



Contents lists available at ScienceDirect

Meta Gene



Splicing variants of porcine synphilin-1[☆]



Knud Larsen^{*}, Lone Bruhn Madsen, Leila Farajzadeh, Christian Bendixen

Department of Molecular Biology and Genetics, Aarhus University, Blichers Allé 20, P.O. Box 50, DK-8830 Tjele, Denmark

ARTICLE INFO

Article history:

Received 19 January 2015
 Revised 29 April 2015
 Accepted 30 April 2015
 Available online 2 June 2015

Keywords:

α -Synuclein
 Lewy body
 Parkinson's disease
 Pig
 Synphilin-1

ABSTRACT

Parkinson's disease (PD), idiopathic and familial, is characterized by degradation of dopaminergic neurons and the presence of Lewy bodies (LB) in the substantia nigra. LBs contain aggregated proteins of which α -synuclein is the major component. The protein synphilin-1 interacts and colocalizes with α -synuclein in LBs. The aim of this study was to isolate and characterize porcine synphilin-1 and isoforms hereof with the future perspective to use the pig as a model for Parkinson's disease. The porcine *SNCAIP* cDNA was cloned by reverse transcriptase PCR. The spatial expression of *SNCAIP* mRNA was investigated by RNAseq. The presented work reports the molecular cloning and characterization of the porcine (*Sus scrofa*) synphilin-1 cDNA (*SNCAIP*) and three splice variants hereof. The porcine *SNCAIP* cDNA codes for a protein (synphilin-1) of 919 amino acids which shows a high similarity to human (90%) and to mouse (84%) synphilin-1. Three shorter transcript variants of the synphilin-1 gene were identified, all lacking one or more exons. *SNCAIP* transcripts were detected in most examined organs and tissues and the highest expression was found in brain tissues and lung. Conserved splicing variants and a novel splice form of synphilin-1 were found in this study. All synphilin-1 isoforms encoded by the identified transcript variants lack functional domains important for protein degradation.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Abbreviations: ORF, open reading frame; LB, Lewy body; *SNCAIP*, α -synuclein interacting protein

[☆] The sequence of the porcine *SNCAIP* cDNA, encoding the synphilin-1 protein and the three splice variants of synphilin-1, synphilin-1tv1, synphilin-1tv2, and synphilin-1tv3 have been submitted to GenBank under the accession numbers, GenBank ID: NM_001105053, GenBank ID: NM_001105054, GenBank ID: EF154192, and GenBank ID: NM_001098601, respectively. The genomic sequences representing the porcine *SNCAIP* were submitted under the accession numbers GenBank ID: JQ916898 and GenBank ID: JQ941708, respectively.

^{*} Corresponding author at: Department of Molecular Biology and Genetics, Aarhus University, Blichers Alle 20, P.O. Box 50, DK-8830 Tjele, Denmark. Tel.: +45 87157701; fax: +45 87154994.

E-mail addresses: Knud.Larsen@mbg.au.dk (K. Larsen), LoneB.Madsen@mbg.au.dk (L.B. Madsen), Leila.Farajzadeh@mbg.au.dk (L. Farajzadeh), Christian.Bendixen@mbg.au.dk (C. Bendixen)

<http://dx.doi.org/10.1016/j.mgene.2015.04.005>

2214-5400/© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Parkinson's disease (PD) is a common progressive neurodegenerative movement disorder. Clinically, PD is characterized by resting tremor, rigidity and bradykinesia along with characteristics as postural instability and marked response to dopaminergic therapy (Hardy et al., 2006). In the past 15 years, genetic studies have revealed that different gene mutations and copy number variants cause familial forms of PD (Crosiers et al., 2011). The pathological hallmark of idiopathic and most monogenic (familial) forms of PD comprises loss of nigrostriatal dopaminergic neurons in the substantia nigra pars compacta and α -synuclein containing Lewy bodies (LB) in the surviving neurons (Dickson et al., 2009). The mechanisms of α -synuclein aggregation in LB and the influence on neurodegeneration are still unresolved.

Synphilin-1 (α -synuclein-interacting protein) is a cytoplasmic protein which was identified by its interaction with α -synuclein in a yeast two-hybrid analysis (Engelender et al., 1999). Also, in vivo interaction between synphilin-1 and α -synuclein has been demonstrated (Engelender et al., 1999; Kawamata et al., 2001; Lee et al., 2004; Marx et al., 2003). Synphilin-1 localizes close to synaptic vesicles (Wheeler et al., 2002) and as the protein constitutes an intrinsic component of LB and it is likely involved in the pathogenesis of PD (Wakabayashi et al., 2000). The coding sequence of the human *SNCAIP* gene spans 10 exons resulting in a 919-amino acid protein. The human *SNCAIP* gene has been mapped to chromosome 5q23.1–23.3 (Engelender et al., 2000).

The human synphilin-1 protein contains several functional domains such as four ankyrin-like repeats (tandemly repeated modules of about 33 amino acids), a coiled-coil domain and an ATP/GTP-binding motif (Fig. 1) (Engelender et al., 1999). All these domains, being highly conserved in the human, mouse (O'Farrell et al., 2001) and porcine (this report) sequence, have been identified in proteins involved in or mediating protein-protein interactions. Although the physiological function of synphilin-1 is still unknown, the presence of the protein domains delivers circumstantial evidence for a critical function. Synphilin-1 interacts with proteins such as α -synuclein (Neystat et al., 2002), parkin (Chung et al., 2001), dorfins, and SIAH-1 and SIAH-2 (Liani et al., 2004; Nagano et al., 2003). These proteins are all E3 ubiquitin ligases and bind in the central part of the synphilin-1 polypeptide sequence. To date only one mutation in the *SNCAIP* gene has been described; an arginine to cysteine substitution in position 621 of the encoded amino acid sequence (Marx et al., 2003). The R621C mutation is located in the fifth ankyrin-like repeat. *SNCAIP* is constitutively expressed in several tissues and at a particularly high level in brain, heart and placenta (Engelender et al., 2000). The expression profiles for human and mouse synphilin-1 are very similar (Engelender et al., 2000; O'Farrell et al., 2001). The synphilin-1 protein is, like α -synuclein, predominantly expressed in neurons and is enriched in presynaptic nerve terminals during development (Ribeiro et al., 2002).

Synphilin-1 has an important function in protein degradation mediated by autophagic clearance of aggresome-like inclusions (Wong et al., 2012). Autophagy is a process that facilitates degradation of intracellular components through the sequestration of portions of the cytosol inside double membrane vesicles that then fuse with lysosomes. Small synphilin-1 aggregates and large aggresomes are differentially targeted by constitutive and inducible autophagy (Wong et al., 2012). Specific regions in synphilin-1 are necessary for its own basal and inducible autophagy and for degradation of other pro-aggregating proteins (Wong et al., 2012).

Synphilin-1 displays trophic and neuroprotective effects (Li et al., 2010). Overexpression of synphilin-1 in mouse neuroblastoma cells leads to promoted neurite outgrowth and to protection against Rotenone-induced toxicity (Li et al., 2010). This suggests that synphilin-1 may have a protective role in PD pathogenesis.

Synphilin-1 is also involved in the control of energy balance. Overexpression of human synphilin-1 in transgenic mice resulted in hyperphagia and obesity (Li et al., 2012, 2014). These mice also displayed hyperinsulinemia, hyperleptinemia and impaired glucose tolerance (Li et al., 2012). Similarly, synphilin-1 overexpression in *Drosophila* positively regulates energy homeostasis (Liu et al., 2012). Synphilin-1 expression in fruitfly neurons induces obesity-like phenotypes including body weight, body fat and food intake (Liu et al., 2012).

To date, a minimum of eight different isoforms of human synphilin-1 have been reported (Humbert et al., 2007). They all arise from the *SNCAIP* gene by alternative splicing. Similarly, four different isoforms of mouse synphilin-1 can be retrieved from sequence databases. Alternatively spliced transcript variants encoding different isoforms of human synphilin-1 have been described, but their full-length nature remains to be determined. A synphilin-1 isoform, named synphilin-1A, lacking exons three and four and containing an insertion in exon nine has been described (Eyal et al., 2006). The synphilin-1A isoform has enhanced aggregatory properties and causes neurotoxicity (Eyal et al., 2006).

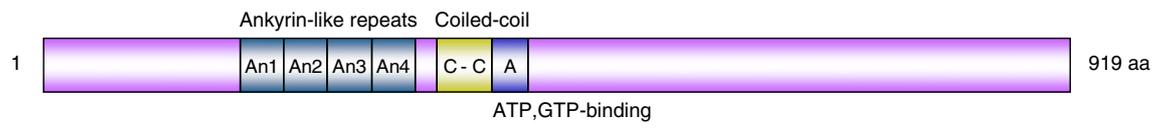


Fig. 1. Schematic presentation of the Synphilin-1 protein. The characteristic domains of synphilin-1 are indicated by their respective names.

Here, we describe and characterize the *SNCAIP* cDNA variants encoding porcine synphilin-1. Furthermore, we mapped the *SNCAIP* gene and determined its spatial expression pattern at the RNA level.

Materials and methods

Animals and tissue collection

Pigs were housed and used in compliance with European Community animal care guidelines. Beforehand, the experimental procedures were submitted to the National Ethical Committee in Denmark. The pig cerebellum used for RT-PCR cloning of *SNCAIP* and various other organs and tissues employed in the expression analysis were obtained from Danish Landrace adult pigs (2–3 years old).

Cloning of the porcine SNCAIP cDNA

In an attempt to isolate the sequence encoding the porcine synphilin-1, we screened the porcine EST data bank at the Institute of Molecular Biology and Genetics, Aarhus University, in silico, with the human *SNCAIP* cDNA sequence. Sequence similarity search was carried out with gapped alignment using NCBI Blastall with options *blastn*, minimum value 10^{-8} . The porcine cDNAs thus identified were used to derive oligonucleotide primers for cloning and as queries for further searches in genomic sequence databases.

The various tissues, frontal cortex and cerebellum, were obtained from an adult pig and subsequently dissected and pulverized in liquid nitrogen. Total RNA was isolated by the RNeasy method (Qiagen) and RNA integrity was verified by ethidium bromide staining of 1% agarose gels. Synthesis of cDNA was conducted with 5 µg of total RNA isolated from pig cerebellum using SuperScript II RNase H⁻ reverse transcriptase (Invitrogen) and oligo (dT)_{12–18} primers according to the manufacturer's recommendations. The RT-PCR reaction mix contained: 2.0 µL cDNA, 1.5 mM MgCl₂, 0.2 mM dNTP, 0.5 µM of each primer *SNCAIP*-F: 5'-ATGG AAGCCCTGAATACCTTGATTG-3' and *SNCAIP*-R: 5'-TTATGCTGCCTATCTTCCTTTGCTAGCGG-3' and 1 U Phusion DNA polymerase (Finnzymes), in a total volume of 25 µL. The PCR profile was as follows: 95 °C for 2 min, 10 touchdown cycles of 95 °C for 20 s, 60 °C for 30 s, 72 °C for 45 s, followed by 25 cycles of 95 °C for 20 s, 55 °C for 30 s, 72 °C for 45 s and finally an elongation at 72 °C for 5 min. Two PCR products of approx. 2800 bp and 1600 bp were visualized and isolated from an ethidium bromide stained 1% agarose gel. The recovered cDNA amplicons were cloned directly into the pCR TOPO 2.1 vector (Invitrogen) and sequenced in both directions. DNA sequencing was performed, as previously described, employing the dideoxy chain termination method using BigDye terminator cycle sequencing kit with AmpliTaq DNA polymerase FS (PE Applied Biosystems (Bjerre et al., 2006)). The sequencing analysis was performed on an automated DNA sequencer (ABI PRISM™ Genetic Analyzer Model 3730xl, PE Applied Biosystems).

Expression analysis

RNA isolation

Total RNA was extracted from 10 different tissues from two boars using standard procedures and RNA integrity was assessed by manual inspection of each RNA sample on a 2% agarose gel. The tissues employed were cerebellum, frontal cortex, occipital cortex, hypothalamus, lung, spleen, liver, heart, kidney, and musculus longissimus dorsi.

Library preparation, sequencing and estimation of FPKM values

The FPKM values for synphilin-1 presented in this study were drawn from a large dataset generated using next generation sequencing. Briefly, mRNA sample preparation was accomplished using the mRNA-seq sample prep kit from Illumina and according to the manufacturer's protocol. Shortly, 10 µg of each total RNA sample was employed in polyA mRNA selection using magnetic beads, followed by thermal fragmentation. Subsequently, the fragmented mRNA was reverse transcribed using reverse transcriptase, SuperScript II, and random primers. The cDNA was size selected on a low melting 2% agarose gel and fragments

corresponding to sizes of 200 nucleotides were excised from the gel and DNA was recovered employing QIAquick gel extraction kit (Qiagen).

In order to enrich the libraries 15 cycles of PCR were employed followed by purification by the QIAquick PCR purification kit (QIAGEN). The purified libraries were diluted to a concentration of 10 nM, and each library was sequenced as a multiplex of 10 samples per lane using 55 bp sequencing at a concentration of 20 pM on a Genome Analyzer (Illumina). Subsequently, all fragments were mapped to build 10.2 including mitochondrial DNA of the *Sus scrofa* genome applying TopHat version 1.3.3 (Trapnell et al., 2009) and the aligned reads were processed and assembled into transcripts by Cufflinks version 0.8.0 (Trapnell et al., 2010). Furthermore, Cufflinks estimates the relative abundance of each transcript and reports it in fragments per kilobase of exon per million fragments mapped (FPKM), and hence the FPKM values for synphilin-1 and synphilin-1tv1 for each animal for all tissues were estimated.

Mapping of porcine SNCAIP

Chromosomal localization of the *SNCAIP* gene was determined using a porcine-rodent somatic cell hybrid panel (Yerle et al., 1996). A PCR primer pair, SNCAIP-EX9F and SNCAIP-EX9R, (Table S1) was designed to amplify a 261 bp genomic fragment spanning PART of exon 9. The PCR was performed in a total volume of 25 μ l containing 15 ng DNA, 1 X PCR buffer, 2.5 mM of each dNTP, 5 pmol of each primer and 1.0 U of DyNAzyme polymerase (Finnzymes) using the following conditions: 95 °C for 3 min; 10 cycles of 95 °C for 30 s, 62 °C for 20 s and 72 °C for 1 min with touchdown of -0.5 °C/cycle; 25 cycles of 95 °C for 30 s, 62 °C for 20 s and 72 °C for 1 min and finally 72 °C for 5 min.

Results and discussion

Characterization of porcine SNCAIP cDNA

Using a combination of bioinformatic tools and RT-PCR we have obtained a cDNA representing the entire *SNCAIP* open reading frame (ORF). The identity of the porcine *SNCAIP* was established by comparison of the deduced protein sequence with human and other isolated synphilin-1 sequences. The cloned porcine *SNCAIP* cDNA (Fig. S1) consisted of 2757 bp in length with the translation start codon at nucleotide 1 and the TAA stop codon at nucleotide 2754. The ORF has a G + C content of 52%.

The deduced amino acid sequence of the cloned synphilin-1 (GenBank ID: EF154190) is shown in Fig. S1. The ORF of the cDNA clone encodes a 919 amino acid polypeptide with a predicted molecular mass of 100 kDa and a pI of 6.24. The porcine synphilin-1 sequence contains multiple ankyrin-like repeats (amino acids 349–489), a coiled-coil domain (amino acids 515–546) and an ATP, GTP-binding domain. The size of the porcine synphilin-1 protein is similar to both human and the mouse proteins. Pig synphilin-1 is one amino acid shorter than the human protein, and three amino acids longer than its mouse counterpart. The absence of one amino acid in exon 3 (amino acid pos. 103 in the human sequence) and one amino acid in exon 9 (amino acid pos. 830 in the human sequence), and also the presence of one additional amino acid in exon 9 (amino acid pos. 901 in the porcine sequence) compared to the human sequence was confirmed by amplification and sequencing of genomic *SNCAIP* sequences covering the relevant exons (3 and 9). Amino acid sequence similarity between porcine synphilin-1 and its human and mouse counterparts was determined by the Clustal method (Fig. 2). The overall identity at the amino acid level was 90% between porcine and human synphilin-1 and 84% between the porcine and the mouse sequences. Human and mouse synphilin-1 were 86% identical. The amino acid identities between the aligned synphilin-1 sequences are even higher (93%) within the central part of the protein covering the sequences for the ankyrin-like1 domain to the coiled-coil domain. Also, the amino acid residue Arg621, a R621C SNP in rare cases of familial PD, is conserved in the pig, mouse and human synphilin-1 polypeptide sequences (Marx et al., 2003).

Splice variants of porcine SNCAIP

The encoded porcine *SNCAIP* gene is alternatively spliced into two forms: a full-length *SNCAIP* mRNA that retains all ten exons, and a short form mRNA, encoding synphilin-1 transcript variant 1 that lacks exons 2, 3, 4 and 5 (Fig. 3A). The deleted interval region of 1239 bp in pig synphilin-1 tv1 matched exactly the exon 2–exon

	ex1	▽ ex2	▽ ex3							
SsSYNP1	MEAPEYLDLDEIDFSDDISYVTSFKTIPELCRRCDQNE	DRPVS	SSSSWNC	GVSTLITNAQKPTGIADVSKFRPVKRV	80					
HsSYNP1	MEAPEYLDLDEIDFSDDISYVTSFKTIPELCRRCDQNE	DRPVS	SSSSWNC	GISTLITNTQKPTGIADVSKFRPVKRV	80					
MmSYNP1	MEAPEYLDLDEIDFSDDISYVTSFKTIPELCRRCDQNE	DRPVS	SSSSWNC	GVSTLITNPQKPTGIADVSKFRPVKRV	80					
	*****	*****	*****	*****						
SsSYNP1	PLKHQPETLENNESDDQKNQKA-EYQKGGSDPAPQPEELS	PRDGE	SPQSKSTEPS	ALLGELEHYDLDMDEILDVPYIK	159					
HsSYNP1	PLKHQPETLENNESDDQKNQKVEYQKGGESDLGPQPEL	GP	DVGGVPPGKSSE	PTSLGELEHYDLDMDEILDVPYIK	160					
MmSYNP1	PLKHQPETLENNENEDQKNNTV-EYQKGGETDQGPQPEELS	PEDG	VGGVLPKGSE	PSQALGELEHYDLDMDEILDVPYIK	159					
	*****	*****	*****	*****						
SsSYNP1	SSQQLASF	TKVTSEKRI	LGLCTTINGLSGKAC	SAGSAENSLSNVTPFCVLS	SPVKSPYLKAPSVVRDQK	STEESETSP	239			
HsSYNP1	SSQQLASF	TKVTSEKRI	LGLCTTINGLSGKAC	STGSESSSSNM	APFCVLS	SPVKSPHLRKASAVIHDQ	HKLSTEETEISP	240		
MmSYNP1	SSQQLAPL	TKVTSEKRI	LGLCTTINGLSAKT	CPIASTENSTPNMTPFCVLS	SPVKSPHLRKAPTALRD	QPKLSTEDSESSP	239			
	*****	*****	*****	*****	*****	*****	*****			
SsSYNP1	PLVKCSAYE	PENQSKDFLNKTFSDPHGRK	VEKTPDCQLRA	FHLQSSAAEPK	LGEQVSGVNW	TGSPRPQK	KRSEYLKVV	319		
HsSYNP1	PLVKCSAYE	PENQSKDFLNKTFSDPHGRK	VEKTPDCQLRA	FHLQSSAAESK	PEEQVSGLN	RNTSSQ	GPPEE-RSEYLKVV	319		
MmSYNP1	ALGKCGPAYE	SENHNSKDFLNKVFSDPHSR	KIEKSGP	DKLRSFHLQSSA	AGAKTEEP	INGMN	WTNTQGTEE-RTEYLKVV	318		
	*****	*****	*****	*****	*****	*****	*****			
		▽ ex4	Ank-like1	Ank-like2	▽ ex5					
SsSYNP1	KSILNIVKEGQISLLPHLAADNLDKIHDH	NGNNLLHVAASQGHAECLQHLTSLMGEDCLNER	NA	DKLTPAGLAIKNGQLE		399				
HsSYNP1	KSILNIVKEGQISLLPHLAADNLDKIHDH	NGNNLLHIAASQGHAECLQHLTSLMGEDCLNER	NT	EKLTPAGLAIKNGQLE		399				
MmSYNP1	RSILNIVNEGQISLLPHLAADNLDKIHDH	NGNNLLHIAASKGHAEC	LQHLTSLMGEDCLNER	NT	EQLTPAGLAIKNGQLE	398				
	*****	*****	*****	*****	*****					
SsSYNP1	CVRWVMVSETEAIAELSCSK	DFPSLIHYAGCYGQEKILLWLLQFMQEQGISL	DEV	DDGNSAVHVASQHG	YLGCIQTLVEY	479				
HsSYNP1	CVRWVMVSETEAIAELSCSK	DFPSLIHYAGCYGQEKILLWLLQFMQEQGISL	DEV	DDGNSAVHVASQHG	YLGCIQTLVEY	479				
MmSYNP1	CVRWVMVSETEAIAELSCSK	DFPSLIHYAGCYGQEKILLWLLQFMQEQGISL	DEV	DREGNSAVHVASQHG	YLGCIQTLVEY	478				
	*****	*****	*****	*****	*****					
		Ank-like3	▽ ex6	Ank-like4	▽ ex7					
SsSYNP1	GANVTMQNHAGEKPSQSAERHGHTLCSRYLVVVE	CMSLASQVVKLTKQLKEQTV	RVTLQNLQFQHI	LEAQKSEARSLLS		559				
HsSYNP1	GANVTMQNHAGEKPSQSAERHGHTLCSRYLVVVE	CMSLASQVVKLTKQLKEQTV	RVTLQNLQFQHI	LEAQKSEARSLLS		559				
MmSYNP1	GANVTMQNHAGEKPSQSAERHGHTLCSRYLVVVE	CMSLASQVVKLTKQLKEQTV	RVTLQNLQFQHI	LEAQKSEARSLLS		558				
	*****	*****	*****	*****	*****					
		Coiled-coil domain	▽ ex8							
SsSYNP1	GANVTMQNHAGEKPSQSAERHGHTLCSRYLVVVE	CMSLASQVVKLTKQLKEQTV	RVTLQNLQFQHI	LEAQKSEARSLLS		559				
HsSYNP1	GANVTMQNHAGEKPSQSAERHGHTLCSRYLVVVE	CMSLASQVVKLTKQLKEQTV	RVTLQNLQFQHI	LEAQKSEARSLLS		559				
MmSYNP1	GANVTMQNHAGEKPSQSAERHGHTLCSRYLVVVE	CMSLASQVVKLTKQLKEQTV	RVTLQNLQFQHI	LEAQKSEARSLLS		558				
	*****	*****	*****	*****	*****					
	▽ ex9									
SsSYNP1	SPSSPSPASRKSQWKS	PDTDEESLAKSKPGVQEGIQV	LSLVSS	RARAKAKDEDS	DKILRQLL	GKDI	ISENVCTQEKLS	639		
HsSYNP1	SPSSPSPASRKSQWKS	PDTDEESLAKSKPGVQEGIQV	LSLVSS	RARAKAKDEDS	DKILRQLL	GKDI	ISENVCTQEKLS	639		
MmSYNP1	SPSSPSPASTKQWALD	TDEESTGKSKVGAQEGIQV	LSLVSS	RARTKKGDEDS	DKILRQLL	GKDI	ISENVCTQEKLS	638		
	*****	*****	*****	*****	*****	*****	*****			
SsSYNP1	LEFQDAQASSRNSKKI	PLEKRELKLARLQLMQ	RSMS	ELDTSN	NSEDPKGT	PVRKADR	PRPQPIVESVESVDS	AEHL	719	
HsSYNP1	LEFQDAQASSRNSKKI	PLEKRELKLARLQLMQ	RSMS	ELDTSN	NSEDPKGT	PVRKADR	PRPQPIVESVESMDS	AEHL	719	
MmSYNP1	LEFQDAQASSRNSKKI	PLEKRELKLARLQLMQ	RSMS	ELDTSN	NSEDPKNT	PVRKADR	PRPQPIVESVENVDS	AEHL	718	
	*****	*****	*****	*****	*****	*****	*****	*****		
SsSYNP1	MIKKHTLASG	-RRFFFGIKASKSLDGHS	PSPTSESSE	PDLSE	QCPGSGTTPPNQ	PPAEP	PPQSPDSTAAQKVATSPKSAL	798		
HsSYNP1	MIKKHTLASG	-RRFFFGIKASKSLDGHS	PSPTSESSE	PDLSE	QCPGSGTTPPNQ	PPAEP	PPQSPDSTAAQKVATSPKSAL	799		
MmSYNP1	MIKKHTLASG	-RRFFFGIKASKSLDGHS	PSPTSESSE	PDLSD	HGPGLGMP	PTQP	STEAQSPDSTAAQKVATSPKSAL	798		
	*****	*****	*****	*****	*****	*****	*****	*****		
SsSYNP1	KSPSSKRRTSQNLKLRVTFE	EPVQMEQSGLEQNGEKDRD	KGRPLPRTS	SAGNESGDQ	LKRPF	GTF	FRSIME	TL	SGNQNNN	878
HsSYNP1	KSPSSKRRTSQNLKLRVTFE	EPVQMEQSGLEQNGEKDRD	KGRPLPRTS	SAGNESGDQ	LKRPF	GTF	FRSIME	TL	SGNQNNN	879
MmSYNP1	KSPSSKRRTSQNLKLRVTFE	EPVQMEQSGLEQNGEKDRD	KGRPLPRTS	SAGNESGDQ	LKRPF	GTF	FRSIME	TL	SGNQNNN	875
	*****	*****	*****	*****	*****	*****	*****	*****	*****	
		▽ ex10								
SsSYNP1	NYQASNQLKTSTLPLTSLGRKT	TD	SKGNP	SGSSASKG	GKNKAA					919
HsSYNP1	NYQASNQLKTSTLPLTSLGRKT	-DAKGN	PASSASKG	GKNKAA						920
MmSYNP1	NYQPASQLKTCTLPLTSLGRKT	TADAKGN	PVSPASKG	GKNKAA						916
	*****	*****	*****	*****	*****	*****	*****	*****	*****	

Fig. 2. Multiple alignment of the porcine synphilin-1 and with synphilin-1 sequences from human and mouse. Alignments of sequences were performed using the Clustal W program on EBI WWW molecular biology server. The numbers represent the position of the amino acids in the respective protein sequences. Identical amino acid residues in all sequences are indicated by asterisks. Arrowheads indicate intron-exon boundaries. The abbreviations for species acronyms and corresponding accession numbers of the sequences used for the alignment are: Hs = *Homo sapiens* (NM_005460); Ss = *Sus scrofa* (NM_001105053); Mm = *Mus musculus*; (NM_026408).

5 of the human *SNCAIP* gene. The deletion does not disrupt the ORF of the 506 amino acid polypeptide, so that the *SNCAIP tv1* mRNA potentially encodes a polypeptide of 55.3 kDa (theoretical pI: 7.64) but lacks a region of 413 amino acids (Tyr20 to Gln432). The calculated molecular mass of tv1 is 55 kDa. Analysis of human brain extracts demonstrated the presence of synphilin-1 predominantly as a 90 kDa polypeptide, but also 120 kDa, 65 kDa and 50 kDa polypeptide bands have been identified (Murray et al., 2003). The different sizes may arise

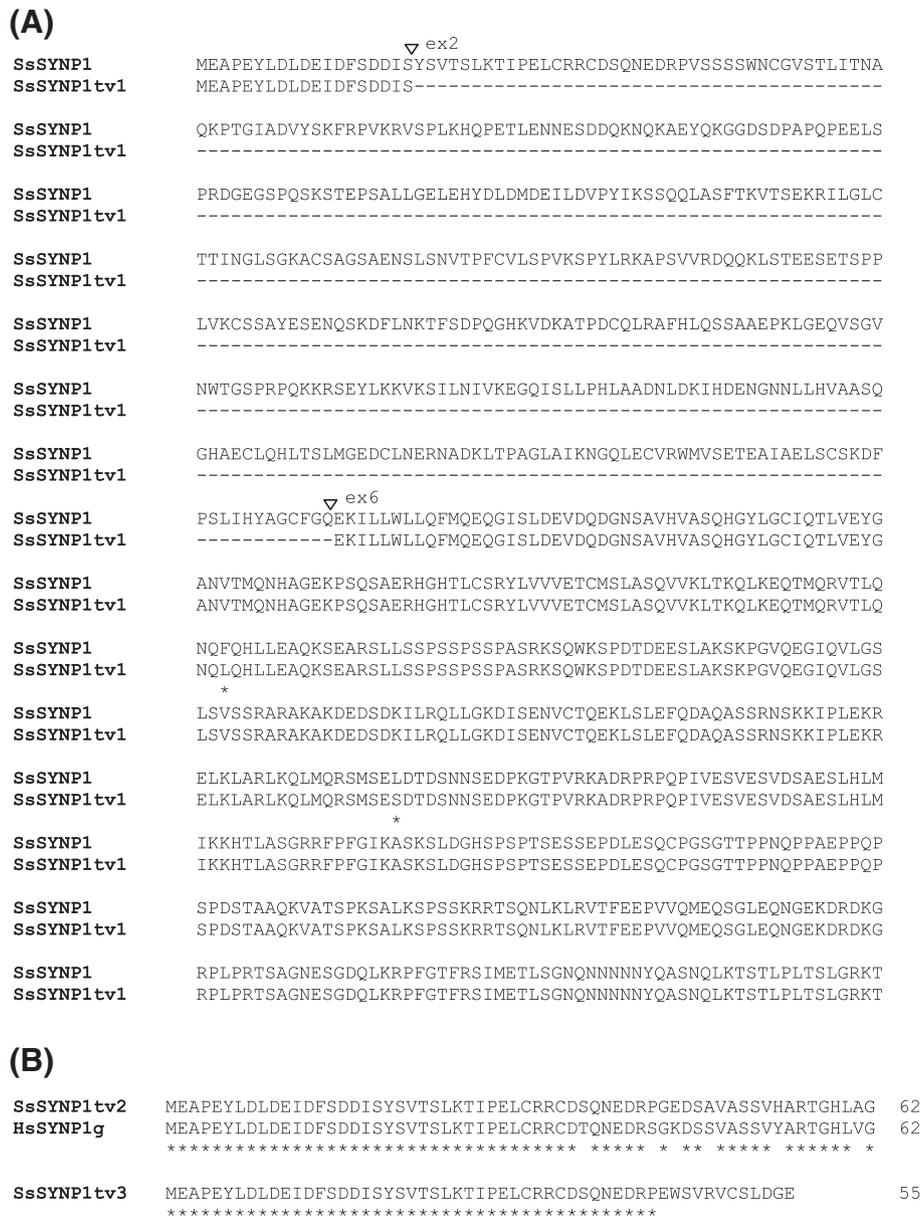


Fig. 3. Splice variants of porcine *SNCAIP*. Amino acid sequence alignment of porcine synphilin-1 and a splice variant, synphilin-1tv1, hereof (A). The synphilin-1tv1 is lacking exon 2–exon 5. Also, a second variant of synphilin-1, synphilin-1tv2, was identified by the cloning (B). This splice variant is homologous with human transcript variant synphilin-1g.

from alternative splicing and post-translational processing. It is not known whether the porcine *SNCAIP* tv1 is represented by any of these band sizes. Distinct regions of synphilin-1 define its ability to form either multiple small aggregates or a single large perinuclear aggresome (Zaarur et al., 2008). The coiled-coil domain and C-terminal ankyrin domains (Fig. 1) determine the formation of small aggregates of the protein whereas its translocation into aggresomes upon proteasome inhibition requires the presence of an N-terminal ankyrin domain. This particular domain is missing in porcine *SNCAIP* tv1. Also, the N-terminal ankyrin domain is important for autophagy of both synphilin-1 aggregate and aggresome (Wong et al., 2012). The tv1 lacks the highly conserved critical domains for the physiological function of synphilin-1. Exons four and five, which are spliced out in tv1, are containing Ank-1, Ank-2 and part of Ank-3 domains, all critical for the interaction with α -synuclein, parkin and dorfins (Chung et al., 2001; Neystat et al., 2002). To our knowledge this particular splice variant of porcine *SNCAIP* has not been identified in human or other species. The deduced protein lacks three ankyrin-like motifs in the amino-terminal end of the protein so it probably has a different function than that of the full length synphilin-1 protein.

By the RT-PCR cloning of porcine *SNCAIP* we also identified a very short transcript variant, tv2, of the *SNCAIP* gene (Fig. 3B). The deduced protein is a truncated version of the full length synphilin-1 polypeptide of 62 amino acids encoded by exon 1 and exon 2. This splice variant of synphilin-1 has also been found in human, where it is named synphilin-1g (GenBank Accession No. AB110794). An even shorter transcript variant, named synphilin-1tv3, was identified (Fig. 3C). This variant also represents a truncated version of synphilin with exon 1 conserved and exon 2 interrupted. A similar transcript variant was seen in human (synphilin-1e). Several attempts to isolate a porcine homologue of human synphilin-1A (Eyal et al., 2006) have failed. However, in silico analysis of sequences from our local database revealed a splice variant resembling synphilin-1A as the exon 10 coding sequence is prolonged. The deduced amino acid sequence of this pig sequence is similar to that reported in GenBank ID: DQ227317 and GenBank ID: BC094759. Also, the potential start codons identified by Eyal et al. (2006) in the human synphilin-1 sequence are also found in the porcine exon 1 sequence (data not shown). To date, a minimum of eight different isoforms of human synphilin-1 have been reported (Humbert et al., 2007). They all arise from the *SNCAIP* gene by alternative splicing. Four of the isoforms represent short C-terminal truncated proteins. The remaining isoforms are longer and contain two or more functional domains. A search in our local porcine protein database did not reveal any peptides representing the three identified splice variants. Hence, it is still not known whether they have a biological significance.

Characterization of porcine SNCAIP genomic sequence

To determine the intron–exon structure of the porcine *SNCAIP* gene we performed a blast search in the Pig Genome v.10 sequence database (<http://www.animalgenome.org/blast/blast.php?bdb=pig10>) using the porcine *SNCAIP* cDNA sequence (GenBank ID: EF1654190). Two sequences of 85 kb and 96 kb covering exon 1 to exon 7 and exon 8 to exon 9, respectively, were retrieved (GenBank IDs: JQ916898 and JQ941708). We did not succeed in finding exon 10. The intron–exon structure is presented in Table S2. The alignment of the genomic clones with the coding region of *SNCAIP* cDNA revealed that the exonic sequences matched 100%. Exon–intron boundaries were initially estimated by alignment of the *SNCAIP* cDNA against the amplified genomic sequences. This revealed a total of 10 exons ranging from 57 to 1069 nucleotides (Table S2). Further sequencing of the flanking regions of the porcine *SNCAIP* gene confirmed the locations of these intron–exon boundaries. All the observed splice acceptor and donor sites were in accordance with the consensus GT-AG rule (Table S2). All exons, except for exon 3 of the porcine *SNCAIP* gene have the same length of coding sequence as those of the human *SNCAIP* sequence. Also the lengths of most introns of the porcine *SNCAIP* gene were comparable to the human counterparts. The length of introns 2 and 9 was not determined.

Sequence analysis of the 5' flanking region of the porcine SNCAIP gene

We have PCR cloned the porcine *SNCAIP* gene promoter and performed a sequence analysis. The nucleotide sequence of the genomic DNA 369 bp upstream from the transcription start site (TSS) of the porcine *SNCAIP* gene was analyzed for transcription factor binding sites using the computer-based MatInspector and TFSEARCH program (<http://molsun1.cbrc.aist.go.jp/htbin/nph-tfsearch>) and using the transfac database. The analysis revealed both a TATA box and two CAAT boxes in the 369 bp 5'-flanking sequence of porcine *SNCAIP*. However, at least two GC-boxes were identified close to the TSS at positions –34 and –84,

respectively. The computer analysis also revealed potential binding sites for the transcription factors AML-1a, STRE and ADR1 (Fig. S2). The porcine and human *SNCAIP* promoter sequences were compared by alignment of 350 nucleotides upstream from the TSS. A high degree of sequence homology (72%) was observed within this region. The high sequence similarity between human and porcine *SNCAIP* could indicate the existence of similar mechanisms for regulation of expression.

Mapping of porcine *SNCAIP*

Initially, to determine the chromosomal location of the *SNCAIP* gene we screened a porcine-rodent hybrid cell panel by PCR (Yerle et al., 1996). The statistical evaluation showed a chromosome probability of 0.1×10^{-3} and correlation of 1.000 to chromosome 2 indicating that the *SNCAIP* gene maps to chromosome 2. The human synphilin-1 gene has been mapped to chromosome 5q23.1–23.3 (Engelender et al., 2000). Our assignment of *SNCAIP* to SSC2 is in agreement with the conservation of synteny between pig and human (Meyers et al., 2005; Rink et al., 2002). Recently, we have used Blat software to localize the *SNCAIP* gene in the *Sus scrofa* 10.2 genome (Groenen et al., 2012). The *SNCAIP* gene maps to SsChr2:131097464–131120771.

Spatial expression of *SNCAIP* mRNA

The *SNCAIP* mRNA expression for the full-length transcript and transcript variant 1 was also examined by RNAseq in various selected organs and tissues from two adult pigs. *SNCAIP* transcripts were detected for both isoforms in all examined organs and tissues and a differential expression was observed (Fig. 4). The highest expression was seen in brain tissues such as frontal cortex, occipital cortex, hypothalamus and cerebellum and also in lung. Both isoforms are expressed at comparable levels. An exception to this is kidney and muscle where *SNCAIP* transcript is only detected for transcript variant 1. The obtained expression results are very similar to those found for human *SNCAIP* (Engelender et al., 2000). Synphilin-1 is expressed constitutively in many human tissues and organs, at high levels in brain, heart and placenta. Furthermore, synphilin-1 expression is identified in subpopulations of neurons and the neuropil in the human brain (Engelender et al., 1999; Wakabayashi et al., 2000). This distribution is similar to that of alpha-synuclein. The expression of synphilin-1 isoform transcripts are altered in Lewy body diseases, generally upregulated compared with healthy controls, and may be involved in pathogenesis (Meyers et al., 2005).

Identification of a potential miRNA recognition site in the 3'UTR of *SNCAIP*

Using the Target Scan (<http://www.targetscan.org>) a recognition site sequence for miR24/24ab/24-3p was identified in the 3'UTR of the human *SNCAIP* gene. The position of the recognition sequence was completely identical to that for the porcine counterpart. The nucleotide identity within a 23 bp stretch of the porcine

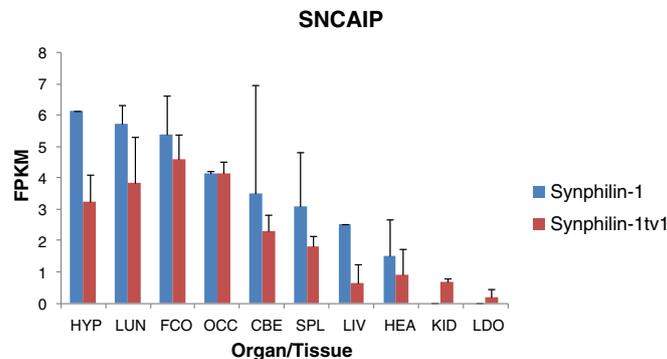


Fig. 4. Expression of porcine *SNCAIP* mRNA determined by RNAseq. Relative abundance of Synphilin-1 and Synphilin-1tv1 in FPKM units. The tissues presented are: hypothalamus (HYP), lung (LUN), frontal cortex (FCO), occipital cortex (OCC), cerebellum (CBE), spleen (SPL), heart (HEA), kidney (KID), liver (LIV) and musculus longissimus dorsii (LDO). The error bars represent the biological variation between the two animals employed.

and the human *SNCAIP* recognition sequence for miR24/24ab/24-3p was 91%. A porcine homologue of miR24 has previously been identified in our laboratory (Nielsen et al., 2010). miR24 and also miR93 directly inhibit vesicular stomatitis virus (VSV) in mammalian cells (Otsuka et al., 2007). All known mammalian miR24 are conserved, suggesting that their ability to restrict VSV proliferation may extend to natural VSV hosts such as cows and sheep. miR24 is ubiquitously expressed in different tissues (Otsuka et al., 2007).

Conclusion

In summary, the present study demonstrates the existence of three splice variants of the porcine *SNCAIP* gene encoding synphilin-1, of which transcript variant 1 is novel. Significant parts of a functional domain, ankyrin repeats, are missing in this variant and this may give rise to a functionally different protein. Transcript variant tv1 lacks exons 2, 3 and 4 of which the last two contain critical ankyrin domains. Exons 3 and 4 are also missing in the human synphilin-1A isoform that has enhanced aggregating properties and causes neurotoxicity (Eyal et al. 2006). Therefore pig synphilin-1 tv1 could have some of the same molecular properties as it lacks the same critical domains. If similar molecular properties with respect to aggregation and inducible autophagy are found in the tv1 variant it could play a role in PD pathogenesis. As no peptides representing either of the splice variants have been identified so far it still remains to be seen if these variants play a functional role in the cells. Further studies will be needed to elucidate the physiological significance of these potential peptides.

Our identification of transcript variants of *SNCAIP* increases the understanding of the function of synphilin-1 and the molecular consequences of alternative splicing. The differential expression pattern of the full-length *SNCAIP* transcript and the tv1 transcript could indicate different physiological roles. Overexpression of the porcine transcript variants of *SNCAIP* should be carried out in cell cultures and in transgenic animals eg. zebrafish in order to examine possible physiological roles of the encoded isoforms. The here presented initial characterization of the porcine synphilin-1 isoforms could be the first step towards a future inclusion of the pig as a model animal to study and elucidate the biological functions of the proteins and their putative role in PD.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.mgene.2015.04.005>.

Acknowledgments

We gratefully acknowledge Dr. Martine Yerle of INRA Toulouse, France, for providing the pig-rodent hybrid panel. The authors wish to thank Connie Jakobsen Juhl, Mette Jacobsen and Bente Flügel for excellent technical assistance and Janne Hansen for critical reading of the manuscript.

References

- Bjerre, D., Madsen, L.B., Bendixen, C., Larsen, K., 2006. Porcine parkin: molecular cloning of *PARK2* cDNA, expression analysis, and identification of a splicing variant. *Biochem. Biophys. Res. Commun.* 347, 803–813.
- Chung, K.K., Zhang, Y., Lim, K.L., Tanaka, Y., Huang, H., Gao, J., Ross, C.A., Dawson, V.L., Dawson, T.M., 2001. Parkin ubiquitinates the alpha-synuclein-interacting protein, synphilin-1: implications for Lewy-body formation in Parkinson disease. *Nat. Med.* 7, 1144–1150.
- Crosiers, D., Theuns, J., Cras, P., Van Broeckhoven, C., 2011. Parkinson disease: insights in clinical, genetic and pathological features of monogenic disease subtypes. *J. Chem. Neuroanat.* 42, 131–141.
- Dickson, D.W., Braak, H., Duda, J.E., Duyckaerts, C., Gasser, T., Halliday, G.M., Hardy, J., Leverenz, J.B., Del Tredici, K., Wszolek, Z.K., Litvan, I., 2009. Neuropathological assessment of Parkinson disease: refining the diagnostic criteria. *Lancet Neurol.* 8, 1150–1157.
- Engelender, S., Kaminsky, Z., Guo, X., Sharp, A.H., Amaravi, R.K., Kleiderlein, J.J., Margolis, R.L., Troncoso, J.C., Lanahan, A.A., Worley, P.F., Dawson, V.L., Dawson, T.M., Ross, C.A., 1999. Synphilin-1 associates with alpha-synuclein and promotes the formation of cytosolic inclusions. *Nat. Genet.* 22, 110–114.
- Engelender, S., Wanner, T., Kleiderlein, J.J., Wakabayashi, K., Tsuji, S., Takahashi, H., Ashworth, R., Margolis, R.L., Ross, C.A., 2000. Organization of the human synphilin-1 gene, a candidate for Parkinson's disease. *Mamm. Genome* 11, 763–766.
- Eyal, A., Szargel, R., Avraham, E., Liani, E., Haskin, J., Rott, R., Engelender, S., 2006. Synphilin-1A: an aggregation-prone isoform of synphilin-1 that causes neuronal death and is present in aggregates from alpha-synucleinopathy patients. *Proc. Natl. Acad. Sci. U. S. A.* 103, 5917–5922.
- Groenen, M.A., Archibald, A.L., Uenishi, H., Tuggle, C.K., et al., 2012. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 491, 393–398.
- Hardy, J., Cai, H., Cookson, M.R., Gwinn-Hardy, K., Singleton, A., 2006. Genetics of Parkinson's disease and parkinsonism. *Ann. Neurol.* 60, 389–398.
- Humbert, J., Beyer, K., Carrato, C., Mate, J.L., Ferrer, I., Ariza, A., 2007. Parkin and synphilin-1 isoform expression changes in Lewy body diseases. *Neurobiol. Dis.* 26, 681–687.
- Kawamata, H., McLean, P.J., Sharma, N., Hyman, B.T., 2001. Interaction of alpha-synuclein and synphilin-1: effect of Parkinson's disease-associated mutations. *J. Neurochem.* 77, 929–934.

- Lee, G., Tanaka, M., Park, K., Lee, S.S., Kim, Y.M., Junn, E., Lee, S.H., Mouradian, M.M., 2004. Casein kinase II-mediated phosphorylation regulates alpha-synuclein/synphilin-1 interaction and inclusion body formation. *J. Biol. Chem.* 279, 6834–6839.
- Li, X., Liu, Z., Tamashiro, K., Shi, B., Rudnicki, D.D., Ross, C.A., Moran, T.H., Smith, W.W., 2010. Synphilin-1 exhibits trophic and protective effects against Rotenone toxicity. *Neuroscience* 165, 455–462.
- Li, X., Tamashiro, K.L., Liu, Z., Bello, N.T., Wang, X., Aja, S., Bi, S., Ladenheim, E.E., Ross, C.A., Moran, T.H., Smith, W.W., 2012. A novel obesity model: synphilin-1-induced hyperphagia and obesity in mice. *Int. J. Obes. (London)* 36, 1215–1221.
- Li, X., Treesukosol, Y., Moghadam, A., Smith, M., Ofeldt, E., Yang, D., Li, T., Tamashiro, K., Choi, P., Moran, T.H., Smith, W.W., 2014. Behavioral characterization of the hyperphagia synphilin-1 overexpressing mice. *PLoS One* 9, e91449.
- Liani, E., Eyal, A., Avraham, E., Shemer, R., Szargel, R., Berg, D., Bornemann, A., Riess, O., Ross, C.A., Rott, R., Engelender, S., 2004. Synphilin-1A: an aggregation-prone isoform of synphilin-1 that causes neuronal death and is present in aggregates from alpha-synucleinopathy patients. *Proc. Natl. Acad. Sci. U. S. A.* 101, 5500–5505.
- Liu, J., Li, T., Yang, D., Ma, R., Moran, T.H., Smith, W.W., 2012. Synphilin-1 alters metabolic homeostasis in a novel *Drosophila* obesity model. *Int. J. Obes. (London)* 36, 1529–1536.
- Marx, F.P., Holzmann, C., Strauss, K.M., Li, L., Eberhardt, O., Cookson, M.R., Hernandez, D., Farrer, M.J., Kachergus, J., Engelender, S., Ross, C.A., Berger, K., Schöls, L., Schulz, J.B., Riess, O., Krüger, R., 2003. Identification and functional characterization of a novel R621C mutation in the synphilin-1 gene in Parkinson's disease. *Hum. Mol. Genet.* 12, 1223–1231.
- Meyers, S.N., Rogatcheva, M.B., Larkin, D.M., Yerle, M., Milan, D., Hawken, R.J., Schook, L.B., Beever, J.E., 2005. Piggy-BACing the human genome II. A high-resolution, physically anchored, comparative map of the porcine autosomes. *Genomics* 86, 739–752.
- Murray, I.J., Medford, M.A., Guan, H.P., Rueter, S.M., Trojanowski, J.Q., Lee, V.M., 2003. Synphilin in normal human brains and in synucleinopathies: studies with new antibodies. *Acta Neuropathol.* 105, 177–184.
- Nagano, Y., Yamashita, H., Takahashi, T., Kishida, S., Nakamura, T., Iseki, E., Hattori, N., Mizuno, Y., Kikuchi, A., Matsumoto, M., 2003. Siah-1 facilitates ubiquitination and degradation of synphilin-1. *J. Biol. Chem.* 278, 51504–51514.
- Neystat, M., Rzhetskaya, M., Kholodilov, N., Burke, R.E., 2002. Analysis of synphilin-1 and synuclein interactions by yeast two-hybrid beta-galactosidase liquid assay. *Neurosci. Lett.* 325, 119–123.
- Nielsen, M., Hansen, J.H., Hedegaard, J., Nielsen, R.O., Panitz, F., Bendixen, C., Thomsen, B., 2010. MicroRNA identity and abundance in porcine skeletal muscles determined by deep sequencing. *Anim. Genet.* 41, 159–168.
- O'Farrell, C., Murphy, D.D., Petrucelli, L., Singleton, A.B., Hussey, J., Farrer, M., Hardy, J., Dickson, D.W., Cookson, M.R., 2001. Transfected synphilin-1 forms cytoplasmic inclusions in HEK293 cells. *Brain Res. Mol. Brain Res.* 97, 94–102.
- Otsuka, M., Jing, Q., Georgel, P., New, L., Chen, J., Mols, J., Kang, Y.J., Jiang, Z., Du, X., Cook, R., Das, S.C., Pattnaik, A.K., Beutler, B., Han, J., 2007. Hypersusceptibility to vesicular stomatitis virus infection in Dicer1-deficient mice is due to impaired miR24 and miR93 expression. *Immunity* 27, 123–134.
- Ribeiro, C.S., Carneiro, K., Ross, C.A., Menezes, J.R., Engelender, S., 2002. Synphilin-1 is developmentally localized to synaptic terminals, and its association with synaptic vesicles is modulated by alpha-synuclein. *J. Biol. Chem.* 277, 23927–23933.
- Rink, A., Santschi, E.M., Eyer, K.M., Roelofs, B., Hess, M., Godfrey, M., Karajusuf, E.K., Yerle, M., Milan, D., Beattie, C.W., 2002. A first-generation EST RH comparative map of the porcine and human genome. *Mamm. Genome* 13, 578–587.
- Trapnell, C., Pachter, L., Salzberg, S.L., 2009. TopHat: discovering splice junctions with RNA Seq. *Bioinformatics* 25, 1105–1111.
- Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., Pachter, L., 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515.
- Wakabayashi, K., Engelender, S., Yoshimoto, M., Tsuji, S., Ross, C.A., Takahashi, H., 2000. Synphilin-1 is present in Lewy bodies in Parkinson's disease. *Ann. Neurol.* 47, 521–523.
- Wheeler, T.C., Chin, L.S., Li, Y., Roudabush, F.L., Li, L., 2002. Regulation of synaptophysin degradation by mammalian homologues of seven in absentia. *J. Biol. Chem.* 277, 10273–10282.
- Wong, E., Bejarano, E., Rakshit, M., Lee, K., Hanson, H.H., Zaarur, N., Phillips, G.R., Sherman, M.Y., Cuervo, A.M., 2012. Molecular determinants of selective clearance of protein inclusions by autophagy. *Nat. Commun.* 3, 1240.
- Yerle, M., Echard, G., Robic, A., Mairal, A., Dubut-Fontana, C., Raquel, J., Pinton, P., Milan, D., Lahbib-Mansais, Y., Gellin, J., 1996. A somatic cell hybrid panel for pig regional gene mapping characterized by molecular cytogenetics. *Cytogenet. Cell Genet.* 73, 194–202.
- Zaarur, N., Meriin, A.B., Gabai, V.L., Sherman, M.Y., 2008. Triggering aggresome formation. Dissecting aggresome-targeting and aggregation signals in synphilin-1. *J. Biol. Chem.* 283, 27575–27584.