: 5'-CAG GAA ACA GCT ATG ACC-3'

SP6, 20-mer: 5'-GGA TTT AGG TGA CAC TAT AG-3'

T7, 20-mer: 5'-TAA TAC GAC TCA CTA TAG GG-3'

3.2.3 Chemikalien und Verbrauchsmaterialien

Die eingesetzten Chemikalien und Verbrauchsmaterialien wurden von den Firmen Applicem (Darmstadt), Amersham Biosciences (Freiburg), Biozym Diagnostik GmbH (Hessisch Oldendorf), Clontech Laboratories, Inc (Palo Alto, USA), Eppendorf Gerätebau (Hamburg), Fluka-Feinchemikalien (Neu-Ulm), GibcoBRL Life Technologies (Heidelberg), Greiner GmbH (Nürtingen), E. Merck AG (Darmstadt), Stratagene (Heidelberg), Roth (Karlsruhe), MWG-Biotech GmbH (Ebersberg), Nunc GmbH (Wiesbaden), QIAGEN GmbH (Hilden), Sigma Chemie (München), Schleicher und Schuell GmbH (Dassel) und Whatman Limited (Springfield Mill, USA) bezogen.

3.3 Laborgeräte

- Bachhofer Laboratoriumsgeräte, Reutlingen:
Hybridisierungsöfen
- Benda, Wiesloch:
UV-Flächenstrahler N-90L 366nm
- Biometra , Göttingen:
BioDocAnalyzer
- Bio-Rad Laboratories, München:
Agarosegel-Elektrophorese-Kammern, Pulsfeldgelelektrophorese-System CHEF DR III,
Power Pac 3000
- B. Braun, Melsungen:
Certomat H Schüttler, Certomat R Inkubatorhaube
- Eppendorf Gerätebau, Hamburg:
Heizblock, *Bio*Photometer, Tischzentrifuge 5415D
- EquiBio, Kent (GB):
Easyject T Basic Elektroporator
- Gesellschaft für Laborgeräte mbH, Burgwedel:
Wasserbäder
- Haake, Karlsruhe:
Wasserbad DC10

- Heidolph, Kehlheim:
REAX top Schüttler, Magnetrührer MR2002

- Heraeus Instruments, Osterode:
Megafuge 1.0R (Rotor BS 4402/A), Biofuge Stratos (Rotor 3331), Inkubator

- IKA-Werke, Staufen:
Ultraturrax T25

- Knick Elektronische Messgeräte, Berlin:
Digital-pH-Meter 643

- Kontron Instruments, Neufahrn:
Centrikon H-401 (Rotor A6.14)

- Millipore, Eschborn:
MilliQ biocell, Vacuum manifold MAVM 096 0R

- MWG-Biotech GmbH, Ebersberg:
LI-COR DNA Analyzer Gene Reader 4200

- MJ Research Inc, Watertown (USA):
PTC-100TM und PTC-200TM Programmable Thermal Controller

- Octurno Medizintechnik GmbH, Oldenburg
Standard-Vernebler

- QIAGEN, Hilden:
QIAvac 96 Vacuum Regulator

- Sartorius AG, Göttingen:
Waage BL1500S, Feinwaage 1801

- Savant, München:
Speed Vac Plus SC110A

- Sigma, Osterode:
Sigma 4-15 (Rotor 09100)

4 Methoden

4.1 Vermehrung und Isolierung von Plasmid-DNA

4.1.1 Bakterienkultur (Ausubel et al., 1990)

4.1.1.1 Flüssigkultur

Autoklaviertes L-Broth Medium (LB-Medium) wurde nach Zugabe der entsprechenden Antibiotika, mit Bakterien einer Plattenkultur oder einer anderen Flüssigkultur angeimpft. Die Flüssigkulturen wurden 8-20 Stunden bei 37 °C im Schüttler inkubiert.

LB-Medium: 10 g Trypton
 5 g Hefeextrakt
 5 g NaCl
 ad 1000 ml H₂O bidest

Nach dem Autoklavieren wurden dem Medium je nach Bedarf Antibiotika in folgenden Endkonzentrationen zugesetzt: 50 µg/ml Ampicillin aus einer Stammlösung (50 mg/ml in H₂O bidest., gelagert bei -20 °C). 12,5 µg/ml Tetracyclin aus einer Stammlösung (12,5 mg/ml in 70 % Ethanol, gelagert bei -20 °C). 25 µg/ml Kanamycin aus einer Stammlösung (50 mg/ml in H₂O bidest, gelagert bei -20 °C). Flüssigkulturen wurden bei 4 °C maximal eine Woche aufbewahrt.

4.1.1.2 Plattenkultur

LB-Agar wurde autoklaviert und auf ca. 50 °C abgekühlt, je nach Selektionsanforderung mit Ampicillin (50 µg/ml) und Tetracyclin (12,5 µg/ml) oder mit Kanamycin (25 µg/ml) sowie einer Endkonzentration von 40 µg/ml X-Gal (Stammlösung 40 mg/ml in N,N-Dimethylformamid (DMF)), versetzt und ca. 5 mm hoch in sterile Petrischalen gegossen. Nach dem Erstarren wurden die Platten 40 Minuten lang offen in einer sterilen Laminarhaube getrocknet und danach steril verpackt.

Mittels einer sterilen Pipette wurden 150-250 µl Flüssigkultur auf eine Platte pipettiert und mit einer sterilen Glaspipette ausgestrichen. Die Platten wurden daraufhin 16-24 Stunden bei 37 °C inkubiert.

LB-Agar: 10 g Trypton
 5 g Hefeextrakt
 5 g NaCl
 15 g Bacto-Agar
 ad 1000 ml H₂O bidest

Plattenkulturen wurden bei 4 °C bis zu einem Monat aufbewahrt.

4.1.2 Isolierung von Plasmid-DNA aus *E. coli*

4.1.2.1 Isolierung größerer Mengen Plasmid-DNA mit QIAGEN Plasmid Kits

Um größere Mengen reiner Plasmid-DNA zu isolieren, wurden QIAGEN Plasmid Kits verwendet. Die Bakterien werden dabei durch eine NaOH/SDS-Behandlung lysiert und Proteine sowie assoziierte genomische DNA durch KOAc gefällt. Die weitere Reinigung der Plasmid-DNA erfolgt durch eine Ionenaustauschchromatographie.

20-200 ml Bakterienkultur wurden 15 Minuten bei 5500 Upm und 4 °C in der Sorvall RC 5B (GSA-Rotor) zentrifugiert. Das Bakterienpellet wurde in einer den Angaben des Herstellers entsprechenden Menge Puffer P1 resuspendiert. Zur Lyse der Bakterien wurde das gleiche Volumen Puffer P2 zugegeben, vorsichtig gemischt und 5 Minuten bei Raumtemperatur inkubiert. Danach wurde das gleiche Volumen Puffer P3 zugegeben und geschüttelt. Nach einer Inkubation von 15 Minuten auf Eis, wurde 30 Minuten bei 6000 Upm und 4 °C in der Sorvall RC 5B (GSA-Rotor) zentrifugiert. Der geklärte Überstand wurde auf eine mit Puffer QBT äquilibrierte Qiagen-Säule aufgetragen, mit Puffer QC gewaschen und die Plasmid-DNA mit Puffer QF eluiert. Unmittelbar nach der Zugabe von 0,7 Volumen Isopropanol wurde die DNA 90 Minuten bei 6000 Upm und 4 °C in der Megafuge 1.0R zentrifugiert. Das Pellet wurde mit 3 ml -20 °C kaltem 70 %igem Ethanol gewaschen und an der Luft getrocknet. Die DNA wurde in einem entsprechenden Volumen TE-Puffer (10/1) gelöst und die Konzentration durch Messung der optischen Dichte bei 260 nm im Spektralphotometer bestimmt.

Puffer P1:	50 mM Tris-HCl pH 8,0 10 mM EDTA 100 µg/ml RNase A, gelagert bei 4 °C
RNase- Stammlösung:	10 mg/ml in H ₂ O bidest., zur Inaktivierung von DNAsen 15 Minuten auf 100°C erwärmt; Lagerung bei -20 °C
Puffer P2:	200 mM NaOH 1 % SDS
Puffer P3:	3 M KOAc, pH 5,5
Puffer QBT:	750 mM NaCl 50 mM Na-MOPS pH 7,0 15 % Ethanol 0,15 % Triton X-100

Puffer QC:	1 M NaCl 50 mM Na-MOPS pH 7,0 15 % Ethanol
Puffer QF:	1,25 M NaCl 50 mM Na-MOPS pH 8,5 15 % Ethanol
TE (10/1):	10 mM Tris-HCl pH 8,0 1 mM EDTA

4.1.2.2 Schnellaufschluß durch alkalische Lyse

Um kleinere Mengen Plasmid-DNA zu erhalten wurden 5 ml LB-Medium mit den entsprechenden Antibiotika mit einer einzelnen Plattenkolonie angeimpft und bei 37 °C 8-20 Stunden geschüttelt. Die Kultur wurde bei 4500 Upm 10 Minuten bei 4 °C in der Megafuge 1.0R zentrifugiert. Das Pellet wurde in 150 µl Puffer P1 resuspendiert und in ein 1,5 ml Eppendorf-Reaktionsgefäß überführt. Es wurde 150 µl Puffer P2 zugegeben und ca. 5 Minuten bis zur vollständigen Lyse der Bakterien inkubiert. Nach Zugabe von 150 µl Puffer P3 wurde der Ansatz gemischt und anschließend bei 22000 Upm 15 Minuten bei 4 °C in der Biofuge Stratos zentrifugiert. Der Überstand wurde in ein neues 1,5 ml Eppendorf-Reaktionsgefäß überführt und mit 1 ml 100 %igem Ethanol versetzt. Nach kurzem Mischen wurde der Ansatz anschließend bei 22000 Upm 15 Minuten bei 4 °C in der Biofuge Stratos zentrifugiert. Das Pellet wurde mit 1 ml -20 °C kaltem 70 %igem Ethanol gewaschen und in der Speed Vac 10 Minuten getrocknet. Das Pellet wurde in 50 µl TE aufgenommen.

4.1.2.3 Isolierung von Plasmid DNA im 96 well Maßstab

Um Plasmid DNA im 96 well Maßstab zu isolieren wurde der Montage Plasmid Miniprep 96 Kit von Millipore verwendet. Je 1,2 ml 2x LB-Medium mit dem benötigten Antibiotikum wurden in 96 well Inkubationsplatten pipettiert und jeweils mit einer einzelnen Plattenkultur angeimpft. Anschließend wurden die Kulturen bei 37 °C und 320 Upm 24 Stunden inkubiert. Die Inkubationsplatten wurden bei 4500 Upm 5 Minuten in der Sigma 4-15 zentrifugiert und der Überstand verworfen. Nach Ausschlagen der Inkubationsplatten auf einem Papierstapel wurde 120 µl Lösung 1 in jede Vertiefung pipettiert und die Pellets durch Vortexen resuspendiert. Nach Zugabe von 120 µl Lösung 2 und erneutem Vortexen wurde 2 Minuten bei Raumtemperatur bis zur vollständigen Lyse der Zellen inkubiert. Zu jeder Vertiefung wurden 120 µl Lösung 3 gegeben und nach 2 minütigem Vortexen 300 µl des Lysates auf eine Reinigungs-Platte übertragen. Die Reinigungsplatte wurde über einer Plasmid-Platte in dem Vakuumblock platziert. Durch Anlegen von 270 mbar Unterdruck wurde das Lysat durch die Reinigungsplatte auf die Plasmid-Platte übertragen. Der anschließend noch vorhandene Überstand in der Plasmid-Platte wurde bei 810 Millibar abfiltriert. Anschließend wurde durch Zugabe von 200 µl Lösung 4 zu jeder Vertiefung und erneutes Absaugen mit 810 Millibar gewaschen und die Plasmid DNA durch Zugabe von 50 µl Lösung 5 und 30 Minuten Inkubation bei Raumtemperatur eluiert. Abschließend wurde die DNA zur Lagerung in eine U-Boden Mikrotiterplatte übertragen.

Die benötigten Lösungen und Geräte wurden vom Hersteller des Kits bezogen und nach dessen Angaben verwendet. Die Zusammensetzung der Lösungen wurden vom Hersteller nicht freigegeben.

4.1.3 Isolierung von BAC-DNA aus *E. coli*

4.1.3.1 Isolierung von BAC-DNA mit QIAGEN Plasmid Kits

Um größere Mengen reiner BAC-DNA zu isolieren, wurden QIAGEN Plasmid Kits verwendet. Da in diesem Fall nur eine Kopie des Plasmides pro Zelle vorhanden ist (single copy plasmid) wurde ein modifiziertes Protokoll von 4.1.2 angewendet: 500 ml Bakterienkultur wurden 10 Minuten bei 4000 Upm und 4 °C in der Sorvall RC 5B (GSA-Rotor) zentrifugiert. Das Bakterienpellet wurde in 50 ml Puffer P1 resuspendiert. Zur Lyse der Bakterien wurde das gleiche Volumen Puffer P2 zugegeben, vorsichtig gemischt und 5 Minuten bei Raumtemperatur inkubiert. Danach wurden 50 ml Puffer P3 zugegeben und geschüttelt. Nach 20 Minuten Inkubation auf Eis wurde 30 Minuten bei 13000 Upm und 4 °C in der Sorvall RC 5B (GSA-Rotor) zentrifugiert. Der Überstand wurde in ein neues Reaktionsgefäß überführt und 15 Minuten bei 13000 Upm und 4 °C in der Sorvall RC 5B (GSA-Rotor) zentrifugiert. Der geklärte Überstand wurde auf eine mit Puffer QBT äquilibrierte Qiagen-Säule aufgetragen. Nachdem zweimal mit 30 ml Puffer QC gewaschen worden war, wurde die Plasmid-DNA mit 65 °C warmen Puffer QF in fünf Schritten zu je 3 ml eluiert. Unmittelbar nach der Zugabe von 0,7 Volumen Isopropanol wurde die DNA 90 Minuten bei 6000 Upm und 4°C zentrifugiert. Das Pellet wurde mit -20 °C kaltem 70 %igem Ethanol gewaschen und an der Luft getrocknet. Die DNA wurde in 300 µl TE-Puffer (10/1) gelöst und die Konzentration durch Messung der optischen Dichte bei 260 nm im Spektralphotometer bestimmt.

4.1.3.2 Isolierung hochreiner BAC-DNA mit dem QIAGEN Large Construct Kit

Um hochreine BAC-DNA ohne Kontamination von genomischer Bakterien DNA zu erhalten, wurde der Large Construct Kit von QIAGEN verwendet. 500 ml Bakterienkultur wurden 10 Minuten bei 4000 Upm und 4 °C in der Sorvall RC 5B (GSA-Rotor) zentrifugiert. Das Bakterienpellet wurde in 50 ml Puffer P1 resuspendiert. Zur Lyse der Bakterien wurde das gleiche Volumen Puffer P2 zugegeben, vorsichtig

gemischt und 5 Minuten bei Raumtemperatur inkubiert. Danach wurden 50 ml Puffer P3 zugegeben und vorsichtig geschüttelt. Nach 20 Minuten Inkubation auf Eis wurde 30 Minuten bei 13000 Upm und 4 °C in der Sorvall RC 5B (GSA-Rotor) zentrifugiert. Der Überstand wurde filtriert und die DNA durch Zugabe von 0,6 Volumen Isopropanol gefällt. Die DNA wurde bei 6000 Upm 60 Minuten in der Megafuge 1.0R zentrifugiert. Das Pellet wurde mit 5 ml -20 °C kaltem 70 %igem Ethanol gewaschen und an der Luft getrocknet. Die DNA wurde in 9,5 ml Puffer EX gelöst und dann nach Zugabe von 200 µl ATP-abhängiger Exonuklease (350 ng/µl) und 300 µl 100 mM ATP eine Stunde bei 37 °C inkubiert. Anschließend wurden 10 ml Puffer QS zu der Probe pipettiert und die DNA mit Hilfe einer Qiagen-Säule gereinigt. Der geklärte Überstand wurde auf eine mit Puffer QBT äquilibrierte Qiagen-Säule aufgetragen. Nachdem zweimal mit 30 ml Puffer QC gewaschen worden war, wurde die Plasmid-DNA mit 65 °C warmen Puffer QF in fünf Schritten zu je 3 ml eluiert. Unmittelbar nach der Zugabe von 0,7 Volumen Isopropanol wurde die DNA 90 Minuten bei 6000 Upm und 4°C zentrifugiert. Das Pellet wurde mit -20 °C kaltem 70 %igem Ethanol gewaschen und an der Luft getrocknet. Die DNA wurde in 100 µl TE-Puffer (10/1) gelöst und die Konzentration durch Messung der optischen Dichte bei 260 nm im Spektralphotometer bestimmt.

4.1.4 Herstellung transformationskompetenter *E. coli* XL1-Blue (Hanahan, 1983)

20 ml TYM-Medium mit 12,5 mg/ml Tetracyclin wurde in einem 250 ml-Kolben mit 50 µl *E. coli* XL1-Blue aus einer Vorkultur angeimpft und unter Schütteln bei 37 °C inkubiert, bis die Bakterien eine optische Dichte von 0,5-0,8 bei 600 nm (OD_{600}) erreicht hatten. Die Kultur wurde mit TYM/tet-Medium auf 100 ml verdünnt und in einem 2000 ml-Schikanekolben erneut bis zu einer OD_{600} von 0,5-0,8 vermehrt. Es wurde auf 500 ml Volumen mit TYM-Medium verdünnt und bis zu einer OD_{600} von 0,6 unter Schütteln inkubiert. Die Bakterienkultur wurde durch vorsichtiges Schütteln in einem NaCl/Eiswasserbad rasch auf 0 °C abgekühlt und bei 4000 Upm und 2 °C 10 Minuten in der Sorvall RC 5B (GSA-Rotor) zentrifugiert. Das Bakterienpellet wurde vorsichtig in 100 ml 0 °C kaltem TfBI resuspendiert und bei 4000 Upm und 2 °C 10 Minuten in der

Sorvall RC 5B (GSA-Rotor) zentrifugiert. Das Pellet wurde in 20 ml 0 °C kaltem TfbII resuspendiert, 200 µl Aliquots in vorgekühlte 1,5 ml Eppendorf-Reaktionsgefäße pipettiert und sofort in flüssigem Stickstoff gefroren. Die kompetenten Bakterien wurden bei -80 °C aufbewahrt.

TYM-Medium: 20 g Trypton
 5 g Hefeextrakt
 5 g NaCl
 2 g Mg SO x 7 H₂O

TfBI: 30 mM KOAc, pH 6,5
 50 mM MnCl₂
 100 mM KCl
 10 mM CaCl₂
 15 % (v/v) Glycerin

TfbII: 10 mM Na-MOPS, pH 7,0
 75 mM CaCl₂
 10 mM KCl
 15 % (v/v) Glycerin

4.1.5 Transformation kompetenter *E. coli*

*4.1.5.1 Transformation chemisch kompetenter *E. coli* XL1-Blue*

50 µl kompetente *E. coli* XL1-Blue wurden auf Eis aufgetaut und zu 30 µl DNA-Lösung pipettiert, die ca. 50-200 ng ligierte Plasmid-DNA oder 10-50 ng Plasmid-DNA enthielt. Der Ansatz wurde 15 Minuten auf Eis inkubiert, dann 2 Minuten bei 42 °C gehalten und

4.1.5.3 Transformation elektrokompetenter TOP10 One Shot Zellen

Zu den 6 μl Ligationsansatz aus der TOPO Klonierungsreaktion (4.2.5.6) wurden 100 μl H_2O bidest gegeben und 10 Minuten bei 37 °C inkubiert. Zur Fällung der DNA wurden 10 μl 3 M Natriumacetat Lösung pH 5,2 , 2 μl 20 mg/ml Glykogen Lösung und 300 μl 100 % iges Ethanol hinzugegeben und durch Vortexen gemischt. Nach 15 Minuten Inkubation bei -80 °C wurde die DNA bei 22000 Upm 15 Minuten bei 4 °C in der Biofuge Stratos zentrifugiert. Das Pellet wurde mit 1 ml 80 % igem Ethanol gewaschen, in der Speed Vac getrocknet und in 10 μl H_2O bidest aufgenommen.

3,3 μl der gefällten DNA wurden zu 50 μl elektrokompetenten TOP10 One Shot gegeben und vorsichtig durch Schwenken gemischt. Die Elektroporation wurde dann mit 2500 Volt und 410 Ampere mit einem 5 Millisekunden Impuls durchgeführt. Der Transformationsansatz wurde mit 1 ml SOC-Medium versetzt und bei 37 °C 60 Minuten mit 225 Upm inkubiert. 25-100 μl dieses Ansatzes wurden auf LB/amp/X-Gal-Platten ausplattiert.

4.1.6 Erstellen einer Shotgun Plasmid Genbank mit Hilfe des TOPO Shotgun Subkloning Kits

4.1.6.1 Scheren der BAC-DNA

25 μg gereinigte BAC-DNA ohne Verunreinigung durch genomische Bakterien DNA, wurden mit 750 μl Shearing Puffer versetzt und mit H_2O auf 2100 μl aufgefüllt. Die Lösung wurde anschließend in einen Standard-Vernebler pipettiert und für 90-120 s bei etwa 1 bar Druck mit Stickstoff geschert. Die DNA wurde durch Zugabe von 80 μl 3 M Natriumacetat, 4 μl Glykogen (20 mg/ml) und 700 μl 100 %igem Isopropanol gefällt, 15 Minuten bei 22000 Upm abzentrifugiert und nach dem Trocknen in H_2O zu einer Konzentration von 500 ng/ μl gelöst.

Shearing Puffer: TE pH 8,0
 10 % Glycerin

4.1.6.2 Klonieren der gescherten DNA

Die überhängenden Enden der gescherten BAC-DNA wurden nach Protokoll des Herstellers mit T4 DNA Polymerase aufgefüllt (4.2.5.5) und durch Zugabe von Kälberdarm-Phosphatase dephosphoryliert (4.2.5.2). Anschließend wurde die DNA mit Hilfe des TOPO Klonierungsverfahrens in den pCR4Blunt-TOPO Vektor ligiert (4.2.5.6) und in chemisch oder elektrokompetente *E. coli* transformiert (4.1.5.3). Eine Verdünnungsreihe des Transformationsansatzes wurde dann auf 9-15 LB/amp/X-Gal Agarplatten ausplattiert.

4.2 Enzymatische Manipulation und Analyse von DNA

4.2.1 Auftrennung von kleinen DNA-Fragmenten

Zur Auftrennung von kleinen DNA-Fragmenten wurden horizontale Agarosegele von ca. 5-10 mm Dicke und 10-25 cm Länge verwendet. Je nach Größe der zu untersuchenden DNA-Fragmente wurden Agarose-Konzentrationen von 0,6-2,0 % gewählt. Die jeweilige Menge Agarose wurde in einen Erlenmeyer-Kolben eingewogen und im entsprechenden Volumen 1x TBE-Laufpuffer aufgeschlämmt. Die Suspension wurde in der Mikrowelle aufgeköcht und solange gerührt bis eine homogene Lösung entstand. Die Lösung wurde auf ca. 50 - 60 °C abgekühlt und mit Ethidiumbromid-Lösung (10 mg/ml in H₂O bidest.) auf eine Ethidiumbromidkonzentration von 0,4 µg/ml eingestellt. Die Lösung wurde in eine Gelform mit einem Kamm der erforderlichen Taschengröße gegossen und bis zum Erstarren der Agarose stehen gelassen. Der Kamm wurde vorsichtig gezogen und das Gel in einer Elektrophoresekammer mit 1x TBE-Laufpuffer bedeckt. Die DNA-Lösung wurde

mit 0,1 Volumen 10x Auftragspuffer versetzt und in eine Tasche des Gels pipettiert. Die Elektrophorese erfolgte bei Feldstärken von 5–10 V/cm. Die DNA-Fragmente wurden durch Fluoreszenz des in die DNA interkalierten Ethidiumbromids unter UV-Licht bei 366 nm oder 312 nm sichtbar gemacht. Die Dokumentation erfolgte mit Hilfe des BioDoc Analyzers (Biometra).

10x TBE-Laufpuffer: 121 g Tris
 51,4 g Borsäure
 3,72 g EDTA
 ad 1000 ml H₂O bidest.

10x Auftragspuffer: 50 % (v/v) Glycerin
 0,1 M EDTA pH 8,0
 0,25 % (w/v) Bromphenolblau
 0,25 % (w/v) Xylencyanol

4.2.2 Pulsfeldgelelektrophorese (Schwartz und Cantor, 1984)

Mit 150 ml 0,5 x TBE wurde ein 1 %iges Agarosegel mit einer Ethidiumbromidkonzentration von 0,4 µg/ml gegossen. Die Elektrophoresekammer wurde mit 2,8 l 0,5x TBE gefüllt und auf 16 °C temperiert. Nach Erstarren des Gels wurden die Geltaschen mit 0,5x TBE gefüllt und mit der zu untersuchenden DNA beladen. Die Taschen wurden anschließend mit 1 %iger Agarose (in 0,5 x TBE) abgedichtet. Das Gel wurde mit einer Feldstärke von 6 V/cm, einem Pulswinkel von 120° und einer Pulszeit von 1 s bis 25 s für 20 h laufen gelassen.

4.2.3 Isolierung und Reinigung von DNA-Fragmenten aus Agarosegelen mittels QIAEX II-Kit

Die zu isolierenden DNA-Fragmente wurden nach der elektrophoretischen Auftrennung auf einem UV-Schirm (366 nm) ausgeschnitten, abgewogen und in ein Eppendorf-Reaktionsgefäß überführt. Es wurden 3 Volumen Puffer QX 1 zugegeben und 10 µl resuspendierte QIAEX II-Lösung zum Reaktionsansatz dazupipettiert. Dieser wurde 10 Minuten bei 50 °C inkubiert und dabei alle 2 Minuten durch starkes Schütteln gemischt. Der Ansatz wurde 30 s bei 13000 Upm in der Tischzentrifuge zentrifugiert und der Überstand sorgfältig entfernt. Das Pellet wurde in 500 µl QX 1 resuspendiert, 30 s bei 13000 Upm in der Tischzentrifuge zentrifugiert und der Überstand entfernt. Das Pellet wurde zweimal auf gleiche Weise mit je 500 µl Puffer PE gewaschen und anschließend 15 Minuten luftgetrocknet. Um die DNA zu eluieren, wurde das Pellet mit 20 µl TE-Puffer (10/1) 5 Minuten bei Raumtemperatur inkubiert und 30 s bei 13000 Upm in der Tischzentrifuge zentrifugiert. Der Überstand wurde sorgfältig in ein neues Reaktionsgefäß überführt.

4.2.4 Isolierung und Reinigung von DNA-Fragmenten aus Lösungen mittels QIAEX II-Kit

Um DNA-Fragmente aus einer Lösung zu reinigen, wurde zu der DNA-Lösung 3 Volumen Puffer QX 1 gegeben, die QIAEX II-Lösung durch starkes Schütteln resuspendiert und 10 µl zum Reaktionsansatz pipettiert. Dieser wurde 10 Minuten bei Raumtemperatur inkubiert und dabei alle 2 Minuten durch starkes Schütteln gemischt. Der Ansatz wurde 30 s bei 13000 Upm in der Tischzentrifuge zentrifugiert und der Überstand sorgfältig entfernt. Das Pellet wurde in 500 µl Puffer PE resuspendiert, 30 s bei 13000 Upm in der Tischzentrifuge zentrifugiert und der Überstand entfernt. Das Pellet wurde noch einmal auf gleiche Weise mit 500 µl Puffer PE gewaschen und anschließend 15 Minuten luftgetrocknet. Um die DNA zu eluieren, wurde das Pellet mit 20 µl TE-Puffer (10/1) 5 Minuten bei Raumtemperatur inkubiert und 30 s bei 13000 Upm in der Tischzentrifuge zentrifugiert. Der Überstand wurde in ein neues Reaktionsgefäß überführt.

4.2.5 Enzymatische Manipulation von DNA

4.2.5.1 Spaltung von DNA mit Restriktionsendonukleasen

Gereinigte DNA wurde mit kommerziell erhältlichen Restriktionsendonukleasen gespalten. Die Reaktionsbedingungen wurden entsprechend den Herstellerangaben eingestellt. Die DNA wurde je nach Reinheitsgrad mit 1 bis 5 Units Enzym pro μg DNA für mindestens 1 h bei 37 °C inkubiert.

4.2.5.2 Dephosphorylierung von DNA-Fragmenten

Nach mechanischem Scheren von DNA oder Spaltungen mit Restriktionsendonukleasen trägt das 5'-Ende der gespaltenen DNA-Fragmente eine Phosphatgruppe. Um eine Ligation der Enden zu verhindern, wurde dieser Phosphatrest mit alkalischer Phosphatase abgespalten.

Reaktionsansatz: 10 μg linearisierte DNA
 10 μl 10x Reaktionspuffer des Herstellers
 5 μl alkalische Phosphatase (1 U/ μl)
 ad 100 μl H₂O bidest.

Der Reaktionsansatz wurde zur Inaktivierung der alkalischen Phosphatase 20 Minuten auf 68 °C erhitzt.

4.2.5.3 5'-Kinasierung von DNA-Fragmenten

Um linearisierte DNA-Fragmente ligieren zu können, müssen diese an ihrem 5'-Ende eine Phosphatgruppe tragen. Daher müssen PCR-Produkte, die an ihren Enden synthetische Oligonukleotide ohne 5'-Phosphatgruppe tragen, vor einer Ligation kinasiert werden, um eine Phosphatgruppe anzuhängen.

Reaktionsansatz: ca. 1 µg DNA
 3 µl 10x Reaktionspuffer des Herstellers
 1 µl T4-Polynukleotidkinase (10 U/µl)
 0,5 µl 10 mM ATP
 ad 30 µl H₂O bidest.

Die T4-Polynukleotidkinase wurde vor einer Ligation durch 20 Minuten Erhitzen auf 68 °C inaktiviert.

4.2.5.4 Ligation von Plasmiden

Um zwei DNA-Enden kovalent zu verbinden, müssen enzymatisch Phosphodiesterbindungen mit Hilfe der Ligase geknüpft werden. Für diese Verknüpfung ist mindestens ein DNA-Ende mit einer 5'-Phosphatgruppe erforderlich.

Reaktionsansatz: 50 ng Vektor-DNA
 x ng Insert-DNA (molares Verhältnis Insert : Vektor ~ 1 : 1)
 3 µl 10x Reaktionspuffer des Herstellers
 0.5 µl T4-DNA-Ligase (5 Weiss-U/µl)
 ad 30 µl H₂O bidest.

Der Ansatz wurde über Nacht bei 16 °C inkubiert.

4.2.5.5 Herstellung glatter Enden mit T4-DNA-Polymerase

Zum Auffüllen von überhängenden Enden an gescherter DNA oder Oligonukleotiden wurde die DNA nach den Angaben des Herstellers mit T4-DNA-Polymerase und Desoxynukleosidtriphosphaten im Reaktionspuffer des Herstellers 30 Minuten bei Raumtemperatur inkubiert. Die Polymerase wurde anschließend durch 20 Minuten Erhitzen bei 68 °C inaktiviert.

4.2.5.6 Ligation von Plasmiden mit Hilfe des TOPO Cloning Systems

Das TOPO Cloning System basiert auf dem linearisierten pCR4Blunt-TOPO Vektor, an dessen 3'-Enden die *Vaccinia Virus* Topoisomerase I kovalent gebunden ist. Dies ermöglicht eine effiziente Ligation von DNA-Fragmenten mit glatten Enden.

Reaktionsansatz: 10 ng pCR4Blunt-TOPO Vektor-DNA
 100 ng Insert-DNA mit glatten Enden
 1 µl Salz-Puffer des Herstellers

Der Reaktionsansatz wurde durch Schwenken vorsichtig gemischt und 25 Minuten bei 37 °C inkubiert.

4.2.5.6 Ligation von PCR-Produkten mit Hilfe des QIAGEN PCR Cloning Systems

Das QIAGEN PCR Cloning System nutzt den A-Überhang an den 3'-Enden von PCR-Produkten, welcher durch die *Taq*-DNA-Polymerase eingeführt wird. Der linearisierte pDrive Cloning Vector weist an beiden Enden einen U-Überhang auf, so dass eine Ligation von PCR-Produkten effizient erfolgen kann. Die Ligation wurde im Puffer

des Herstellers durchgeführt, der alle Kofaktoren und Reagenzien für die Ligation beinhaltet.

Reaktionsansatz: 50 ng pDrive Cloning Vector
 100 ng PCR-Produkt
 5 µl Ligationspuffer des Herstellers

Der Ansatz wurde gemischt und anschließend 30 Minuten bei 16 °C inkubiert.

4.2.6 Alkalischer Transfer von DNA auf Trägermembranen (Reed und Mann, 1985)

Die DNA wurde im Agarosegel elektrophoretisch getrennt und zweimal 15 Minuten in 0,4 M NaOH geschwenkt. Anschließend wurde das Gel mit der Unterseite nach oben auf eine Frischhaltefolie gelegt und eine passend zurechtgeschnittene positiv geladene Nylonmembran luftblasenfrei auf das Gel aufgebracht. Zwei Lagen passend zurechtgeschnittenes Whatmanpapier wurden angefeuchtet und ebenfalls luftblasenfrei aufgelegt. Zuletzt wurde ein ca. 10 cm hoher Stapel Kleenex-Papier aufgelegt und mit etwa 2 kg beschwert. Nach 4-24 h wurde die Nylonmembran kurz in 2 x SSC geschwenkt und zwischen zwei Blättern Whatmanpapier luftgetrocknet.

20 x SSC Stocklösung: 175 g NaCl
 88 g Na₃-Citrat
 ad 1000 ml H₂O

4.2.7 Hybridisierung und Detektion von DNA-Fragmenten mit ECL direct labelling system

Als Hybridisierungslösung wurden zu 25 ml Gold Puffer 0,73 g NaCl gegeben und gemischt. Anschließend wurden 1,25 g Blocking Reagent hinzugegeben und bei 42 °C eine Stunde gerührt. Die getrocknete Membran aus dem alkalischen Transfer (4.2.6) wurde in

eine Hybridisierungsrolle überführt und mit der vorbereiteten Hybridisierungslösung 1 h bei 42 °C vorhybridisiert. Zur Markierung der spezifischen Sonde wurden 100 ng DNA in 10 µl H₂O in einem verschraubbaren 1,5 ml Reaktionsgefäß 3 Minuten im kochendem Wasserbad denaturiert und sofort auf Eis gekühlt. Anschließend wurden 10 µl Glutaraldehyd und 10 µl Labelling Reagent zur Sonde zugegeben und 10 Minuten bei 37 °C inkubiert. Die markierte Sonde wurde unmittelbar danach zur Hybridisierungslösung pipettiert. Die Hybridisierung wurde bei 42 °C über Nacht durchgeführt. Am nächsten Tag wurde die Membran 5 Minuten bei 55 °C in 0,4x SSC/0,4 % SDS und zweimal 5 Minuten bei Raumtemperatur in 2x SSC gewaschen. Dies entspricht einer mittleren Stringenz. Die Membran wurde 1 Minute in 6 ml Detektionsflüssigkeit (3 ml Detection Reagent 1 und 3 ml Detection Reagent 2) geschwenkt und auf einem Kodak XOMAT-LS Röntgenfilm 1 Minute bis 24 h exponiert.

4.3 Polymerase-Kettenreaktion (PCR) (Saiki *et al.*, 1988)

Ein Standardansatz enthielt in 50 µl Gesamtvolumen ca. 100 ng Matrizen-DNA, 5 µl 10x *Taq*-Puffer des Herstellers, je 50 pmol Primer, 200 µM Desoxyribonukleosidtriphosphate und 1 U *Taq*-DNA-Polymerase. Die Komponenten des Ansatzes wurden auf Eis pipettiert, gemischt und die Reaktionsgefäße in einen auf 94 °C vorgeheizten Thermocycler überführt. Nach 5-10 Minuten Denaturierung bei 94 °C wurden 30 Zyklen mit 1 Minute Denaturierung bei 94 °C, 1 Minute Annealing bei der jeweiligen Annealingtemperatur des Primerpaares und 1 Minute Polymerisation bei 72 °C durchgeführt. Anschließend wurden die PCR-Produkte durch Agarose-Gelelektrophorese analysiert.

4.4 DNA-Sequenzierung

4.4.1 Nichtradioaktive Sequenzierung doppelsträngiger DNA

Sequenzreaktionen wurden mit dem Thermo Sequenase Fluorescent Labelled Cycle Sequencing Kit durchgeführt. Es wurden 200 fmol Plasmid-DNA mit 2 pmol IRD markiertem Primer und 1 µl Dimethylsulfoxid (DMSO) versetzt und mit H₂O bidest. auf 25 µl aufgefüllt. Je 6 µl von diesem Ansatz wurden zu 2 µl der Terminationsmische pipettiert und mit Mineralöl überschichtet. Die Sequenzreaktion wurde auf einem PTC-100TM Programmable Thermal Controller mit 45 s Denaturierung bei 94 °C, 45 s Annealing bei 58 °C und 45 s Extension bei 68 °C durchgeführt.

4.4.2 Gelelektrophorese von Sequenzprodukten auf einem automatischen DNA-Sequencer

Für die Auftrennung von Sequenzprodukten wurde das LICOR Modell 4200 und die Steuersoftware BaseImageIR V4.10 verwendet. Als Gelmatrix für die Auftrennung von bis zu max. 1000 bp wurde Sequagel XR unter Zusatz von 1 % DMSO und Verwendung von 41 cm langen Glasplatten verwendet. Die Auftrennung erfolgte bei 1500 V, 35 mA und 31 W für ca. 14 Stunden. Als Laufpuffer wurde 1x TBE verwendet. Vor dem Auftrag wurden die Proben mit 5 µl Ladepuffer versetzt.

Ladepuffer:	97,5 % deionisiertes Formamid
	10 mM EDTA pH 8,0
	0,1 % Bromphenolblau
10 x TBE long run:	1,34 M Tris
	450 mM Borsäure
	25 mM EDTA

4.5 Isolierung und Analyse von RNA

Beim Umgang mit RNA wurden zur Vermeidung von RNase-Kontaminationen alle Lösungen neu angesetzt und soweit möglich steriles Einweg-Material verwendet.

4.5.1 Isolierung von Gesamt-RNA aus Gewebe

1 g Gewebe wurde in 5 ml TRIZOL™ (Total RNA Isolation Reagent) mit einem Ultra-Turrax homogenisiert. Der Ansatz wurde 5 Minuten bei Raumtemperatur inkubiert. Nach Zugabe von 1 ml Chloroform wurde das Röhrchen 15 Sekunden kräftig geschüttelt und weitere 2 Minuten bei Raumtemperatur inkubiert. Der Ansatz wurde bei 6000 Upm und 4 °C 15 Minuten in der Megafuge 1.0R zentrifugiert. Die obere wässrige Phase wurde in 7 Aliquots zu 1 ml aufgeteilt und in neue 1,5 ml Eppendorf-Reaktionsgefäße überführt. Die RNA wurde durch Zugabe von je 500 µl Isopropanol präzipitiert und bei -20 °C gelagert. Unmittelbar vor einer weiteren Verwendung wurde die RNA 10 Minuten bei 12000 Upm und 4 °C in der Biofuge Stratos zentrifugiert und das Pellet mit -20 °C kaltem 75 %igem Ethanol gewaschen. Das Pellet wurde 10 Minuten an der Luft getrocknet und in einem geeigneten Volumen H₂O bidest. gelöst.

4.5.2 Amplifikation von cDNA mit Hilfe des First Choice RLM-RACE Kit

Um vollständige cDNAs von Volllänge-mRNA zu amplifizieren, wurde der First Choice RLM-RACE (RNA-ligation mediated rapid amplification of cDNA ends) Kit von Ambion verwendet. Die Gesamt RNA wurde mit Kälberdarm Phosphatase dephosphoryliert und die Cap-Struktur der Volllänge-mRNA mit saurer Tabak-Pyrophosphatase entfernt. Anschließend wurde ein Oligonukleotid-Adapter an das 5'-Ende der mRNA ligiert. Da nur Volllänge-mRNAs eine 5'-Phosphatgruppe tragen, wurde so gewährleistet, dass der Adapter ausschließlich an die Volllänge-mRNA ligierte. Mit zwei verschränkten („nested“) Gen-spezifischen und Adapter-spezifischen Primerpaaren wurde dann eine reverse

Transkription-Polymerase Kettenreaktion (RT-PCR) durchgeführt, um das komplette 5'-Ende der cDNAs zu amplifizieren. In gleicher Weise wurde das 3'-Ende unter Verwendung von zwei Paaren verschränkter Gen-spezifischer und Adapter-spezifischer Primer amplifiziert. Die Reaktionsansätze der einzelnen Schritte wurden nach den Angaben des Herstellers pipettiert und inkubiert. Die 5'- und 3'-RACE Produkte wurden zur weiteren Analyse kloniert und sequenziert.

4.6 Bioinformatik

4.6.1 Programme (lokal)

4.6.1.1 LICOR Steuersoftware BaseImageIR V4.10

Die LICOR Sequenzdaten wurden mit der Data Collection V4.10 Software gesammelt und mit dem Programm Image Analysis V4.10 ausgewertet. Anschließend wurden die Daten mit dem Programm SCF File Creation V4.10 in Standard Chromatogramm Format (SCF) Dateien umgewandelt.

4.6.1.2 Sequencher 4.0.5

Die ermittelten Randsequenzen der „Shotgun“-Plasmidklone wurden mit dem Programm Sequencher 4.0.5 automatisch zu Contigs zusammengesetzt, manuell kontrolliert und nach Einfügen der Walking-Primer Sequenzen zur Gesamtsequenz der BAC-Klone zusammengesetzt.

4.6.1.3 DNASTAR Programmpaket

Mit dem Programm EditSeq wurden Nukleotidsequenzen bearbeitet und zu Proteinsequenzen übersetzt. Anschließend wurde mit diesem Programm nach offenen Leserahmen in den Proteinsequenzen gesucht. Des weiteren wurden Nukleotid- und Proteinsequenzen mit Hilfe dieses Programms bearbeitet und für die Eingabe in Internetprogramme formatiert.

Das Programm MegAlign wurde verwendet um Nukleotid- und Proteinsequenzen paarweise zu vergleichen.

4.6.2 Internetseiten

4.6.2.1 Entrez-Browser

Die Suchen von Daten in Nukleotid- und Proteindatenbanken, sowie der PubMed-Datenbank wurden mit dem Entrez-Browser durchgeführt.

<http://www3.ncbi.nlm.nih.gov/Entrez/>

4.6.2.2 BLAST

Um homologe Sequenzabschnitte in Datenbanken zu ermitteln, wurde das Programm BLAST verwendet.

<http://www.ncbi.nlm.nih.gov/BLAST/>

4.6.2.3 Repeatmasker

Da die in genomischen Sequenzen auftretenden repetitiven Sequenzabschnitte eine Datenbanksuche erschweren, wurden diese Abschnitte mit Hilfe des Programms Repeatmasker identifiziert und maskiert.

<http://repeatmasker.genome.washington.edu/>

4.6.2.4 Gene Fisher

Geeignete Primerpaare wurden mit dem Programm Gene Fisher ermittelt

(<http://bibiserv.techfak.uni-bielefeld.de/genefisher>)

4.6.2.5 PipMaker

Das Programm PipMaker wurde verwendet, um homologe Sequenzabschnitte zwischen zwei Sequenzen grafisch darzustellen. Das Programm ermöglichte außerdem die Darstellung von codierenden Bereichen, repetitiven Elementen und CpG-Islands.

<http://bio.cse.psu.edu/pipmaker/>

4.6.2.6 EBI Toolbox

Um Alignments von mehreren Proteinsequenzen zu erstellen, wurde das Programm ClustalW verwendet.

Die Programme CpG Plot und CpG Report wurden verwendet, um den GC-Gehalt von Sequenzen grafisch darzustellen und die Position von CpG-Islands zu lokalisieren.

<http://www.ebi.ac.uk/Tools/index.html>

5 Ergebnisse

5.1 cDNA der equinen *CRISP* Gene

5.1.1 Sequenzanalyse der equinen *AEGI* cDNA und der abgeleiteten Proteinsequenz

Mit Hilfe eines modifizierten Protokolls für die schnelle Amplifikation von cDNA Enden (sog. RLM-RACE) wurde eine Volllänge-*AEGI* cDNA aus epididymidaler Gesamt-RNA des Pferdes synthetisiert. Bei diesem Protokoll werden Volllänge-mRNAs mit einem Cap am 5'-Ende angereichert und nach Entfernen des Cap ein RNA-Adapter an das 5'-Ende ligiert. Diese Strategie ermöglicht die Isolation von kompletten cDNA 5'-Enden. Im Gegensatz hierzu verliert man bei konventionellen RACE Protokollen einige Nukleotide am 5'-Ende des Transkriptes während der cDNA Synthese. Die 3'-Enden der cDNA wurden durch ein konventionelles RACE Experiment erhalten. Die überlappenden 5'- und 3'-Enden wurden kloniert und die vollständige 1475 bp Sequenz der längsten *AEGI* cDNA ermittelt. Die Sequenzdaten der equinen *AEGI* cDNA wurden unter der Accession AJ315379 in der EMBL Nucleotide Sequence Database hinterlegt. Es wurden fünf identische 5'-RACE Klone gefunden, die 2 bp später, und acht identische 5'-RACE Klone, die 115 bp später als die Volllänge-*AEGI* cDNA beginnen. Die durchgeführten RACE Experimente weisen darauf hin, dass die Transkription an den Positionen -226, -224 und -111 der Volllänge-*AEGI* cDNA initiiert wird.

Am 3'-Ende der *AEGI* cDNA wurde ein variiertes Polyadenylierungssignal gefunden, das die Sequenz AATAA aufwies und 10 bp stromaufwärts der Polyadenylierungsstelle lag. In der längsten cDNA Isoform erstreckte sich der 5'-untranslatierte Bereich über eine Länge von 226 bp, während der 3'-untranslatierte Bereich 499 bp umfasste. Unter der Annahme, dass das homologe ATG-Startcodon wie bei anderen Säugetieren verwendet wird, kodieren alle equinen *AEGI* cDNA Isoformen für ein Protein von 250 aa Länge. Im 5'-untranslatierten Bereich der *AEGI* cDNA wurden zwei weitere ATG Startcodons gefunden, die sich im Leserahmen befinden (Abbildung 2).

↓ ↓	ATGATTTGCA ATCTTCACAG GCTCTAGATT TGTTCTGATT ACT <u>ATGCGTT</u> TGTGCCCTTT	60
	GCCTGCTGTA AATTCCTGGA <u>TGTCTTGTA</u> GCATTTTCCT TGCCTGACT TCTTCATAGA	120
	GCAAATCTTT CCTGTAGGGG GCTGTTGCGA GGCATTAGCA CACTATCTAG CCAAAGATCT	180
	GGGTCTCATC GGGACACTGA ACTTTACAAA TTGTGTGAAC TTGGAG <u>ATGA</u> CAATGAAACA	240
	M T M K H	5
	CTCTTTGTTT TTGGCTGCTG CTGCTGCTGG CTTCTGCCT GTTGTGATCA TAAGAGCTAA	300
	F L F L A A A A A G F L P V V I I R A K	25
	ACCAGCTAGC GTTCCATATC ATACGCTTCT CACTGAGTTG GCAACTGTCC AAGAAGAAAT	360
	P A S V P Y H T L L T E L A T V Q E E I	45
	TGTTACTGTA CATAATACTC TCAGAAGAGG AGTAGTTCCG CCAGCCAGCA ACATGCTGAA	420
	V T V H N T L R R G V V P P A S N M L K	65
	GATGAATTGG AGTGAAGAAG CTGCACAAAA CGCCAGAAGG TTGTCTGCGC AGTGTGAATT	480
	M N W S E E A A Q N A R R L S A Q C E L	85
	GATAGAGAGC AAGGCAGTTG AGAGGAGAAT TACAATACC TTTGTGGAG AAAATATGCA	540
	I E S K A V E R R I T N T F C G E N M H	105
	TCTGACATCT TATCCTATCT CATGGTCAAA TGTAATTGGA ATCTGGTTCA GTGAGTCTAA	600
	L T S Y P I S W S N V I G I W F S E S K	125
	ATATTTCAAG TATGGCAAT GGATATCAAC AGATGATGAA GTAATAATG AGCATTATAC	660
	Y F K Y G Q W I S T D D E V I I E H Y T	145
	TCAGCTTGTT TGGGCCACTT CTTACCTCAT TGGCTGTGGG ATATCATCAT GCTCCAAAAG	720
	Q L V W A T S Y L I G C G I S S C S K R	165
	AAGGTCAACT CAGTATCTCT ACGTTTGTCA CTACTGTTCAT GAGGGAAATG ATCCCGACAA	780
	R S T Q Y L Y V C H Y C H E G N D P D K	185
	AAAGAACGTG CCTTATAATA TGGGAACTCC GTGTAAAGAC TGTCCAAATG ACTGTGAAGA	840
	K N V P Y N M G T P C K D C P N D C E D	205
	TAGACTTTGT ACTAACCCCT GCCTCTACTA TGATGAATCC AATAAGTGTA AGACACAAAA	900
	R L C T N P C L Y Y D E S N K C K T Q K	225
	AGAAGCTCTT GGATGTAGCC ACTTGTCACT TAACTGTTC TGCAAAGCTA CTGTCTGTG	960
	E A L G C S H L S V K L F C K A T C L C	245
	TGACACTGAG ATAAAA <u>TAA</u> C CTTTGTATT CGCAACTGTA ATGTGCTGTT GGGGAAACGT	1020
	D T E I K	250
	CTACTAGATT ATATCTTTA CTTTACTTGC TTCAGTAGCT TCTGCTGAAT TTCTTTGGT	1080
	TTAATTGCA CTCAAAATCC TAATTTTCAC TAATCATCTA TGGGCATCAG ATATTTACAT	1140
	CAATGCCCC AACTTCTGTC CTGATACATC CTCATGTAAA GTCTTAAACT TTCTTCTGTC	1200
	TGGAGTGAAA GATCCATTC AACACCTTC TCATCTTTAA ATGATGTGCC TTGTCTGCA	1260
	TAAGTTGTCA TAACTCAACT GATGGTATGA CGGGCCCCCT CAAGAACCAC CTGTGCTGGG	1320
	GGAGGGTCTT CCATGTTTAA CTCCTCCCTC TGCTCTCTG TGTCTCCAGC CACCTGATA	1380
	TTGTCTTCC CTGCACAACT TAACATGCAC TAGCAATTCT TCACTTCTTA ACTTTGCTTC	1440
	TATTTTAGAT CAATCATTAA <u>AATAA</u> TCCAC AAAGC	1475

Abbildung 2: AEG1 cDNA mit Translation. Nichtcodierende Bereiche sind kursiv dargestellt. Der dicke Pfeil markiert den Transkriptionsstart. Die dünnen Pfeile markieren die beiden anderen Transkriptionsstartstellen. Das ATG-Startcodon sowie das TAA-Stopcodon sind fett und unterstrichen hervorgehoben. Die beiden anderen im Leserahmen befindlichen ATG-Codons sind unterstrichen gekennzeichnet. Das geringfügig von der Consensus-Sequenz abweichende Polyadenylierungssignal ist durch einen Kasten hervorgehoben.

Das equine AEG1 Protein zeigt eine Übereinstimmung von 63 % zu dem orthologen Protein vom Mensch, 37 % zu dem der Maus und 33 % zu dem der Ratte (Abbildung 3). Obwohl die Sequenzhomologie relativ niedrig ist, zeigen die AEG1 Proteine dieser Spezies die für die Mitglieder der CRISP-Genfamilie charakteristischen 16 konservierten Cystein-Reste.

		↓	
Pferd	MTMKHFLFLAAAAAGFLPVVIRAKPASVPYHTLLTELATVQEEIVTVHNTLRRGVVPPA		60
Mensch	MEIKHLLFLVAAAC-LLPMLSMKKKSARDQFNKLVTDLPNVQEEIVNIHNALRRRVVPPA		59
Maus	MALMLVLFFLAAVL-PPSLLQDSSQENR--LEKLSTTKMSVQEEIVSKHNQLRRMVSPSG		57
Ratte	MALMLVLLFLAAVL-PPSLLQDSTDDEWRDLENLSTTKLSVQEEIINKHNQLRRTVSPSG		59
	* : .*: : ** . : : * * . *****: . ** * * * * *		
		↓ ↓	
Pferd	SNMLKMNWSEEAQNARRLSAQCELIIESKAVERRITNTFCGENMHLTSYPISWSNVIGIW		120
Mensch	SNMLKMSWSEEAQNARIFSKYCDMTESNPLERRLPNTFCGENMHMTSYPVSWSSVIGVW		119
Maus	SDLKMEWNYDAQVNAQQWADKCTFSHS-PIELRTTNLRCGENLFMSSYLASWSSAIQGW		116
Ratte	SDLLRVEWDHDAYVNAQKWANRCIYNHS-PLQHRTTTLKCGENLFMANYPASWSSVIQDW		118
	*::: : . * : * ** : : * . * : : * . . *****: : . * * * * . * *		
		↓ ↓ ↓ ↓	
Pferd	FSESKYFKYGQWISTDDEVIIEHYTQLVWATSYLIGCGISSCSKRRSTQYLYVCHYCHEG		180
Mensch	YSESTSFKHGEWTTTDDITTDHYTQIVWATSYLIGCAIASCRQQGSPRYLYVCHYCHEG		179
Maus	YNEYKDLTYDVGP-KQPDSVVGHYTQVWVNSTFQVACGVAECPKN-PLRYYYVCHYCPVG		174
Ratte	YDESLDFVFGFGP-KKVGKVGHYTQVWVNSTFVACGVAECPDQ-PLKYFVCHYCPG		176
	: . * : ***** ** : : : . * . : : . * . . . : * ***** *		
		↓ ↓ ↓ ↓ ↓ ↓ ↓	
Pferd	NDPDKKNVPYNGTPECKDCPNDCEDRLCTNPCLYDESNKCKTQKEALGCSHLSVKLFCK		240
Mensch	NDPETKNEPYKTGVPCAECPNCEKLNCTNPCIYYDEYFDCDIQVHYLGCNHSTTILFCK		239
Maus	NYQGRLYTPYTAGEPCASCPDHCEGLCTNSCGHEDKYTNCKYLKMLSCHEHELLKKGCK		234
Ratte	NYVGRLYSPYTEGEPDCSCPNCEGLCTNSCEYEDNYSNCGDLKMKVSCDDPLLKKGCR		236
	* ** . * ** ** . ** * * * * * : : * : . * . : . * . . *		
		↓ ↓	
Pferd	ATCLCDTEIK		250
Mensch	ATCLCDTEIK		249
Maus	ATCLCEGKIH		244
Ratte	ASCFCEDKIH		246
	*::: : : * :		

Abbildung 3: Alignment des equinen AEG1 Proteins mit den Sequenzen der bekannten orthologen AEG1 Proteine von Mensch, Maus und Ratte. Die Sequenzen wurden aus den Datenbankeinträgen mit den Accessions P54107 (AEG1 Mensch), Q03401 (Aeg1 Maus), und AAB59716 (Crisp-1 oder Protein DE Ratte) entnommen. Der dicke Pfeil (↓) markiert die Position der in einigen Pferden auftretende Aminosäuresubstitution ³⁷E → ³⁷G. Die 16 konservierten Cysteine sind durch dünne Pfeile (↓) hervorgehoben. Identische Aminosäuren sind durch Sterne unter dem Alignment gekennzeichnet, während Punkte und Doppelpunkte ähnliche oder sehr ähnliche Aminosäuren bezeichnen.

5.1.2 Sequenzanalyse der equinen *TPXI* cDNA und der abgeleiteten Proteinsequenz

Mit Hilfe des RLM-RACE Protokolls wurde eine Volllänge-*TPXI* cDNA aus testiculärer Gesamt-RNA vom Pferd synthetisiert. Die überlappenden 5'- und 3'-RACE Produkte wurden kloniert und die vollständige 1386 bp DNA Sequenz der längsten *TPXI* cDNA ermittelt. Zehn von sechzehn analysierten 5'-RACE Klonen begannen an Position -224 der Volllänge-*TPXI* cDNA. Die anderen sechs 5'-RACE Klone repräsentieren verschiedene abgebrochene cDNA Produkte. Das weist darauf hin, dass die Transkription des equinen *TPXI* Gens primär an der Position -224 der ermittelten *TPXI* cDNA beginnt. Die Sequenzdaten der equinen *TPXI* cDNA wurden unter der Accession AJ459963 in der EMBL Nucleotide Sequence Database hinterlegt.

Unter der Annahme, dass das gleiche ATG wie in anderen Säugetierspezies als Startcodon verwendet wird, weist die equine *TPXI* cDNA einen offenen Leserahmen von 735 bp auf und kodiert für ein Protein von 244 aa. Der 5'-untranslatierte Bereich ist 224 bp lang, während der 3'-untranslatierte Bereich sich über 427 bp erstreckt. Im 3'-Bereich des *TPXI* Gens wurde kein konserviertes Polyadenylierungssignal gefunden. Die ermittelte *TPXI* cDNA enthielt drei zusätzliche ATG-Codons stromaufwärts des ermittelten Translationsstartcodons, die sich nicht im Leserahmen befinden. Das *TPXI* Protein umfasst 244 aa inklusive eines 22 aa langen hydrophoben Signalpeptids für die Translokation in das endoplasmatische Retikulum. Die Position des Signalpeptids wurde durch Vergleiche mit Konsensussequenzen von Signalpeptiden festgelegt (Simon und Blobel, 1993).

↓									
	ATCTTATTCT	CGCTGTAGGA	GGGGAGAGTA	GCCGTTAGTC	TCTGGCTCCG	GACCCGCAGC			60
	AGGAAGATTT	GGCAGCTGGA	GCCAGGGATT	<u>ATGAAACCAT</u>	ACACCATCCT	CTTCCTGTAG			120
	<u>ATGCCCTGGG</u>	<u>AGACCCATGT</u>	TATCCTGGTC	TTAAAAACACC	TGCAAGAAAAG	GGCACTTTTC			180
	CATTACTGAA	GCAGATACTT	CATCCTGCTC	TGAGAACCAA	<u>CACAATG</u> GCT	TTCCTCCAG			240
					M A F L P				5
	TGGTGATGTT	TTTGGCTGCT	GTGCTACTTC	CATCTTTACC	CACAGAAGGA	AAGGATCCAG			300
	V V M F	L A A	V L L	P S L P	T E G	K D P			25
	CTTTTTCTGC	TTTGTAAACC	ACTCAAACCC	AAGTCCAAAG	AGAGATTGTA	AATAAACACA			360
	A F S A	L L T	T Q T	Q V Q R	E I V	N K H			45
	ATGAACTAAG	GAAATCAGTC	TCTCCACCTG	CCAGCAACAT	GCTAAAGATG	GAATGGAGCA			420
	N E L R	K S V	S P P	A S N M	L K M	E W S			65
	GAGAAGCAAC	AGCAAATGCC	CAAAAGTGGG	CAAACAAGTG	CACTTTAGAA	CACAGTAGTG			480
	R E A T	A N A	Q K W	A N K C	T L E	H S S			85
	CAGATGACCG	GAAAACCAAGT	ACGAGATGTG	GTGAGAATAT	CTATATGTCA	AGCGACCCTA			540
	A D D R	K T S	T R C	G E N I	Y M S	S D P			105
	CTCCCTGGTC	AGATGCAATC	CAAAGCTGGT	ATGATGAGAG	CCTTGACTTT	ACCTATGGTG			600
	T P W S	D A I	Q S W	Y D E S	L D F	T Y G			125
	TAGGACCAA	GAGTGCCGGT	TCAGTAGTTG	GACATTACAC	TCAGGCTGTT	TGGTACTCAT			660
	V G P K	S A G	S V V	G H Y T	Q A V	W Y S			145
	CTTACCRCGT	TGGATGTGGA	ATTGCCTACT	GTCCCAATCA	AGAAAAGCCTA	AAATACTACT			720
	S Y R V	G C G	I A Y	C P N Q	E S L	K Y Y			165
	ATGTTTGCCA	ATATTGTCCT	GTTGGTAATA	ATGTGAATAA	AAAGAACACC	CCTTACCAGC			780
	Y V C Q	Y C P	V G N	N V N K	K N T	P Y Q			185
	AAGGAACACC	TTGTGCCAGT	TGCCCTGGTA	ACTGTGACAA	TGGACTATGC	ACCAATAGCT			840
	Q G T P	C A S	C P G	N C D N	G L C	T N S			205
	GCGAGTATGA	AGATCTCCTT	AGTAACTGTG	ATTCTTGAA	GAAAACAGCT	GGCTGTGAAC			900
	C E Y E	D L L	S N C	D S L K	K T A	G C E			225
	ATGAATGCT	CAAGGAAAAG	TGCAAGGCTA	CTTGTCGATG	TGAAAACAAA	<u>ATTTACTGAA</u>			960
	H E L L	K E K	C K A	T C R C	E N K	I Y			244
	<i>GTTCCAGTGT</i>	<i>GCATGCGCA</i>	<i>AGACTAAGTG</i>	<i>GAGAAGGGCT</i>	<i>GCATCCTTTT</i>	<i>GTTGACACAT</i>			1020
	<i>ACCAGAGGGG</i>	<i>AAATTATAGG</i>	<i>CATGCTAGTT</i>	<i>GCAAATTTGA</i>	<i>TTCCAGACAG</i>	<i>CAATGCATCT</i>			1080
	<i>TTTTCTGTTG</i>	<i>GATCTTTACA</i>	<i>GAAATCTCTT</i>	<i>CCATACAATG</i>	<i>ATTTACAAAA</i>	<i>GCAGAGTAGT</i>			1140
	<i>CTGTGATGAC</i>	<i>AACATGGGGC</i>	<i>TTTGATGTGA</i>	<i>ATTTGATACT</i>	<i>TTAAATGTAA</i>	<i>TGAATTGAGT</i>			1200
	<i>CAAGTTGAAG</i>	<i>ATTTGAATGC</i>	<i>TCTACAACCA</i>	<i>TAAGACTTAG</i>	<i>GTCTAGAAC</i>	<i>TTTGGATTAA</i>			1260
	<i>AATGAAGAAT</i>	<i>TACGTGTTTC</i>	<i>CTGAAACAAC</i>	<i>ATGTCAACAA</i>	<i>AAACCAAAAA</i>	<i>CACTGTAATA</i>			1320
	<i>TCAATGTCAT</i>	<i>TCCTACTGCA</i>	<i>TGATTTTTAA</i>	<i>CATGTATAAA</i>	<i>GAATGAATTC</i>	<i>AAAGATTGAC</i>			1380
	<i>ATCTAC</i>								1386

Abbildung 4: *TPX1* cDNA mit Translation. Nichtcodierende Bereiche sind kursiv dargestellt. Der dicke Pfeil markiert den Transkriptionsstart. Das ATG-Startcodon sowie das TGA-Stopcodon sind fett und unterstrichen hervorgehoben. Die drei anderen nicht im Leserahmen befindlichen ATG-Codons sind unterstrichen gekennzeichnet.

Das equine TPX1 Protein weist 84 % Übereinstimmung zu dem orthologen Protein vom Mensch, 63 % zu Maus, 59 % zu Ratte und 70 % zu Meerschweinchen auf (Abbildung 5). Die für die CRISP Familie charakteristischen 16 konservierten Cysteine sind in allen bisher beschriebenen orthologen TPX1 Proteinen vertreten.

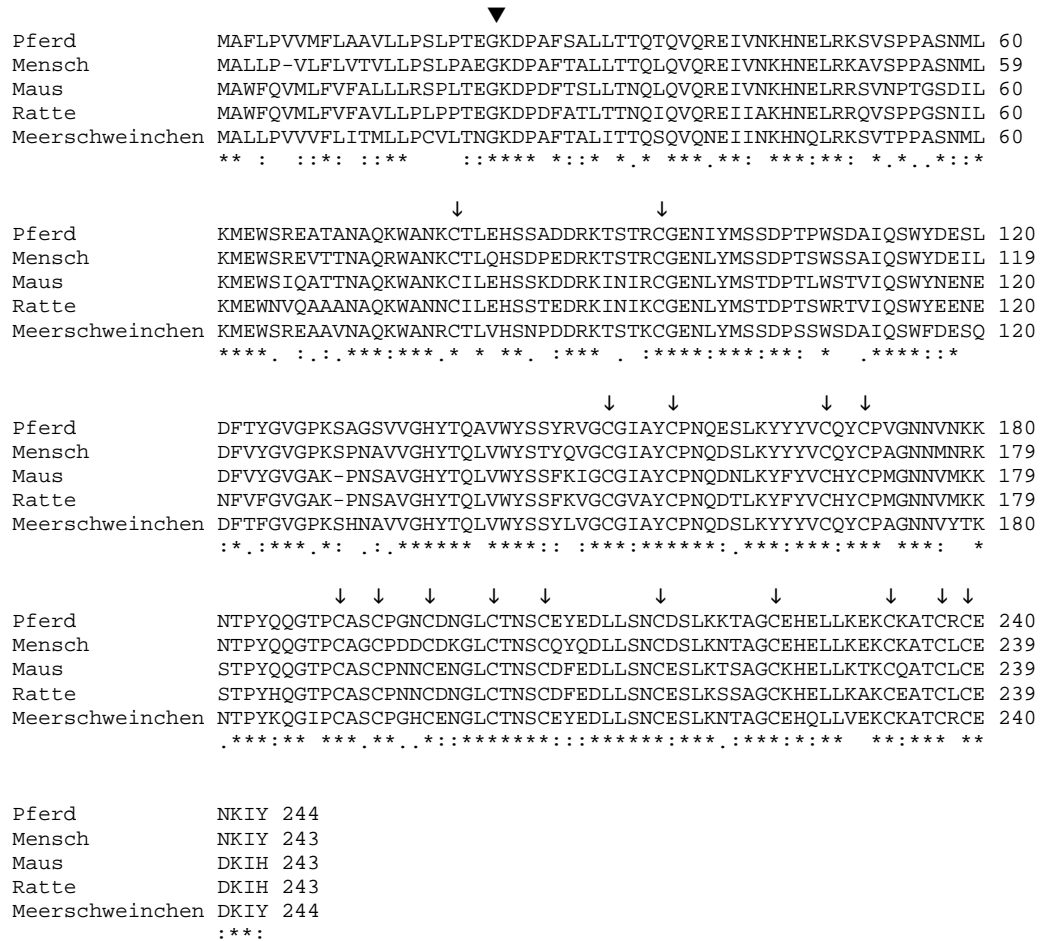


Abbildung 5: Alignment des equinen TPX1 Proteins mit den Sequenzen der bekannten orthologen TPX1 Proteine von Mensch, Maus Ratte und Meerschweinchen. Die Sequenzen wurden aus den Datenbankeinträgen mit den Zugangs Nummern NP_003287.1 (TPX1 Mensch), NP_033446.1 (Tpx1 Maus), JE0204 (Tpx1 Ratte) and Q60477 (AA1 Meerschweinchen) entnommen. Das Dreieck (▼) markiert die Spaltstelle des Signalpeptides. Die 16 konservierten Cysteine sind durch Pfeile (↓) hervorgehoben. Identische Aminosäuren sind durch Sterne unter dem Alignment gekennzeichnet, während Punkte und Doppelpunkte ähnliche oder sehr ähnliche Aminosäuren bezeichnen.

5.1.3 Sequenzanalyse der equinen *AEG2* cDNA und der abgeleiteten Proteinsequenz

Unter Verwendung der gleichen RLM-RACE Strategie wie bei den *AEG1* und *TPX1* cDNAs, wurde eine 1287 bp Volllänge-cDNA Sequenz des equinen *AEG2* Gens ermittelt. Hier wurden aus den RACE Experimenten zwei PCR-Produkte erhalten. Bei der Gelelektrophorese nach der PCR zeigte sich eine starke Bande, die dem Volllänge-*AEG2* Transkript entspricht, und eine schwache Bande eines alternativ gespleißten Transkriptes. Beim Vergleich der Sequenz des alternativ gespleißten Transkriptes mit dem Volllänge-Transkript wurde festgestellt, dass das Exon 2 des *TPX1* Gens in diesem Transkript fehlt. Die cDNA des alternativen Transkriptes umfasst 1216 bp. Unter der Annahme, dass das erste im gleichen Leserahmen befindliche Startcodon verwendet wird, kodiert die alternative Spleißvariante für ein Protein von 187 Aminosäuren. Stromaufwärts von dem vermutlichen ATG-Startcodon wurde ein weiteres ATG-Codon gefunden, dem aber nach zwei Codons ein Stop-Codon folgt. Da durch den alternativen Spleißvorgang das Signalpeptid für die Translokation durch die Membran verloren geht, kann das so entstehende Protein nicht sezerniert werden.

Alle 16 analysierten 5'-Klone des langen Transkripts wiesen als Anfang Position -58 der Volllänge-*AEG2* cDNA auf. Daher wird die Transkription des equinen *AEG2* Gens mit großer Wahrscheinlichkeit an dieser Position initiiert. Am 3'-Ende der *AEG2* cDNA befindet sich ein Polyadenylierungssignal (AATAAA) 15 bp stromaufwärts der Polyadenylierungsstelle. Der 5'-untranslatierte Bereich ist 58 bp lang und der 3'-untranslatierte Bereich erstreckt sich über eine Länge von 491 bp (Abbildung 6). Die Sequenzdaten der equinen *AEG2* cDNA wurden unter der Accession AJ459964 in der EMBL Nucleotide Sequence Database hinterlegt. Gegenüber einer früher beschriebenen *AEG2* cDNA (Schambony et al., 1998b) fiel eine Insertion von drei Basen unmittelbar vor dem Startcodon auf.

↓	ACCTTTCTTC	TTCCTACAAG	TAAAGCAGAT	<u>GCTTCACCTT</u>	GCTCTGAGAA	CCAAC	<u>TCAAT</u>	60
							M	1
	<u>GGCATTATTA</u>	CCAGTGTTC	TGTTTCTGGC	TGCCGTGTTG	CTTCCATTCT	TTCCTGCAAG		120
	A L L	P V L	L F L A	A V L	L P F	F P A S		21
▼								
	TGGACAGGAT	CCAGGTTTTG	CTGCTTTGTC	AATCACCAAA	AGCGAAGTCC	AAAAAGAGAT		180
	G Q	D P G F	A A L S	I T K	S E V	Q K E I		41
	TGTAATAAAA	CACAATGACC	TAAGGAGAAC	AGTCTCTCCA	CTTGCCAGTA	<u>ACATGCTAAA</u>		240
	V N K	H N D	L R R T	V S P	L A S	N M L K		61
	GATGCAATGG	GACAGCAAGA	CAGCAACAAA	TGCCCAAAAC	TGGGCCAACA	AGTGCCTTCT		300
	M Q W	D S K	T A T N	A Q N	W A N	K C L L		81
	CCAACACAGT	AAAGCAGAAG	ATCGCGCAGT	CGGTACAATG	AAATGTGGTG	AGAATCTCTT		360
	Q H S	K A E	D R A V	G T M	K C G	E N L F		101
	TATGTCGAGT	ATACCTAATT	CCTGGTCAGA	TGCAATCCAA	AATGGCATG	ATGAGGTCCA		420
	M S S	I P N	S W S D	A I Q	N W H	D E V H		121
	TGATTTTAAA	TACGGTGTGG	GGCCAAAGAC	TCCCAATGCA	GTAGTTGGAC	ATTATACCCA		480
	D F K	Y G V	G P K T	P N A	V V G	H Y T Q		141
	GGTTGTTTGG	TACTCATCTT	ACCGTGTGG	ATGTGGAATA	GCCTATTGTC	CCAAACAAGG		540
	V V W	Y S S	Y R V G	C G I	A Y C	P K Q G		161
	AACTCTAAAA	TACTACTATG	TTTGCCAATA	TTGTCTGCT	GGTAATTATG	TGAATAAAAT		600
	T L K	Y Y Y	V C Q Y	C P A	G N Y	V N K I		181
	AAATACTCCT	TATGAACAAG	GAACACCTTG	TGCCCGTTGC	CCTGGTAACT	GTGACAATGG		660
	N T P	Y E Q	G T P C	A R C	P G N	C D N G		201
	ACTATGCACC	AATAGCTGCG	AGTATGAAGA	TCTCGTTAGT	AACTGTGATT	CCTTGAAGAA		720
	L C T	N S C	E Y E D	L V S	N C D	S L K K		221
	AATAGCTGGC	TGTGAACATG	AATTGCTCAA	GGAAAACATG	AAGGCTACTT	GTCAATGTGA		780
	I A G	C E H	E L L K	E N C	K A T	C Q C E		241
	AAACAAAATT	TACTGAAGTT	CCAGTGTGCA	TTGCGCAAGA	CAAAGTGGAG	AAGGGCTGCA		840
	N K I	Y						245
	<i>TCCTTTTGGT</i>	<i>GACACATACC</i>	<i>AGAGGGGAAA</i>	<i>TTATAGGCAT</i>	<i>GCTAGTTGCA</i>	<i>AAGCTGATT</i>		900
	<i>CAAACCGTAA</i>	<i>TGCATCTTTT</i>	<i>TCTCTGGAT</i>	<i>CTTTACAGAA</i>	<i>ATCTCTTCCA</i>	<i>CACAATGATT</i>		960
	<i>TACAAAAGCA</i>	<i>GAGTAGTCTG</i>	<i>TGATGACAAC</i>	<i>TTGGGCTTT</i>	<i>GATATAAATT</i>	<i>TGGTACTTTA</i>		1020
	<i>AATGTAATAA</i>	<i>ATTGAATCAA</i>	<i>TTGGAGATTT</i>	<i>TGAAAGTTGT</i>	<i>ATAACCATAA</i>	<i>GACTTAGGTC</i>		1080
	<i>CCTAGAACTT</i>	<i>TGGATTAATA</i>	<i>CGAAGAATTA</i>	<i>CATGTTTCCT</i>	<i>GAAACAACAT</i>	<i>GCCCCCCCAA</i>		1140
	<i>AATTTACAGA</i>	<i>CTAACAAATC</i>	<i>ATTGTTATTC</i>	<i>CTAATACATG</i>	<i>ATTTTTTACC</i>	<i>TGTATAAAGA</i>		1200
	<i>ATGAATTCAA</i>	<i>AGATTGACAT</i>	<i>GTACTTTCT</i>	<i>CTTATGGCCT</i>	<i>CCTCTTTAAA</i>	<i>CAAAAGTAAT</i>		1260
	<i>CAGTAAG</i>	<u><i>AAT</i></u>	<u><i>AAA</i></u>	<i>GAACCTTA</i>	<i>CCTAAGC</i>			1287

Abbildung 6: *AEG2* cDNA mit Translation. Nichtcodierende Bereiche sind kursiv dargestellt. Der dicke Pfeil markiert den Transkriptionsstart. Das Dreieck (▼) markiert die Spaltstelle des Signalpeptides. Die ATG-Startcodons der beiden Spleißvarianten sowie das TGA-Stopcodon sind fett und unterstrichen hervorgehoben. Das zusätzliche im Leserahmen befindlichen ATG-Codon des *AEG2* Gens sowie das nicht im Leserahmen befindliche ATG-Codon der alternativen Spleißvariante sind unterstrichen gekennzeichnet. Die in der alternativen Spleißvariante fehlenden Nukleotide sind grau unterlegt. Die drei in der bisher beschriebenen equinen *AEG2* cDNA (Schambony et al., 1998b) fehlenden Nukleotide sowie das Polyadenylierungssignal ist durch einen Kasten hervorgehoben.

Es konnte bestätigt werden, dass die equine *AEG2* cDNA für ein Protein von 245 Aminosäuren codiert (Schambony et al., 1998). Im Gegensatz zum Menschen und der Maus enthält die equine *AEG2* cDNA im 5'-untranslatierten Bereich ein zusätzliches ATG Codon, das sich im Leserahmen befindet. Sollte dieses ATG als Startcodon verwendet werden, hätte das equine CRISP3 Protein ein längeres Signalpeptid und eine Gesamtlänge von 255 Aminosäuren. Das equine *AEG2* Protein weist 66 % Übereinstimmung zum orthologen Protein von Mensch und 48 % zur Maus auf (Abbildung 7).



Abbildung. 7: Alignment des equinen *AEG2* Proteins mit den Sequenzen der bekannten orthologen *AEG2* Proteine von Mensch und Maus. Die Sequenzen wurden aus den Datenbankeinträgen mit den Accessions S68691 (SGP28 Mensch) und NP_033769.1 (*AEG2* Maus) entnommen. Das Dreieck (▼) markiert die Spaltstelle des Signalpeptides. Die 16 konservierten Cystein Reste sind durch Pfeile (↓) hervorgehoben. Identische Aminosäuren sind durch Sterne unter dem Alignment gekennzeichnet, während Punkte und Doppelpunkte ähnliche oder sehr ähnliche Aminosäuren bezeichnen.

5.2. Genomstruktur der equinen *CRISP* Gene

5.2.1 Sequenzanalyse des *AEG1* Gens

Um das equine *AEG1* Gen zu klonieren und zu sequenzieren, wurde eine genomische Pferde-BAC-Genbank (Godard *et al.*, 1998) mit Hilfe der PCR durchmustert und ein positiver BAC Klon mit der Bezeichnung 106F4 isoliert. Durch Pulsfeldgelelektrophorese wurde eine Insertgröße von etwa 90 kb ermittelt und die Identität des Klons durch Hybridisierung mit einer genspezifischen Sonde betätigt. Die Sonde wurde durch eine PCR mit den für die Durchmusterung verwendeten Primern auf genomischer Pferde DNA erstellt.

Von dem isolierten BAC-Klon wurde eine „Shotgun“ Plasmid Genbank angelegt. Dazu wurde die DNA des BAC-Klons mechanisch zu einer durchschnittlichen Größe von 1 kb bis 2,5 kb geschert und in den pCR4Blunt-TOPO Plasmidvektor ligiert. Die Plasmide wurden anschließend in *E. coli* transformiert. Auf diese Weise wurden etwa 300 verwendbare Plasmid-Subklone erhalten. Zur Ermittlung der Sequenz des BAC-Inserts wurden zunächst „Shotgun“ Plasmid-Subklone jeweils von beiden Rändern ansequenziert bis etwa 360 kb Rohdaten bzw. eine vierfache Abdeckung der Gesamtsequenz erhalten worden war. Anschließend wurden die noch vorhandenen Lücken durch eine „Primer Walking“ Strategie geschlossen. Auf diese Weise wurde die gesamte Sequenz des 90631 bp langen BAC-Inserts auf beiden DNA-Strängen bestimmt. Die Sequenzdaten des equinen *AEG1* Gens wurden unter der Accession AJ315378 in der EMBL Nucleotide Sequence Database hinterlegt.

Durch Datenbankvergleiche der Gesamtsequenz des BAC-Klons wurde 39 kb stromabwärts vom *AEG1* Gen das intronlose *PGK2* der testis-spezifischen Phosphoglyceratkinase lokalisiert. Bei diesen Datenbankvergleichen wurde außerdem eine hohe Homologie zum humanen *AEG1* Gen festgestellt. Daher wurde eine Analyse der Gesamtsequenz mit dem Programm PIP-Maker durchgeführt, um durch mögliche Homologien in nichtkodierenden Regionen regulatorische DNA-Bereiche identifizieren zu können.

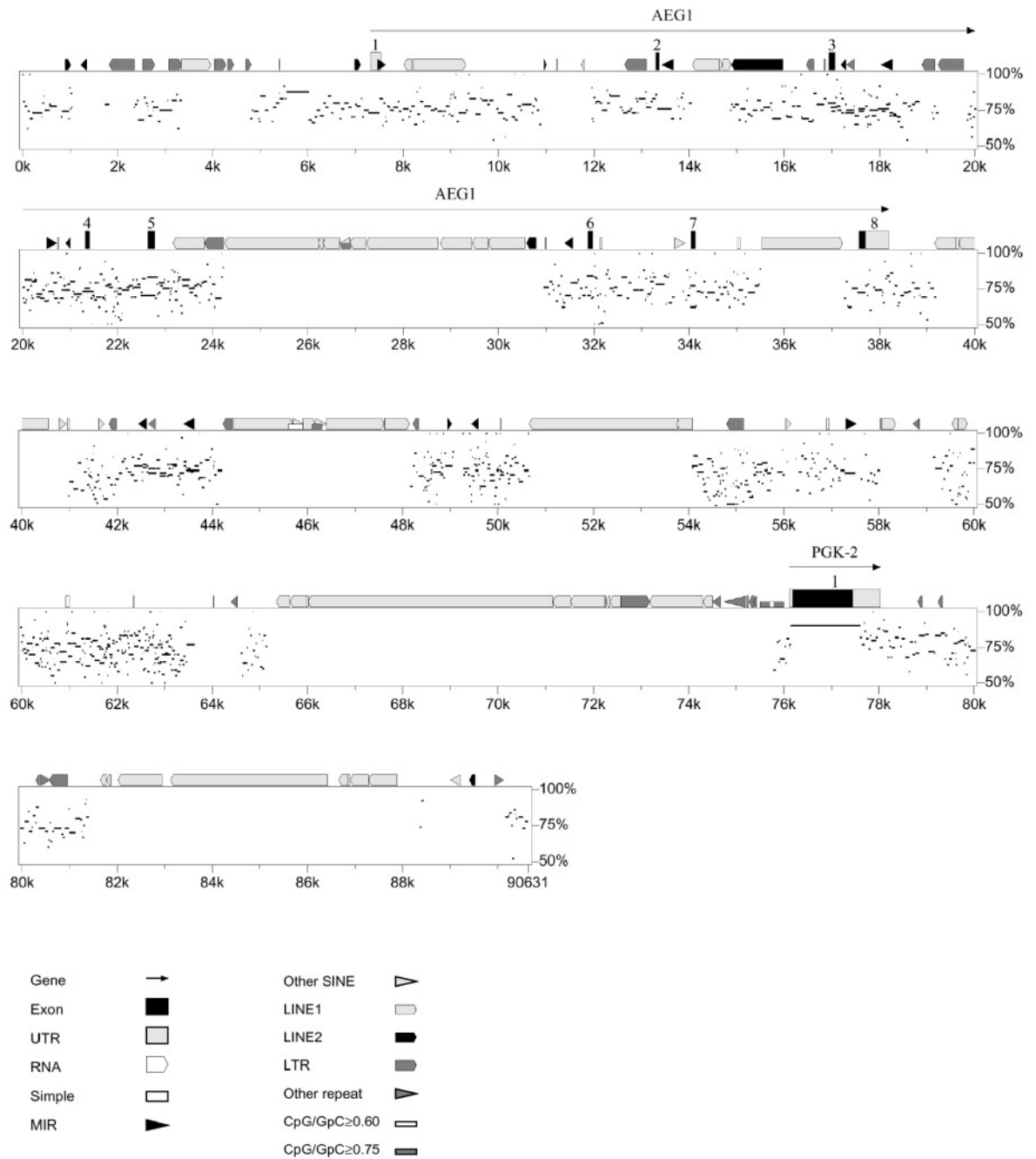


Abbildung 8: Genomische Organisation des equinen *AEG1* Locus. Oberhalb der Kästen sind wichtige Sequenzinformationen wie die Lage der Exons und repetitiver Elemente dargestellt. Unter diesen Sequenzinformationen ist die Übereinstimmung der equinen genomischen Sequenz mit der entsprechenden humanen genomischen Sequenz (abgeleitet aus den Datenbankeinträgen AL121974 und AL356586) grafisch dargestellt. Die Abbildung wurde mit Hilfe der PipMaker Software (Schwartz et al., 2000) erstellt.

Die Exons des *AEGI* Gens wurden durch einen Vergleich der genomischen equinen Sequenz mit der equinen *AEGI* cDNA ermittelt (Abbildung 8). Es wurden insgesamt acht Exons mit einer Größe von 71 bis 624 bp identifiziert. Die Längen der Introns betragen 1225 bis 9095 bp. Das equine *AEGI* Gen hat somit eine Gesamtlänge von etwa 31 kb. Genauere Angaben über die einzelnen Exon/Intron-Übergänge sind in Tabelle 2 zusammengefasst. Alle Exon/Intron Übergänge weisen die AG/GT Dinukleotide auf, die von U1/U2-snRNP-haltigen Spleißosomen erkannt werden. Die Transkriptionsstartstelle wurde durch einen Vergleich mit dem längsten 5'-RACE Produkt und der genomischen *AEGI* Sequenz ermittelt. Sie liegt 20 bp stromabwärts von einer TATA-Box.

Tabelle 2: Exon/Intron Übergänge des equinen *AEGI* Gens. Exon-Sequenzen sind in Großbuchstaben und Intron-Sequenzen in Kleinbuchstaben dargestellt. Untranslatierte Bereiche sind kursiv gekennzeichnet. Das ermittelte ATG Translationsstartcodon wurde am Anfang des zweiten Exon lokalisiert und ist unterstrichen gekennzeichnet. Die konservierten GT/AG Dinukleotide an Exon/Intron Übergängen sind fett hervorgehoben. Die Position +1 entspricht dem Adenin des Translationsstartcodons ATG. Für das erste Exon sind anstelle einer 3'-Spleißstelle die drei gefundenen Transkriptionsstartstellen an den Positionen -226, -224 und -111 unterstrichen dargestellt. Für das letzte Exon ist das geringfügig von der Consensus-Sequenz abweichende Polyadenylierungssignal an Stelle eines Exon/Intron-Übergangs in einem Kasten dargestellt.

3'-Spleißstelle	Exon	5'- Spleißstelle	Intron-phase	Intron-größe
-226		-3		
aatgcatccacaatc <u>ATG</u> <i>GATT</i> ... (Exon 1, 224 bp	224 bp	... <i>ACTTGGg</i> taagttatataatg	1	5766 bp
tgcatccacaatcat <u>GATT</u> ...	222 bp			
tgacttcttc <u>ATAGA</u> ...	109 bp)			
-2		+69		
tacttttcttttc <u>ag</u> <u>AGATGA</u> ... (Exon 2, 71 bp)	71 bp)	... <i>ATAAGAg</i> tatgtgaggattgt	0	3573 bp
+70		+198		
atthttctgataaac <u>ag</u> GCTAAA ... (Exon 3, 129 bp)	129 bp)	... <i>AAGATGg</i> taagattaattaaa	0	4246 bp
+199		+289		
actctccttttgt <u>ag</u> AATTGG ... (Exon 4, 91 bp)	91 bp)	... <i>TTACAAG</i> taggtatatgtcta	1	1225 bp
+290		+438		
ctttcctcctttc <u>ag</u> ATACCT ... (Exon 5, 149 bp)	149 bp)	... <i>ACTCAGg</i> taaactgtgtactc	0	9095 bp
+439		+536		
tctttatcttttgc <u>ag</u> CTTGTT ... (Exon 6, 98 bp)	98 bp)	... <i>TCATGAg</i> tatgtattatacag	2	2066 bp
+537		+625		
ttattctttttta <u>ag</u> GGGAAA ... (Exon 7, 89 bp)	89 bp)	... <i>TTTGTAg</i> taagttttacatct	1	3451 bp
+626		+1249		
gttttgttctttc <u>ag</u> CTAACCC ... (Exon 8, 624 bp)	624 bp)	... <i>TCATTA</i> AATAA <i>TCCACAAAGC</i>		

5.2.2 Sequenzanalyse des *TPX1* Gens

Um die equinen *TPX1* und *AEG2* Gene zu analysieren, wurden wie beim *AEG1* Gen aus der equinen BAC-Genbank zwei Klone isoliert. Für das *TPX1* Gen wurde der Klon mit der Bezeichnung 399F10 und für das *AEG2* Gen der Klon mit der Bezeichnung 304F9. Die Größe der beiden Klone wurde durch Pulsfeldgelelektrophorese mit 75 kb für den *TPX1* Klon und 95 kb für den *AEG2* Klon ermittelt. Nach Hybridisierung der Klone mit genspezifischen Sonden wurde bestätigt, dass sich Teile der entsprechenden Gene auf den Klonen befanden. Bei einer Kontroll-PCR mit den genspezifischen Primern der Genbankdurchmusterung wurde festgestellt, dass sich die BAC Klone für das *TPX1* und *AEG2* Gen in weiten Bereichen überlappen.

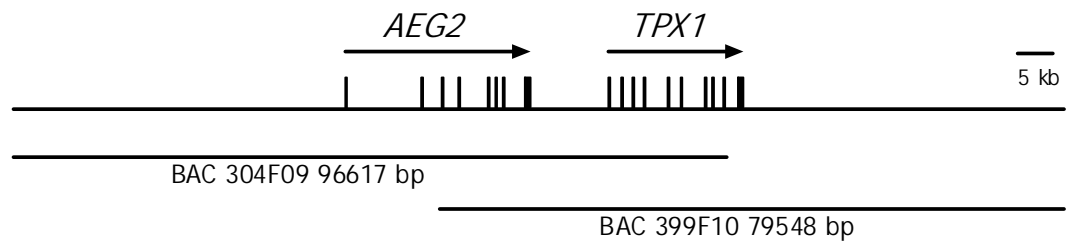


Abbildung 9: Schematische Darstellung des *AEG2* und *TPX1* BAC-Contigs. Die BAC-Klone sowie die genomische DNA sind durch horizontale Linien dargestellt. Die Exons sind durch senkrechte Linien dargestellt und die Lage der Gene durch horizontale Pfeile markiert.

Die 137520 bp Sequenz des gesamten Contigs wurde mit der gleichen Strategie wie beim *AEG1* BAC-Klon 106F4 ermittelt und unter der Zugangs Nummer AJ459965 in der EMBL Nucleotide Sequence Database hinterlegt. Die *TPX1* Exons wurden durch einen Vergleich der genomischen Sequenzen mit der equinen *TPX1* cDNAs ermittelt (Abbildung 10).

Für das equine *TPX1* Gen wurden so 10 Exons mit einer Größe von 37 bp bis 555 bp identifiziert. Die Introngrößen wurde mit 427 bp bis 3594 bp ermittelt. Hieraus ergibt sich für das equine *TPX1* Gen eine Gesamtgröße von 17 kb. In der Tabelle 3 sind die Details der einzelnen Exon/Intron-Übergänge dargestellt. Die Dinukleotide AG am Ende der Introns (3'-Spleißstelle) und GT am Beginn der Introns (5'-Spleißstelle) sind bei allen Exon/Intron-Übergängen konserviert und stimmen mit den Konsensussequenzen von

anderen eukaryontischen Genen überein. Durch einen Vergleich der genomischen *TPXI* Sequenz mit dem längsten 5'-RACE Produkt wurde die Transkriptionsstartstelle ermittelt. Im Gegensatz zu den beiden anderen Vertretern der CRISP Genfamilie, *AEG1* und *AEG2*, konnte im Promotorbereich des equinen *TPXI* keine TATA-Box identifiziert werden. Der *TPXI* Promotor weist jedoch einen hohen Anteil an CpG Dinukleotiden auf ($CpG_{obs}/CpG_{exp} > 0.6$). Der gesamte GC-Gehalt des *TPXI* Promotors liegt allerdings unter 50 % und repräsentiert daher keinen typischen GC-reichen Promotor.

Tabelle 3: Exon/Intron Übergänge des equinen *TPXI* Gens. Exon-Sequenzen sind in Großbuchstaben und Intron-Sequenzen in Kleinbuchstaben dargestellt. Untranslatierte Bereiche sind kursiv gekennzeichnet. Das vermutliche ATG-Translationsstartcodon befindet sich am Anfang des vierten Exon und ist unterstrichen gekennzeichnet. Die konservierten GT/AG Dinukleotide an Exon/Intron Übergängen sind fett hervorgehoben. Die Position +1 entspricht dem Adenin des Translationsstartcodons ATG.

3'- Spleißstelle	Exon	5'- Spleißstelle	Intron-phase	Intron-größe
-224		-139		
cgtagtagccccgggtATCTTA...	(Exon 1, 86 bp)	...GCCAGGgtaagaattagaact		870 bp
-138		-40		
ttttccccattctagGATTAT...	(Exon 2, 99 bp)	...CCATTAgtaagtagcaaggaa		1152 bp
-39		-3		
ttcctcttcctacagCTGAAG...	(Exon 3, 37 bp)	...CCAACAgtaagtgataattca		1708 bp
-2		+69		
tgacttgcttccccagCAATGG...	(Exon 4, 71 bp)	...GGAAAGgtaaggttaaagtgg	0	3414 bp
+70		+186		
cttatttcccttttagGATCCA...	(Exon 5, 117 bp)	...AAGATGgtaagaggcagcaat	0	780 bp
+187		+274		
tcctgtgacatttagGAATGG...	(Exon 6, 88 bp)	...AAACCAGtacgtaaataagtc	1	3594 bp
+275		+420		
tctgtttcttcttagGTACGA...	(Exon 7, 146 bp)	...ACTCAGgtaagggaaacaaat	0	427 bp
+421		+518		
tttatggattttagGCTGTT...	(Exon 8, 98 bp)	...TCCTGTgtaagtatacactta	2	1907 bp
+519		+607		
tatttcaaatttagTGGTAA...	(Exon 9, 89 bp)	...TATGCAGtaagtttgaagtga	1	1949 bp
+608		+1162		
ctgtttttgttacagCCAATA...	(Exon 10, 555 bp)	...AATTCAAAGATTGACATCTAC		

5.2.3 Sequenzanalyse des *AEG2* Gens

Die Daten für das equine *AEG2* Gen wurden nach der gleichen Strategie wie für die beiden anderen Vertreter der CRISP Familie, *AEG1* und *TPXI*, ermittelt. Nach dem Vergleich der equinen genomischen *AEG2* Sequenz mit der Sequenz der *AEG2* cDNA wurden acht

Exons mit einer Größe von 56 bp bis 619 bp ermittelt. Die Größe der Introns liegt zwischen 424 bp und 10433 bp. Daraus ergibt sich eine Gesamtlänge von 24 kb für das equine *AEG2* Gen. Die Tabelle 4 zeigt die Einzelheiten der Exon/Intron-Übergänge des *AEG2* Gens. Wie bei den beiden anderen Vertretern der CRISP Familie sind die Dinukleotide AG/GT an den Exon/Intron Übergängen konserviert. Im Promotorbereich des *AEG2* Gens wurde 24 bp stromaufwärts des Transkriptionsstartes eine TATA-Box identifiziert.

Tabelle 4: Exon/Intron Übergänge des equinen *AEG2* Gens. Exon Sequenzen sind in Großbuchstaben und Intron Sequenzen in Kleinbuchstaben dargestellt. Untranslatierte Bereiche sind kursiv gekennzeichnet. Das ermittelte ATG Translationsstartcodon ist unterstrichen gekennzeichnet. Die konservierten GT/AG Dinukleotide an Exon/Intron Übergängen sind fett hervorgehoben. Für das letzte Exon ist das Polyadenylierungssignal an Stelle eines Exon/Intron Überganges in einem Kasten dargestellt. Die Position +1 entspricht dem Adenin des Translationsstartcodon ATG.

3'- Spleißstelle	Exon	5'- Spleißstelle	Intron-phase	Intron-größe
	acettctctctctctccACCTTT... (Exon 1, 56 bp)...	CCAAC g taagtgatagctcc		10433 bp
-2		+69		
	cttttgtcttccc ag CAATGG... (Exon 2, 71 bp)...	GGACAG g tatggttgaggttg	0	2902 bp
+70		+186		
	ttgactttccttt ag GATCCA... (Exon 3, 117 bp)...	AAGAT g taagaggcagtaac	0	1475 bp
+187		+274		
	tccttgatTTTT ag CAATGG... (Exon 4, 88 bp)...	CAGTC g atgtaaattcata	1	3957 bp
+275		+423		
	tctgtttcttctt ag GTACAA... (Exon 5, 149 bp)...	ACCCAG g taaggggaacaaat	0	424 bp
+424		+521		
	cttacggatttgc ag GTTGTT... (Exon 6, 98 bp)...	TCCTGC g taagtataacttt	2	1894 bp
+522		+610		
	tattttaaattcc ag TGGTAA... (Exon 7, 89 bp)...	TATGC ag taagtttgaagtga	1	1949 bp
+611		+1229		
	ctgtttttgttac ag CCAATA... (Exon 8, 619 bp)...	TAAG AATAAA GAACTTACCTAAGC		

Die genomische Organisation der equinen *TPXI* und *AEG2* Gene, sowie die Verteilung der repetitiven Elemente sind in einem Vergleich mit der orthologen humanen Sequenz in der Abbildung 10 dargestellt.

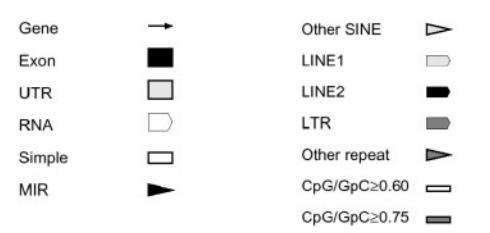
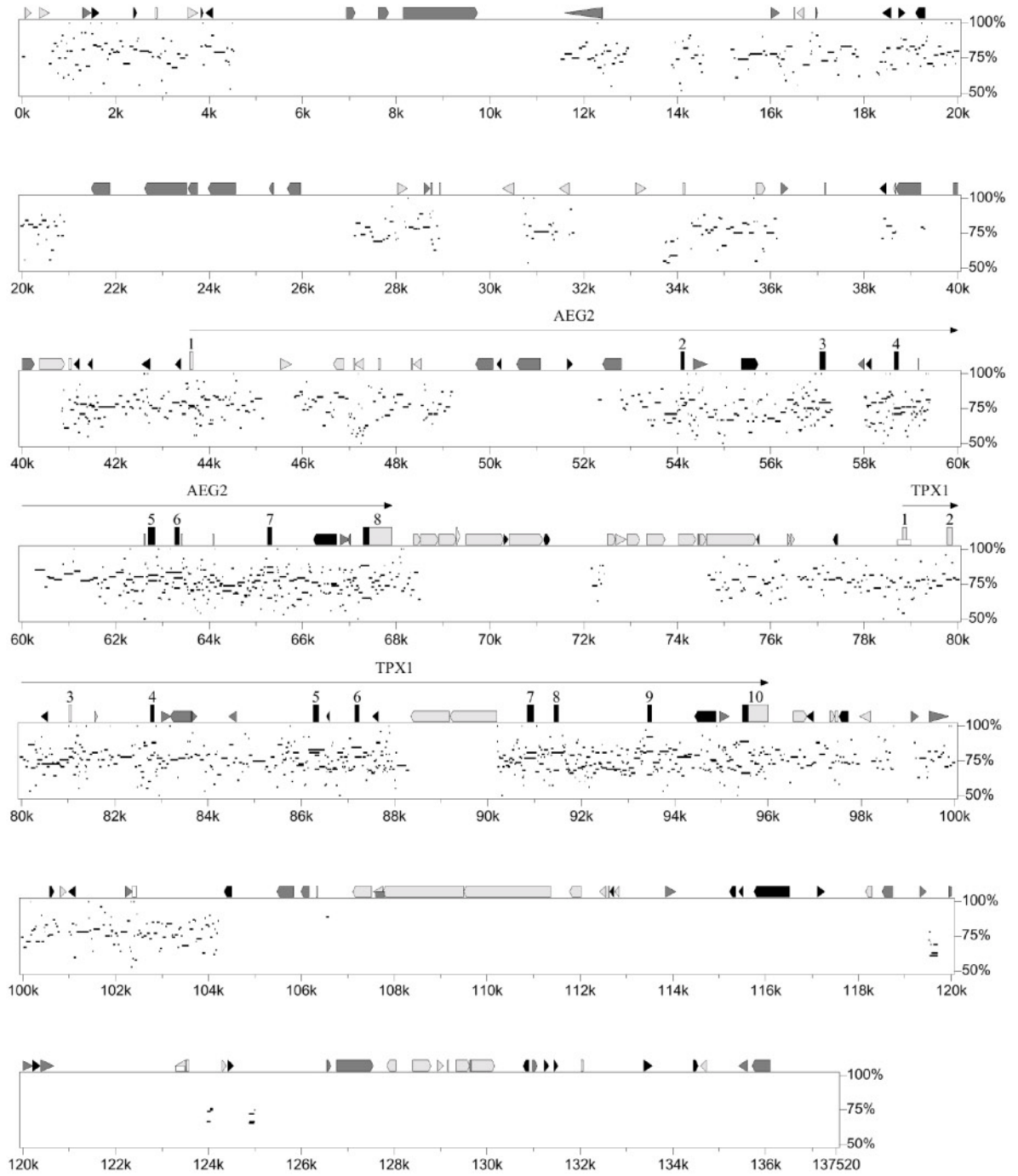


Abbildung 10: Genomische Organisation des equinen *TPX1* und *AEG2* Locus. Oberhalb der Kästen sind wichtige Sequenz Informationen wie die Lage der Exons und repetitiver Elemente dargestellt. Unter diesen Sequenzinformationen ist die Übereinstimmung der equinen genomischen Sequenz mit der entsprechenden humanen genomischen Sequenz (abgeleitet aus den Datenbankeinträgen AL121950 und AL121974) grafisch dargestellt. Die Abbildung wurde mit Hilfe der PipMaker Software (Schwartz et al., 2000) erstellt.

5.3 Sequenzanalyse der flankierender Regionen

5.3.1 Sequenzanalyse des equinen *PGK2* Gens

Die ermittelte 90631 bp Sequenz des BAC-Klons 106F04 wurde mit Hilfe des BLAST Programms mit der Nukleotid Datenbank verglichen. Dabei konnte bestätigt werden, dass sich das equine *AEGL1* Gen vollständig auf diesem Klon befand. Des weiteren wurde das *PGK2* Gen für die Testis spezifische Phosphoglycerat Kinase etwa 39 kb stromabwärts des *AEGL1* Gens identifiziert (Abbildung 8). Das equine *PGK2* Gen enthält keine Introns und entspricht vermutlich einem Retroposon, das für ein funktionelles Protein kodiert. Am 5'-Ende des equinen *PGK2* Genes ist eine CpG-Insel lokalisiert. Die *PGK2* Gene von Mensch (Accession X05246) und Pferd weisen eine Sequenzübereinstimmung von 89 % auf. Das 417 Aminosäuren lange Protein der equinen Phosphoglyceratkinase ist zu 96 % identisch mit der humanen Testis-spezifische Phosphoglyceratkinase.

Durch Analyse der equinen Sequenz mit Hilfe des BLAST Programms wurden Homologien zu vier humanen Sequenzen gefunden. Bei diesen Sequenzen handelt es sich um das *PGK1* Gen auf HSA Xq13, das *PGK2* Gen auf HSA 6p12.3 sowie die Pseudogene *PGKIP1* und *PGKIP2* auf HSA Xq12-q13 und HSA 19p13.3. Sowohl das *PGK2* Gen als auch die beiden Pseudogene *PGKIP1* und *PGKIP2* stellen Retroposons dar, wobei allerdings nur *PGK2* noch einen offenen Leserahmen enthält und für ein funktionelles Protein kodiert (McCarrey und Thomas, 1987). Da das beim Pferd aufgefundene Gen ebenfalls die Struktur eines intronlosen Retroposons sowie einen intakten Leserahmen besitzt und darüber hinaus auf ECA 20q22 in der syntänischen Region zu HSA 6p12 lokalisiert ist, kann davon ausgegangen werden, dass es sich bei dem equinen Gen tatsächlich um das orthologe Gen zum humanen *PGK2* Gen und nicht um ein anderes Mitglied der *PGK*-Genfamilie handelt.

5.3.2 Sequenzanalyse der flankierenden Regionen der equinen *CRISP* Gene

Um Informationen über die flankierenden Bereiche der equinen *CRISP* Gene zu erhalten, wurden die Sequenzdaten der drei BAC-Klone mit verschiedenen Programmen analysiert. Die Sequenz des BAC-Klons 106F04, der das equine *AEG1* und *PGK2* Gen enthielt, wies einen durchschnittlichen GC-Gehalt von 36 % auf. Im 5'-Bereich des *AEG1* Gens konnte keine CpG-Insel gefunden werden. 53 % der analysierten Sequenz bestehen aus repetitiven Elementen, von denen die Mehrzahl zur Familie der „Long interspersed repetitive elements“ (LINE) gehört. Das stimmt mit dem relativ geringen GC-Gehalt der Sequenz überein, der im Zusammenhang mit der Familie der LINE beschrieben wird. Die Verteilung der repetitiven Elemente im Bereich des equinen *AEG1* Locus ist in Abbildung 8 dargestellt.

Die Analyse der Gesamtsequenz des equinen *AEG2/TPX1* BAC-Contigs aus den BAC-Klonen 399F10 und 304F09 ergab, dass beide Gene vollständig auf dem Contig vorhanden waren. Das *TPX1* Gen befindet sich stromabwärts vom *AEG2* Gen und als Abstand zwischen den beiden Genen wurden 10912 bp festgelegt.

Der durchschnittliche GC-Gehalt der Sequenz dieses BAC-Contigs ist 37 %. In der gesamten Sequenz wurde keine CpG-Insel gefunden, die mit den Kriterien von Gardiner-Garden und Frommer (1987) übereinstimmt. Passend zum relativ geringen GC-Gehalt gehört die Mehrzahl der repetitiven Elemente zur Familie der „Long interspersed repetitive elements“ (LINE). Abbildung 10 zeigt eine Darstellung der Verteilung der repetitiven Elemente über die Gesamtlänge des BAC-Contigs.

5.3.3 Analyse der Mikrosatelliten im Bereich der equinen *CRISP* Gene

In den analysierten Sequenzdaten wurden 12 Mikrosatelliten lokalisiert. Vier befanden sich in der Region des *AEG1* Gens und acht weitere um den *TPX1* und *AEG2* Locus. Die Tabelle 5 fasst die Informationen über die gefundenen Mikrosatelliten zusammen.

Tabelle 5: Verteilung der Mikrosatelliten im Bereich der equinen *CRISP* Gene.

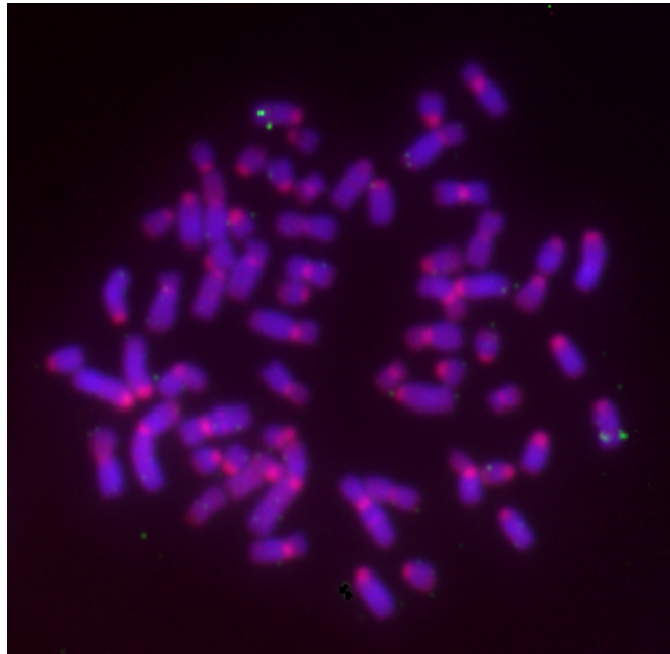
Repetitive Sequenz	Mikrosatelliten Typ	Position
AEG1		EMBL Accession AJ315378
(CA) ₁₁ /(TA) ₂ /(CA) ₆	zusammengesetzt	32135-32172
(AT) ₁₈	perfekt	39630-39665
(TA) ₁₀	perfekt	47609-47628
(TA) ₂ /(AA)/(TA) ₁₀	imperfekt	65987-66012
TPX1/AEG2		EMBL Accession AJ459965
(TG) ₁₅	perfekt	2855-2884
(TG) ₃ /(TA)/(TG) ₈	imperfekt	28923-28951
(TTC)/(TTT)/(TTC) ₆	imperfekt	47092-47115
(CA) ₂ /(CT)/(CA) ₇ /(TA)/(CA) ₄	imperfekt	47609-47638
(CA)/(TA)/(CA) ₁₃	imperfekt	62601-62630
(TG) ₁₁	perfekt	74438-74460
(TA) ₇ /(TG) ₁₉	zusammengesetzt	106315-106352
(TA) ₁₁	perfekt	129633-129655

Für das menschliche Genom wird eine Häufigkeit von einem Mikrosatellit auf etwa 15 kb angegeben (International Human Genome Sequencing Consortium, 2001). In der untersuchten equinen Sequenz lag die Häufigkeit bei einem Mikrosatellit auf 19 kb. Bei einer Untersuchung des homologen Sequenzbereichs des Menschen ergab sich eine Häufigkeit von einem Mikrosatellit auf etwa 11 kb. Daher scheint es, dass das Auftreten von Mikrosatelliten beim Pferd geringer als beim Menschen ist.

5.4 Chromosomale Lokalisation der equinen *CRISP* Gene

Bei der Durchmusterung der equinen BAC-Genbank mit genspezifischen Primern wurden die Klone 106F04, 399F10 und 304F09 isoliert. Durch Hybridisierung mit genspezifischen Sonden wurde bestätigt, dass der Klon 106F04 das *AEG1* Gen beinhaltet (Abbildung 8). Die Klone 399F10 und 304F09 enthalten das *TPX1* und *AEG2* Gen (Abbildung 10). Um die chromosomale Lokalisation der equinen *CRISP* Gene und des *PGK2* Gens zu ermitteln, wurden daher die genomischen BAC Klone 106F4 und 399F10 als Proben für eine Fluoreszenz *in situ* Hybridisierung (FISH) verwendet. Beide BAC-Klone ergaben je zwei Hybridisierungssignale auf den beiden equinen Chromosomen 20q22 (Abbildung 11).

A



B

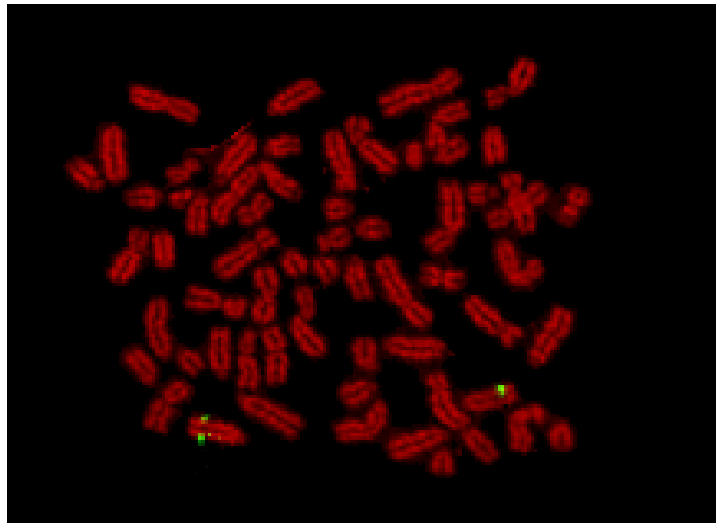


Abbildung 11: Chromosomale Lokalisation der untersuchten Gene durch FISH Analyse. A: *AEG1* und *PGK2*. B: *TPX1* und *AEG2*. Die Digoxigenin markierten BAC Klone mit Teilen der equinen Gene wurden mit equinen Methaphasechromosomen hybridisiert. Die Chromosomen wurden mit Propidiumiodid gegengefärbt und anschließend durch DAPI-Färbung identifiziert (Die Abbildung A wurde von Frau Dr. Heidi Kuiper zur Verfügung gestellt und die Abbildung B von Dr. Bhanu Chowdhary).

6 Diskussion

6.1 Interpretation der Sequenzdaten

6.1.1 Diskussion der Sequenzdaten des *AEGL* Gens

Durch die Verwendung eines modifizierten RACE-Protokolls wurden Volllänge-cDNAs amplifiziert und das Auftreten von Abbruchprodukten deutlich reduziert. Trotzdem wurden mehrere Klone isoliert, die an den Positionen 2 bp beziehungsweise 115 bp stromabwärts der längsten *AEGL* cDNA begannen. Dies deutet darauf hin, dass die Transkription des equinen *AEGL* Gens an drei verschiedenen Stellen initiiert wird, die den Positionen -226, -224 und -111 der ermittelten *AEGL* cDNA entsprechen.

Die equine *AEGL* cDNA weist stromaufwärts des vermuteten Translationsstartcodons innerhalb des 5'-untranslatierten Bereichs, zwei weitere ATG-Codons an den Positionen -147 und -183 der Volllänge-cDNA auf, die sich im Leserahmen befinden. Würde das erste ATG als Startcodon verwendet werden, würde die cDNA für ein Protein von 311 Aminosäuren kodieren. Western Blot Analysen weisen jedoch darauf hin, dass die Größe des *AEGL* Proteins wahrscheinlich mit einem Translationsprodukt von 250 Aminosäuren übereinstimmt (Giese et al., 2002). Daher scheint das equine *AEGL* Gen zu der Gruppe von Genen zu gehören, die untranslatierte Startcodons im 5'-untranslatierten Bereich enthalten. (Peri and Pandey, 2001).

Dass die meisten der gefundenen 5'-RACE Klone an Position -111 der ermittelten *AEGL* cDNA begannen, spricht für die bevorzugte Initiation der Transkription an dieser Position. In diesem Fall könnte die Translation nur am dritten ATG Startcodon beginnen, da die beiden weiter 5'-gelegenen ATG Startcodons nicht mehr enthalten wären. Die Anzahl der Klone ist jedoch nicht aussagekräftig, da eine „nested“-PCR mit insgesamt 60 Zyklen durchgeführt wurde. Eine quantitative Aussage ist daher nicht möglich, da kürzere Produkte bei einer so hohen Zyklenzahl bevorzugt werden. Für eine physiologisch relevante Initiation der Transkription an Position -226 spricht dagegen die Lokalisation der TATA-Box 20 bp stromaufwärts dieser Position.

A

```

acgtctgaaa gagggcatgg ttcactcttag tacctcaata aatatgtagt gaatgcatga      60
gtgcatgcgt gcgtgtaata aagctgataa gttgcttctc ttcttcaggc catggggctca    120
ctttcccttt cagggctctt gctatctgat tgtggctcca ctgaatcatg caggtgatct      180
ctgcagcttg ttccaaacat ggaagtttaa ggctggtgac gtccactgat gacgggtgg[at]    240
                                     ↓ ↓
tataaa ccag aatgcatcc acaatcATGA TTTGCAATCT TCACAGGCTC TAGATTTGTT      300
CTGATTACTA A TCCGTTTGTG CCCTTTGCCT GCTGTAAATT CCTGGAATGTC TTGTAAGCAT    360
                                     ↓
TTTCCTTGCA CTGACTTCTT CATAGAGCAA ATCTTTCCTG TAGGGGGCTG TTGCGAGGCA    420
TTAGCACACT ATCTAGCCAA AGATCTGGGT CTCATCGGGA CACTGAACTT TACAAATTGT      480
GTGAACTTGG gtaagttata                                                    500

```

B

```

tagaaattgt ccaatcccct gcagttacaa gcaatagtaa cacaccatac atagctagaa      60
gaaacaaaaa tcattattct atgtaagagc aggaagtact ctaaatacaa gtcaccttgg    120
tagccccctt aaaagaaaaa aaaatoctca cagctctgat aagataaaaa ttatatgatt      180
aaaagttttt gaagcattta atggcaaata tatacacaca ttttattcac tagatgagca    240
tgaaagagcc agtaaaattt gtttctaata tacagtcatc acatgatgtc aaaatgtcag    300
ttcgatgacc agcaataaaa tctggcatga caaatcctgt gtcatacctg gttaaccaag    360
accagcttgc acaagcagat tgtgcaaaaa caggaactat gcagag[tata aat]agacagc    420
                                     ↓
tgagactctt gtctgATGTC ACCTTCTTC TTCCTGCAGA ACAACGCCTC ATTCTACTCT      480
GAAGCCAGCA gtaagtagtt                                                    500

```

Abbildung 12: Darstellung der *AEG1* Promotorregion von Pferd und Maus. A: 5'-Bereich und Exon 1 des equinen *AEG1* Gens. B: 5'-Bereich und Exon 1 des murinen *Aeg1* Gens. Exon-Sequenzen sind in Großbuchstaben und nichtkodierende Bereiche in Kleinbuchstaben dargestellt. Die ermittelten ATG Codons im untranslatierten Bereich des equinen *AEG1* Gens sind unterstrichen und fett gekennzeichnet. Das GT Dinukleotid der 5'-Spleißstellen der ersten Introns ist fett hervorgehoben. Die dicken Pfeile (▼) bezeichnen die Positionen der Transkriptionsstarts und die dünnen Pfeile (↓) die Positionen der zusätzlichen möglichen Transkriptionsstartstellen im Bereich des equinen *AEG1* Gens. Die TATA-Boxen sind in Kästen dargestellt. Die AREs im Bereich des murinen *Aeg1* Promotors sind grau unterlegt. Aufgrund der hohen Diversität der Promotorbereiche wurden die Sequenzen nicht als Alignment dargestellt.

Ein Vergleich mit dem orthologen *AEG1* Gen des Menschen, dessen Sequenz beim humanen Genom Projekt ermittelt wurde (Accession AL359458), zeigt eine hohe Übereinstimmung zum equinen *AEG1* Gen. Das menschliche Gen erstreckt sich über 37 kb

und ist ebenfalls in acht Exons unterteilt. Im Promotor des humanen *AEGI* Gens wurde jedoch keine TATA-Box lokalisiert. Die im equinen *AEGI* ermittelte TATA-Box, die 20 bp stromaufwärts vom längsten Transkriptionsstart gefunden wurde, stimmt mit der Organisation im orthologen murinen *AegI* Gen überein. Die Sequenz der equinen TATA-Box weicht jedoch geringfügig von der Konsensussequenz ab und unterscheidet sich dadurch von der murinen TATA-Box. Im *AegI* Promotor von Maus und Ratte wurden einige Androgenrezeptor Bindestellen (androgen response elements, AREs) identifiziert (Roberts et al., 2001). Die Homologie zwischen den Nager- und Pferde- *AEGI* Promotoren ist sehr gering und es wurden im equinen *AEGI* Gen keine AREs ermittelt.

Die deutlichen Unterschiede in der Promotororganisation von Pferd, Mensch, Maus und Ratte unterstreichen die bestehende Diversität der *AEGI* Gene und sind vermutlich verantwortlich für die speziesspezifischen Expressionsprofile des *AEGI* Gens.

6.1.2 Diskussion der Sequenzdaten des *TPXI* Gens

Der Hauptanteil der isolierten *TPXI* 5'-RACE Klone begann an der Position -224 der Volllänge-cDNA und alle 3'-RACE Klone endeten an der gleichen Position. Deshalb wurde die Länge der equinen *TPXI* cDNA auf 1386 bp festgelegt. Diese Länge stimmt exakt mit der des orthologen humanen *TPXI* Gens überein. Die wenigen kürzeren 5'-RACE Klone sind vermutlich PCR- oder Klonierungsartefakte, da alle Klone an unterschiedlichen Positionen beginnen und sich keine Häufung einer bestimmten Länge feststellen ließ.

Interessanterweise wurden in allen bisher untersuchten *TPXI* cDNAs von Säugetieren mehrere ATG-Codons im 5'-untranslatierten Bereich gefunden. In der ermittelten equinen *TPXI* cDNA wurden drei zusätzliche ATG-Codons im 5'-untranslatierten Bereich gefundenen, was mit der Anzahl der zusätzlichen ATG-Startcodons in der menschlichen cDNA übereinstimmt. Bei der Maus wurden im 5'-untranslatierten Bereich zwei zusätzliche ATG Startcodons gefunden.

Die im 5'- untranslatierten Bereich der equinen *TPXI* cDNA gefundenen ATG-Codons befinden sich alle nicht im TPX1-Leserahmen und gehen kurzen Leserahmen voran. Der

offene Leserahmen des ersten ATG umfasst 30 bp an der Position -134 bis -105 der Volllänge-cDNA und endet mit dem Codon TAG. Das zweite ATG steht am Anfang des längsten offenen Leserahmens mit 69 bp an der Position -104 bis -35 und endet mit dem Codon TGA. Das letzte der drei ATG-Codons öffnet einen Leserahmen von 18 bp an der Position -88 bis -71 der mit dem Codon TAA endet. Daher ist es unwahrscheinlich, dass diese ATG-Codons als Translationsstartcodons fungieren.

Die genomische Organisation der *CRISP* Gene ist bei Säugern konserviert, was sich auch beim equinen *TPXI* Gen zeigt. Beim Vergleich mit dem orthologen humanen *TPXI* Gen, dessen Sequenz während des humanen Genomprojekts ermittelt wurde (Accession AL121950), wurde eine vollständig konservierte genomische Struktur festgestellt.

In der Promotorregion des equinen *TPXI* Gens konnte keine TATA-Box identifiziert werden. Die stimmt mit den regulatorischen Bereichen bei Mensch und Maus überein. Die Homologie zwischen den *TPXI* Promotoren von Mensch, Maus und Pferd ist sehr gering. Ebenso wie im equinen *AEGL* Promotor wurden im equinen *TPXI* Gen keine AREs ermittelt.

```

accoctgagtg cgctgttggc accacactga tgttttgctg caaacatgtc tatggacttt      60
tccaaacagc tctttcaatt ctaactcgca ttaatttagg agattttgtc tgaaaaggaa      120
cgctccttt ttcagccacg ccctgggcat ttcggatcct gatattatct tcacctcca      180
tctagtgttc aggtgtgtga gcagacatct accttcctca tcaaatgaaa gtaatgagag      240
aaacgcagaa ctctgctttt cacaaggaa gtgcgtaaga ttttctcggg gcagccctag      300
agggcgttgg ggtagtgtca ggaagaattc agttagccaa tgactgctgct gaaggcataa      360
agccctatga cgtcgcacgc acacacgcac gtggtagccc gggtATCTTA TTCTCGCTGT      420
AGGAGGGGAG AGTAGCCGTT AGTCTCTGGC TCCGGCCCCG CAGCAGGAAG ATTTGGCAGC      480
TGGAGCCAGG gtaagaatta                                                    500

```

Abbildung 13: Darstellung der equinen *TPXI* Promotorregion. Exon-Sequenzen sind in Großbuchstaben und nichtkodierende Bereiche in Kleinbuchstaben dargestellt. Das GT Dinukleotid am Beginn des ersten Introns ist fett hervorgehoben. Der Pfeil (↓) bezeichnet die Position des Transkriptionsstarts.

6.1.3 Diskussion der Sequenzdaten des *AEG2* Gens

Bei der Klonierung und Sequenzierung der 5'-RACE Produkte des equinen *AEG2* Gens zeigten sich zwei Produkte mit einem Längenunterschied von 71 bp. Bei einem Vergleich der kürzeren *AEG2* cDNA Sequenz mit der genomischen *AEG2* Sequenz wurde festgestellt, dass das zweite Exon vollständig fehlt. Daraus resultiert eine alternative cDNA mit einer Länge von 1216 bp. Bei Verwendung eines ATG-Startcodons an der Position +233 der Vollängen-*AEG2* cDNA würde die alternativ gespleißte cDNA mit einem offenen Leserahmen von 564 bp für 187 Aminosäuren kodieren. Die Bedeutung dieses alternativ gespleißten Produktes ist zur Zeit noch unklar. Das von der alternativ gespleißten *AEG2* cDNA kodierte Protein besitzt kein Signalpeptid für die Translokation in das endoplasmatische Retikulum. Sollte dieses alternative *AEG2* Protein synthetisiert werden, müsste es daher intrazellulär lokalisiert sein.

Der Vergleich mit der früher veröffentlichten equinen *AEG2* cDNA Sequenz (Schambony et al., 1998b) zeigte, dass die ermittelte equine *AEG2* Volllänge-cDNA drei zusätzliche Nukleotide aufwies, die genau vor dem vermutlichen ATG Startcodon an den Positionen -3 bis -1 liegen.

Sowohl die 16 analysierten 5'-RACE Klone des längeren Transkriptes als auch die acht 5'-RACE Klone des alternativ gespleißten Transkriptes beginnen an der Position -58 der ermittelten Volllänge-cDNA des *AEG2* Gens. Da keine kürzeren 5'-RACE Klone gefunden wurden, wird die Transkription des equinen *AEG2* Gens mit sehr hoher Wahrscheinlichkeit an der Position -58 der Volllänge-cDNA initiiert.

Im Gegensatz zur Organisation der orthologen *AEG2* Gene bei Mensch und Maus befindet sich im 5'-untranslatierten Bereich des equinen *AEG2* Gens ein zusätzliches ATG Codon, das sich innerhalb des Leserahmens befindet. Dieses zusätzliche ATG Codon entspricht jedoch nicht der von Kozak beschriebenen Konsensussequenz, die für eine Initiierung der Translation notwendig wäre (Kozak 1996). Sollte das erste ATG als Startcodon verwendet werden, würde das equine *AEG2* Vorläuferprotein aus 255 aa mit einem längeren Signalpeptid bestehen. Nach der Spaltung des Signalpeptides würde jedoch das gleiche reife Protein aus beiden Translationprodukten entstehen.

Im Promotorbereich des equinen *AEG2* Gens konnte 23 bp stromaufwärts vom Transkriptionsstart eine TATA-Box identifiziert werden. Dies stimmt mit der

Promotororganisation des orthologen *Aeg2* Gens der Maus überein. Obwohl die Sequenzen beider TATA-Boxen mit der Konsensussequenz übereinstimmen, unterschieden sie sich jedoch untereinander. Im murinen *Aeg2* Promotor wurden mehrere AREs beschrieben (Schwidetzky et al., 1995). Im Promotor des equinen *AEG2* konnten dagegen, wie auch beim *AEG1* Gen, keine AREs identifiziert werden.

```

gcatacagtt gtacaatcac tatgcctaca tgaataccca gaaatgccca ccagtgttat    60
gtatgtgctg cttatgattt attaattttat actccctcat ttataaaaac tcagtatttg    120
aaatcattac attttagagg tagaaagaat ctgaaatcat tagtccaagc ctcttcatta    180
tcaagacgtg gaaacgagtc tggaatgatt aaataaagct ttccatagat gaacagaatc    240
aaatgatgat cactgaggag aatcacaaga tactgccaaa actgcagtta gacaaccgac    300
aatgaaattt gagactattg gtatgacaaa tcctgtttca gacctgttga cttaggctta    360
cttgcacaaac cagaatttgc caaaacagga aataggtgta atgTatataT accctctctaa    420
          ↓
ccttctctct ctccACCTTT CTTCTTCCTA CACGTAAAGC AGATGCTTCA CCTTGCTCTG    480
AGAACCAACT gtaagtgata                                                    500

```

Abbildung 14: Darstellung der equinen *AEG2* Promotorregion. Exon-Sequenzen sind in Großbuchstaben und nichtkodierende Bereiche in Kleinbuchstaben dargestellt. Das zusätzliche ATG Codon ist unterstrichen und fett gekennzeichnet. Das GT Dinukleotid am Beginn des ersten Introns ist fett hervorgehoben. Die TATA-Box ist in einem Kasten dargestellt. Der Pfeil (↓) bezeichnet die Position des Transkriptionsstarts.

6.2 Analyse der chromosomalen Lokalisation der equinen *CRISP* Gene

Die chromosomale Lokalisation der equinen *AEG1* und *PGK2* Gene wurde auf ECA 20q22 ermittelt. Beide Gene wurden beim Menschen auf HSA 6p21 lokalisiert (Raudsepp et al., 1996). In der Maus wurde das *Aeg1* Gen auf MMU 17 kartiert (Kasahara et al., 1995; Yoshino et al., 1998). Damit bestätigt die Lokalisation des equinen *AEG1* die bisher etablierten Übereinstimmungen zwischen dem equinen Chromosom 20, dem menschlichen Chromosom 6 und dem murinen Chromosom 7 (Raudsepp et al., 1996).

Die chromosomale Lokalisation der equinen *TPX1* und *AEG2* Gene ergab wie für das equine *AEG1* Gen ein deutliches Hybridisierungssignal auf ECA 20q22. Dies stimmt mit der kürzlich veröffentlichten Lokalisation des equinen *AEG2* Gens überein. (Lindgren et al., 2001). Die orthologen humanen *TPX1* und *AEG2* Gene wurden auf dem kurzen Arm

des humanen Chromosom 6 (HSA 6p21) lokalisiert (Raudsepp et al., 1996). Bei der Maus wurde das *TPX1* Gen auf MMU 17 kartiert (Kasahara et al., 1989a). Damit wurde die Syntänie zwischen ECA 20, HSA6 und MMU7 durch Lokalisation der equinen *TPX1* und *AEG2* Gene betätigt (Raudsepp et al., 1996). Zusammenfassend wurde hiermit gezeigt, dass alle drei Vertreter der equinen *CRISP* Familie wie in anderen Säugetierspezies eng benachbart auf einer chromosomalen Region lokalisiert sind, wobei sich allerdings die Genreihenfolge von Mensch und Pferd von der Genreihenfolge bei der Maus unterscheidet.

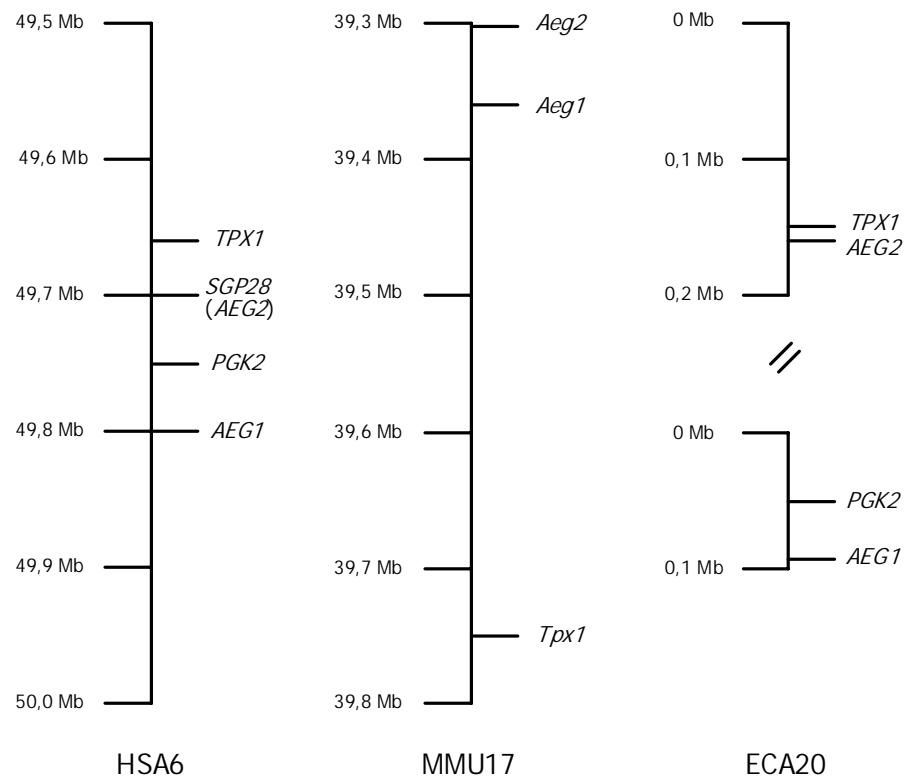


Abbildung 15: Vergleichende Genkarte der *CRISP* Genregion zwischen Mensch, Maus und Pferd. Die Informationen der humanen Lokalisationen wurden aus dem Build 30 der NCBI Datenbank entnommen (<http://www.ncbi.nlm.nih.gov/genome/guide/human>). Die Daten der murinen Lokalisationen wurden aus dem Build 3 der NCBI Datenbank entnommen (<http://www.ncbi.nlm.nih.gov/genome/guide/mouse>). Die angegebenen Genpositionen bei Mensch und Maus entsprechen den Abständen vom Telomer von HSA 6p beziehungsweise der Position auf MMU 17.

6.3 Analyse der Zuverlässigkeit der Sequenzen

Bei der Sequenzierung beider DNA-Stränge der BAC-Klone wurde davon ausgegangen, dass in 10.000 Basen ein Fehler vorkommt. Dies entspricht einer Zuverlässigkeit von 99,99 % (EU-Standard). Das Auftreten von zwei Fehlern in 10.000 Basen würde einer Zuverlässigkeit von 99,98 % entsprechen. In Sequenzbereichen bei denen nur ein DNA-Strang sequenziert wurde, liegt die angenommene Fehlerrate bei 1 bis 2 auf 1.000 Basen. Dies entspricht einer mittleren Zuverlässigkeit von 99,8 bis 99,9 %.

Bei der Sequenzierung des BAC-Klons 106F4, der das *AEGL* Gen beinhaltet, wurden 90631 bp durch mindestens einmalige Sequenzierung beider Stränge ermittelt. In keinem Teil der Gesamtsequenz lagen einzelsträngige Sequenzbereiche vor. Bei 96,1 % der Sequenz lagen zusätzlich zu den Daten beider DNA-Stränge von unabhängigen Subklonen die Daten von einem der beiden DNA-Stränge von mindestens einem weiteren unabhängigen Subklon vor. Bei 3,9 % der Sequenz lag eine einmalige Sequenzierung beider DNA-Stränge vor, bei der die Sequenzdaten aus eindeutigen Bereichen bis maximal 500 bp hinter den Primerbindungsstellen ermittelt wurden. Danach kann für die ermittelte Sequenz eine Zuverlässigkeit von 99,99 % angenommen werden.

Die 137520 bp Gesamtsequenz des BAC-Contigs aus den Klonen 399F10 mit den kodierenden Bereichen des *TPXI* Gens und 304F9 mit den kodierenden Bereichen des *AEGL2* Gens, wurde durch mindestens einmalige Sequenzierung beider DNA-Stränge ermittelt. 98,2 % der ermittelten Sequenzdaten wiesen zusätzlich zu den Sequenzen beider DNA-Stränge von unabhängigen Subklonen die Sequenzen eines der beiden DNA-Stränge von mindestens einem weiteren unabhängigen Subklon auf. Bei 1,8 % der Sequenz lag eine einmalige Sequenzierung beider DNA-Stränge vor, bei der die Sequenzen in einem eindeutigen Bereich bis maximal 500 bp hinter den Primerbindungsstellen lagen. Danach kann für die ermittelte Sequenz eine Zuverlässigkeit von 99,99 % angenommen werden.

6.4 Bedeutung der Arbeit für das Pferde-Genomprojekt

Das Pferde-Genomprojekt (<http://www.uky.edu/Ag/Horsemap/welcome.html>) umfasst zur Zeit 213 Gene auf 153 verschiedenen Klonen. Die längste bisher in der EMBL Nukleotidsequenz Datenbank veröffentlichte genomische Sequenz des Pferdes umfasst 15 kb (Giffard J.M., 1998). Die weiteren zur Verfügung stehenden Sequenzinformationen des Pferdes sind unbearbeitete „working draft“ Sequenzen. Diese stammen aus „Shotgun“ Sequenzierungen von BAC-Klonen. Die vorrangigen Bemühungen des Pferde-Genomprojekts beziehen sich auf die Kartierung von Markern mit Hilfe der humanen Genomkarte. Die Klonierung und Sequenzierung der equinen *CRISP* Gene, bei der insgesamt 227 kb genomische DNA-Sequenz in der Datenbank hinterlegt wurden, ist daher ein weiterer wichtiger Beitrag zum internationalen Pferde-Genomprojekt.

6.5 Ausblick auf weiterführende Untersuchungen

Die in dieser Arbeit ermittelten Sequenzdaten bieten die Grundlage zu weiteren molekularbiologischen Untersuchungen zur Fruchtbarkeit von Hengsten. Im Rahmen eines Folgeprojekts wird eine Mutationsanalyse an 16 verschiedenen Tieren mit definiertem Fruchtbarkeitsstatus durchgeführt. Durch Amplifikation der Exons aller *CRISP* Gene der einzelnen Tiere ist es möglich nach Sequenzunterschieden einzelner Basen (single nucleotide polymorphisms, SNPs) oder anderen Sequenzunterschieden zu suchen. Innerhalb der acht Exons des equinen *AEGI* Gen wurden bisher 17 Polymorphismen gefunden, von denen einer zu einem Austausch einer Aminosäure führt (Jude et al., 2002). Anschließend soll anhand eines größeren Tiermaterials von 116 Tieren des Hannoverschen Warmblutes untersucht werden, inwieweit die gefundenen Sequenzunterschiede mit der Befruchtungsrate korrelieren.

Diese Arbeit ist der erste Ansatz zur molekularbiologischen Untersuchung der Fruchtbarkeit beim Pferd. Bisherige Forschungsprojekte zu diesem Thema wurden im Bereich der Pferdehaltung, der Fütterung, des Besamungsmanagements und bei klinischen Untersuchungen des Genitaltrakts unternommen.

7 Literaturverzeichnis

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402.

Brooks, D.E., Tiver, K. (1983) Localization of epididymal secretory proteins on rat spermatozoa. *J. Reprod. Fert.* 69, 651-657.

Brooks, D.E., Means, A.R., Wriqth, E.J., Singh, S.P., Tiver K.K. (1986) Molecular cloning of the cDNA for androgen-dependent sperm-coating glycoproteins secreted by the rat epididymis. *Eur. J. Biochem.* 161,13-18.

Brooks, D.E. (1987) Androgen-regulated epididymal secretory proteins associated with post-testicular sperm development. *Ann. N. Y. Acad. Sci.* 513, 179-194.

Cameo, M.S., Blaquier, J.A. (1976) Androgen controlled specific proteins in rat epididymis. *J. Encocrinol.* 69, 47-55.

Charest, N.J., Joseph, D.R., Wilson, E.M., French, F.S. (1988) Molecular cloning of complementary deoxyribonucleic acid for an androgen-regulated epididymal protein: sequence homology with metalloproteins. *Mol. Endocrinol.* 2, 999-1004.

Christmas, R. (2001) Management and Evaluation Considerations for Range Beef Bulls.
International Veterinary Information Service A0508.1101

Cohen, D.J., Ellerman, D.A., Cuasnicú, P.S. (2000) Mammalian sperm-egg fusion: evidence that epididymal protein DE plays a role in mouse gamete fusion. *Biol. Reprod.* 63, 462-468.

Cooper, T.G. (1992) Epididymal proteins and sperm maturation. In: *Spermatogenesis, Fertilisation, Contraception. Molecular, Cellular and Endocrine Events in Male Reproduction.* (ed. Nieschlag, E., Habenicht, U.F.), Springer Verlag Berlin. 285-318.

Dohms, T. (2001) Züchterische Nutzung von Fruchtbarkeitsdaten in der Pferdezucht
Dissertation, Universität Göttingen

Eberspaecher, U., Roosterman, D., Krätzschar, J., Haendler, B., Habenicht, U.F., Becker, A., Quensel, C., Petri, T., Schleuning, W.D., Donner, P. (1995) Mouse androgen-dependent epididymal glycoprotein CRISP-1

- (DE/AEG): isolation, biochemical characterization, and expression in recombinant form. *Mol. Reprod. Dev.* 42, 157-172.
- Fornes, M.W., Burgos, M.H. (1994) Epididymal glycoprotein involved in rat sperm association. *Mol. Reprod. Dev.* 38, 43-47.
- Foster, J.A., Gerton, G.L. (1996) Autoantigen 1 of the guinea pig sperm acrosome is the homologue of mouse Tpx-1 and human TPX1 and is a member of the cysteine-rich secretory protein (CRISP) family. *Mol. Reprod. Dev.* 44, 221-229.
- Gardiner-Garden, M., Frommer, M. (1987) CpG islands in vertebrate genomes. *J. Mol. Biol.* 196, 261-282.
- Giese, A., Jude, R., Kuiper, H., Piumi, F., Schambony, A., Guérin, G., Distl, O., Töpfer-Petersen, E., Leeb, T. (2002a) Molecular characterization of the equine *AEGI* locus. *Gene*, 292, 65-72.
- Giese, A., Jude, R., Kuiper, H., Raudsepp, T., Piumi, F., Schambony, A., Guérin, G., Chowdhary, B.P., Distl, O., Töpfer-Petersen, E., Leeb, T. (2002b) Molecular characterization of the equine *TPX1* and *AEG2* genes encoding members of the cysteine-rich secretory protein (CRISP) family. Zur Veröffentlichung angenommen in *Gene*.
- Godard, S., Schibler, L., Oustry, A., Cribiu, E.P., Guérin, G. (1998) Construction of a horse BAC library and cytogenetical assignment of 20 type I and type II markers. *Mamm. Genome* 9, 633-637.
- Haendler, B., Habenicht, U.F., Schwidetzky, U., Schuttke, I., Schleuning, W.D. (1997) Differential androgen regulation of the murine genes for cysteine-rich secretory proteins (CRISP). *Eur. J. Biochem.* 250, 440-446.
- Haendler, B., Krätzschar, J., Theuring, F., Schleuning, W.D. (1993) Transcripts for the cysteine-rich secretory protein-1 (CRISP-1; DE/AEG) and the novel related CRISP-3 are expressed under androgen control in the mouse salivary gland. *Endocrinology* 133, 192-198.
- Haendler, B., Schuttke, I., Schleuning, W.D. (2001) Androgen receptor signalling: comparative analysis of androgen response elements and implication of heat-shock protein 90 and 14-3-3eta. *Cell Endocrinol.* 173, 63-73.
- Hardy, M.D., Huang, T.T.F., Driscoll, W.J., Tung, K.S.K., Wild, G.C. (1988) Purification and characterization of the primary acrosomal autoantigen of guinea pig epididymal spermatozoa. *Biol. Reprod.* 38, 423-437.

- Hayashi, M., Fujimoto, S., Takano, H., Ushiki, T., Abe, K., Ishikura, H., Yoshida, M.C., Kirchhoff, C., Ishibashi, T., Kasahara, M. (1996) Characterization of a human glycoprotein with a potential role in sperm-egg fusion: cDNA cloning, immunohistochemical localization, and chromosomal assignment of the gene (AEGL1). *Genomics* 32, 367-374.
- Hinton, B.T., Palladino, M.A. (1995) Epididymal epithelium: 1st contribution to the formation of a luminal fluid microenvironment. *Microsc. Res. Tech.* 30, 67-81.
- International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.
- Jacobs, K.-O. (1977) Abfohlergebnisse, die durch Hengste des Celler Landgestütes von 1815 bis 1973 erzielt wurden. Dissertation, Tierärztliche Hochschule Hannover
- Jude, R., Giese, A., Piumi, F., Guérin G., Sieme, H., Töpfer-Petersen, E., Distl, O., Leeb, T. (2002) Molecular characterization of the *CRISPI* gene - a candidate for stallion fertility. *Theriogenology* 58, 417-420.
- Kasahara M., Figueroa F., Klein J. (1987) Random cloning of genes from mouse chromosome 17. *Proc. Natl. Acad. Sci. U.S.A.*, 84, 3325-3328.
- Kashara, M., Passmore, H.C., Klein, J. (1989a) A testis-specific gene *Tpx-1* maps between *Pgk-2* and *Mep-1* on mouse chromosome 17. *Immunogenetics* 29, 61-63.
- Kasahara, M., Gutknecht, J., Brew, K., Spurr, N., Goodfellow, P.N. (1989b) Cloning and mapping of a testis-specific gene with sequence similarity to a sperm-coating glycoprotein gene. *Genomics* 5, 527-534.
- Kasahara, M., Hayashi, M., Yoshida, M.C., Nadeau, J.H., Fujimoto, S., Ishibashi, T. (1995) Mapping of acidic epididymal glycoprotein (Aeg) genes to mouse chromosome 17. *Mamm. Genome* 6, 52-54.
- Kierszenbaum, A.L. (1994) Mammalian spermatogenesis *in vivo* and *in vitro*: A partnership of spermatogenic and somatic cell lineages. *Endocr. Rev.* 15, 116-134.
- Kjeldsen, I., Cowland, J.B., Johnson, A.H., Borregard, N. (1996) Sgp28, a novel matrix glycoprotein in specific granules of human neutrophils with similarity to a human testis-specific gene product and to a rodent sperm-coating protein. *FEBS Lett.* 380, 246-250.

- Klemme, L.M., Roberts, K.P., Hoffman, L.B., Ensrud, K.M., Siiteri, J.E., Hamilton, D.W. (1999) Cloning and characterization of the rat *Crisp-1* gene. *Gene* 240, 279-288.
- Klug, E. (2002) Untersuchung von Stute und Hengst auf Zuchttauglichkeit *Pferdeheilkunde* 18, 225-229.
- Kozak, M. (1996) Interpreting cDNA sequences: some insights from studies on translation. *Mamm. Genome* 7, 563-574.
- Krätzschmar, J., Haendler, B., Eberspaecher, U., Roosterman, D., Donner, P., Schleuning, W.D. (1996) The human cysteine-rich secretory protein (CRISP) family. Primary structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3. *Eur. J. Biochem.* 236, 827-836.
- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.
- Lea, O.A., Petrusz, P., French, F.S. (1978) Purification and localization of acidic epididymal glycoprotein (AEG): A sperm coating protein secreted by the rat epididymis. *Int. J. Androl. Suppl.* 2, 592-607.
- Lindgren, G., Breen, M., Godard, S., Bowling, A., Murray, J., Scavone, M., Skow, L., Sandberg, K., Guérin, G., Binns, M., Ellegren, H. (2001) Mapping of 13 horse genes by fluorescence in-situ hybridization (FISH) and somatic cell hybrid analysis. *Chromosome Res.* 9, 53-59.
- Maeda, T., Sakashita, M., Ohba, Y., Nakanishi, Y. (1998) Molecular cloning of the rat Tpx-1 responsible for the interaction between spermatogenic and Sertoli cells. *Biochem. Biophys. Res. Commun.* 248, 140-146.
- Maeda, T., Nishida, J., Nakanishi, Y. (1999) Expression pattern, subcellular localization and structure-function relationship of rat Tpx-1, a spermatogenic cell adhesion molecule responsible for association with Sertoli cells. *Dev. Growth Differ.* 41, 715-722.
- McCarrey, J.R., Thomas, K. (1987) Human testis-specific PGK gene lacks introns and possesses characteristics of a processed gene. *Nature* 326, 501-505.
- Merkt, H. (1995) Fruchtbarkeit und Fortpflanzung des Pferdes. In: *Handbuch Pferd* (ed. von Thein, P.), BLV Verlagsgesellschaft mbH München. 728-756.
- Merkt, H. und E. Klug (1989) Gesundheitliche und geschlechtliche Mindestanforderungen an Zuchthengste. *Deutsche tierärztliche Wochenschrift* 96, 433-472.

- Mizuki, N., Sarapata, D.E., Garcia Sanz, J.A., Kasahara, M. (1992a) The mouse male germ cell-specific gene *Tpx-1*: molecular structure, mode of expression in spermatogenesis, and sequence similarity to two non-mammalian genes. *Mamm Genome*. 3, 274-80.
- Mizuki, N., Kasahara, M. (1992b) Mouse submandibular glands express an androgen-regulated transcript encoding an acidic epididymal glycoprotein-like molecule. *Mol. Cell. Endocrinol.* 89, 25-32.
- O'Bryan, M.K., Loveland, K.L., Herszfeld, D., McFarlane, J.R., Hearn, M.T., de Kretser, D.M. (1998) Identification of a rat testis-specific gene encoding a potential rat outer dense fibre protein. *Mol. Reprod. Dev.* 50, 313-322.
- Peri, S., Pandey, A. (2001) A reassessment of the translation initiation codon in vertebrates. *Trends Genet.* 17, 685-687.
- Pfisterer, P., König, H., Hess, J., Lipowsky, G., Haendler, B., Schleuning, W.D., Wirth, T. (1996) CRISP-3, a protein with homology to plant defense proteins, is expressed in mouse B cells under the control of Oct2. *Mol. Cell. Biol.* 16, 6160-6168.
- Rankin, T.I., Tsuruta, K.J., Holland, M.K., Griswold, M.D., Orgebin-Crist, M.C. (1992) Isolation, immunolocalization, and sperm association of three proteins of 18, 25 and 29 kilodaltons secreted by the mouse epididymis. *Biol. Reprod.* 46, 747-766.
- Raudsepp, T., Fronicke, L., Scherthan, H., Gustavsson, I., Chowdhary, B.P. (1996) Zoo-FISH delineates conserved chromosomal segments in horse and man. *Chromosome Res.* 4, 218-225.
- Rochwerger, I., Cohen, D.J., Cuasnieú, P.S. (1992) Mammalian sperm egg fusion: the rat egg has complementary sites for a sperm protein that mediates gamete fusion. *Dev. Biol.* 153, 83-90.
- Roberts, K.P., Hoffman, L.B., Ensrud, K.M., Hamilton, D.W. (2001) Expression of *crisp-1* mRNA splice variants in the rat epididymis, and comparative analysis of the rat and mouse *crisp-1* gene regulatory regions. *J. Androl.* 22, 157-163.
- Russel, L.D., Petersen, R.N. (1985) Sertoli cell junctions: Morphological and functional correlates. *Int. Rev. Cytol.* 94, 177-211.
- Schambony, A., Hess, O., Gentzel, M., Töpfer-Petersen, E. (1998a) Expression of CRISP proteins in the male equine genital tract. *J. Reprod. Fertil. Suppl.* 53, 67-72.

- Schambony, A., Gentzel, M., Wolfes, H., Raida, M., Neumann, U., Töpfer-Petersen, E. (1998b) Equine CRISP-3: primary structure and expression in the male genital tract. *Biochim. Biophys. Acta.* 1387, 206-216.
- Schwartz, S., Zhang, Z., Frazer, K.A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R., Miller, W. (2000) PipMaker--a web server for aligning two genomic DNA sequences. *Genome Res.* 10, 577-586.
- Schwidetzky, U., Haendler, B., Schleuning, W.D. (1995) Isolation and characterization of the androgen-dependent mouse cysteine-rich secretory protein-3 (CRISP-3) gene. *Biochem. J.* 309 (Pt 3), 831-836.
- Sieme, H. (1996) Optimierung der Fruchtbarkeit und Besamung. Vortragsveranstaltung der deutschen Gesellschaft für Züchtungskunde e.V. Neumünster, Oktober 1996
- Simon SM, Blobel G. (1993) Mechanisms of translocation of proteins across membranes. *Subcell. Biochem.* 21, 1-15.
- Sivashanmugam, P., Richardson, R.T., Hall, S., Hamil, K.G., French, F.S., O'Rand, M.G. (1999) Cloning and characterization of an androgen-dependent acidic epididymal glycoprotein/CRISP1-like protein from the monkey. *J. Androl.* 20, 384-393.
- Vierula, M.E., Rankin, T.L., Orgebin-Crist, M.C. (1995) Electron microscopic immunolocalization of the 18 and 29 kilodalton secretory proteins in the mouse epididymis: evidence for differential uptake by clear cells. *Microsc. Res. Tech.* 30, 24-36.
- Wang, K., Boysen, C., Shizuya, H., Simon, M.I., Hood, L. (1997) Complete nucleotide sequence of two generations of a bacterial artificial chromosome cloning vector. *Biotechniques*, 23, 992-994.
- Yanagimachi, R. (1994) Mammalian Fertilization. In: *The Physiology of Reproduction.* (ed. Knobil E., Neill J.D.) Raven Press New York. 189-317.
- Yoshino, M., Xiao, H., Amadou, C., Jones, E.P., Lindahl, K.F. (1998) BAC clones and STS markers near the distal breakpoint of the fourth t-inversion, In(17)4d, in the H2-M region on mouse chromosome 17. *Mamm. Genome.* 9, 186-192.

8 Anhang

Verwendete Oligonukleotide:

PCR-Primer

Name	Sequenz	Länge	T _M	Produkt	equines Gen
eC1_1	AAC CCC TGC CTC TAC TAT G	19	58	226	<i>CRISP1</i>
eC1_2	CCA AAA GAA ATT CAG CAG A	19	52	226	<i>CRISP1</i>
eC2_1	CGA GAT GTG GTG AGA ATA TC	20	58	114	<i>CRISP2</i>
eC2_2	CAC TCT TTG GTC CTA CAC C	19	58	114	<i>CRISP2</i>
eC3_1	CTT GAA GAA AAT AGC TGG C	19	54	140	<i>CRISP3</i>
eC3_2	CAA CAA AAG GAT GCA GC	17	50	140	<i>CRISP3</i>

Sequenzprimer

Name	Sequenz	Länge	equines Gen
pCRISP1.1rv	5'-ATC TCC ATC TCA GAG CCT C-3'	19	<i>CRISP1</i>
pCRISP1.2fw	5'-TAA ATG GCA GTG AAT TGT GG-3'	20	<i>CRISP1</i>
pCRISP1.3rv	5'-TGA AAG TAG CCA CAC AGA G-3'	19	<i>CRISP1</i>
pCRISP1.4fw	5'-TCA GAT TAT TCA CTA CAA GTG-3'	21	<i>CRISP1</i>
pCRISP1.5rv	5'-TCA TCA TAG TAG AGG CAG G-3'	19	<i>CRISP1</i>
pCRISP1.6fw	5'-TAC AGA TTC AAT GCA ATC CC-3'	20	<i>CRISP1</i>
pCRISP1.7rv	5'-TAG CCT TCC AAG TCT GAA G-3'	19	<i>CRISP1</i>
pCRISP1.8fw	5'-CTC TAC CCT TAA GTA CTT CC-3'	20	<i>CRISP1</i>
pCRISP1.9rv	5'-TGA CTT TGG ATA AGC TAC CT-3'	20	<i>CRISP1</i>
pCRISP1.10fw	5'-GTT TGG GCC ACT TCT TAC C-3'	19	<i>CRISP1</i>
pCRISP1.11rv	5'-TCA TGA CAG TAG TGA CAA AC-3'	20	<i>CRISP1</i>
pCRISP1.12rv	5'-TTT CTA GTA ACA TTG GTT GAC-3'	21	<i>CRISP1</i>
pCRISP1.13fw	5'-TGC TCC TAG AGA TGG AAG C-3'	19	<i>CRISP1</i>
pCRISP1.14rv	5'-GCT CAT TGA ATC ATA TGT GTC-3'	21	<i>CRISP1</i>
pCRISP1.15fw	5'-CTC TCA TCC GAA ACA ACT G-3'	19	<i>CRISP1</i>
pCRISP1.16rv	5'-AGT TGG CTC TCA ATC AAC G-3'	19	<i>CRISP1</i>
pCRISP1.17fw	5'-TTG GTT TCT CAC TCT CAG C-3'	19	<i>CRISP1</i>
pCRISP1.18rv	5'-TAC TGA AGC CAA AGA CAC C-3'	19	<i>CRISP1</i>
pCRISP1.19fw	5'-CAT ACA TAG GAG GTA GTC TC-3'	20	<i>CRISP1</i>
pCRISP1.20rv	5'-CGC ACT GTC CTA GGG TAC-3'	18	<i>CRISP1</i>
pCRISP1.21fw	5'-AGT GAG GAG TGC ACC AAG-3'	18	<i>CRISP1</i>
pCRISP1.22rv	5'-GTA CAA GGT ATC CAG CTA G-3'	19	<i>CRISP1</i>
pCRISP1.23fw	5'-CTC TCA AGA AGT GGC TTC C-3'	19	<i>CRISP1</i>

Name	Sequenz	Länge	equines Gen
pCRISP1.24fw	5'-GGC ACA GAT GTT AGC TCA G-3'	19	<i>CRISP1</i>
pCRISP1.25rv	5'-GAG TGA TTA GCA GTC AGA TG-3'	20	<i>CRISP1</i>
pCRISP1.26rv	5'-CCA TGT ATC TGC ACT TGT G-3'	19	<i>CRISP1</i>
pCRISP1.27fw	5'-GTG AGA GAG AGC TGT GTG-3'	18	<i>CRISP1</i>
pCRISP1.28rv	5'-CAA ATT TAT CAC TCT GTA CAC-3'	21	<i>CRISP1</i>
pCRISP1.29fw	5'-CCA AGA GCA GAA TTA TTG AG-3'	20	<i>CRISP1</i>
pCRISP1.30rv	5'-CTA TGA GTG GCG TTA TGT G-3'	19	<i>CRISP1</i>
pCRISP1.31fw	5'-TGT ATT GTG ACC TCT TGA G-3'	19	<i>CRISP1</i>
pCRISP1.32fw	5'-TAC CAG GTA AGA CCT CAC-3'	18	<i>CRISP1</i>
pCRISP1.33fw	5'-GTT TGT TCA GTA AGG AAG-3'	19	<i>CRISP1</i>
pCRISP1.34rv	5'-CCT CTC CTA GAA ACT CTC-3'	18	<i>CRISP1</i>
pCRISP1.35rv	5'-GAT TGA TGG TCA AAT ACG GTG-3'	21	<i>CRISP1</i>
pCRISP1.36fw	5'-ACA GAT TTC CCA TCT ATT TCC-3'	21	<i>CRISP1</i>
pCRISP1.37rv	5'-AAC TCT TTA CAG TGA ACT ACC -3'	21	<i>CRISP1</i>
pCRISP1.38rv	5'-CCC ATT CTC TTA CTT TAT CCT C-3'	22	<i>CRISP1</i>
pCRISP1.39fw	5'-AAG ATA ATG GTC ATA CAC ACC-3'	21	<i>CRISP1</i>
pCRISP1.40rv	5'-ATC CAG GAT TGT GCA AGT CC -3'	20	<i>CRISP1</i>
pCRISP1.41fw	5'-AGC TGT TAC CAC TGG AGT C-3'	19	<i>CRISP1</i>
pCRISP1.42fw	5'-AAC CTA TGA GAG GCA GAA G-3'	19	<i>CRISP1</i>
pCRISP1.43fw	5'-AAC AGT GCA TAG GAC ACA G-3'	19	<i>CRISP1</i>
pCRISP1.44fw	5'-ATG AAG CCA ACA TTA GTC TG-3'	20	<i>CRISP1</i>
pCRISP1.45fw	5'-TTG ACACACCAA CTC TTG G-3'	19	<i>CRISP1</i>
pCRISP1.46fw	5'-ATG ACA GAA CAT GGT TAG AG-3'	20	<i>CRISP1</i>
pCRISP1.47rv	5'-GGC TGC CAA ATC TGA GTG-3'	18	<i>CRISP1</i>
pCRISP1.48rv	5'-GAA TTT CTG CAA TGC TAT CC-3'	20	<i>CRISP1</i>
pCRISP1.49rv	5'- GCT GTA ACA CTC ACT ATG C-3'	19	<i>CRISP1</i>
pCRISP1.50fw	5'-GAT GAA GGG CAA GGA AGC-3'	18	<i>CRISP1</i>
pCRISP1.51rv	5'- TAA ATA GAG AGA GGT TGG-3'	18	<i>CRISP1</i>
pCRISP1.52rv	5'-CAA ATT CGG TAA AGT CAC AG-3'	20	<i>CRISP1</i>
pCRISP1.53fw	5'-GCT TGT TGT TAG TAT GTA G-3'	19	<i>CRISP1</i>
pCRISP1.54rv	5'-AGA AAT GTG AAA GGC ATG G-3'	19	<i>CRISP1</i>
pCRISP1.55fw	5'-TCA CAT GCA TTC AGA ACA C-3'	19	<i>CRISP1</i>
pCRISP1.56rv	5'-CTC TGG ACA GAT TTG TTC C-3'	19	<i>CRISP1</i>
pCRISP1.57rv	5'-TTT GA TGA GGA CTT CCA TG-3'	19	<i>CRISP1</i>
pCRISP1.58fw	5'-GAG CAT CCT GAT TTC AGA C-3'	19	<i>CRISP1</i>
pCRISP1.59rv	5'-AGG TAT GGG CAT AAA ATG G-3'	19	<i>CRISP1</i>
pCRISP1.60fw	5'-TTG ATA GCA TTG TTT GAT GC-3'	20	<i>CRISP1</i>
pCRISP1.61rv	5'-ATG CTA TTG AAG TTA AGC TG-3'	20	<i>CRISP1</i>
pCRISP1.62fw	5'-GTT TCC TAT ATG TCT TAT AGC-3'	21	<i>CRISP1</i>
pCRISP1.63fw	5'-TCA GCC ATT TGT CCA AGG-3'	18	<i>CRISP1</i>
pCRISP1.64fw	5'-TTC ATG GCC TTA TGC TTC C-3'	19	<i>CRISP1</i>
pCRISP1.65rv	5'-CAA GCC ATG AAA AGA TAT GG-3'	20	<i>CRISP1</i>
pCRISP1.66fw	5'-CCA CAT CCT CAC CAA CTG-3'	18	<i>CRISP1</i>
pCRISP1.67rv	5'-CAG CAA CAA TCA GAA AAC AG-3'	20	<i>CRISP1</i>
pCRISP1.68fw	5'-ATA TTC ATG TCG ATT CTT CC-3'	20	<i>CRISP1</i>
pCRISP1.69rv	5'-GTG AGA AGT AAC TTG GTA G-3'	19	<i>CRISP1</i>

Name	Sequenz	Länge	equines Gen
pCRISP1.70fw	5'-AGT TTC CCT CTT CAG ATG C-3'	19	<i>CRISP1</i>
pCRISP1.71rv	5'-CTG TTT GTT GGT GTG AAC C-3'	19	<i>CRISP1</i>
pCRISP1.72rv	5'-AAC TTA ACC ACG ACA TCA C-3'	19	<i>CRISP1</i>
pCRISP1.73rv	5'-TTA GCA TAA TAC CCT CAA GG-3'	20	<i>CRISP1</i>
pCRISP1.74fw	5'-TTC TTC ACT TCT TAA CTT TGC-3'	21	<i>CRISP1</i>
pCRISP1.75rv	5'-CTT GCG ATT TGA AAA ACA TGG-3'	21	<i>CRISP1</i>
pCRISP1.76rv	5'- TGT ATA ATT CCA TTC ATA TGA C -3'	22	<i>CRISP1</i>
pCRISP2.1rev	5'-ATG GAT TCT GCT GAG AAA TCC-3'	21	<i>CRISP2</i>
pCRISP2.2fwd	5'-GTC TGA TAT AAG TAT ATC TAC-3'	21	<i>CRISP2</i>
pCRISP2.3fwd	5'-CTT AGC ATC CCA GAT CTG TG-3'	20	<i>CRISP2</i>
pCRISP2.4fwd	5'-GGG AGC TAG TTG GTA ATG TC-3'	20	<i>CRISP2</i>
pCRISP2.5rev	5'-CAG ACA AAC CTC CAT CTG AG-3'	20	<i>CRISP2</i>
pCRISP2.6fwd	5'-TCT CAT GGT TAT TAA GAT TAC-3'	22	<i>CRISP2</i>
pCRISP2.7rev	5'-CAA TTT AAC TAG ATA GAG AC-3'	20	<i>CRISP2</i>
pCRISP2.8fwd	5'-ACA TTC TCA AAG CGA CTG TG-3'	20	<i>CRISP2</i>
pCRISP2.9rev	5'-CAG CTC CCT ACA GTG TTT CC-3'	20	<i>CRISP2</i>
pCRISP2.10fwd	5'-CAT ATC AAT TAT TAC TCT AAG TG-3'	23	<i>CRISP2</i>
pCRISP2.11rev	5'-TAC TGG GCC TGG CAA AGT G-3'	19	<i>CRISP2</i>
pCRISP2.12fwd	5'-TGA TGG CAT TTT ACC CAC AG-3'	20	<i>CRISP2</i>
pCRISP2.13fwd	5'-GTT CTG ATT TGC TGA TTA AGG-3'	21	<i>CRISP2</i>
pCRISP2.14rev	5'-AGT ACA AAA GTC TGG AAT CAG-3'	21	<i>CRISP2</i>
pCRISP2.15rev	5'-GTG GAA TTC ATT TGT GTA G-3'	19	<i>CRISP2</i>
pCRISP2.16fwd	5'-TCA TCC ATG TTG TCG CAT G-3'	19	<i>CRISP2</i>
pCRISP2.17fwd	5'-CAG CTT TAT AGC TGA TAA GG-3'	20	<i>CRISP2</i>
pCRISP2.18rev	5'-GTT TTC AAA GCA GGA GTA GG-3'	20	<i>CRISP2</i>
pCRISP2.19rev	5'-CAG GCT GAA TAA CTA AAT TGG-3'	21	<i>CRISP2</i>
pCRISP2.20rev	5'-CTA TTG TAT AAA AGT TGA AGG-3'	21	<i>CRISP2</i>
pCRISP2.21fwd	5'-TTG TGA AAC TGT ATC TGC TG-3'	20	<i>CRISP2</i>
pCRISP2.22rev	5'-TTG GAT ATT GGT GCC AAC C-3'	19	<i>CRISP2</i>
pCRISP2.23rev	5'-TCA ACC AAT TTG TGT ATG AC-3'	20	<i>CRISP2</i>
pCRISP2.24fwd	5'-CAA CTT TTC GGA CAG GAT C-3'	19	<i>CRISP2</i>
pCRISP2.25rev	5'-TGA CTG TCT CTT ACT TCT TC-3'	20	<i>CRISP2</i>
pCRISP2.26fwd	5'-CTT GCT GAA AGA GGA AAC TC-3'	20	<i>CRISP2</i>
pCRISP2.27fwd	5'-TTC GGG AAA ATA ATT CAT CAG-3'	21	<i>CRISP2</i>
pCRISP2.28rev	5'-CTG GAC ATG AGT ATA TTC TC-3'	20	<i>CRISP2</i>
pCRISP2.29fwd	5'-AGC GCT ATA GTT GTA AGA G-3'	19	<i>CRISP2</i>
pCRISP2.30fwd	5'-TGA TTT TAT AGA GCT CTA GC-3'	20	<i>CRISP2</i>
pCRISP2.31rev	5'-GTG CAT ATC CTC ATC AAT GC-3'	20	<i>CRISP2</i>
pCRISP2.32fwd	5'-TTT CAC AGA TGA AAG CTT GC-3'	20	<i>CRISP2</i>
pCRISP2.33fwd	5'-AAT GCC TCA TTG CAG ATG C-3'	19	<i>CRISP2</i>
pCRISP3.1fwd	5'-CTC AGA TGT GAC AAG ACA G-3'	19	<i>CRISP3</i>
pCRISP3.2rev	5'-TTC CCT GGG CAA TCA ACC-3'	18	<i>CRISP3</i>
pCRISP3.3fwd	5'-GGA CAG CAT GTA GAG TCT G-3'	19	<i>CRISP3</i>
pCRISP3.4rev	5'-CTG CAG AAA ACT CAA GAT GG-3'	20	<i>CRISP3</i>
pCRISP3.5rev	5'-GGG AAG TCT TCA ACC TCT C-3'	19	<i>CRISP3</i>
pCRISP3.6fwd	5'-AAT CTA ATA TCC ATG AAT GTG-3'	21	<i>CRISP3</i>

Name	Sequenz	Länge	equines Gen
pCRISP3.7rev	5'-GAA GCA CGT GTA CTT AGC-3'	18	<i>CRISP3</i>
pCRISP3.8rev	5'-CAG TCA GTC ATT AGC TGT G-3'	19	<i>CRISP3</i>
pCRISP3.9fwd	5'-TGT GCT TAG CAA ACA CCA G-3'	19	<i>CRISP3</i>
pCRISP3.10fwd	5'-GCT TTG GTC AGG ACA TCC-3'	18	<i>CRISP3</i>
pCRISP3.11rev	5'-TTA TCT TCA TGT GGT AAG AC-3'	20	<i>CRISP3</i>
pCRISP3.12fwd	5'-TGA AGA CAC TGG TTG CAG-3'	18	<i>CRISP3</i>
pCRISP3.13rev	5'-TCT ACC TAC TAC GTA GCA G-3'	19	<i>CRISP3</i>
pCRISP3.14rev	5'-CAC CAT TCA TTC TTA GTT CC-3'	20	<i>CRISP3</i>
pCRISP3.15rev	5'-TGT AGG AAT TCA GCT TTC TG-3'	20	<i>CRISP3</i>
pCRISP3.16fwd	5'-TTA AGC GGC TGT GTA CTG-3'	18	<i>CRISP3</i>
pCRISP3.17rev	5'-TTC TGC ATA CTT TAG CAG AG-3'	20	<i>CRISP3</i>
pCRISP3.18rev	5'-GGA CAC TTA GTC CCT GAG-3'	18	<i>CRISP3</i>
pCRISP3.19fwd	5'-GTT CAC ATT TCT CAG ATA AGC-3'	21	<i>CRISP3</i>
pCRISP3.20rev	5'-TCT CAC TAC CAC CCT ATG G-3'	19	<i>CRISP3</i>
pCRISP3.21fwd	5'-ATG AAA TTG TCA GAT GAT TGC-3'	21	<i>CRISP3</i>
pCRISP3.22fwd	5'-TAT GTA TGT ATC AAC AGT TGG-3'	21	<i>CRISP3</i>
pCRISP3.23fwd	5'-AAT CTT CTG TAA CTA CTT CTG-3'	21	<i>CRISP3</i>
pCRISP3.24rev	5'-TAG ACT AGG GAG AAT TTG C-3'	19	<i>CRISP3</i>
pCRISP3.25fwd	5'-GCT AAC AGA AGA TTA AGA TC-3'	20	<i>CRISP3</i>
pCRISP3.26rev	5'-CAT GTT GTA ACA TGG CAT C-3'	19	<i>CRISP3</i>

RACE-Primer

Name	Sequenz	Länge	equines Gen
C1_5'RACE_in	5'-GAG ATA GGA TAA GAT GTC AGA TGC -3'	24	<i>CRISP1</i> Exon 5
C1_5'RACE_out	5'-TCA TCA TCT GTT GAT ATC CAT TGC-3'	24	<i>CRISP1</i> Exon 5
C1_3'RACE_in	5'-CAA GGC AGT TGA GAG GAG AAT TAC-3'	24	<i>CRISP1</i> Exon 4
C1_3'RACE_out	5'-TGC ACA AAA CGC CAG AAG GTT GTC-3'	24	<i>CRISP1</i> Exon 4
C2.5'RACE.in	5'-TGG GTT TGA GTG GTT AAC AAA GCA-3'	24	<i>CRISP2</i> Exon 5
C2.5'RACE.out	5'-TGG AGA GAC TGA TTT CCT TAG TTC-3'	24	<i>CRISP2</i> Exon 5
C2.3'RACE.in	5'-TGT GCT ACT TCC ATC TTT ACC CAC-3'	24	<i>CRISP2</i> Exon 4
C2.3'RACE.out	5'-TGG CTT TCC TCC CAG TGG TGA TGT-3'	24	<i>CRISP2</i> Exon 4
C3.5'RACE.in	5'-CGA CAT AAA GAG ATT CTC ACC ACA-3'	24	<i>CRISP3</i> Exon 5
C3.5'RACE.out	5'-TGT CCA ACT ACT GCA TTG GGA GTC-3'	24	<i>CRISP3</i> Exon 5
C3.3'RACE.in	5'-CCA ACA CAG TAA AGC AGA AGA TCG-3'	24	<i>CRISP3</i> Exon 4
C3.3'RACE.out	5'-GGA CAG CAA GAC AGC AAC AAA TGC-3'	24	<i>CRISP3</i> Exon 4

Danksagung

Mein besonderer Dank gilt Herrn Prof. Dr. Tosso Leeb für die Überlassung des Themas, die Bereitstellung eines Laborplatzes und die Finanzierung dieser Arbeit. Ich danke aber auch für die hervorragende Anleitung und Betreuung, sowie das außergewöhnlich gute Arbeitsklima in seiner Arbeitsgruppe. Er hat sich trotz seiner neuen Pflichten als Professor immer Zeit für die Probleme im Labor genommen und nie den Kontakt zur Laborarbeit verloren.

Herrn Prof. Dr. Gossler möchte ich für die Übernahme des Korreferats dieser Arbeit danken.

Herrn Prof. Dr. Ottmar Distl danke ich für die Aufnahme in seinem Institut und für die Einblicke in die tierzüchterischen Aspekte der Tiermedizin, sowie die unterhaltsamen Exkursionen.

Mein herzlicher Dank gilt allen Mitarbeiterinnen und Mitarbeitern des Institutes für Tierzucht und Vererbungsforschung, für das ausgezeichnete Arbeitsklima und die freundschaftliche Zusammenarbeit. Besonders danke ich Frau Heike Klippert, Herrn Stefan Neander und Herrn Jörn Wrede für den gemeinsamen „Kampf“ gegen die Tücken der Sequenzierung und der Datenverarbeitung. Frau Dr. Birgit Alsleben und Frau Kathrin Löhring möchte ich für die entspannenden Spaziergänge in den Mittagspausen und die gegenseitigen Einladungen zum Essen danken. Vielen Dank auch an Herrn Andreas Spötter für die angenehme Atmosphäre im Schreibraum und die fruchtbaren Diskussionen molekularbiologischer und grafischer Probleme.

Frau Dr. Heidi Kuiper und Herrn Dr. Bhanu Chowdhary danke ich für die erfolgreiche Zusammenarbeit und die Durchführung der *in situ* Hybridisierungen.

Des weiteren danke ich Herrn Dr. Gérard Guérin für die Bereitstellung der equinen BAC-Klone.

Meiner Familie danke ich besonders für die Unterstützung über all die Jahre und den Halt den ich dadurch hatte.

Abschließend geht ein großer Dank an Frau Dr. Claudia Goldmann für die Durchsicht dieser Arbeit, und die moralische Unterstützung. Ihr Verständnis und ihre Gesellschaft waren mir eine große Hilfe.

Lebenslauf

Name	Alexander Giese
Adresse	Menschingstrasse 3 30173 Hannover
Geburtsdatum	22.07.1971
Geburtsort	Kassel
1978-1982	Dorothea-Viehmann-Grundschule, Kassel
1982-1988	Gymnasium Heinrich-Schütz-Schule, Kassel
1988-1991	Oberstufengymnasium Jacob-Grimm-Schule, Kassel
07. Juni 1991	Allgemeine Hochschulreife
1991-1999	Biologie-Studium an der Georg-August-Universität Göttingen
15. Mai 1998	Diplomprüfungen
Mai 1998-Mai 1999	Diplomarbeit im Tierärztlichen Institut der Georg-August-Universität-Göttingen unter Anleitung von Prof. Dr. B. Brenig und PD Dr. T. Leeb: "Klonierung und strukturelle Analyse des Promotors des porcinen Ryanodin-Rezeptor 3 Gens (<i>RYR3</i>)"
Mai 1999-Oktober 1999	Wissenschaftlicher Mitarbeiter am Tierärztlichen Institut der Georg-August-Universität-Göttingen
Oktober 1999-September 2002	Doktorarbeit am Institut für Tierzucht und Vererbungsforshung der Tierärztlichen Hochschule Hannover unter Anleitung von Prof. Dr. Tosso Leeb: "Molekulargenetische Analyse der equinen <i>CRISP</i> Gene und ihrer Rolle in der Fruchtbarkeit des Hengstes"

Vollständige Liste der wissenschaftlichen Publikationen

Originalartikel:

Leeb, T., Giese, A., Al Bayati, H., Rettenberger, G., Brenig, B. (1998) Assignment of the porcine ryanodine receptor 3 gene (RYR3) to chromosome 7q22-->q23. *Cytogenet. Cell Genet.* 83, 244-245.

Leeb, T., Giese, A., Pfeiffer, I., Brenig, B. (1999) Two highly polymorphic microsatellites within the porcine ryanodine receptor 3 gene (RYR3). *Anim. Genet.* 30, 321-322.

Giese, A., Deppe, A., Brenig, B., Leeb, T. (2000) Genomic structure of the 5' end of the porcine ryanodine receptor 3 gene (RYR3). *DNA Sequence* 11, 175-179.

Giese, A., Jude, R., Kuiper, H., Piumi, F., Schambony, A., Guérin, G., Distl, O., Töpfer-Petersen, E., Leeb, T. (2002) Molecular characterization of the equine *AEGI* locus. *Gene* 292, 65-72.

Giese, A., Jude, R., Kuiper, H., Raudsepp, T., Piumi, F., Schambony, A., Guérin, G., Chowdhary, B.P., Distl, O., Töpfer-Petersen, E., Leeb, T. Molecular characterization of the equine *TPXI* and *AEG2* genes encoding members of the cystein-rich secretory protein (CRISP) family. *Gene* 299, 101-109.

Poster und Kongressbeiträge:

Jude, R., Giese, A., Piumi, F., Guérin, G., Sieme, H., Töpfer-Petersen, E., Distl, O., Leeb, T. Poster auf dem 8. International Symposium on Equine Reproduction in Fort Collins, Co, USA, 21-26.07.2002 und *Theriogenology* 58, 417-420.

Giese, A., Jude, R., Kuiper, H., Piumi, F., Schambony, A., Sieme, H., Guérin, G., Distl, O., Töpfer-Petersen, E., Leeb, T. Characterization of the equine *AEGI* gene and its role in stallion fertility. Poster auf dem 28. International Society of Animal Genetics (ISAG) Meeting in Göttingen, Deutschland, 11.-15.08.2002.