Universidade de Aveiro
2012

Departamento de Biologia

Luis Korrodi Mineiro Marques Gregório

Characterization of PPP1 interacting proteins in male reproduction

Caracterização de proteínas que interagem com a PPP1 na reprodução masculina

Universidade de Aveiro
2012

Departamento de Biologia

Luis Korrodi Mineiro Marques Gregório

Characterization of PPP1 interacting proteins in male reproduction

Caracterização de proteínas que interagem com a PPP1 na reprodução masculina

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Bioquímica, realizada sob a orientação científica da Doutora Margarida Sâncio da Cruz Fardilha, Professora Auxiliar Convidada da Secção Autónoma das Ciências da Saúde da Universidade de Aveiro e co-orientação científica da Doutora Odete A.B. da Cruz e Silva, Professora Auxiliar da Secção Autónoma das Ciências da Saúde da Universidade de Aveiro.

Esta dissertação teve o financiamento da FCT através da bolsa SFRH/BD/42334/2007.

Dedico este trabalho ao meu avô, João Alves Mineiro, que é um exemplo para mim a nível pessoal e professional. Estás sempre presente e nunca serás esquecido.

o júri

presidente

Prof. Dr. Amadeu Soares
Professor Catedrático, Departamento de Biologia, Universidade de Aveiro

Prof. Dr. Georg H. Lüers
Professor Associado da Universidade de Hamburg-Eppendorf

Prof. Dr. Srinivasan Vijayaraghavan
Professor Associado da Universidade de Kent

Prof. Dr. Odete Abreu Beirão da Cruz e Silva
Professora Auxiliar com Agregação da Universidade de Aveiro

Prof. Dr. Pedro José Esteves
Professor Coordenador do Instituto Politécnico de Saúde do Norte da Cooperativa de Ensino Superior Politécnico e Universitário

Prof. Dr. Margarida Sâncio da Cruz Fardilha
Professora Auxiliar Convidada da Universidade de Aveiro

agradecimentos

Este trabalho não teria sido possível sem a ajuda de muitas pessoas às quais agradeço o apoio dado:

Ao meu orientador, Prof. Edgar da Cruz e Silva, por ter sido durante a primeira parte do meu doutoramento um mentor para mim.

À minha co-orientadora e mais tarde orientadora, Margarida Fardilha que me acompanhou neste percurso, aguentou as minhas teimas e teve a paciência para me animar sempre que alguma coisa não corria pelo melhor. A ela agradeço o meu crescimento científico.

À minha co-orientadora, Prof[a]. Odete da Cruz e Silva pela disponibiliadade, força e incentivo prestados, em especial na recta final.

A todo o grupo do Laboratório de Transdução de Sinais em especial à Sara Esteves por todo o seu inestimável apoio desde o inicio deste trabalho, pela camaradagem e por ter razão. Às novas estupendas adições, Mónica, João e Joana pela ajuda na recta final. Sem vocês este trabalho não seria tão condimentado.

Ao grupo de Neurociências em especial à Sandra Vieira pelo seu grande apoio em parte do trabalho.

To Dr. Vijay for the collaboration in the transgenic mice work, for all the fruitful scientific discussions and all the support during the one year stay in Kent, USA. Also to all lab members, Nilam, Ramdas, Teja, Sromona, David and Meenakshi for the support and friendship. Also to Cheryl, for being my borrowed mother and for helping me in the american journey.

Ao Dr. Pedro Esteves pelas diversas discussões futebolísticas mas sempre científicas e também à Joana e à Ana pela ajuda na parte final do trabalho.

Aos meus amigos por animarem esta etapa e tornarem-na mais fácil, em especial ao Gonçalo pelo companheirismo e aventura por diferentes terras e à Sofia por me acompanhar nos momentos laboratoriais mais hilariantes. Aos colegas de casa passados e presentes, em especial ao Tiago, o eterno comparsa nesta academia.

À minha namorada, Rita, por todo apoio prestado a todos os níveis e por ser uma pessoal especial. Obrigado por me aguentares.

À minha família, em especial aos meus pais e às minhas avós por me terem sempre apoiado em todas as escolhas que fiz e por serem quem são. Sem vocês isto não teria sido possível.

Obrigado a todos.

palavras-chave

PPP1, fosfatase, Tctex1d4, citoesqueleto, microtúbulos, tubulina, fosforilação, dineína, inibidor 2, PPP1R2, PPP1R2P3, espermatozoide, testículo

resumo

A fosforilação reversível de proteínas é um importante mecanismo de controlo em eucariotas. A fosfoproteína fosfatase 1 (PPP1) é uma fosfatase de serina/treonina envolvida em vários processos celulares. Existem três isoformas da subunidade catalítica (α/CA, δ/β/CB e γ/CC) com pequenas diferenças nos terminais amino e carboxílico. O gene PPP1CC sofre ainda *splicing* alternativo para produzir duas isoformas, a PPP1CC1 ubíqua e a PPP1CC2 enriquecida em testículo e específica de esperma. A localização e especificidade de substratos da PPP1 está dependente da formação de complexos oligoméricos com proteínas que interagem com a PPP1 (PIPs).

O objetivo principal desta tese foi estudar novas PIPs, específicas de testículo e esperma, a fim de melhor caracterizar o papel desta fosfatase e dos respetivos complexos na reprodução em mamíferos. Com este fim, estudou-se a presença, localização e possíveis funções de uma PIP previamente conhecida, PPP1R2, e de duas novas PIPs, PPP1R2P3 e Tctex1d4.

PPP1R2 e PPP1R2P3 estão presentes em esperma humano colocalizando com a PPP1CC2, na cabeça e na cauda. A hipótese é que as holoenzimas localizadas na cabeça terão um papel na reação acrossómica, enquanto que as holoenzimas presentes no axonema são relevantes para o controlo da motilidade flagelar. De seguida foram estudados os pseudogenes da PPP1R2, em termos de história evolutiva e de possíveis funções. Na espécie humana, a PPP1R2 tem 10 pseudogenes, 7 deles específicos de primatas. Estudos de bioinformática e dados de expressão mostram que os PPP1R2P1/P3/P9 são os pseudogenes com maior probabilidade de serem transcritos e traduzidos. Também identificámos o PPP1R2P9 em esperma humano e mostrámos que alguns pseudogenes poderão estar associados a estados fisiopatológicos. Isto indica que o processo de evolução poderá estar ligado á formação de novos genes ou ao controlo do mRNA da PPP1R2. A sobre-expressão da PPP1R2 ou PPP1R2P3 em testículo de ratinho também foi realizada, para caracterizar os mecanismos envolvidas na função dos complexos PPP1R2/PPP1R2P3-PPP1CC2 na espermatogénese e fisiologia dos espermatozoides.

A dineína de cadeia leve, Tctex1d4, foi encontrada como interagindo com a PPP1C e como estando presente em testículo de ratinho e em esperma humano. Demonstrámos que a Tctex1d4 e a PPP1 colocalizam no centro organizador de microtúbulos e nos microtúbulos e que o motivo de ligação à PPP1 presente na Tctex1d4 parece ser importante para manter a PPP1 no centro organizador de microtúbulos e/ou para disromper ou atrasar o seu movimento ao longo dos microtúbulos emergentes. Estes resultados abrem novos caminhos para os possíveis papéis do complexo Tctex1d4-PPP1 na dinâmica dos microtúbulos, motilidade do esperma, reação acrossómica e na regulação da barreira hemato-testicular, provavelmente, através da via de sinalização do TGFß. A análise do motivo de ligação à PPP1 mostra que este é altamente conservado entre os mamíferos, com exceção das Pikas, sugerindo que esta perda aconteceu antes da radiação das Pikas, há 6-20 milhões de anos atrás. Através de um rastreio por mutações demonstrámos que a capacidade da Tctex1d4 se ligar à PPP1 é mantida nas Pikas, embora o motivo de ligação à PPP1 esteja disrompido.

Este estudo abre portas para novas descobertas na área da reprodução mostrando o papel da PPP1CC2 na espermatogénese e fisiologia do esperma.

keywords

PPP1, phosphatase, Tctex1d4, cytoskeleton, microtubules, tubulin, phosphorylation, dynein, inhibitor 2, PPP1R2, PPP1R2P3, spermatozoa, testis

abstract

Reversible phosphorylation of proteins is an important intracellular control mechanism in eukaryotes. Phosphoprotein Phosphatase 1 (PPP1) is a major serine/threonine protein phosphatase involved in a wide range of cellular processes. Three closely related catalytic subunit isoforms (α/CA, δ/β/CB and γ/CC) exist with only minor differences at their N- and C-terminus. PPP1CC gene can also undergo tissue-specific processing to yield a ubiquitously expressed PPP1CC1 and the testis-enriched and sperm-specific PPP1CC2 isoforms. PPP1C exists in the cell as an oligomeric complex binding to a spectrum of PPP1 interacting proteins (PIPs), which modulate both its intracellular localization and substrate specificity.

The main goal of this thesis was to study novel PIPs in testis and sperm, in order to further characterize the role of PPP1CC2 and the respective complexes in mammalian reproduction. To this end we addressed the presence, localization and putative roles of a previously known PIP, PPP1R2, in testis and sperm, and two novel PPP1CC2 testis/sperm specific PIPs, PPP1R2P3 and Tctex1d4.

PPP1R2/PPP1R2P3 were shown to be present in human sperm co-localizing with PPP1CC2, in the head and tail. It was shown that PPP1R2P3 is a heat stable inhibitor of PPP1CC that cannot be phosphorylated by GSK3. We hypothesize that the holoenzymes localized in the head may have a role in the acrosome reaction while the axoneme bound holoenzymes are relevant for the control of flagellar motility. To further address the PPP1R2 significance, its pseudogenes were described in terms of evolutionary history and putative functions. In human specie, PPP1R2 has ten pseudogenes most of them primate-specific. Besides PPP1R2P3, bioinformatic studies and expression data show that PPP1R2P1, PPP1R2P2 and PPP1R2P9 are the pseudogenes with more probability of being transcribed and eventually translated. Moreover, we identified PPP1R2P9 in human sperm and showed that several pseudogenes appear to be associated with physiological and pathological states. This indicates that evolution processes might be in part related with the formation of new genes or in the control of the parental PPP1R2 message. Overexpression of human PPP1R2 or PPP1R2P3 in mouse testis was also pursued to provide the molecular tools to initiate the characterization of the mechanisms behind PPP1R2/PPP1CC2 and PPP1R2P3/PPP1CC2 role in spermatogenesis and sperm physiology.

Dynein light chain, Tctex1d4, was found to bind to PPP1C and to be present in mouse testis and human sperm. Tctex1d4-PPP1CC complex was shown to co-localize in the microtubule organizing centre and in microtubules. Moreover, the Tctex1d4 PPP1 binding motif seems to be important to retain PPP1CC in the microtubule organizing centre, and also to disrupt or delay its movement along microtubules. These results open new avenues to the possible roles of Tctex1d4-PPP1 complex in microtubule dynamics, sperm motility, acrosome reaction and in regulation of the blood testis barrier possibly via TGFß signaling. Moreover, PPP1 binding motif is highly conserved among mammals, except in Pikas, suggesting that this event happened before the Pikas radiation, 6-20 Million years ago. Mutational screening shows that the ability of Tctex1d4 to bind to PPP1 is maintained in Pikas, although the PPP1 binding motif is disrupted.

This work opens doors to new discoveries in male reproduction and unravels the roles of PPP1CC2 and its PIPs in spermatogenesis and sperm physiology.

Index

Abbreviations

ACN	acetonitrile
AKAP	A-kinase-anchoring protein
APC/C	anaphase-promoting complex
Arg	arginine
ASPP1/2	apoptosis-stimulating of p53 protein 1 and 2, also known as PPP1R13B and PPP1R13A respectively
ATP	adenosine triphosphate
Bad	Bcl2-associated agonist of cell death
BCA	bicinchoninic acid assay
Bcl-2	B-cell CLL/lymphoma 2
BEB	bayes empirical bayes
BLAST	basic local alignment search tool
bp	base pair
BRCA1	breast cancer 1
BSA	bovine serum albumin
BTB	blood testis barrier
Ca^{2+}	calcium
cAMP	cyclic adenosine monophosphate
CASA	computer assisted sperm analyser
CASK	calcium/calmodulin-dependent serine protein kinase
CDC25C	cell division cycle 25 homolog C
CDH	cadherin
CDK	cyclin-dependent kinase
cDNA	complementary DNA
CDS	coding sequence

CHAPS	3[(3-Cholamidopropyl)dimethylammonio]-propanesulfonic acid
CID	collision-induced dissociation
CKII	casein kinase II
CMV-IE	cytomegalovirus immediate early
COS-7	monkey kidney fibroblast cell line
CPI-17	protein kinase C potentiated inhibitor 17, also known as PPP1R14A
CT	C-terminal
CTD	carboxy-terminal domain
Cre	causes recombination
CReP	constitutive repressor of eIF2alpha phosphorylation, also known as PPP1R15B
DAPI	4,6-diamidino-2-phenylindole
DARPP32	dopamine and cAMP regulated phosphoprotein, also known as PPP1R1B
ddH_2O	double distilled water
DIC	differential interference contrast
DMEM	Dulbecco's Modified Eagle medium
dN	non- synonymous substitutions per site
DNA	deoxyribonucleic acid
dS	synonymous substitutions per site
DTA	diphtheria toxin A
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EF1α	elongation factor 1α
EGTA	ethylene glycol tetraacetic acid
ELM	eucaryotic linear motif
ERK	extracellular signal-regulated kinase
ES	embryonic stem

EST	expressed sequence tags
FA	formic acid
FAK	focal adhesion kinase
FCP	TFIIF-associating CTD phosphatase
FDR	false discovery rate
FEL	fixed-effect likelihood
FITC	fluorescein isothiocyanate
FLP	yeast-derived Flip
FRT	flipase recognition target
FSIP	fibrous sheath interacting protein
FSP95	fibrous sheath protein of 95 kDa
GADD34	growth arrest and DNA damage-inducible protein 34, also known as PPP1R15A
gDNA	genomic DNA
GBPI1	gastrointestinal and brain-specific PP1-inhibitory protein 1, also known as PPP1R14D
GC	guanine-cytosine
GC2	germ cells
GEO	gene expression omnibus
GL	hepatic glycogen-targeting protein phosphatase 1 regulatory subunit GL also known PPP1R3B
Glu	glutamic acid
GM	protein phosphatase 1 regulatory subunit GM, also known PPP1R3A
GSK3	glycogen synthase kinase 3
hAAT	human α1-antitrypsin
Hb2E	protein phosphatase 1 regulatory subunit 3F, also known as PPP1R3F
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HLA	human leukocyte antigen

hsp90	heat shock protein 90
I1	inhibitor 1, also known as PPP1R1A
I2	inhibitor 2, also known as PPP1R2
I2-L	inhibitor 2-like, also known as PPP1R2P3
IACUC	institutional animal care and use committee
ICs	intermediate chains
IFT	intraflagellar transport
IHC	immunohistochemistry
IPP5	inhibitor-5 of protein phosphatase 1, also known as PPP1R1C
IPTG	isopropyl-β-D-thio-galactopyranoside
IUP	intrinsically unstructured protein
JHDM1D	jumonji C domain containing histone demethylase 1 homolog D
KEPI	kinase-enhanced PP1 inhibitor, also known as PPP1R14C
KIAA1443	protein phosphatase 1, regulatory subunit 3E, also known as PPP1R3E
LB	loading buffer
LCs	light chains
Leu	leucine
LICs	light intermediate chains
LINE	long interspersed nuclear elements
LRT	likelihood ratio test
LTQ	linear ion trap
LTR	long terminal repeat
MAOA	monoamine oxidase A
MAPK	mitogen activated protein kinase
Mg^{2+}	magnesium
MHC	major histocompatibility complex

ML	maximum-likelihood
Mn^{2+}	manganese
MoMLV	moloney murine leukemia virus
mPrm	mouse protamine promoter
mRNA	messenger RNA
MTOC	microtubule organizing center
MYPTs	myosin phosphatase targeting subunit, also known as PPP1R12A/B. MYPT3 is known as PPP1R16B
NCBI	National Center for Biotechnology Information
Nek2	nimA-related protein kinase 2
Neurabins I/II	neuronal binding proteins, also known as PPP1R9A/B
NIPP1	nuclear inhibitor of protein phosphatase 1, also known as PPP1R8
NJ	neighbor-joining
NOM1	nucleolar protein with MIF4G domain 1
NT	N-terminal
ORF	open reading frame
p85	protein phosphatase 1 myosin-binding subunit of 85 kDa, also known as PPP1R12C
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline with Tween
PCDH	protocadherin
PCR	polymerase chain reaction
PDHA2	pyruvate dehydrogenase (lipoamide) alpha 2
PGK1	phosphoglycerate kinase 1
PH	phase contrast
PHACTR	phosphatase and actin regulator
PHI1	phospholipase C beta-3 neighboring gene protein, also known as PPP1R14B

PIP	phosphatase interacting protein
PKA	cAMP-dependent protein kinase
PMSF	phenylmethylsulfonyl fluoride
PNUTS	protein phosphatase 1 nuclear targeting subunit, also known as PPP1R10
PP	protein phosphatase
PPM	protein phosphatase Mg^{2+} or Mn^{2+} dependent
PPP	phosphoprotein phosphatase
PPP1BM	PPP1 binding motif
PPP1C	PPP1 catalytic subunit
PPP1CA/B/C	PPP1 catalytic subunit alpha/beta/gamma isoform
PPP1R	PPP1 regulatory subunit
Pro	proline
PTG	protein targeting to glycogen, also known as PPP1R3C
RAP-1	ras related protein 1
RAP1GDS1	GTP-GDP dissociation stimulator 1
REL	random effect likelihood
RIPA	radio-immunoprecipitation assay
RIPP1	ribosomal inhibitor of PP1
RNA	ribonucleic acid
RSV-LTR	rous sarcoma virus long terminal repeat
RT	room temperature
RT-PCR	reverse transcriptase PCR
RT-qPCR	real time quantitative PCR
RUNX1	Runt-related transcription factor 1
SARP	several ankyrin repeat protein
SCP	small CTD phosphatase

SDS	sodium dodecyl sulfate
sds22	protein phosphatase 1 regulatory subunit 22, also known as PPP1R7
SDS-PAGE	SDS polyacrylamide gel electrophoresis
Ser	serine
SETD4	SET domain containing 4
SGCD	sarcoglycan delta
SI	spectral index
SINE	Short interspersed nuclear elements
SLAC	single likelihood ancestor counting
SLC5A7	solute carrier family 5 choline transporter, member 7
SLC37A3	solute carrier family 37 (glycerol-3-phosphate transporter), member 3
ST6GAL2	ST6 beta-galactosamide alpha-2,6-sialyltranferase 2
STPP	serine-threonine protein phosphatase
SV-40	simian virus 40
Szp1	spermatogenic zip protein 1
TAKAP-80	testis-specific A-kinase-anchoring-protein
TAP	transporter , ATP-binding cassette, sub-family B
TAT-1	transactivating regulatory protein 1
TAU	microtubule-associated protein tau
TBS	tris buffered saline
TBS-T	tris buffered saline with tween
Tctex5	t-complex-associated-testis-expressed 5, also known as PPP1R11
TFIIF	transcription factor IIF
TGFβ	transforming growth factor beta
Thr	threonine
TIMAP	transforming growth factor β inhibited membrane-associated protein, also known as PPP1R16B

TIMD4	T-cell immunoglobulin and mucin domain containing 4
tk	thymidine kinase
TLRR	testis leucine-rich repeat
TPCK	tosyl phenylalanyl chloromethyl ketone
Tsga2	testis specific gene A2
Tyr	tyrosine
URI	unconventional prefoldin RPB5 interactor
UTR	untranslated region
WAVE 1	Wiskott-Aldrich syndrome protein family member 1

Chapter I

Introduction

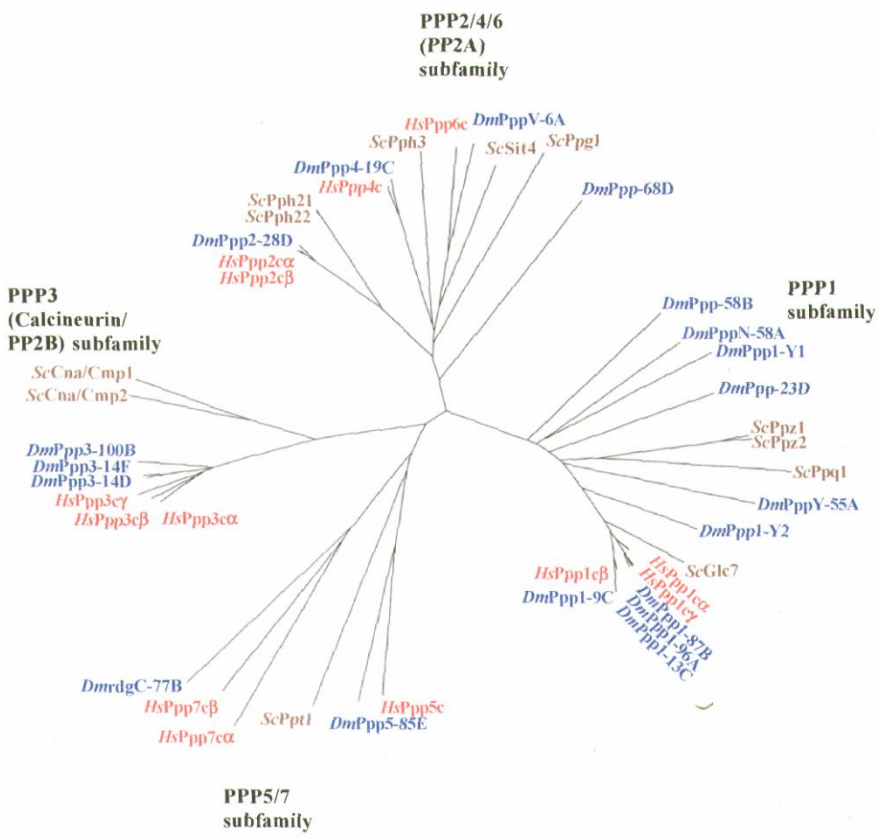
Introduction

Protein phosphorylation - kinases and phosphatases

Reversible phosphorylation of structural and regulatory proteins is an important intracellular control mechanism in eukaryotes affecting up to 30% of the proteome. It plays a central role in the control of almost all cellular functions including metabolism, signal transduction, cell division and memory [1-3]. Intracellularly, protein phosphorylation regulates a variety of important functions, including subcellular localization, protein degradation and stabilization, as well as biochemical activities [4].

The phosphorylation state of a protein is a dynamic reversible process involving both protein kinases (add the phosphate), and protein phosphatases (PPs, remove the phosphate). Unlike protein kinases, that all belong to a single gene family, PPs are divided into several distinct and unrelated protein/gene families. The Tyr-specific PP family, as well as including the Tyr-specific PPs, also comprises the so-called dual specificity PPs (capable of dephosphorylating Ser, Thr and Tyr residues). Within the Ser/Thr-specific protein phosphatases (STPPs) three distinct gene families have been described: the PPMs (Mg2+ or Mn2+ dependent protein phosphatases), the FCPs (TFIIF-associating CTD phosphatases) and the PPPs (phosphoprotein phosphatases). The PPM family comprises the Mg2+-dependent PPs, such as pyruvate dehydrogenase, PP2C, and relatives [5]. The FCP family comprises FCP1 and SCPs 1-3 PPs [6, 7]. Sequence analysis using three well characterized genomes, allowed the construction of a phylogenetic tree and the division of PPP family in four subfamilies: PPP1 (Fig I.1A right branch), PPP2 (PPP2/PP2A, PPP4 and PPP6, Fig I.1A upper branch), PPP3/PP2B/Calcineurin (Fig I.1A, left branch) and PPP7 (PPP5 and PPP7, Fig I.1A, lower branch). These gene subfamilies share high homology in the catalytic domains (Fig I.1B, dark blue bars) but differ in the N- and C-terminal domains (Fig I.1B) [5, 8, 9]. Besides these intracellular PPs involved in signal transduction, there are also unrelated non-specific alkaline and acidic PPs that are usually found either in specialized intracellular compartments or in the extracellular milieu [10].

A



B

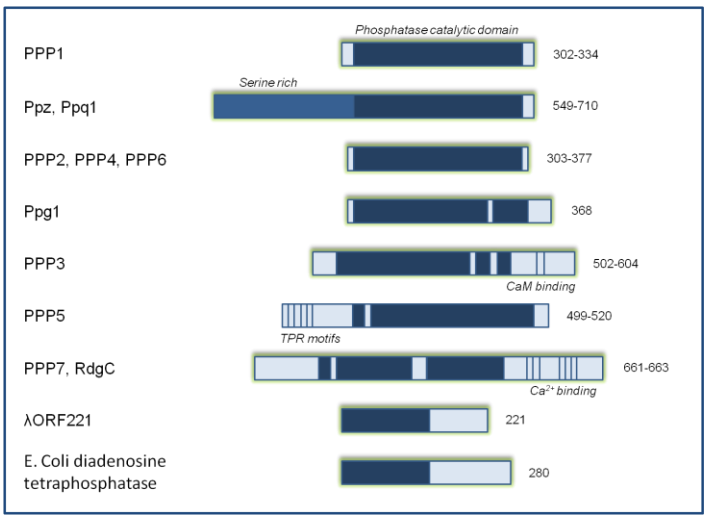


Figure I. 1: **Evolutionary and structural relationship between the different PPP families. A.** *Phylogenetic tree, depicting the relationships between Homo sapiens (Hs, red), Drosophila melanogaster (Dm, blue) and budding yeast Saccharomyces cerevisiae (Sc, brown) protein serine/threonine phosphatases of the PPP family (figure taken from* [4]*).* **B.** *Domain organization of PPP family members. The size in amino acids is shown on the right (adapted from* [11]*).*

The sequencing of entire genomes has revealed that approximately 3% of all eukaryotic genes encode protein kinases or PPs [12]. Surprisingly, there appear to be 2-5 times fewer PPs than protein kinases. This imbalance is even more pronounced when the analysis is limited to STPPs and Ser/Thr-kinases, particularly in vertebrates. The human

genome, for instance, encodes approximately 20 times fewer STPPs than Ser/Thr-kinases. Thus, it is often concluded that, whereas, the diversity of the Ser/Thr-protein kinases has kept pace with the increasing complexity of evolving organisms, the STPPs apparently have not. However, in the past two decades it has become evident that the diversity of STPPs is achieved not only by the evolution of new catalytic subunits, but also by the ability of a single catalytic subunit to interact with multiple regulatory subunits [1, 13].

Phosphoprotein Phosphatase 1 (PPP1)

Phosphoprotein Phosphatase 1 (PPP1) is a major STPP involved in a wide range of cellular processes such as cell cycle progression, protein synthesis, muscle contraction, glycogen metabolism, cytokinesis, and neuronal signaling [1, 3].

The PPP1 catalytic subunit (PPP1C) is expressed from one (*Saccharomyces cerevisiae*) to eight (*Arabidopsis thaliana*) different isoforms in the eukaryotic genomes. In mammalian genomes three separate genes encode for three closely related isoforms (α/A, δ/β/B and γ/C). These isoforms are > 90% identical in amino acid sequence, with minor differences, primarily at their N- and C-terminus [1]. *PPP1CC* gene also undergoes tissue-specific alternative splicing to yield a ubiquitously expressed PPP1CC1 (γ1) isoform and the testis-enriched and sperm-specific PPP1CC2 (γ2) isoform. [14-18] (Fig I.2).

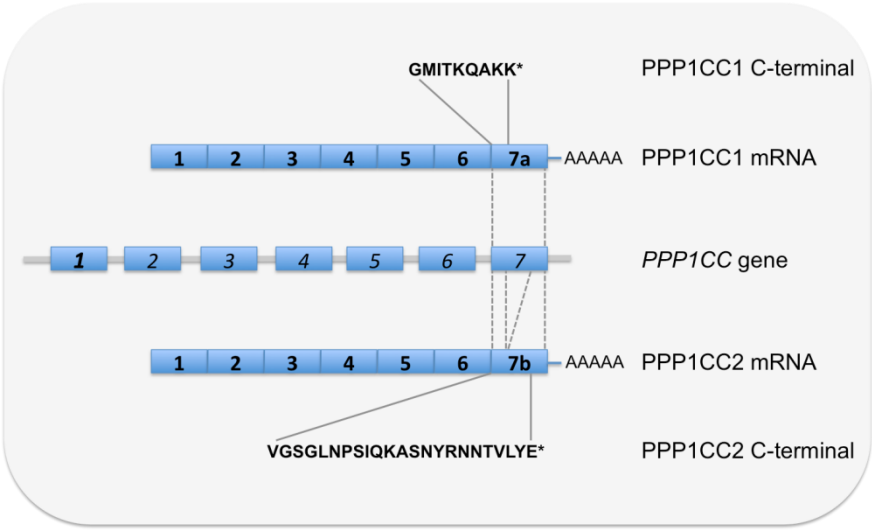


***Figure I. 2:* Schematic representation of PPP1CC gene exon-intron organization.** *Exons are shown in blue boxes with numbers. Alternative splicing of exon 7 originates the testis-enriched and sperm-specific PPP1CC2 isoform. Amino acids corresponding to the exon 7 specific C-terminal are shown. * denotes the termination codon.*

It has already been shown that PPP1 isoforms are expressed in a variety of mammalian cells, although they localize intracellularly in a distinct and characteristic manner. PPP1CA was found to be ubiquitous in all mouse tissues except in skeletal and heart muscles. PPP1CB is also ubiquitous in all mouse tissues except in skeletal muscle and PPP1CC1 has higher levels in brain, small intestine and lung compared to other tissues and not detectable in heart and spleen. PPP1C2 is the isoform in higher quantity in testis and virtually the only one expressed in spermatozoa but with low amounts in brain, lung, spleen and thymus [19]. Specifically in brain, the different PPP1 isoforms were shown to be present in different regions and revealed also specific subcellular localization [18, 20]. While PPP1CB is the predominant isoform associated with microtubules in the neuronal cell body, PPP1CC1 and PPP1CA are preferentially concentrated in the dendritic spines [20, 21]. Tissue expression of PPP family members is shown in Table 1.

Table I. 1: *Tissue expression of the main PPP family members.*

Protein Phosphatase	Tissues
PPP1CA	Ubiquitous; more predominantly in brain [19]
PPP1CB	Ubiquitous; more predominantly in liver and kidney [19]
PPP1CC1	Ubiquitous; more predominantly in brain, small intestine and lung [19]
PPP1CC2	Yes, low abundance [20]
PPP2	Ubiquitous; more predominantly in brain [22]
PPP3	Ubiquitous; more predominantly in brain [23]
PPP4	Ubiquitous; more predominantly lung, liver and kidney [24]
PPP5	Ubiquitous; more predominantly in brain [25, 26]
PPP6	Ubiquitous; more predominantly in heart and skeletal muscle [27]
PPP7	Ubiquitous; more predominantly in sensory organs [28]

During the cell cycle, phosphorylation status, activity, and subcellular localization of PPP1 changes. Studying PPP1 localization during the cell cycle, Andreassen *et al.* demonstrated that the distribution of PPP1 isoforms in cells was highly dynamic. PPP1CA/B/C1 localize to distinct subcellular compartments during both interphase and mitosis [29]. PPP1 is expressed in various cellular compartments, but it is most abundant in the nucleus. Within the nucleus, PPP1CA associates with the nuclear matrix, PPP1CC1 localizes to the nucleolus, and PPP1CB is associated with whole chromatin. During mitosis, PPP1CA localizes to centrosomes, while PPP1CC1 is associated with microtubules of the mitotic spindle. In contrast, PPP1CB is strongly localized to chromosomes [30-32].

Although most biochemical studies have not directly addressed the significance of the different isoforms, it is now well established that these distinct subcellular localizations, activity and phosphorylation status are due in part to the interacting proteins [17, 29, 31]. Moreover PPP1C tissue specificity could be a determinant of which subset of different regulators are available to. Also, during evolution, other proteins present in the same tissues, and/or with the same subcellular localization, might have gained the ability to bind to a specific isoform, giving rise to new functions, and extending the repertoire of known regulators [33].

Phosphoprotein phosphatase 1 – Phosphatase Interacting Proteins (PIPs)

PPP1C exists in the cell as an oligomeric complex. The PPP1C binds a spectrum of interacting proteins, PPP1 interacting proteins (PIPs), also known as PPP1 regulatory subunits (PPP1R), which modulate both PPP1C intracellular localization and substrate specificity, and may also function as target subunits [1, 3, 34]. This implies differences in the specificity of interaction of a particular PPP1 isoform with a particular PIP, which may in turn exhibit subcellular compartmentalization or tissue specific enrichment. For example, PPP1R9A/neurabin I, targets PPP1C to the actin-rich post-synaptic density, where the complex regulates the dendritic spine morphogenesis and maturation. In contrast PPP1R9B/neurabin II preferentially binds to the PPP1CC1 isoform over the other two isoforms, PPP1CA and PPP1CB [35-38].

During the past two decades, a variety of approaches have identified more than two hundred PIPs. However, considering the number of phosphatases and phosphoprotein substrates encoded for by the human genome many more remain unknown [2, 5, 29, 31,

38, 39]. PIPs are divided in four major categories: substrates, substrate specifiers, targeting subunits or inhibitors of the catalytic activity [39].

Some PIPs, like BRCA1 [40], the protein phosphatase CDC25C [41], the apoptotic protein Bad [42], caspase 2 [43] and protein kinase Nek2 [44] are PPP1 substrates. In contrast to the others, the last two are maintained in an inactive state by PPP1. While PPP1R2 [45] and PPP1R14A [46] are both substrates and inhibitors of PPP1, Nek2 [47] and PPP1R16B [48] are dephosphorylated by PPP1 but also target other proteins to PPP1 mediating their dephosphorylation. PIPs can also target PPP1 to specific structures, such as the nucleus (PPP1R10) [49], nuclear membrane (AKAP149) [50], nucleoli (NOM1) [51], chromatin (Repo-man) [52], centrosome (Nek2) [44], plasma membrane (integrin αIIB) [53], actin cytoskeleton (PPP1R9 subfamily) [54], microtubules (TAU), myosin (PPP1R12 subfamily and PPP1R16A) [55], glycogen particles (PPP1R3 subfamily) [56], endoplasmic reticulum (PPP1R15A) [57], mitochondria (URI, Bcl2) [58, 59].

Finally, some PIPs are true PPP1 inhibitors, as they block the access to the active site and inhibit the dephosphorylation of all substrates, for instance PPP1R2 and PPP1R11 [45]. Additionally, several regulators can show preferential binding to a specific PPP1 isoform. Some examples are given in Table I.2.

Table I. 2: PPP1 regulatory subunits. *Regulatory subunit genes nomenclature and the function of the proteins coded by these genes are given. Also, the specific PPP1C isoform that was found to bind to these regulatory subunits is shown (adapted from* [2]*).*

Regulatory Subunit		PPP1C isoform	Family	Holoenzyme function	References supporting function
Gene	Alternatives names				
PPP1R1A	Inhibitor 1, I1			PPP1 inhibitor, glycogen metabolism, synaptic plasticity and muscle	[60, 61]
PPP1R1B	DARPP32	PPP1CA		Neurotransmission	[62, 63]
PPP1R1C	IPP5	PPP1CA		Apoptosis	[64]
PPP1R2	Inhibitor 2, I2	PPP1CA PPP1CB PPP1CC		PPP1 inhibition, phosphorylated by Pro-directed kinases	[65, 66]
PPP1R3A	GM, RGL, PPP1R3, PP1G			Glycogen metabolism	[56, 67, 68]
PPP1R3B	GL, PPP1R4	PPP1CA		Glycogen metabolism	[69]
PPP1R3C	PTG, PPP1R5			Glycogen metabolism	[70-72]
PPP1R3D	PPP1R6	PPP1CC		Glycogen metabolism	[68]
PPP1R3E	KIAA1443, FLJ00089			Glycogen metabolism	[73]
PPP1R3F	Hb2E			Glycogen metabolism; Depletion resulted in G2 M arrest	[74]
PPP1R3G				Glycogen metabolism	[33]

Regulatory Subunit		PPP1C isoform	Family	Holoenzyme function	References supporting function
Gene	Alternatives names				
PPP1R7	Sds22	PPP1CB		Mitosis, regulation of sperm function and epithelial cell polarity and shape	[75-77]
PPP1R8	NIPP1, ARD-1	PPP1CA PPP1CB PPP1CC		RNA splicing	[78, 79]
PPP1R9A	Neurabin I	PPP1CA	Neurabin family	Dendritic spine signaling, synaptic plasticity and synaptic transmission	[37, 38, 80, 81]
PPP1R9B	Spinophilin, Neurabin II	PPP1CA PPP1CB PPP1CC	Neurabin family	Dendritic spine signaling, synaptic plasticity and synaptic transmission	[37, 38, 81, 82 , 83]
PPP1R10	PNUTS, p99, CAT53	PPP1CA		RNA splicing, chromosome decondensation, apoptosis, proteasomal degradation and retinal synaptic activity	[84-89]
PPP1R11	Inhibitor 3, HCG-V, TCTE5, TCTEX5, IPP3	PPP1CA PPP1CC		Inhibits PPP1, apoptosis, sperm function	[90-92]
PPP1R12A	MYPT1, M110, MBS, M130	PPP1CB	MYPT family	Myosin/actin targeting	[55]
PPP1R12B	MYPT2, PP1bp55, M20 splice form	PPP1CB	MYPT family	Myosin/actin targeting; target subunit of myosin phosphatase in heart	[93]
PPP1R12C	p85, LENG3	PPP1CB	MYPT family	Myosin/actin targeting	[94]
PPP1R13A	TP53BP2, p53BP2, ASPP2	PPP1CA PPP1CC		Apoptosis	[95, 96]
PPP1R13B	ASPP1, p53BP2-like	PPP1CA		Apoptosis	[97]
PPP1R14A	CPI-17		PHI family	Inhibits smooth muscle myosin phosphatase increasing muscle contraction	[46, 98]
PPP1R14B	PHI-1		PHI family	Modulates retraction of endothelial and epithelial cells	[46, 99]
PPP1R14C	KEPI, CPI-17like		PHI family	PKC-dependent PPP1 inhibitor regulated by morphine; regulation of signaling pathways important for drug reward and addiction	[46, 100, 101]
PPP1R14D	shorter isoform, GBPI-1, CPI17like		PHI family	Inhibits PPP1 when phosphorylated (Brain/Stomach) - activated by PKC and inactivated by PKA	[46, 102]
PPP1R15A	GADD34	PPP1CA PPP1CB PPP1CC	GADD34 and related	Protein synthesis, regulation of calreticulin exposure, TGFbeta signaling	[57, 103, 104];
PPP1R15B	CReP	PPP1CA	GADD34 and related	Protein synthesis	[105 , 106]
PPP1R16A	MYPT3	PPP1CB	MYPT family	Myosin/actin targeting, translocation of nuclear receptors	[107 , 108]
PPP1R16B	TIMAP, ANKRD4			Regulation of pulmonary endothelial barrier	[48]

Phosphoprotein phosphatase 1 binding motif

PIPs are structurally unrelated, but most of them share a short, degenerate RVXF-type docking motif that binds to a hydrophobic groove located on a surface behind the PPP1C active site [109]. Frequently, this motif is flanked N-terminally by four or five basic residues and C-terminally by four or five acidic residues. However the binding of this motif does not change the PPP1 conformation and functions only to anchor the PIPs to PPP1. This binding is essential, as it brings PPP1 into close proximity with its PIPs and promotes secondary interactions that will contribute to PPP1 isoform selection and regulates the activity and substrate specificity of the holoenzyme [39]. Presently, the Hendrickx pattern of the RVxF motif is the one in which both the specificity and the sensitivity are relatively high compared with the Wakula and Meiselback previous definitions [109-111]

Other PPP1 binding motifs have also been described. The SILK motif, present in PPP1R2 and other regulators, occurs N-terminal to the RVxF motif and binds also to a PPP1 hydrophobic groove, different from RVxF, which faces opposite to the catalytic site. This motif doesn't change the conformation of PPP1 and serves for anchoring [65, 112]. A motif present in the MYPT family of proteins (PPP1R12 subfamily) is the myosin phosphatase N-terminal helicoidal element or MyPhoNE. This motif is also present N-terminally of RVxF, binds to a shallow hydrophobic cleft of PPP1 and contributes to substrate selection [55]. Other motifs already described, but not well studied, include the anti-apoptotic family members of Bcl2 motif juxtaposed the RVxF motif that confers an apoptotic signature [113, 114] and the C-terminal RARA motif, which is also required, besides RVxF, for the binding of PPP1R15A to PPP1C [57] (Table I.3).

These PPP1 signatures in the different PIPs permit the identification and characterization of new PIPs that interact with PPP1, and are a key to understanding the myriad of functions of PPP1 and its subcellular role.

Table I. 3: PPP1 binding motifs. *Information about each PPP1 binding motif known to date is presented, as well as, the specific pattern. Examples of PIPs for each motif are also shown. aa, amino acid.*

	Motif	PIP	Reference
RVxF motif	[RK]-X(0,1)-[VI]-{P}-[FW] X(0,1) is any aa, present or absent {P} represents any aa except P	PPP1R8 PPP1R10	[110]
	[HKR]-[ACHKMNQRSTV]-V-[CHKNQRST]-[FW]	PPP1R10	[111]
	$[K_{54}R_{34}L_4]$-$[K_{28}R_{26}S_{10}T_9A_8M_3V_3H_4N_3Q_3]$-$[V_{94}I_6]$-{FIMYDP}-$[F_{83}W_{17}]$ {FIMYDP} represents any aa except F/I/M/Y/D/P numbers show the respective percentage of each aa calculated from all the known PIPs	PPP1R8 PPP1R10	[109]
SILK motif	K-[GS]-I-L-[RK] -X(7-107)-[RK]-X(0,1)-[VI]-{P}-[FW] X(7-107) means that SILK motif needs to be from 7 to 107 aa of distance from the RVxF motif; X(0,1) is any aa, present or absent; {P} represents any aa except P	NOM-1 WBP11	[51] [115, 116]
MyPhoNE motif	R-X-X-Q-[VIL]-[KR]-X-[YW] X is any aa	PPP1R12A PPP1R12B	[55]
PPP1R2 degenerate motif	R-[KR]-X-H-Y X is any aa	PPP1R2	[112]
	K-S-Q-K-W	PPP1R2	[45]
Other motifs	[RK]-X(0,1)-[VI]-X-F-X-X-[RK]-X-[RK] X(0,1) is any aa, present or absent X is any aa	Bcl-2 Bad Bcl-X Bcl-W	[114] [42]
	R-A-R-A	PPP1R15A	[57]
	R-N-Y-F	iASPP	[117]

Protein phosphatase 1 role in testis and sperm

Several PPP family members have shown to be expressed in cells from testis and/or spermatozoa, suggesting an important function in spermatozoa formation (Table I.3). All PPP1C isoforms (A/B/C1/C2), are expressed in mammalian testis [118], whereas to date only PPP1CA and PPP1CC2 were shown to be present in spermatozoa, with the latter being more abundant. *Ppp1CC* gene null male mice were shown to be infertile due to impaired spermatogenesis, leading to the absence of epididymal spermatozoa [119]. Although PPP1CA expression was increased and its localization altered, it could not substitute for PPP1CC, further suggesting a specific role for the sperm-specific PPP1CC2 in sperm differentiation and morphogenesis [118].

Table I. 4: *Testis and sperm expression of the main PPP family members.*

Protein Phosphatase	Testis	Sperm
PPP1CA	Yes [120]	Yes [121] Fardilha, unpublished data)
PPP1CB	Yes [120]	No
PPP1CC1	Yes [120]	No
PPP1CC2	Yes, highly abundant [15, 16, 120, 122]	Yes [14, 15, 123]
PPP2	Yes [124]	Yes [123, 125]
PPP3	Yes [126, 127]	Yes [128]
PPP4	Yes, highly abundant [24, 124]	ND
PPP5	Yes [26]	ND
PPP6	Yes, highly abundant [27, 124]	ND
PPP7	Yes, highly abundant [28]	ND

Legend: ND, Not determined.

Testes contain hundreds of tightly packed seminiferous tubules, each one composed of several layers of peritubular myoid cells. The peritubular myoid cells are responsible for the irregular contractions of the seminiferous tubules which propel fluid secreted by the supporting Sertoli cells, together with testicular spermatozoa into the lumen and through the tubular network [129].

Spermatogenesis takes place in the seminiferous tubules and is the process by which spermatozoa are produced. It can be divided into three major phases: (1) proliferation and differentiation of spermatogonia, (2) meiosis and (3) spermiogenesis (transformation of round spermatids into spermatozoa) and spermiation (release of spermatozoa from the supporting Sertoli cells) [130, 131]. Interstitial Leydig cells are responsible for testosterone production, which is essential for maintenance of spermatogenesis.

In testis, PPP1CC2 is localized in the cytoplasm of secondary spermatocytes and round spermatids, as well as elongating spermatids and testicular and epididymal spermatozoa, while PPP1CC1 expression is observed mainly in Leydig cells but also weakly in all stages of spermatogenesis in both the cytoplasm and nuclei and PPP1CA in spermatogonia, peritubular cells, pachytene spermatocytes and interstitial Leydig cells [118].

After spermiation, spermatozoa exit the seminiferous tubules through a system of genital ducts and enter the first part of the epididymis. The epididymides are divided, morphologically and functionally into *caput*, *corpus* and *cauda*. The sperm mature during their passage through the *caput* and *corpus*, whereas the *cauda* functions predominantly for storage. Epididymal sperm maturation involves a series of modifications: (1) remodeling of the sperm plasma membrane, (2) changes in composition and cellular localization of the proteins, (3) alteration of the glycoproteins content and (4) changes in pH and in the levels of Ca^{2+} and cAMP [125, 132, 133]. Sperm maturation involves the interaction with proteins that are synthesized and secreted in a region-dependent manner from the epididymal epithelium. These maturation steps allow spermatozoa to acquire the progressive motility.

Given that spermatozoa are terminally differentiated cells, essentially devoid of transcriptional and translational activity, they are an ideal model system to study the regulation of PPP1 in relation to motility and metabolism. Of note, several lines of evidence have demonstrated that PPs are direct players in the acquisition of sperm movement. Previous results show that PPP1CC2 activity is correlated with motility since phosphatase inhibitors were able to induce motility in completely immotile bovine caput

epididymal sperm and to stimulate the kinetic activity of mature caudal sperm. Intriguingly, these effects were completely independent of calcium and cAMP meaning that PPP1R1 is not involved [14, 123].

Several studies have demonstrated that PPP1CC2 is important in regulating sperm motility [14, 76, 134, 135]. In sperm, PPP1CC2 is present along the entire flagellum including the midpiece, consistent with a role in sperm motility, but it is also found in the posterior and equatorial regions of the head, suggesting a role in the acrosome reaction [135]. In *Chlamydomonas* PPP1C is primarily, but not exclusively, anchored in the central pair apparatus, associated with the C1 microtubule and at less extent to the outer doublet microtubules, suggesting that PPP1 can control both dynein arms and thereby flagellar motility [136].

The PPP1 driven endogenous regulation of protein phosphorylation and sperm motility, could represent an important mechanism for physiological regulation of a cell that encounters dramatically different environments, as it journeys through the seminiferous tubules and the female reproductive tract. Previous results [14, 123] provide strong support for a novel unifying hypothesis, based on the observation that PPP1 is present in sperm and that pharmacological modulation of its activity profoundly affects sperm motility. In other cell types, PPP1 has been implicated in the control of diverse processes such as cellular metabolism, muscle contraction, mitosis, neurotransmitter release, etc. [1, 3, 137]. Regulation of these processes involves complex intracellular pathways, initiated by activation of distinct receptors and second messenger systems [138]. However, the precise role played by STPPs and their regulation, have only recently started to be elucidated. The available data demonstrate their highly regulatable nature, contrary to previous opinions. One particularly interesting mechanism for controlling PPP1 activity involves its inhibition by heat-stable PP inhibitors PPP1R1, and PPP1R2, phosphoproteins whose state of phosphorylation controls their inhibitory activity. PPP1R1 is phosphorylated by cAMP-dependent protein kinase and dephosphorylated by calcium/calmodulin-dependent PPP3 [139]. Thus, PPP1 is also involved in the cross-talk between the intracellular messengers, calcium and cAMP [138].

Furthermore, other maturation steps are acquired by spermatozoa in the female tract and prepare them for the fertilization process: capacitation, hyperactivation and acrosome reaction [140, 141]. It has already been shown the involvement STPPs in hyperactivated motility cAMP-dependent phosphorylation [142]. Also, Visconti and co-workers have recently shown that inhibition of STPPs induces capacitation-associated signaling by tyrosine phosphorylation [143].

Testis and sperm PPP1CC/PIP complexes

The diversity of PPP1 function is achieved by its capacity to form functionally distinct multimeric complexes. Significantly, some testis/sperm-specific PIPs have been identified. For example, the spermatogenic zip protein 1 (Szp1), a member of the basic helix-loop-helix family of transcription factors, which binds to PPP1CC2 in mouse testis [144]. Overexpression of Szp1 and loss of PPP1CC in the testis show similar phenotypes, such as spermatogenic arrest and germ cell apoptosis [145]. Another example is endophilin B1t, this testis enriched isoform of endophilin B1a was shown to bind PPP1CC2 but did not interact with a mutant form of PPP1CC2, lacking the specific C-terminus, nor with PPP1CA [146]. Moreover, the somatic isoform did not interact with any of the PPP1C isoforms and the characteristic punctuate expression pattern of endophilin in testis, was absent in PPP1CC null mice. Also, endophilin B1t was able to inhibit a recombinant PPP1CC2 activity [146]. Other proteins have also been implicated in the regulation of PPP1CC2 in testis/sperm either by protecting (14-3-3/YWHA) or inhibiting (PPP1R2-like, PPP1R7 and PPP1R11) PPP1CC2 activity during sperm maturation correlating with increased spermatozoa motility [15, 125, 135, 144, 147, 148].

PPP1R2 is capable of inhibiting the catalytic subunit of PPP1 leading to the production of a stable PPP1C-PPP1R2 complex. GSK-3 phosphorylates PPP1R2 in the complex, relieving the inhibition and producing active PPP1 (Fig I.3). This biochemical pathway is likely to be operative in mammalian sperm since preliminary studies have identified a PPP1R2-like activity and also the presence of GSK-3 in mammalian sperm [14, 123].

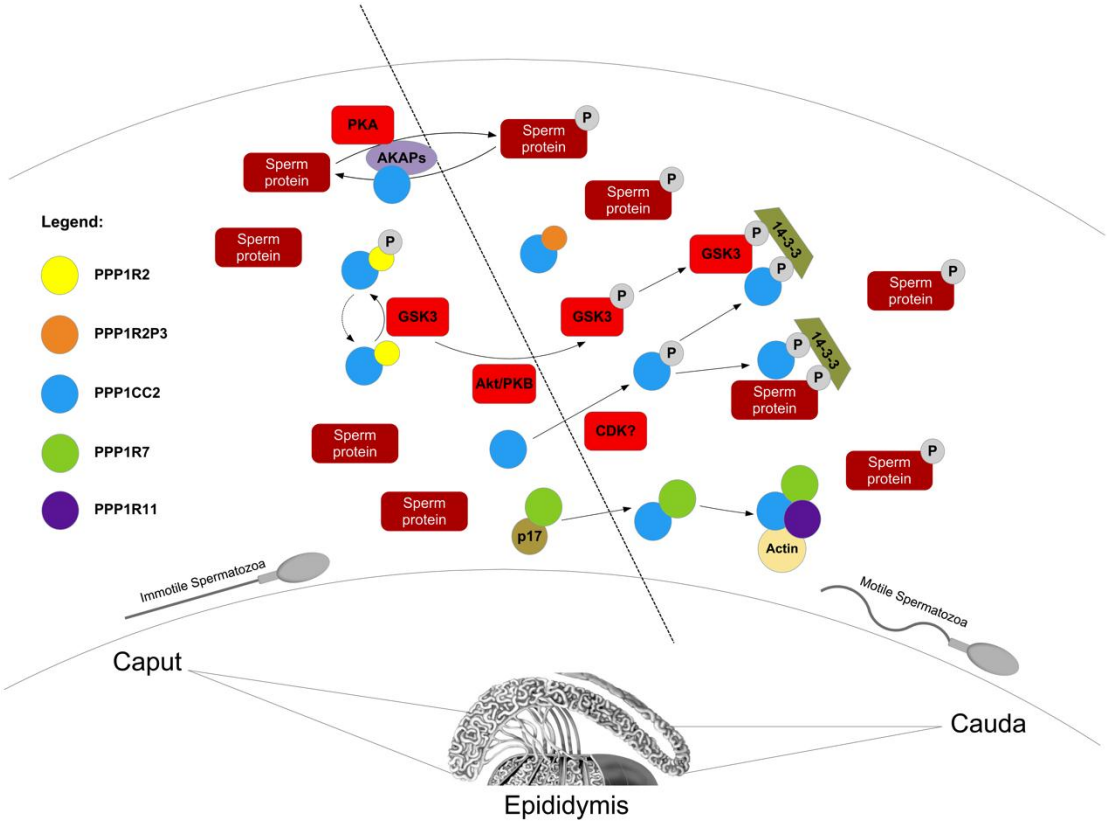


Figure I. 3: Illustrative scheme of the proteins involved in the acquisition of sperm motility based on PPP1CC2 regulation*. AKAPs anchoring seem to modulate PPP1CC2 and PKA activities inducing sperm motility and phosphorylation of several proteins. PPP1CC2 and GSK3 phosphorylation by AKT and an unknown , respectively, lead to their inactivation and their possible binding to the bridging molecule 14-3-3, in epididymal cauda. PPP1CC2 is maintained active in caput due to the phosphorylation of a PPP1R2-like protein by GSK3. In cauda, GSK3 activity lowers and therefore a PPP1R2-like protein inhibits PPP1CC2 leading to sperm motility. PPP1R7 binds to PPP1CC2 in cauda sperm also inhibiting the phosphatase. In caput, PPP1R7 is prevented to inhibit PPP1CC2 due to a p17 unknown protein. Also, a multimeric complex has been identified composed by PPP1CC2, PPP1R7, actin and PPP1R11, where PPP1CC2 was inactive. (adapted from [149]).*

PPP1CC/PPP1R11 complex

PPP1R11/I3 is a potent heat-stable PPP1 inhibitor [90, 150] and is a human homologue of the mouse t-complex expressed protein 5 (Tctex5), being genetically linked to the male sterility phenotypes of impaired sperm tail development and poor sperm motility in t complex mice [151-153]. Tctex5 gene synergistically with Tsga2 gene is a candidate for the “curlicue” and for the “stop” sperm phenotypes. The first phenotype is a chronic, negative bend of both the flagellar middle-piece and principal piece of spermatozoon, while the second phenotype prevents sperm from t-haplotype homozygotes from penetrating zona-free egg [154-156]. Tctex5 was shown to be present

in sperm protein lysates localizing to nuclei of pachytene spermatocytes, round spermatocytes, cytoplasm of Sertoli cells, in testis; cilia, secretion bodies and nuclei of epithelial cells and interstitium smooth muscle cells in the epididymis. In epididymal mouse spermatozoa Tctex5 is present in the head and principal piece of the tail [154] (Fig I.4). These are also the locations where PPP1CC2 is expressed [147].

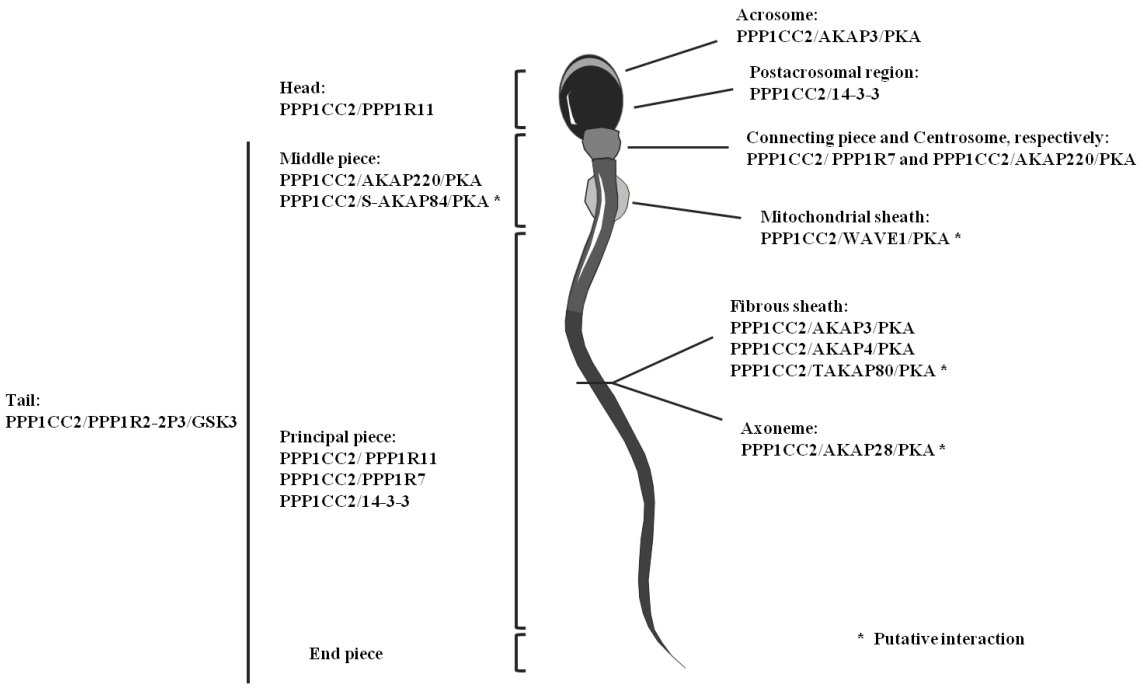


Figure I. 4: *Schematic representation of the subcellular localization of PPP1-PIP complexes in spermatozoon (adapted from [149]).*

PPP1CC/PPP1R7 complex

A yeast sds22 homologue, PPP1R7, was identified in sperm [76], and inhibits the PPP1 catalytic subunit in rat liver nuclei [157]. Consistently, a PPP1R7 homologue was also identified in rat testis in association with PPP1CC2 [158]. The expression pattern of rat PPP1R7 matches that of PPP1CC2, suggesting that its involvement in spermatogenesis is correlated with the control of PPP1CC2 activity. Furthermore PPP1R7 was also identified in motile caudal spermatozoa as a regulator of PPP1CC2 catalytic activity [135]. Additionally, sds22 has consensus sites for phosphorylation by GSK3, PKA and CDK2 (calmodulin dependent kinase II), all present in sperm.

In male germ cells PPP1CC2, PPP1R11, PPP1R7 and actin form a multimeric complex in which PPP1CC2 is inactive [159]. The stability of the complex depended on

functional PPP1 interaction sites in PPP1R7 and PPP1R11, indicating that PPP1 mediates the interaction between these two proteins, forming a catalytically inactive complex in the germ cell [160]. The function of this complex in sperm motility, if any, still needs to be elucidated (Fig I.3).

PPP1CC/14-3-3 complex

PPP1CC2, as well as other PPP1 isoforms, have at the C-terminus a consensus TPPR amino acid sequence containing a threonine residue (T311) that can be phosphorylated by CDK2, reducing its activity [161-163]. The proportion of phosphorylated PPP1CC2 in caudal sperm is higher than in caput epididymal sperm and is localized to the posterior region of the sperm head, the equatorial region, implicated in sperm-egg binding, and in the principal piece of the sperm tail [164]. Interestingly, CDK2 knockout mice are viable but male and female are sterile [165, 166]. Vijayaraghavan *et al.* proposed that regulation of PPP1CC2 activity by CDK2 phosphorylation might be a mechanism for developing sperm motility. This might be achieved through binding of PPP1CC2 to the bridging molecule 14-3-3. In sperm 14-3-3 binds PPP1CC2 [125] (Fig I.3). The 14-3-3 protein is highly conserved among eukaryotic cells and acts as an adaptor protein in cellular signaling and metabolism. More than 100 binding partners have been identified, using affinity chromatography coupled with proteomic analysis [167-170]. 14-3-3 and its binding partners are regulators of protein-protein interactions during spermatogenesis [171]. It is consistent that 14-3-3 appears to regulate diverse cellular events such as cell cycle, apoptosis, protein trafficking, cytoskeleton rearrangements and metabolism [172]. The study of 14-3-3 interactome in bull sperm identified many proteins involved in different cellular events, from acrosome reaction to metabolism [173]. In particular, GSK3 was found to bind 14-3-3 [173, 174] (Fig I.3). Sperm 14-3-3 protein is present in the post-acrosomal region of the head and the principal piece, similar to PPP1CC2 [125] (Fig I.4). As already stated, changes in tyrosine and serine phosphorylation of GSK3 occur in parallel with motility stimulation in sperm [175]. The exact functions of 14-3-3 through binding to sperm phosphoproteins are still subject of extensive research.

PPP1CC/AKAPs complexes

The cyclic AMP (cAMP)-dependent protein kinase (PKA) is a ubiquitous, multifunctional enzyme involved in the regulation of several cellular events. PKA

holoenzyme consists of four subunits, two catalytic and two regulatory (RI and RII). Subcellular targeting to the vicinity of preferred substrates is a means of restricting the specificity of these enzymes [176, 177]. Compartmentalization of PKA is mediated through association of its regulatory subunits with A-kinase anchoring proteins (AKAPs) [176]. To date, over 40 AKAPs have been identified, and in testis/sperm there are three AKAPs that have been directly related to PPP1CC2 (AKAP220, AKPA3 and AKAP4) and many more show a similar localization.

AKAP220/AKAP11 binds PKA and PPP1, being a competitive inhibitor of PPP1 [178]. In testis, AKAP220 associates with PKA, where it may target the kinase to peroxisomes [179]. *AKAP220* mRNA is expressed at high levels in human testis and in isolated human pachytene spermatocytes and round spermatids [180]. AKAP220 is present in human male germ cells and mature sperm and like RIIα, is located in the midpiece and is probably associated with cytoskeletal structures [180] (Fig I.4). The midpiece associated AKAP220 could serve to anchor PKA and/or PPP1CC2, directly regulating the contractile machinery in the sperm axoneme. Furthermore, it has been shown that disruption of RII interaction with AKAPs, by membrane-permeable peptides, causes the arrest of sperm motility [181].

AKAP4/AKAP82 cDNA was first isolated from a mouse testis cDNA-expression library [182, 183]. Mouse Akap4 expression was only detected in testis and it was determined that transcription is initiated at 20-22 days after birth and the mRNA is present in spermatids but not in pachytene spermatocytes [182, 184]. In RNA from testes extracts of hamster, guinea pig, rabbit, ram, and human, transcripts that hybridized to the mouse *Akap4* cDNA were found [184]. AKAP4 protein was synthesized as a precursor present throughout the principal piece in testicular sperm. A small peptide corresponding to the precursor fragment and a higher molecular weight protein, possibly a phosphorylated form of the precursor, are present in epididymal sperm in low amounts. In mouse sperm, AKAP4 was detected throughout both the longitudinal columns and the semi-circumferential ribs of the fibrous sheath [185] (Fig I.4). All these findings correlate with a restricted temporal and spatial expression of the AKAP4 protein, present only in spermatogenetic cells and the predominant protein in the fibrous sheath of the sperm flagellum. Targeted disruption of the *Akap4* gene causes absence of sperm motility together with a complete lack of fibrous sheath on the principal piece of mature mice sperm, causing male mice to be infertile [186]. In spermatozoa, *Akap4* gene knockout mice that lack flagellar movement, exhibit a significant change in the activity and

phosphorylation of PPP1CC2 [147]. This suggests the involvement of AKAP4 in the regulation of PPP1CC2 activity in the principal piece of mouse spermatozoa.

AKAP3/AKAP110, also referred to as FSP95, undergoes tyrosine phosphorylation during *in vitro* capacitation of human sperm. Northern blot analysis of RNA from 50 human tissues determined that the transcripts are found only in testis, more specifically in round spermatids. Using a rat antiserum to recombinant protein, FSP95 was found only in the fibrous sheath and localized to the circumferential ribs of human sperm [187] (Fig I.4). Brown *et al.* reported that AKAP4 anchors AKAP3 and two novel spermatogenetic cells specific proteins, Fibrous sheath interacting proteins 1 and 2 (FSIP1 and FSIP2) [188].

Given that many AKAPs have been shown to be present in germ cells and localized to compartments related to motility where PPP1CC2 is also present they might be putatively involved in motility acquisition.

S-AKAP84 localizes to the midpiece of mouse elongating spermatids and co-localizes with mitochondria [189] (Fig I.4). Its splice variants D-AKAP-1 [116] and AKAP121 [190] were also detected in mouse testis and GC2 germ cells, respectively. TAKAP-80, was isolated by screening a rat testis expression library [191]. The corresponding protein was detected in rat testis and in purified fibrous sheath fractions from rat epididymal sperm (Fig I.4). The levels of protein were higher in mature compared to immature rat testis, correlating with the mRNA levels [191]. AKAP28, was detected in human testis, and is highly enriched in the axoneme structure (Fig I.4). It is likely to play a role in signaling mechanisms necessary for ciliar beating frequency [192]. WAVE1 localization in spermatocytes and round spermatids coincided with Golgi apparatus, whereas in elongated spermatids and testicular sperm localized to the mitochondrial sheath [193] (Fig I.4).

Recent data showed that inhibition of protein phosphatases with calyculin A resulted in an enhancement of the phosphorylated state at the activation loop of the PKA catalytic subunit in the mouse sperm principal and midpieces [194]. Also, PKA RII and PPP1CC2 are co-localized in the principal and midpieces. PPP1 and PPP3 suppress full activation of PKA, as well as enhancement of the phosphorylated states of other flagellar proteins, in order to prevent precocious changes of flagellar movement from the progressive type to hyperactivation [194].

Together, these findings suggest that the AKAP/PKA/PPP1 complex is really important for regulation of sperm motility (Fig I.3).

Of marked interest is the fact that each PPP1/PIP complex has a specific sperm subcellular location (Fig I.4). Clearly PPP1CC2/PIP complexes are essential regulatory

components in the signaling transduction cascades involved in sperm motility acquisition during epididymal transit. Defects in any component of these signaling cascades will give rise to pathological anomalies, leading to male infertility. Therefore, the study of new PPP1CC2/PIP complexes in testis and sperm are extremely important in a physiological and pathological point of view.

Objectives

The aim of this work was to study new PPP1CC2/PIP complexes in testis and sperm and to characterize their important physiological role.

Therefore, this thesis addresses a previously well known somatic PIP in testis/sperm, named PPP1R2 and two new PPP1CC2 testis/sperm specific PIPs that were found in a yeast two-hybrid technique using a human testis cDNA library.

Chapter II comprises two sections in paper format and one section in thesis format. In the first section (paper format), the objective was to study the presence of PPP1R2 and PPP1R2P3 (PPP1R2 pseudogene 3) proteins in human sperm as well as their specific sperm localization. PPP1R2P3 is a new PIP, also designated as I2-L (Inhibitor 2-Like/PPP1R2P3, NCBI Id: NM_206858) [149, 195]. Further, phosphorylation studies in PPP1R2P3 *in vitro* were also pursued. In the second section (paper format), the objective was to study the PPP1R2 pseudogenes using bioinformatics to understand their evolution and significance in mammals. In the third section (thesis format), transgenic mice expressing human PPP1R2 or PPP1R2P3 were produced to understand the role of these proteins in spermatogenesis and sperm maturation.

Chapter III comprises two sections. The first section (paper format) addresses the identification and characterization of a novel dynein light chain, Tctex1d4, in testis and sperm as a new binding partner of PPP1CC. In the second section (paper format) bioinformatics and molecular biology were integrated to understand Tctex1d4 evolution and the RVxF modification that occurred in Pika.

By unraveling these three new proteins that interact with PPP1CC and are present in testis and sperm the thesis aim to provide new insights about PPP1CC2 function in these tissues.

References

1. Cohen, P.T., *Protein phosphatase 1--targeted in many directions.* J Cell Sci, 2002. **115**(Pt 2): p. 241-56.
2. Fardilha, M., et al., *The Physiological Relevance of Protein Phosphatase 1 and its Interacting Proteins to Health and Disease.* Current Medicinal Chemistry, 2010. **17**(33).
3. Ceulemans, H. and M. Bollen, *Functional diversity of protein phosphatase-1, a cellular economizer and reset button.* Physiol Rev, 2004. **84**(1): p. 1-39.
4. Philip, C., *The regulation of protein function by multisite phosphorylation – a 25 year update.* Trends in biochemical sciences, 2000. **25**(12): p. 596-601.
5. Barford, D., A.K. Das, and M.-P. Egloff, *The Structure and Mechanism of Protein Phosphatases: Insights into Catalysis and Regulation.* Annual Review of Biophysics and Biomolecular Structure, 1998. **27**(1): p. 133-164.
6. Yeo, M., et al., *A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5.* J Biol Chem, 2003. **278**(28): p. 26078-85.
7. Gallego, M. and D.M. Virshup, *Protein serine/threonine phosphatases: life, death, and sleeping.* Curr Opin Cell Biol, 2005. **17**(2): p. 197-202.
8. Cohen, P.T., *Novel protein serine/threonine phosphatases: variety is the spice of life.* Trends Biochem Sci, 1997. **22**(7): p. 245-51.
9. Honkanen, R.E. and T. Golden, *Regulators of serine/threonine protein phosphatases at the dawn of a clinical era?* Curr Med Chem, 2002. **9**(22): p. 2055-75.
10. Sood, P.P. and M.A. Majid, *Qualitative and quantitative changes of acid and alkaline phosphatases in the testis and epididymis of mice in relation to single high dose of alpha-chlorohydrin.* Acta Eur Fertil, 1987. **18**(1): p. 33-8.
11. Fardilha, M., *Characterization of PP1 Interactome from human testis*, 2004, Universidade de Aveiro.
12. Plowman, G.D., et al., *The protein kinases of Caenorhabditis elegans: A model for signal transduction in multicellular organisms.* Proceedings of the National Academy of Sciences, 1999. **96**(24): p. 13603-13610.
13. David, B., *Molecular mechanisms of theprotein serine/threonine phosphatases.* Trends in biochemical sciences, 1996. **21**(11): p. 407-412.
14. Smith, G.D., et al., *Primate sperm contain protein phosphatase 1, a biochemical mediator of motility.* Biol Reprod, 1996. **54**(3): p. 719-727.
15. Shima, H., et al., *Protein phosphatase 1 gamma 2 is associated with nuclei of meiotic cells in rat testis.* Biochem Biophys Res Commun, 1993. **194**(2): p. 930-7.
16. Kitagawa, Y., et al., *Protein phosphatases possibly involved in rat spermatogenesis.* Biochem Biophys Res Commun, 1990. **171**(1): p. 230-5.
17. Fardilha, M., Da Cruz e Silva, O.A.B., Da Cruz e Silva, E.F, *A importância do mecanismo de "splicing" alternativo para a identificação de novos alvos terapêuticos.* Acta Urológica, 2008. **25**: p. 1: 39-47.
18. da Cruz e Silva, E.F., et al., *Differential expression of protein phosphatase 1 isoforms in mammalian brain.* J Neurosci, 1995. **15**(5 Pt 1): p. 3375-89.
19. Takizawa, N., et al., *Tissue distribution of isoforms of type-1 protein phosphatase PP1 in mouse tissues and its diabetic alterations.* J Biochem, 1994. **116**(2): p. 411-5.
20. Strack, S., et al., *Differential cellular and subcellular localization of protein phosphatase 1 isoforms in brain.* The Journal of Comparative Neurology, 1999. **413**(3): p. 373-384.
21. Ouimet, C.C., E.F. da Cruz e Silva, and P. Greengard, *The alpha and gamma 1 isoforms of protein phosphatase 1 are highly and specifically concentrated in*

dendritic spines. Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(8): p. 3396-3400.
22. Khew-Goodall, Y. and B.A. Hemmings, *Tissue-specific expression of mRNAs encoding alpha- and beta-catalytic subunits of protein phosphatase 2A.* FEBS Lett, 1988. **238**(2): p. 265-8.
23. Wallace, R.W., E.A. Tallant, and W.Y. Cheung, *High levels of a heat-labile calmodulin-binding protein (CaM-BP80) in bovine neostriatum.* Biochemistry, 1980. **19**(9): p. 1831-7.
24. Hu, M.C., et al., *Genomic structure of the mouse PP4 gene: a developmentally regulated protein phosphatase.* Gene, 2001. **278**(1-2): p. 89-99.
25. Chinkers, M., *Targeting of a distinctive protein-serine phosphatase to the protein kinase-like domain of the atrial natriuretic peptide receptor.* Proc Natl Acad Sci U S A, 1994. **91**(23): p. 11075-9.
26. Becker, W., et al., *Molecular cloning of a protein serine/threonine phosphatase containing a putative regulatory tetratricopeptide repeat domain.* J Biol Chem, 1994. **269**(36): p. 22586-92.
27. Bastians, H. and H. Ponstingl, *The novel human protein serine/threonine phosphatase 6 is a functional homologue of budding yeast Sit4p and fission yeast ppe1, which are involved in cell cycle regulation.* J Cell Sci, 1996. **109 (Pt 12)**: p. 2865-74.
28. Andreeva, A.V. and M.A. Kutuzov, *PPEF/PP7 protein Ser/Thr phosphatases.* Cell Mol Life Sci, 2009. **66**(19): p. 3103-10.
29. Andreassen, P.R., et al., *Differential Subcellular Localization of Protein Phosphatase-1 α, γ1, and δ Isoforms during Both Interphase and Mitosis in Mammalian Cells.* The Journal of Cell Biology, 1998. **141**(5): p. 1207-1215.
30. Anderson, K., V. Nisenblat, and R. Norman, *Lifestyle factors in people seeking infertility treatment – A review.* Australian and New Zealand Journal of Obstetrics and Gynaecology, 2010. **50**(1): p. 8-20.
31. Trinkle-Mulcahy, L., J.E. Sleeman, and A.I. Lamond, *Dynamic targeting of protein phosphatase 1 within the nuclei of living mammalian cells.* J Cell Sci, 2001. **114**(23): p. 4219-4228.
32. Trinkle-Mulcahy, L., et al., *Time-lapse Imaging Reveals Dynamic Relocalization of PP1gamma throughout the Mammalian Cell Cycle.* Mol. Biol. Cell, 2003. **14**(1): p. 107-117.
33. Ceulemans, H., W. Stalmans, and M. Bollen, *Regulator-driven functional diversification of protein phosphatase-1 in eukaryotic evolution.* BioEssays, 2002. **24**(4): p. 371-381.
34. Virshup, D.M. and S. Shenolikar, *From promiscuity to precision: protein phosphatases get a makeover.* Mol Cell, 2009. **33**(5): p. 537-45.
35. Terry-Lorenzo, R.T., et al., *Neurabin/protein phosphatase-1 complex regulates dendritic spine morphogenesis and maturation.* Mol Biol Cell, 2005. **16**(5): p. 2349-62.
36. Allen, P.B., C.C. Ouimet, and P. Greengard, *Spinophilin, a novel protein phosphatase 1 binding protein localized to dendritic spines.* Proc Natl Acad Sci U S A, 1997. **94**(18): p. 9956-61.
37. Terry-Lorenzo, R.T., et al., *The neuronal actin-binding proteins, neurabin I and neurabin II, recruit specific isoforms of protein phosphatase-1 catalytic subunits.* J Biol Chem, 2002. **277**(31): p. 27716-24.
38. Terry-Lorenzo, R.T., et al., *Neurabins recruit protein phosphatase-1 and inhibitor-2 to the actin cytoskeleton.* J Biol Chem, 2002. **277**(48): p. 46535-43.
39. Bollen, M., et al., *The extended PP1 toolkit: designed to create specificity.* Trends in biochemical sciences, 2010. **35**(8): p. 450-458.

40. Liu, Y., et al., *Regulation of BRCA1 Phosphorylation by Interaction with Protein Phosphatase 1α.* Cancer Research, 2002. **62**(22): p. 6357-6361.
41. Margolis, S.S., et al., *PP1 control of M phase entry exerted through 14-3-3-regulated Cdc25 dephosphorylation.* EMBO J, 2003. **22**(21): p. 5734-45.
42. Ayllon, V., et al., *The anti-apoptotic molecules Bcl-xL and Bcl-w target protein phosphatase 1alpha to Bad.* Eur J Immunol, 2002. **32**(7): p. 1847-55.
43. Dessauge, F., et al., *Identification of PP1alpha as a caspase-9 regulator in IL-2 deprivation-induced apoptosis.* J Immunol, 2006. **177**(4): p. 2441-51.
44. Helps, N.R., et al., *NIMA-related kinase 2 (Nek2), a cell-cycle-regulated protein kinase localized to centrosomes, is complexed to protein phosphatase 1.* Biochem J, 2000. **349**(Pt 2): p. 509-18.
45. Hurley, T.D., et al., *Structural basis for regulation of protein phosphatase 1 by inhibitor-2.* J Biol Chem, 2007. **282**(39): p. 28874-83.
46. Eto, M., *Regulation of Cellular Protein Phosphatase-1 (PP1) by Phosphorylation of the CPI-17 Family, C-kinase-activated PP1 Inhibitors.* Journal of Biological Chemistry, 2009. **284**(51): p. 35273-35277.
47. Wu, W., et al., *Alternative splicing controls nuclear translocation of the cell cycle-regulated Nek2 kinase.* J Biol Chem, 2007. **282**(36): p. 26431-40.
48. Csortos, C., et al., *TIMAP is a positive regulator of pulmonary endothelial barrier function.* Am J Physiol Lung Cell Mol Physiol, 2008. **295**(3): p. L440-450.
49. Allen, P.B., et al., *Isolation and Characterization of PNUTS, a Putative Protein Phosphatase 1 Nuclear Targeting Subunit.* Journal of Biological Chemistry, 1998. **273**(7): p. 4089-4095.
50. Steen, R.L., et al., *AKAP149 is a novel PP1 specifier required to maintain nuclear envelope integrity in G1 phase.* J Cell Sci, 2003. **116**(Pt 11): p. 2237-46.
51. Gunawardena, S.R., et al., *NOM1 Targets Protein Phosphatase I to the Nucleolus.* Journal of Biological Chemistry, 2008. **283**(1): p. 398-404.
52. Vagnarelli, P., et al., *Condensin and Repo-Man-PP1 co-operate in the regulation of chromosome architecture during mitosis.* Nat Cell Biol, 2006. **8**(10): p. 1133-42.
53. Vijayan, K.V., et al., *Protein phosphatase 1 associates with the integrin alphaIIb subunit and regulates signaling.* J Biol Chem, 2004. **279**(32): p. 33039-42.
54. MacMillan, L.B., et al., *Brain Actin-associated Protein Phosphatase 1 Holoenzymes Containing Spinophilin, Neurabin, and Selected Catalytic Subunit Isoforms.* Journal of Biological Chemistry, 1999. **274**(50): p. 35845-35854.
55. Terrak, M., et al., *Structural basis of protein phosphatase 1 regulation.* Nature, 2004. **429**(6993): p. 780-4.
56. Johnson, D.F., et al., *Identification of Protein-Phosphatase-1-Binding Domains on the Glycogen and Myofibrillar Targetting Subunits.* European Journal of Biochemistry, 1996. **239**(2): p. 317-325.
57. Brush, M.H., D.C. Weiser, and S. Shenolikar, *Growth Arrest and DNA Damage-Inducible Protein GADD34 Targets Protein Phosphatase 1{alpha} to the Endoplasmic Reticulum and Promotes Dephosphorylation of the {alpha} Subunit of Eukaryotic Translation Initiation Factor 2.* Mol. Cell. Biol., 2003. **23**(4): p. 1292-1303.
58. Theurillat, J.P., et al., *URI is an oncogene amplified in ovarian cancer cells and is required for their survival.* Cancer Cell, 2011. **19**(3): p. 317-32.
59. Brichese, L. and A. Valette, *PP1 phosphatase is involved in Bcl-2 dephosphorylation after prolonged mitotic arrest induced by paclitaxel.* Biochem Biophys Res Commun, 2002. **294**(2): p. 504-8.
60. Rodriguez, P., et al., *Identification of a Novel Phosphorylation Site in Protein Phosphatase Inhibitor-1 as a Negative Regulator of Cardiac Function.* Journal of Biological Chemistry, 2006. **281**(50): p. 38599-38608.

61. Nicolaou, P., R.J. Hajjar, and E.G. Kranias, *Role of protein phosphatase-1 inhibitor-1 in cardiac physiology and pathophysiology.* Journal of Molecular and Cellular Cardiology, 2009. **47**(3): p. 365-371.
62. Svenningsson, P., et al., *DARPP-32: An Integrator of Neurotransmission.* Annual Review of Pharmacology and Toxicology, 2004. **44**(1): p. 269-296.
63. Svenningsson, P., A. Nairn, and P. Greengard, *DARPP-32 mediates the actions of multiple drugs of abuse.* The AAPS Journal, 2005. **7**(2): p. E353-E360.
64. Zeng, Q., et al., *Effect of IPP5, a novel inhibitor of PP1, on apoptosis and the underlying mechanisms involved.* Biotechnol Appl Biochem, 2009. **54**(4): p. 231-8.
65. Huang, H.-b., et al., *Characterization of the Inhibition of Protein Phosphatase-1 by DARPP-32 and Inhibitor-2.* Journal of Biological Chemistry, 1999. **274**(12): p. 7870-7878.
66. Li, M., D.L. Satinover, and D.L. Brautigan, *Phosphorylation and functions of inhibitor-2 family of proteins.* Biochemistry, 2007. **46**(9): p. 2380-9.
67. Suzuki, Y., et al., *Insulin Control of Glycogen Metabolism in Knockout Mice Lacking the Muscle-Specific Protein Phosphatase PP1G/RGL.* Mol. Cell. Biol., 2001. **21**(8): p. 2683-2694.
68. Toole, B.J. and P.T.W. Cohen, *The skeletal muscle-specific glycogen-targeted protein phosphatase 1 plays a major role in the regulation of glycogen metabolism by adrenaline in vivo.* Cellular Signalling, 2007. **19**(5): p. 1044-1055.
69. Montori-grau, M., et al., *Expression and glycogenic effect of glycogen-targeting protein phosphatase 1 regulatory subunit GL in cultured human muscle.* Biochem J, 2007. **405**(1): p. 107-113.
70. Brady, M.J., et al., *Role of Protein Targeting to Glycogen (PTG) in the Regulation of Protein Phosphatase-1 Activity.* Journal of Biological Chemistry, 1997. **272**(32): p. 20198-20204.
71. Printen, J.A., M.J. Brady, and A.R. Saltiel, *PTG, a protein phosphatase 1-binding protein with a role in glycogen metabolism.* Science, 1997. **275**(5305): p. 1475-8.
72. Browne, G.J., et al., *The level of the glycogen targetting regulatory subunit R5 of protein phosphatase 1 is decreased in the livers of insulin-dependent diabetic rats and starved rats.* Biochem. J., 2001. **360**(2): p. 449-459.
73. Munro, S., et al., *A novel glycogen-targeting subunit of protein phosphatase 1 that is regulated by insulin and shows differential tissue distribution in humans and rodents.* FEBS Journal, 2005. **272**(6): p. 1478-1489.
74. Mukherji, M., et al., *Genome-wide functional analysis of human cell-cycle regulators.* Proceedings of the National Academy of Sciences, 2006. **103**(40): p. 14819-14824.
75. Peggie, M.W., et al., *Essential functions of Sds22p in chromosome stability and nuclear localization of PP1.* Journal of Cell Science, 2002. **115**(1): p. 195-206.
76. Mishra, S., et al., *Binding and inactivation of the germ cell-specific protein phosphatase PP1gamma2 by sds22 during epididymal sperm maturation.* Biol Reprod, 2003. **69**(5): p. 1572-9.
77. Grusche, F., et al., *Sds22, a PP1 phosphatase regulatory subunit, regulates epithelial cell polarity and shape [Sds22 in epithelial morphology].* BMC Developmental Biology, 2009. **9**(1): p. 14.
78. Trinkle-Mulcahy, L., et al., *Nuclear organisation of NIPP1, a regulatory subunit of protein phosphatase 1 that associates with pre-mRNA splicing factors.* Journal of Cell Science, 1999. **112**(2): p. 157-168.
79. Tanuma, N., et al., *Nuclear Inhibitor of Protein Phosphatase-1 (NIPP1) Directs Protein Phosphatase-1 (PP1) to Dephosphorylate the U2 Small Nuclear Ribonucleoprotein Particle (snRNP) Component, Spliceosome-associated Protein 155 (Sap155).* Journal of Biological Chemistry, 2008. **283**(51): p. 35805-35814.

80. Oliver, C.J., et al., *Targeting Protein Phosphatase 1 (PP1) to the Actin Cytoskeleton: the Neurabin I/PP1 Complex Regulates Cell Morphology.* Mol. Cell. Biol., 2002. **22**(13): p. 4690-4701.
81. Kelker, M.S., et al., *Structural basis for spinophilin-neurabin receptor interaction.* Biochemistry, 2007. **46**(9): p. 2333-44.
82. Feng, J., et al., *Spinophilin regulates the formation and function of dendritic spines.* Proceedings of the National Academy of Sciences, 2000. **97**(16): p. 9287-9292.
83. Sarrouilhe, D., et al., *Spinophilin: from partners to functions.* Biochimie, 2006. **88**(9): p. 1099-1113.
84. Kreivi, J.-P., et al., *Purification and characterisation of p99, a nuclear modulator of protein phosphatase 1 activity.* FEBS Letters, 1997. **420**(1): p. 57-62.
85. Udho, E., et al., *PNUTS (phosphatase nuclear targeting subunit) inhibits retinoblastoma-directed PP1 activity.* Biochem Biophys Res Commun, 2002. **297**(3): p. 463-467.
86. Kim, Y.-M., et al., *PNUTS, a Protein Phosphatase 1 (PP1) Nuclear Targeting Subunit.* Journal of Biological Chemistry, 2003. **278**(16): p. 13819-13828.
87. Landsverk, H.B., et al., *PNUTS enhances in vitro chromosome decondensation in a PP1-dependent manner.* Biochem. J., 2005. **390**(3): p. 709-717.
88. Lee, S.J., et al., *Protein phosphatase 1 nuclear targeting subunit is a hypoxia inducible gene: its role in post-translational modification of p53 and MDM2.* Cell Death Differ, 2007. **14**(6): p. 1106-1116.
89. Rose, M., et al., *PNUTS forms a trimeric protein complex with GABAC receptors and protein phosphatase 1.* Molecular and Cellular Neuroscience, 2008. **37**(4): p. 808-819.
90. Zhang, J., et al., *Identification and characterization of the human HCG V gene product as a novel inhibitor of protein phosphatase-1.* Biochemistry, 1998. **37**(47): p. 16728-34.
91. Zhang, L., et al., *Identification of the interaction sites of Inhibitor-3 for protein phosphatase-1.* Biochem Biophys Res Commun, 2008. **377**(2): p. 710-713.
92. Han, Y., et al., *Mutations of t-complex testis expressed gene 5 transcripts in the testis of sterile t-haplotype mutant mouse.* Asian J Androl, 2008. **10**(2): p. 219-26.
93. Okamoto, R., et al., *Characterization and function of MYPT2, a target subunit of myosin phosphatase in heart.* Cellular Signalling, 2006. **18**(9): p. 1408-1416.
94. Tan, I., et al., *Phosphorylation of a Novel Myosin Binding Subunit of Protein Phosphatase 1 Reveals a Conserved Mechanism in the Regulation of Actin Cytoskeleton.* Journal of Biological Chemistry, 2001. **276**(24): p. 21209-21216.
95. Helps, N.R., et al., *Protein phosphatase 1 interacts with p53BP2, a protein which binds to the tumour suppressor p53.* FEBS Letters, 1995. **377**(3): p. 295-300.
96. Sullivan, A. and X. Lu, *ASPP: a new family of oncogenes and tumour suppressor genes.* Br J Cancer, 2007. **96**(2): p. 196-200.
97. Liu, Z.-J., X. Lu, and S. Zhong, *ASPP—Apoptotic specific regulator of p53.* Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 2005. **1756**(1): p. 77-80.
98. Eto, M., et al., *Molecular cloning of a novel phosphorylation-dependent inhibitory protein of protein phosphatase-1 (CPI17) in smooth muscle: its specific localization in smooth muscle.* FEBS Letters, 1997. **410**(2-3): p. 356-360.
99. Tountas, N., et al., *Juxtamembrane localization of the protein phosphatase-1 inhibitor protein PHI-1 in smooth muscle cells.* Histochemistry and Cell Biology, 2004. **121**(4): p. 343-350.
100. Liu, Q.-R., et al., *KEPI, a PKC-dependent Protein Phosphatase 1 Inhibitor Regulated by Morphine.* Journal of Biological Chemistry, 2002. **277**(15): p. 13312-13320.

101. Gong, J.P., et al., *Mouse brain localization of the protein kinase C-enhanced phosphatase 1 inhibitor KEPI (Kinase C-Enhanced PP1 Inhibitor).* Neuroscience, 2005. **132**(3): p. 713-727.
102. Liu, Q.-R., et al., *GBPI, a novel gastrointestinal- and brain-specific PP1-inhibitory protein, is activated by PKC and inactivated by PKA.* Biochem. J., 2004. **377**(1): p. 171-181.
103. Shi, W., et al., *GADD34-PP1c recruited by Smad7 dephosphorylates TGFbeta type I receptor.* J Cell Biol, 2004. **164**(2): p. 291-300.
104. Kepp, O., et al., *Disruption of the PP1/GADD34 complex induces calreticulin exposure.* Cell Cycle, 2009. **8**(23): p. 3971-3977.
105. Jousse, C., et al., *Inhibition of a constitutive translation initiation factor 2alpha phosphatase, CReP, promotes survival of stressed cells.* J Cell Biol, 2003. **163**(4): p. 767-75.
106. Harding, H.P., et al., *Ppp1r15 gene knockout reveals an essential role for translation initiation factor 2 alpha (eIF2α) dephosphorylation in mammalian development.* Proceedings of the National Academy of Sciences, 2009. **106**(6): p. 1832-1837.
107. Vereshchagina, N., et al., *The Essential Role of PP1β in Drosophila Is to Regulate Nonmuscle Myosin.* Molecular Biology of the Cell, 2004. **15**(10): p. 4395-4405.
108. Sueyoshi, T., et al., *PPP1R16A, The Membrane Subunit of Protein Phosphatase 1β, Signals Nuclear Translocation of the Nuclear Receptor Constitutive Active/Androstane Receptor.* Molecular Pharmacology, 2008. **73**(4): p. 1113-1121.
109. Hendrickx, A., et al., *Docking Motif-Guided Mapping of the Interactome of Protein Phosphatase-1.* Chemistry & biology, 2009. **16**(4): p. 365-371.
110. Wakula, P., et al., *Degeneracy and Function of the Ubiquitous RVXF Motif That Mediates Binding to Protein Phosphatase-1.* Journal of Biological Chemistry, 2003. **278**(21): p. 18817-18823.
111. Meiselbach, H., H. Sticht, and R. Enz, *Structural Analysis of the Protein Phosphatase 1 Docking Motif: Molecular Description of Binding Specificities Identifies Interacting Proteins.* Chemistry & Biology, 2006. **13**(1): p. 49-59.
112. Yang, J., T.D. Hurley, and A.A. DePaoli-Roach, *Interaction of Inhibitor-2 with the Catalytic Subunit of Type 1 Protein Phosphatase.* Journal of Biological Chemistry, 2000. **275**(30): p. 22635-22644.
113. Godet, A.N., et al., *The Combinatorial PP1-Binding Consensus Motif (R/K)x_(0,1)V/IxFxx(R/K)x(R/K) Is a New Apoptotic Signature.* PLoS One, 2010. **5**(4): p. e9981.
114. Ayllón, V., et al., *Bcl-2 Targets Protein Phosphatase 1α to Bad.* The Journal of Immunology, 2001. **166**(12): p. 7345-7352.
115. Llorian, M., et al., *SIPP1, a novel pre-mRNA splicing factor and interactor of protein phosphatase-1.* Biochem J.(Pt 1). 2004. **378**: p. 229-238.
116. Huang, L.J., et al., *NH2-Terminal targeting motifs direct dual specificity A-kinase-anchoring protein 1 (D-AKAP1) to either mitochondria or endoplasmic reticulum.* J Cell Biol, 1999. **145**(5): p. 951-9.
117. Llanos, S., et al., *The Inhibitory Member of the Apoptosis Stimulating Proteins of p53 Family (iASPP) Interacts with Protein Phosphatase 1 (PP1) via a Non-Canonical Binding Motif.* J Biol Chem, 2011.
118. Chakrabarti, R., et al., *Analysis of Ppp1cc-Null Mice Suggests a Role for PP1gamma2 in Sperm Morphogenesis.* Biology of Reproduction, 2007. **76**(6): p. 992-1001.
119. Varmuza, S., et al., *Spermiogenesis is impaired in mice bearing a targeted mutation in the protein phosphatase 1cgamma gene.* Dev Biol, 1999. **205**(1): p. 98-110.

120. Shima, H., et al., *Identification of PP1 catalytic subunit isotypes PP1 gamma 1, PP1 delta and PP1 alpha in various rat tissues.* Biochem Biophys Res Commun, 1993. **192**(3): p. 1289-96.
121. Suzuki, T., et al., *Regulation of hyperactivation by PPP2 in hamster spermatozoa.* Reproduction, 2010. **139**(5): p. 847-56.
122. Sasaki, K., et al., *Identification of members of the protein phosphatase 1 gene family in the rat and enhanced expression of protein phosphatase 1 alpha gene in rat hepatocellular carcinomas.* Jpn J Cancer Res, 1990. **81**(12): p. 1272-80.
123. Vijayaraghavan, S., et al., *Sperm motility development in the epididymis is associated with decreased glycogen synthase kinase-3 and protein phosphatase 1 activity.* Biol Reprod, 1996. **54**(3): p. 709-18.
124. Kloeker, S., et al., *Parallel purification of three catalytic subunits of the protein serine/threonine phosphatase 2A family (PP2A(C), PP4(C), and PP6(C)) and analysis of the interaction of PP2A(C) with alpha4 protein.* Protein Expr Purif, 2003. **31**(1): p. 19-33.
125. Huang, Z., et al., *Protein 14-3-3zeta binds to protein phosphatase PP1gamma2 in bovine epididymal spermatozoa.* Biol Reprod, 2004. **71**(1): p. 177-84.
126. Ueki, K., T. Muramatsu, and R.L. Kincaid, *Structure and expression of two isoforms of the murine calmodulin-dependent protein phosphatase regulatory subunit (calcineurin B).* Biochem Biophys Res Commun, 1992. **187**(1): p. 537-43.
127. Chang, C.D., et al., *cDNA cloning of an alternatively spliced isoform of the regulatory subunit of Ca2+/calmodulin-dependent protein phosphatase (calcineurin B alpha 2).* Biochim Biophys Acta, 1994. **1217**(2): p. 174-80.
128. Tash, J.S., et al., *Identification, characterization, and functional correlation of calmodulin-dependent protein phosphatase in sperm.* J Cell Biol, 1988. **106**(5): p. 1625-33.
129. Ross, M.H. and I.R. Long, *Contractile Cells in Human Seminiferous Tubules.* Science, 1966. **153**(3741): p. 1271-1273.
130. McLachlan, R.I., et al., *Identification of Specific Sites of Hormonal Regulation in Spermatogenesis in Rats, Monkeys, and Man.* Recent Prog Horm Res, 2002. **57**(1): p. 149-179.
131. Yoshida, S., M. Sukeno, and Y.-i. Nabeshima, *A Vasculature-Associated Niche for Undifferentiated Spermatogonia in the Mouse Testis.* Science, 2007. **317**(5845): p. 1722-1726.
132. Vijayaraghavan, S. and D.D. Hoskins, *Regulation of bovine sperm motility and cyclic adenosine 3',5'-monophosphate by adenosine and its analogues.* Biology of Reproduction, 1986. **34**(3): p. 468-477.
133. Jones, R.C., *To store or mature spermatozoa? The primary role of the epididymis.* Int J Androl, 1999. **22**(2): p. 57-67.
134. Smith, G.D., et al., *Motility potential of macaque epididymal sperm: the role of protein phosphatase and glycogen synthase kinase-3 activities.* J Androl, 1999. **20**(1): p. 47-53.
135. Huang, Z., et al., *Sperm PP1gamma2 is regulated by a homologue of the yeast protein phosphatase binding protein sds22.* Biol Reprod, 2002. **67**(6): p. 1936-42.
136. Yang, P., et al., *Protein phosphatases PP1 and PP2A are located in distinct positions in the Chlamydomonas flagellar axoneme.* J Cell Sci, 2000. **113**(1): p. 91-102.
137. da Cruz e Silva, O.A., et al., *Signal transduction therapeutics: relevance for Alzheimer's disease.* J Mol Neurosci, 2004. **23**(1-2): p. 123-42.
138. Greengard, P., P.B. Allen, and A.C. Nairn, *Beyond the dopamine receptor: the DARPP-32/protein phosphatase-1 cascade.* Neuron, 1999. **23**(3): p. 435-47.
139. Cohen, P., *The structure and regulation of protein phosphatases.* Annu Rev Biochem, 1989. **58**: p. 453-508.

140. Visconti, P.E., et al., *Ion channels, phosphorylation and mammalian sperm capacitation.* Asian J Androl, 2011. **13**(3): p. 395-405.
141. Reid, A.T., et al., *Cellular mechanisms regulating sperm-zona pellucida interaction.* Asian J Androl, 2011. **13**(1): p. 88-96.
142. Ahmad, K., et al., *Regulation of human sperm motility and hyperactivation components by calcium, calmodulin, and protein phosphatases.* Arch Androl, 1995. **35**(3): p. 187-208.
143. Krapf, D., et al., *Inhibition of Ser/Thr phosphatases induces capacitation-associated signaling in the presence of Src kinase inhibitors.* J Biol Chem, 2010. **285**(11): p. 7977-85.
144. Hrabchak, C. and S. Varmuza, *Identification of the spermatogenic zip protein Spz1 as a putative protein phosphatase-1 (PP1) regulatory protein that specifically binds the PP1cgamma2 splice variant in mouse testis.* J Biol Chem, 2004. **279**(35): p. 37079-86.
145. Hsu, S.H., H.M. Hsieh-Li, and H. Li, *Dysfunctional spermatogenesis in transgenic mice overexpressing bHLH-Zip transcription factor, Spz1.* Exp Cell Res, 2004. **294**(1): p. 185-98.
146. Hrabchak, C., H. Henderson, and S. Varmuza, *A testis specific isoform of endophilin B1, endophilin B1t, interacts specifically with protein phosphatase-1c gamma2 in mouse testis and is abnormally expressed in PP1c gamma null mice.* Biochemistry, 2007. **46**(15): p. 4635-44.
147. Huang, Z., et al., *Changes in intracellular distribution and activity of protein phosphatase PP1gamma2 and its regulating proteins in spermatozoa lacking AKAP4.* Biol Reprod, 2005. **72**(2): p. 384-92.
148. Jurisicova, A., et al., *DNA damage in round spermatids of mice with a targeted disruption of the Pp1cgamma gene and in testicular biopsies of patients with non-obstructive azoospermia.* Mol Hum Reprod, 1999. **5**(4): p. 323-30.
149. Fardilha, M., et al., *Protein phosphatase 1 complexes modulate sperm motility and present novel targets for male infertility.* Molecular Human Reproduction, 2011. **17**(8): p. 466-477.
150. Giffon, T., et al., *Cloning of a human homologue of the mouse Tctex-5 gene within the MHC class I region.* Immunogenetics, 1996. **44**(5): p. 331-9.
151. Pilder, S.H., M.F. Hammer, and L.M. Silver, *A novel mouse chromosome 17 hybrid sterility locus: implications for the origin of t haplotypes.* Genetics, 1991. **129**(1): p. 237-46.
152. Pilder, S.H., et al., *Hybrid sterility-6: a mouse t complex locus controlling sperm flagellar assembly and movement.* Dev Biol, 1993. **159**(2): p. 631-42.
153. Cebra-Thomas, J.A., et al., *Allele- and haploid-specific product generated by alternative splicing from a mouse t complex responder locus candidate.* Nature, 1991. **349**(6306): p. 239-41.
154. Pilder, S.H., et al., *The molecular basis of "curlicue": a sperm motility abnormality linked to the sterility of t haplotype homozygous male mice.* Soc Reprod Fertil Suppl, 2007. **63**: p. 123-33.
155. Hui, L., et al., *The mouse T complex gene Tsga2, encoding polypeptides located in the sperm tail and anterior acrosome, maps to a locus associated with sperm motility and sperm-egg interaction abnormalities.* Biol Reprod, 2006. **74**(4): p. 633-43.
156. Han, Y.B., et al., *Expression of a novel T-complex testis expressed 5 (Tctex5) in mouse testis, epididymis, and spermatozoa.* Mol Reprod Dev, 2007. **74**(9): p. 1132-40.
157. Dinischiotu, A., et al., *Identification of sds22 as an inhibitory subunit of protein phosphatase-1 in rat liver nuclei.* FEBS Lett, 1997. **402**(2-3): p. 141-4.

158. Chun, Y.S., et al., *A sds22 homolog that is associated with the testis-specific serine/threonine protein phosphatase 1gamma2 in rat testis.* Biochem Biophys Res Commun, 2000. **273**(3): p. 972-6.
159. Cheng, L., et al., *PP1gamma2 and PPP1R11 are parts of a multimeric complex in developing testicular germ cells in which their steady state levels are reciprocally related.* PLoS One, 2009. **4**(3): p. e4861.
160. Lesage, B., et al., *A complex of catalytically inactive protein phosphatase-1 sandwiched between Sds22 and inhibitor-3.* Biochemistry, 2007. **46**(31): p. 8909-19.
161. Brautigan, D.L., *Flicking the switches: phosphorylation of serine/threonine protein phosphatases.* Semin Cancer Biol, 1995. **6**(4): p. 211-7.
162. Kwon, Y.G., et al., *Cell cycle-dependent phosphorylation of mammalian protein phosphatase 1 by cdc2 kinase.* Proc Natl Acad Sci U S A, 1997. **94**(6): p. 2168-73.
163. Liu, C.W., et al., *Inhibitory phosphorylation of PP1alpha catalytic subunit during the G(1)/S transition.* J Biol Chem, 1999. **274**(41): p. 29470-5.
164. Huang, Z. and S. Vijayaraghavan, *Increased phosphorylation of a distinct subcellular pool of protein phosphatase, PP1gamma2, during epididymal sperm maturation.* Biol Reprod, 2004. **70**(2): p. 439-47.
165. Berthet, C., et al., *Cdk2 knockout mice are viable.* Curr Biol, 2003. **13**(20): p. 1775-85.
166. Ortega, S., et al., *Cyclin-dependent kinase 2 is essential for meiosis but not for mitotic cell division in mice.* Nat Genet, 2003. **35**(1): p. 25-31.
167. Milne, F.C., et al., *Affinity purification of diverse plant and human 14-3-3-binding partners.* Biochem Soc Trans, 2002. **30**(4): p. 379-81.
168. Jin, J., et al., *Proteomic, functional, and domain-based analysis of in vivo 14-3-3 binding proteins involved in cytoskeletal regulation and cellular organization.* Curr Biol, 2004. **14**(16): p. 1436-50.
169. Meek, S.E., W.S. Lane, and H. Piwnica-Worms, *Comprehensive proteomic analysis of interphase and mitotic 14-3-3-binding proteins.* J Biol Chem, 2004. **279**(31): p. 32046-54.
170. Pozuelo Rubio, M., et al., *14-3-3-affinity purification of over 200 human phosphoproteins reveals new links to regulation of cellular metabolism, proliferation and trafficking.* Biochem J, 2004. **379**(Pt 2): p. 395-408.
171. Sun, S., et al., *14-3-3 and its binding partners are regulators of protein-protein interactions during spermatogenesis.* J Endocrinol, 2009. **202**(3): p. 327-36.
172. Gardino, A.K., S.J. Smerdon, and M.B. Yaffe, *Structural determinants of 14-3-3 binding specificities and regulation of subcellular localization of 14-3-3-ligand complexes: a comparison of the X-ray crystal structures of all human 14-3-3 isoforms.* Semin Cancer Biol, 2006. **16**(3): p. 173-82.
173. Puri, P., et al., *Proteomic analysis of bovine sperm YWHA binding partners identify proteins involved in signaling and metabolism.* Biol Reprod, 2008. **79**(6): p. 1183-91.
174. Puri, P., et al., *Identification of 14-3-3 Binding Proteins in Mouse Testis by Tandem Affinity Tag Purification* The FASEB Journal, 2008. **22**.
175. Somanath, P.R., S.L. Jack, and S. Vijayaraghavan, *Changes in sperm glycogen synthase kinase-3 serine phosphorylation and activity accompany motility initiation and stimulation.* J Androl, 2004. **25**(4): p. 605-17.
176. Faux, M.C. and J.D. Scott, *Molecular glue: kinase anchoring and scaffold proteins.* Cell, 1996. **85**(1): p. 9-12.
177. Hubbard, M.J. and P. Cohen, *On target with a new mechanism for the regulation of protein phosphorylation.* Trends Biochem Sci, 1993. **18**(5): p. 172-7.

178. Schillace, R.V., et al., *Multiple interactions within the AKAP220 signaling complex contribute to protein phosphatase 1 regulation.* J Biol Chem, 2001. **276**(15): p. 12128-34.
179. Lester, L.B., et al., *Cloning and characterization of a novel A-kinase anchoring protein. AKAP 220, association with testicular peroxisomes.* J Biol Chem, 1996. **271**(16): p. 9460-5.
180. Reinton, N., et al., *Localization of a novel human A-kinase-anchoring protein, hAKAP220, during spermatogenesis.* Dev Biol, 2000. **223**(1): p. 194-204.
181. Vijayaraghavan, S., et al., *Protein kinase A-anchoring inhibitor peptides arrest mammalian sperm motility.* J Biol Chem, 1997. **272**(8): p. 4747-52.
182. Carrera, A., G.L. Gerton, and S.B. Moss, *The major fibrous sheath polypeptide of mouse sperm: structural and functional similarities to the A-kinase anchoring proteins.* Dev Biol, 1994. **165**(1): p. 272-84.
183. Carrera, A., et al., *Regulation of protein tyrosine phosphorylation in human sperm by a calcium/calmodulin-dependent mechanism: identification of A kinase anchor proteins as major substrates for tyrosine phosphorylation.* Dev Biol, 1996. **180**(1): p. 284-96.
184. Fulcher, K.D., et al., *Characterization of Fsc1 cDNA for a mouse sperm fibrous sheath component.* Biol Reprod, 1995. **52**(1): p. 41-9.
185. Johnson, L.R., et al., *Assembly of AKAP82, a protein kinase A anchor protein, into the fibrous sheath of mouse sperm.* Dev Biol, 1997. **192**(2): p. 340-50.
186. Miki, K., et al., *Targeted disruption of the Akap4 gene causes defects in sperm flagellum and motility.* Dev Biol, 2002. **248**(2): p. 331-42.
187. Mandal, A., et al., *FSP95, a testis-specific 95-kilodalton fibrous sheath antigen that undergoes tyrosine phosphorylation in capacitated human spermatozoa.* Biol Reprod, 1999. **61**(5): p. 1184-97.
188. Brown, P.R., et al., *A-kinase anchoring protein 4 binding proteins in the fibrous sheath of the sperm flagellum.* Biol Reprod, 2003. **68**(6): p. 2241-8.
189. Lin, R.Y., S.B. Moss, and C.S. Rubin, *Characterization of S-AKAP84, a novel developmentally regulated A kinase anchor protein of male germ cells.* J Biol Chem, 1995. **270**(46): p. 27804-11.
190. Feliciello, A., et al., *Expression of a kinase anchor protein 121 is regulated by hormones in thyroid and testicular germ cells.* J Biol Chem, 1998. **273**(36): p. 23361-6.
191. Mei, X., et al., *Cloning and characterization of a testis-specific, developmentally regulated A-kinase-anchoring protein (TAKAP-80) present on the fibrous sheath of rat sperm.* Eur J Biochem, 1997. **246**(2): p. 425-32.
192. Kultgen, P.L., et al., *Characterization of an A-kinase anchoring protein in human ciliary axonemes.* Mol Biol Cell, 2002. **13**(12): p. 4156-66.
193. Rawe, V.Y., et al., *WAVE1, an A-kinase anchoring protein, during mammalian spermatogenesis.* Hum Reprod, 2004. **19**(11): p. 2594-604.
194. Goto, N. and H. Harayama, *Calyculin A-sensitive protein phosphatases are involved in maintenance of progressive movement in mouse spermatozoa in vitro by suppression of autophosphorylation of protein kinase A.* J Reprod Dev, 2009. **55**(3): p. 327-34.
195. Fardilha, M., et al., *Identification of the human testis protein phosphatase 1 interactome.* Biochemical Pharmacology, 2011. **82**(10): p. 1403-1415.

Chapter II

Chapter II

Introduction

PPP1CC2 is the only PPP1 isoform highly enriched in sperm [1, 2] present along the entire flagellum including the mid-piece, consistent with a role in sperm motility [3, 4]. PPP1CC2 direct involvement in sperm motility was consistent with the fact that PPP1 has two-fold higher activity in immotile bovine caput epididymal sperm when compared to mature motile caudal sperm [2]. Moreover, inhibition of PPP1CC2 activity by okadaic acid and calyculin A causes both initiation and stimulation of motility in caput and cauda sperm, respectively [2].

A mechanism of PPP1 regulation likely to be directly implicated in sperm motility involves the long-known PPP1 inhibitor 2 (PPP1R2) and GSK3. A PPP1R2-like activity was detected in sperm and was reversed by purified GSK3 [2]. PPP1 and PPP1R2 form a stable, catalytically inactive complex. Phosphorylation of rabbit PPP1R2 at Thr73 by GSK3 releases PPP1 inhibition [5]. In sperm, GSK3 activity was detected by activation of purified PPP1-PPP1R2 holoenzyme, which demonstrated that immotile caput sperm contained six-fold higher GSK3 activity than motile caudal sperm [2]. The presence and activity of GSK3 in sperm has been further characterized sustaining its role in sperm motility regulation [6, 7]. However, the presence and role of the PPP1R2-like entity, which would make the regulating bridge between PPP1 and GSK3, has never been confirmed.

In this Chapter the presence of PPP1R2 in human sperm is for the first time demonstrated and its localization characterized (Chapter II.A). Also, a new PPP1R2 related protein from testis and sperm, named PPP1R2P3, is described for the first time. Further, PPP1R2 related processed pseudogenes are extensively characterized being their functional relevance discussed in detail (Chapter II.B). Finally, we addressed the PPP1R2 and PPP1R2P3 function in spermatogenesis and sperm maturation using the mouse transgenic technique (Chapter II.C).

The results presented clearly demonstrate the relevance of PPP1CC2/PPP1R2-related protein complexes in sperm function. We have proposed a novel model to explain sperm motility acquisition in the epididymis (the structure where sperm matures and acquires motility and the capacity to fertilize the oocyte) based on the above mentioned complexes.

References

1. Smith, G.D., et al., *Primate sperm contain protein phosphatase 1, a biochemical mediator of motility.* Biol Reprod, 1996. **54**(3): p. 719-727.
2. Vijayaraghavan, S., et al., Sperm motility development in the epididymis is associated with decreased glycogen synthase kinase-3 and protein phosphatase 1 activity. Biol Reprod, 1996. **54**(3): p. 709-18.
3. Huang, Z., et al., Sperm PP1gamma2 is regulated by a homologue of the yeast protein phosphatase binding protein sds22. Biol Reprod, 2002. **67**(6): p. 1936-42.
4. Fardilha, M., et al., Protein phosphatase 1 complexes modulate sperm motility and present novel targets for male infertility. Molecular Human Reproduction, 2011. **17**(8): p. 466-477.
5. Hemmings, B.A., T.J. Resink, and P. Cohen, Reconstitution of a Mg-ATP-dependent protein phosphatase and its activation through a phosphorylation mechanism. FEBS Letters, 1982. **150**(2): p. 319-324.
6. Somanath, P.R., S.L. Jack, and S. Vijayaraghavan, Changes in sperm glycogen synthase kinase-3 serine phosphorylation and activity accompany motility initiation and stimulation. J Androl, 2004. **25**(4): p. 605-17.
7. Vijayaraghavan, S., et al., A Role for Phosphorylation of Glycogen Synthase Kinase-3α in Bovine Sperm Motility Regulation. Biology of Reproduction, 2000. **62**(6): p. 1647-1654.

Chapter II.A

Discovery and characterization of two forms of PPP1R2 in human sperm

Luis Korrodi-Gregório*, Mónica Ferreira*, Ana Paula Vintém*, Wenjuan Wu*, Thorsten Muller[†], Katrin Marcus[†], Srinivasan Vijayaraghavan[‡], David L. Brautigan[§], Odete A. B. da Cruz e Silva[ll], Margarida Fardilha[¶1] and Edgar F. da Cruz e Silva*[2]

**Laboratory of Signal Transduction, Centre for Cell Biology, Biology Department, University of Aveiro, Aveiro, Portugal*

[†]*Functional Proteomics Department, Medizinisches Proteom-Center, Ruhr-University Bochum, Bochum, Germany*

[‡]*Anatomy and Cell Biology Department, Kent State University, Kent, Ohio, USA*

[§]*Center for Cell Signaling, Microbiology Department, Immunology, and Cancer Biology, University of Virginia School of Medicine, Charlottesville, Virginia, USA*

[ll]*Laboratory of Neurosciences, Centre for Cell Biology, Biology Department; Health Sciences Department, University of Aveiro, Aveiro, Portugal*

[¶]*Laboratory of Signal Transduction, Centre for Cell Biology, Biology Department; Health Sciences Department, University of Aveiro, Aveiro, Portugal*

[2]*deceased on 2nd of March, 2010*

Running Title: PPP1R2 proteins in spermatozoa.

Keywords: PP1gamma2, PPP1CC2, phosphorylation, PP1 interacting protein, PPP1R2P3

Submitted to: Biochem J (5.06 IF, A)

[1]**Corresponding author:** Margarida Fardilha, Centro de Biologia Celular, Universidade de Aveiro, 3810-193 Aveiro, Portugal, Tel.: 00351 91 8143947; Fax: 00351 234 370 200 E-mail: mfardilha@ua.pt

Abstract

Protein Ser/Thr Phosphatase PPP1CC2 is an alternatively spliced isoform of PPP1CC that is highly enriched in testis and selectively expressed in sperm. Inhibition of PPP1CC2 by addition of the toxins okadaic acid and calyculin A causes both initiation and stimulation of motility in caput and cauda sperm, respectively, making it interesting to define endogenous mechanisms for inhibition. A protein inhibitor was previously detected in sperm and was neutralized by GSK3 phosphorylation, suggesting it is related to PPP1R2. Here we show unequivocally evidence for the presence of PPP1R2 plus a related protein, PPP1R2P3, in human ejaculated sperm. Further, in sperm, PPP1R2/PPP1R2P3 are phosphorylated in serines residues, Ser121 and Ser122. Also, we identified a novel phosphorylated residue, Ser127, by mass spectrometry. Moreover, in PPP1R2P3 the key phosphosites are substituted to non-phosphorylated residues, T73P and S87R. We show that the novel protein PPP1R2P3 is a heat stable inhibitor of PPP1CC, with similar potency as PPP1R2. Furthermore, we show by immunocytochemistry and Western blot localization of PPP1CC2 holoenzymes containing PPP1R2 or PPP1R2P3 in the head and tail of sperm and in both soluble and insoluble fractions. The holoenzymes localized in the head might have an important role in the acrosome reaction while the axoneme bound holoenzymes are likely important for the control of flagellar motility. A novel model for sperm motility acquisition is proposed in which PPP1CC2 is irreversibly inhibited by PPP1R2P3 when it substitutes for PPP1R2 in caudal sperm to trigger motility or in the female reproductive tract to initiate hyperactivated motility. Thus this work proposes a mechanism for epididymal initiation of sperm motility, involving PPP1CC2/PPP1R2 and PPP1CC2/PPP1R2P3 complexes.

Introduction

The reversible phosphorylation of proteins is a major control mechanism involved in a wide range of eukaryotic cellular responses. The phosphorylation events occur mostly on serine, threonine and tyrosine residues, and the pivotal players are kinases and phosphatases. Spermatozoa are specialized cells as they are highly compartmentalized, transcriptionally inactive and unable to synthesize new proteins. Phosphorylation plays a crucial role in sperm physiology controlling motility, capacitation, motility and acrosome reaction [1, 2]. Sperm motility defects are one of the main underlying causes of male infertility [2]. The biochemical mechanisms essential for sperm maturation and the development of motility are still far from understood, however serine/threonine phosphoprotein phosphatase 1 (PPP1) and glycogen synthase kinase-3 (GSK3), are part of the regulatory process [3, 4]. Three separate genes (α/A, β/δ/B and γ/C) encode the catalytic subunit of PPP1 (PPP1C). PPP1CC undergoes an alternative splicing event giving rise to a ubiquitous isoform PPP1CC1 and a testis-enriched and sperm specific-PPP1CC2 [3, 5]. PPP1CC2 is the only PPP1 isoform highly enriched in bovine, rhesus monkey and human sperm [3, 5] being present along the entire flagellum including the mid-piece, consistent with a role in sperm motility, and also in the posterior and equatorial regions of the head, suggesting a role in the acrosome reaction [2, 6]. PPP1CC2 direct involvement in sperm motility was consistent with the fact that PPP1C had two-fold higher activity in immotile bovine caput epididymal sperm compared to mature motile caudal sperm [3]. Moreover, inhibition of PPP1CC2 activity by okadaic acid and calyculin A causes both initiation and stimulation of motility in caput and cauda sperm respectively [3]. Homozygous knockout mice for PPP1CC gene (a deletion of both isoforms) lead to sterility of male but not female mice. This sterility results from a combination of gross structural defects in spermatids that cause apoptosis and lack of spermiation [7, 8]. The evolutionary conservation and the importance of serine/threonine phosphatases in regulating flagellar motility is highlighted by the involvement of a PPP1 homolog in the regulation of rooster sperm motility [9, 10] and by the involvement of a serine/threonine phosphatase in the regulation of microtubule sliding velocity in *Paramecium* and *Chlamydomonas* [11-15]. PPP1 activity towards different substrates is mediated via binding to specific regulatory proteins, the PPP1 interacting proteins (PIPs) [16, 17]. The increasing diversity of such PIPs and their tissue specificity makes them attractive pharmacological targets [17]. A mechanism of PPP1 regulation likely to be directly involved in sperm motility involves the long-known PPP1 inhibitor 2 (PPP1R2) and GSK3. A PPP1R2-like activity was detected in sperm and was neutralized by phosphorylation

with purified GSK3 [3]. PPP1 and PPP1R2 are known to form a stable, catalytically inactive complex. Phosphorylation of rabbit PPP1R2 at Thr73 by GSK3 releases the inhibition and the complex becomes active [18]. In sperm, assays showed that immotile caput sperm contained six-fold higher GSK3 activity than motile caudal sperm [3]. The presence and activity of GSK3 in sperm has been further characterized emphasizing its role in sperm motility regulation [19, 20]. However, the presence and role of PPP1R2, which would make the connection between PPP1 and GSK3, has not yet been confirmed. In this work the presence of PPP1R2 in human sperm is demonstrated and its localization characterized. Also, a new PPP1R2 isoform, PPP1R2P3, previously assigned as an intron-less pseudogene [21], is recovered as a protein for the first time. Further, the relevance of these different holoenzymes PPP1CC2/PPP1R2 and PPP1CC2/PPP1R2P3 in human sperm are discussed.

Material and Methods

Yeast two-hybrid

Methods for yeast two-hybrid screening of a human testis cDNA library using human PPP1C have been previously described [22-24]. DNA sequence analysis was performed using an ABI PRISM 310 Genetic Analyser (Portugal Applied Biosystems, Porto, Portugal). The DNA sequences obtained were compared to the NCBI database, using the BLAST algorithm (http://BLAST.ncbi.nlm.nih.gov/). The multiple sequence alignments were performed using the ClustalW program from Ensembl [25].

PPP1R2 and PPP1R2P3 cloning, expression and purification

The cDNA of PPP1R2 pseudogene 3 (PPP1R2P3) was inserted into the pET28c (Novagen, Madison, Wisconsin, USA) expression vector in EcoRI and XhoI restriction sites, adding a histidine tag (His-tag) to the N-terminal of the protein. The pET-PPP1R2P3 construct sequence was verified and transformed into *E. coli* strain *Rosetta* (DE3) (Novagen, Madison, Wisconsin, USA). The expression of His-tag PPP1R2P3 was induced with 1mM IPTG for 3hrs at 37ºC and the protein purified using a Ni-NTA resin (QIAGEN, Dusseldorf, Germany) according to the supplier's instructions. Briefly, the cells were lysed in 10mM imidazole, sodium phosphate buffer, pH 8.0, centrifuged at 15000g for 30min at 4ºC, and the supernatant was applied to the resin. The resin was washed with 20mM imidazole and the His-tag PPP1R2P3 was eluted with 500mM imidazole. The eluate was further purified from a 12% SDS-PAGE gel. A portion of the lane containing PPP1R2P3 was stained with coomassie blue and the corresponding band containing the remaining protein cut out. The gel band was washed three times with water, cut into smaller pieces and 1ml of 100mM Tris-HCl, pH 8.5, 0.1% SDS was added before frozen at -20ºC overnight. The slurry was frozen and thawed three times and then passed through a 0.22µm filter membrane. The gel-free filtrate was then dialyzed against 4x500ml of 10mM Tris-HCl, pH 7.5 buffer for 24hrs at 4ºC. The protein concentration of recombinant His-PPP1R2P3 was determined by BCA® assay (Fisher Scientific, Loures, Portugal). PPP1R2P3 were also inserted into the pTACTAC expression vector [26] in the NdeI and XbaI restriction sites, the sequence verified and then transformed into Rosetta. The expression of PPP1R2P3 was induced with 0.4mM IPTG for 3hrs at 37ºC. The protein was partially purified to a bacterial heat extract (recombinant PPP1R2P3) as previously described for PPP1R2 [27]. PPP1R2 was also inserted in pTACTAC expression vector

[26] and recombinant protein purified same way as for PPP1R2P3 (recombinant PPP1R2).

Yeast co-transformation with plasmid DNA

Yeast competent AH109 cells were co-transformed with pACT-PPP1R2P3 and pAS2-PPP1CA, pAS2-PPP1CC1, pAS2-PPP1CC2 or pAS2-PPP1CC2end, by the lithium acetate method [23, 24, 28]. Afterwards, the transformation mixture was plated on selective media containing X-α-Gal and incubated at 30ºC to check for MEL1 expression as indicated by the appearance of a blue color (Clontech, Saint Germain-en-Laye, France).

Blot overlay analysis

For blot overlay analysis, 0.3µg of commercial PPP1R2 (NEB, New England Biolabs (UK), Herts, UK) and recombinant His-PPP1R2P3 were resolved by SDS-PAGE and then transferred to a nitrocellulose membrane. Blots were overlaid with purified PPP1CC1 or PPP1CC2 diluted in Tris buffer saline with Tween-20/bovine serum albumin (25pmol/mL) [22, 29] and detected with the antibodies CBC3C [against the C-terminal of PPP1CC, detects both isoforms [30]] or CBC502 (specific for the C-terminal of PPP1CC2), both raised in rabbit. Immunoreactive bands were revealed by incubating with horseradish peroxidase conjugated anti-rabbit secondary antibody and developed by ECL (GE Healthcare, Amersham Biosciences Europe GmbH, Freiburg**,** Germany).

Phosphatase activity assays

The IC50 values of PPP1R2 and PPP1R2P3 for purified PPP1CC1 and PPP1CC2 isoforms were determined using [32P]phosphorylase a as substrate. The substrate was prepared from phosphorylase b (Sigma-Aldrich Química, S.A., Sintra, Portugal) using [γ-32P]ATP (3000 Ci/mmol, GE Healthcare, Amersham Biosciences Europe GmbH, Freiburg**,** Germany) and phosphorylase kinase (Sigma-Aldrich Química, S.A., Sintra, Portugal) as previously described [31]. An appropriate range of concentrations of commercial PPP1R2 and His-PPP1R2P3 were incubated with the purified PPP1C subunits and the phosphatase activity determined. The IC50 was calculated using the BioDataFit 1.02 software (Chang Bioscience, Castro Valley, California, USA).

Phosphorylation of PPP1R2 and PPP1R2P3

Recombinant PPP1R2, PPP1R2P3 or human sperm heat extract of native PPP1R2/PPP1R2P3 (about 200ηg of inhibitor protein in 50mM Tris-HCl pH 7.5, 0.1mM EGTA and 0.03% Brij-35) were phosphorylated by GSK-3β or CK2 (Calbiochem, MERCK, Darmstadt, Germany) or both as previously described [32]. The phosphorylation reaction was run at 30ºC for 90min and then terminated with the addition of 4X SDS loading buffer. In separate 12% gels 1/10 of the reaction volume and the remaining were ran. The 1/10 reaction was analyzed by Western blot and the remaining reaction gel was dried and autoradiographed.

Sperm extracts

Ejaculated sperm was collected from healthy donors by masturbation to an appropriate sterile container (informed consents were signed allowing samples to be used to scientific purposes). Spermograms were performed by experienced technicians and only samples with normal parameters were used [33]. For all methods, sperm was washed three times in 1xPBS. For immunoprecipitation sperm was lysed in 1xRIPA buffer (radioimmunoprecipitation buffer, Millipore Iberica S.A.U., Madrid, Spain) supplemented with protease inhibitors (10mM benzamidine, 1.5μM aprotinin, 5μM pepstatin A, 2μM leupeptin, 1mM PMSF), sonicated 3x10sec and centrifuged at 16000g for 20min, at 4ºC. The supernatant was collected and heat stable extracts were prepared by immersing the sample in a boiling water bath for 30min, chilling on ice for 2min and centrifuging at 16000g for 20min, at 4ºC. The final supernatant was used in the subsequent steps. For Western blot both the supernatant (soluble fraction) and the pellet (insoluble fraction) were resuspended in 1% SDS. For the preparation of heads and tails, washed sperm was briefly sonicated and the detachment checked by phase contrast microscopy (PH) using an Olympus IX81 epifluorescence microscope equipped with appropriate software (Olympus Portugal - Opto-Digital Tecnologias, S.A., Lisboa, Portugal). Sperm were directly applied onto a sucrose gradient (1.8M, 2.02M and 2.2M) to separate heads from tails. In brief, gradient was centrifuged at 5000g for 1hr and fractions corresponding to tails and heads collected. Subsequently fractions were centrifuged at 16000g, 10min, to collect the pellet free of sucrose and proteins dissolved in 1% SDS. For immunocytochemistry, washed sperm were used directly on the coverslips. Human testis extracts were also prepared using 1% SDS using the same protocol as for sperm samples with the exception of using a tissue homogenizer.

Immunoprecipitation

RIPA supernatant sperm extracts were pre-cleared using Dynabeads® Protein G (Life Technologies S.A., Madrid, Spain). Using a direct immunoprecipitation approach, 1µg of sheep anti-PPP1R2 or rabbit anti-PPP1R2 was pre-incubated with Dynabeads® Protein G during 1hr at 4ºC with rotation. After incubation, sperm pre-cleared extracts were applied to the antibody-dynabeads complex and then incubated overnight with rotation at 4ºC. After washing three times with 1xPBS in 3%BSA for 10min with rotation at 4ºC, beads were resuspended in loading buffer and boiled.

Mass spectrometry

For mass spectrometry analysis, immunoprecipitates were resolved by 10% SDS-PAGE along with purified positive controls. Gels were stained with Coomassie blue colloidal (Sigma-Aldrich Química, S.A., Sintra, Portugal). In brief, gels were fixed by a fixation solution (40% methanol and 10% acetic acid) during 1hr, washed with distilled water and then transferred to the Coomassie blue colloidal staining solution for 1hr. After staining, the gels were washed with distilled water and afterwards destained with 25% methanol until bands were visualized.

Bands were excised directly from the gel using a spatula and completely destained. In-gel digestion was performed overnight at 37°C with trypsin (Promega, Madison, Wisconsin, USA) in 10mM HCl and 50mM ammonium hydrogen carbonate (NH_4HCO_3) at pH 7.8. Resulting peptides were extracted once with 100µl of 1% formic acid (FA), and twice with 100µl of 5% FA, 50% acetonitrile (ACN). Extracts were combined and ACN was removed *in vacuo*. For LC-MS analysis, a final volume of 40µl was prepared by addition of 1% FA. Electrospray tandem mass spectrometry (ESI-MS/MS) was performed on a Orbitrap Velos instrument (Thermo Scientific, Bremen, Germany) Fragment ions were generated by low-energy collision-induced dissociation (CID) on isolated ions with a fragmentation amplitude of 0.5 V. MS spectra were summed from four individual scans ranging from m/z 300-1500 with a scanning speed of 8.100 (m/z)/s. MS/MS spectra were a sum of two scans ranging from m/z 100-2800 at a scan rate of 26.000 (m/z)/s. Generated data were imported to ProteinScape™, a proteomics data platform (Bruker Daltonik GmbH, Bremen, Germany, [34]) and analyzed using MASCOT (version 2.2.0, Matrix Science, London, UK, [35]) search algorithm with search parameters as follows: precursor ion tolerance of 1.2 and 0.3 Da tolerance for MS/MS spectra. Proteins were considered to be identified if the Mascot score (ProteinScape™) was higher than 65.

Western blotting

Extracts were mass normalized using BCA® assay (Fisher Scientific, Loures, Portugal). Immunoprecipitates and extracts were resolved by 10% SDS-PAGE. Proteins were subsequently electrotransferred onto nitrocellulose membranes and immunodetected with the appropriate antibodies, using ECL chemiluminescence (GE Healthcare Spain, Madrid, Spain). The primary antibodies used in this study included sheep polyclonal anti-PPP1R2 (1:100), the rabbit CBC502 (1:2000, against the C-terminal of PPP1CC2), the rabbit CBC3C (1:1000, against the C-terminal and detects both PPP1CC isoforms) and the loading controls mouse monoclonal anti-β-tubulin (1:500, Life Technologies S.A., Madrid, Spain) and mouse monoclonal acetylated-α-tubulin (1:2000, Life Technologies S.A., Madrid, Spain). The secondary antibodies used were horseradish peroxidase-conjugated anti-rabbit (1:5000), anti-sheep (1:1000) and anti-mouse (1:5000) IgGs for enhanced chemiluminescence (ECL, GE Healthcare, Amersham Biosciences Europe GmbH, Freiburg**,** Germany) detection.

Immunocytochemistry

An aliquot of washed sperm (25µl) was placed onto a glass coverslip pre-coated with 100µg/ml poly-L-ornithine, and dried at room temperature, in a six-well plate containing one coverslip per well. To each well 1ml of 4% paraformaldehyde in 1xPBS was gently added and left to stand for 10min. Subsequently, sperm were washed twice with 1ml 1xPBS for 10min. For permeabilization, 1ml of 1:1 methanol/acetone solution was added for 2min and then the specimens washed twice with 1ml 1xPBS for 10min and blocked for 1hr with 3% BSA in 1xPBS, before incubation with primary antibodies (rabbit CBC502, 1:250 and sheep anti-PPP1R2, 1:100) for 2hrs at RT. After three washes with 1xPBS, the fluorescent labeled secondary antibodies anti-rabbit Texas-Red, 1:300 (MolecularProbes, Eugene, USA) and anti-sheep FITC, 1:50 (DAKO, Glostrup, Denmark) were added and the coverslips incubated for 2hrs. Finally, three washes with 1xPBS were performed and coverslips were mounted on microscope glass slides with one drop of anti-fading reagent containing DAPI for nucleic acid staining (Vectashield, Vector Laboratories Burlingame, California, USA). Images were acquired using an Olympus IX81 epifluorescence microscope and digital camera, equipped with appropriate software (Olympus Portugal - Opto-Digital Tecnologias, S.A., Lisboa, Portugal).

2D-PAGE analysis

The human sperm heat stable extracts (hsPPP1R2/PPP1R2P3) or hsPPP1R2/PPP1R2P3 plus 10ηg of recombinant PPP1R2 were acetone precipitated and the pellets were resuspended in 250µl of 2D rehydration solution (8M Urea/ 2M Thiurea/ 2% CHAPS/ 0.002% of bromophenol blue) and supplemented with 2.5µl of IPG buffer (in the 4-7 pH range) and 14mg of DTT. The samples were pipetted into a strip holder and the electrophoresis was started (1h at 30V, 2h at 150V, 1h at 500V, 1h at 1000V and 2h at 8000V). After the second dimension on 12% gels, samples were analyzed by Western blot.

Dephosphorylation of human sperm PPP1R2/PPP1R2P3

Human sperm PPP1R2 extracts were incubated overnight with either protein tyrosine phosphatase 1B (Upstate, Millipore Iberica S.A.U., Madrid, Spain) at 37ºC, calf intestinal phosphatase (NEB, New England Biolabs (UK), Herts, UK) at 37ºC or PPP1CC1, at 30ºC, with the respective assay buffers. The reactions were stopped with the addition of 4x SDS loading buffer and analyzed by Western blot.

Results and Discussion

Identification of PPP1R2P3 - a novel PPP1R2 isoform in testis

To identify regulatory subunits of PPP1CC1 and PPP1CC2 these isoforms were used as baits to screen a human testis cDNA library by yeast two-hybrid. From the PPP1CC1 screening, one of the 120 positive clones encoded the complete coding sequence of a novel isoform of the known PPP1C regulator known as PPP1R2 or inhibitor 2 (I2). While PPP1R2 is present in chromosome 3 and is encoded by 6 exons, PPP1R2P3 is an intronless gene and therefore was designated as a probable pseudogene by both NCBI and Ensembl databases. This clone has been deposited in the GenBank database under the ID:JF438008.1 [2, 24]. The clone is located in chromosome 5 and aligns to the sequence ID:NR002168, classified as PPP1R2 pseudogene 3 (PPP1R2P3). PPP1R2P3 mRNA was also present in human testis cDNA libraries by the Mammalian Gene Collection (MGC) program (nucleotide ID: BC066922; protein ID: Q6NXS1) [36]. By searching Unigene from NCBI for PPP1R2P3 specific expressed sequence tags (ESTs) seven ESTs were obtained, three of them from testis, two from brain and one from lung. Taken together, our results, the ESTs, and the MGC results [36], so far, five independent ESTs were found which support the fact that PPP1R2P3 is indeed a testis-expressed gene and not a pseudogene.

PPP1R2P3 has only 16 nucleotide substitutions (92.2%, identity), which correspond to 9 amino acid changes (95%, identity) in the translated sequence compared to PPP1R2. Comparing the PPP1R2P3 protein sequence with that of the PPP1R2 in ClustalW2, we clearly show that all the PPP1C binding regions are maintained (Figure II.A. 1). PPP1R2 interaction with PPP1C involves tow primary motifs, ^{145}KLHY^{148} and ^{43}KSQKW^{47}. The first motif interacts at the catalytic center of PPP1C, displacing the essential metal ions, inducing rapid inhibition and slower inactivation of PPP1C. The later motif binds to the PPP1C hydrophobic groove. Other points of contact in PPP1R2 are the N-terminal SILK motif (^{12}KGILK^{16}, in humans) that possibly initiates the binding, and the C-terminal acidic stretch, required for reaction with GSK3 [32, 37, 38]. GSK3 functions in the indirect activation of PPP1C holoenzyme [39].

We found important differences between PPP1R2 and PPP1R2P3, in that phosphorylation sites for GKS3 (Thr73) and CK2 kinases (Ser87) are altered to Pro and Arg, respectively (Figure II.A. 1). These residues have key roles in PPP1 regulation by PPP1R2. Phosphorylation of rabbit PPP1R2 at Thr73 by GSK3, releases the inhibition

and the complex PPP1R2/PPP1C becomes active [18]. Additionally, the phosphorylation of PPP1R2 by GSK3 is enhanced when Ser87, 121 and 122 are also phosphorylated by CK2, particularly Ser87 [32, 40].

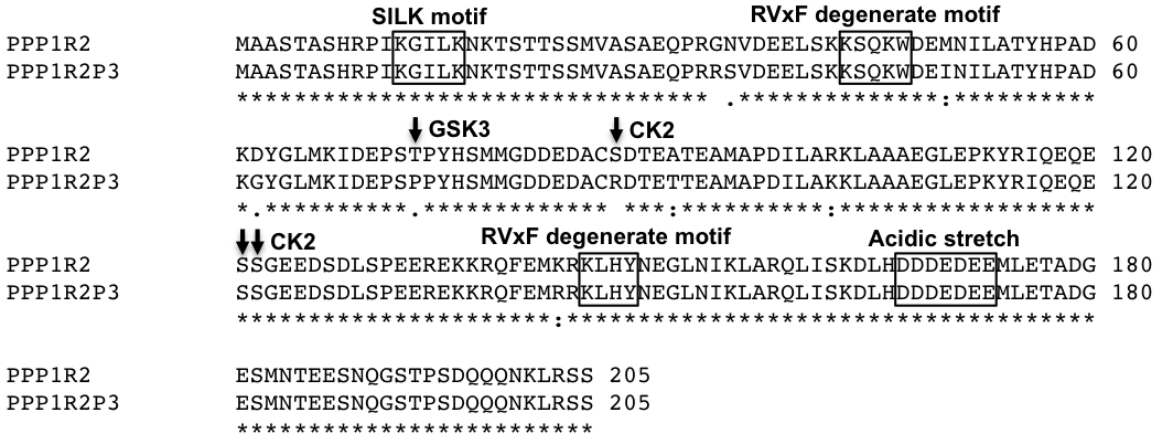


Figure II.A. 1: ClustalW2 alignment of PPP1R2 and PPP1R2P3. *Protein sequences for PPP1R2 and PPP1R2P3 were obtained from Uniprot database. Sequences were submitted to a ClustalW2 alignment. Important motifs/regions for PPP1R2/PPP1R2P3 binding to PPP1C are shown in boxes. Important phosphorylation sites are indicated with black arrows above the residues and with the respective known kinase. * represent high conservation, : and . represent low conservation in which the substituted residue has respectively more and less similar properties.*

PPP1R2P3 interacts and inhibits PPP1C

To rigorously validate the yeast two-hybrid result we re-confirmed the PPP1R2P3 interaction with different PPP1C isoforms. The sequential yeast co-transformation, showed the interaction of PPP1R2P3 with PPP1CA, PPP1CC1 and PPP1CC2 and the unique C-terminal of PPP1CC2 (Figure II.A. 2A). Since in yeast co-transformation the interaction is forced and happens inside the nucleus this might explain the apparent unspecificity toward the PPP1 isoforms. *In vivo*, the PPP1 isoforms have different tissue/cellular and subcellular localizations [41, 42] that may determine the fate of PPP1R2P3 interaction with specific PPP1C isoforms. In order to compare the binding of PPP1R2 and PPP1R2P3 to PPP1CC isoforms, and demonstrate the interaction as direct, a blot overlay was performed (Figure II.A. 2B). A blot containing the same amount of commercial PPP1R2 or His-PPP1R2P3 was incubated with PPP1CC1 or PPP1CC2 and the interaction detected using the CBC3C and the CBC502 antibodies, respectively (Figure II.A. 2B). The binding of PPP1R2P3 to both PPP1CC isoforms was confirmed and the binding was similar as to PPP1R2.

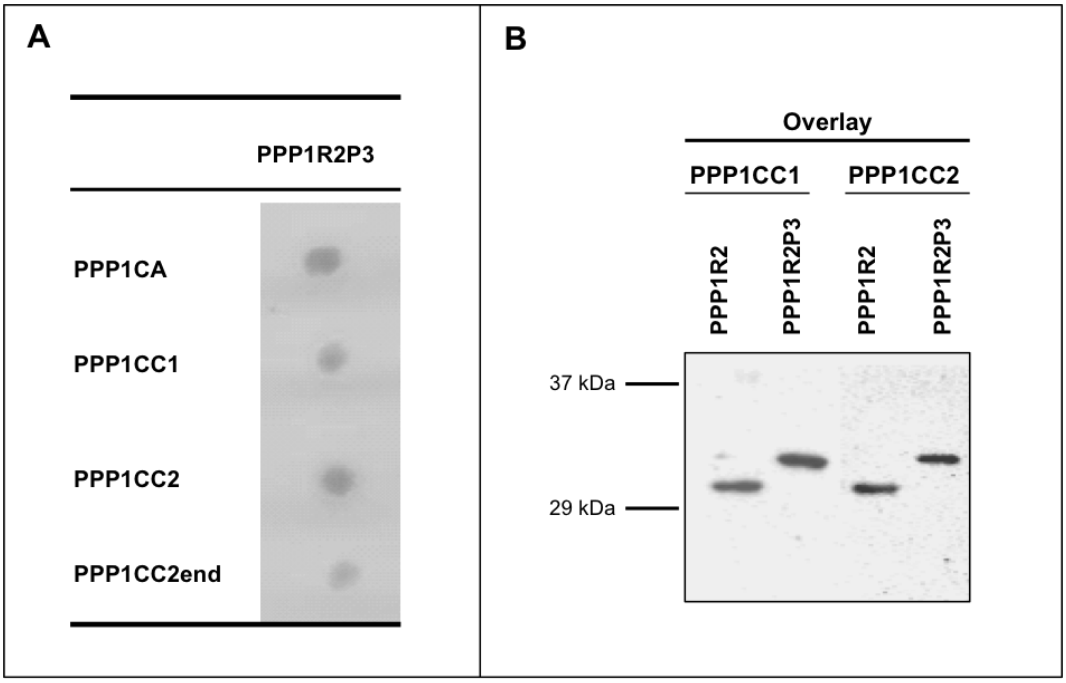
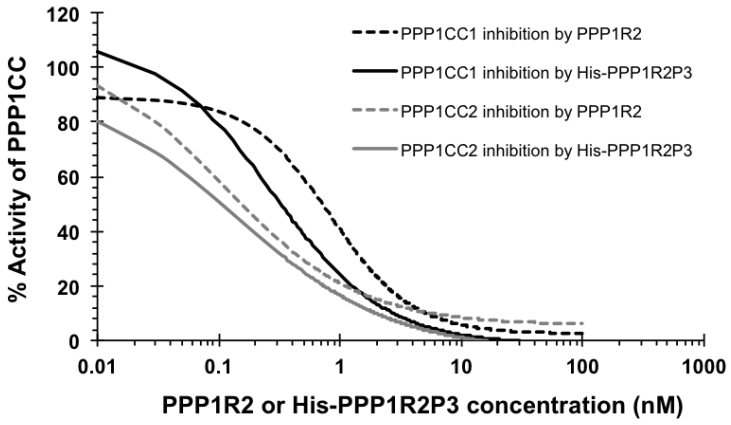


Figure II.A. 2: Interaction of PPP1R2P3 with different PPP1 isoforms. A. *Sequential transformation of yeast AH109 with bait plasmid (pAS2-PPP1CA, pAS2-PPP1CC1, pAS2-PPP1CC2, or pAS2-PPP1CC2end) and the prey plasmid pACT2-PPP1R2P3. pAS2-PPP1CC2end is the unique C-terminal tail of PPP1CC2 produced by alternative splicing of the PPP1CC gene.* **B.** *Overlay Western blot detection of PPP1R2 and PPP1R2P3 and PPP1CC1 and PPP1CC2. Commercial PPP1R2 and His-PPP1R2P3 were separated by SDS-PAGE, transferred to nitrocellulose and overlaid with recombinant PPP1CC1 or PPP1CC2, as indicated. Western blotting was performed with the respective specific antibodies.*

Since PPP1R2 is a potent heat stable inhibitor of PPP1 in the nanomolar range [27], we decided to determine if PPP1R2P3 is heat stable and if it exhibited PPP1 inhibitory activity, and if so, what was the potency compared to that of PPP1R2. We determined the IC50 values of recombinant His-PPP1R2P3 for the PPP1CC isoforms using the standard phosphorylase phosphatase assay [31]. Results show that PPP1R2P3 is a potent heat stable inhibitor of PPP1C with a dose-response curve similar to the commercial PPP1R2 (Figure II.A. 3A) and IC50 values in the subnanomolar range (Figure II.A. 3B, PPP1CC1, 0.73 ± 0.10 nM and PPP1CC2, 0.09 ± 0.08 nM).

A



B

	IC50 (nM)[a]	
Isoforms	**PPP1R2**	**His-PPP1R2P3**
His-PPP1CC1	0.27± 0.09	0.73± 0.10
His-PPP1CC2	0.14± 0.04	0.09± 0.08

Figure II.A. 3: PPP1R2 and PPP1R2P3 inhibit PPP1CC. A. *Graphical representation of PPP1CC1 and PPP1CC2 inhibition curves by commercial PPP1R2 and His-PPP1R2P3, using phosphorylase a as substrate.* **B.** *Table showing the comparison of PPP1CC isoforms IC50s by PPP1R2 and PPP1R2P3 using the phosphorylase phosphatase assay.* [a] *The values are expressed as the mean ± S.E.M. of at least three independent experiments.*

PPP1R2P3 is not phosphorylated by GSK3

It is known that phosphorylation of PPP1R2 at Thr73 (in rabbit) by GSK3, releases the inhibition of PPP1C and the complex PPP1C/PPP1R2 becomes active [18]. Also, phosphorylation by CK2 in the serines 87, 121 and 122 enhances Thr73 phosphorylation [32, 40]. The absence of the GSK3 phosphorylation site, Thr73, and of the CK2 phosphorylation site, Ser87, in PPP1R2P3 led us to investigate the phosphorylation of PPP1R2P3 by these kinases. Both PPP1R2 and PPP1R2P3 were incubated with GSK3 or CK2, or both, in the presence of radioactively-32P-labeled ATP. PPP1R2 is known to be a poor substrate for GSK3 *in vitro*, but if phosphorylated by CK2, synergistic phosphorylation by GSK3 can be observed [32]. In our hands this occurred with recombinant PPP1R2 (Figure II.A. 4, first panel). However, under the same conditions, no PPP1R2P3 phosphorylation by GSK3 was detected and the phosphorylation in the presence of both kinases was the same as with CK2 alone (Figure II.A. 4, second panel). The results are consistent with the T73P substitution and indicate there is no other site for GSK3 phosphorylation in PPP1R2P3.

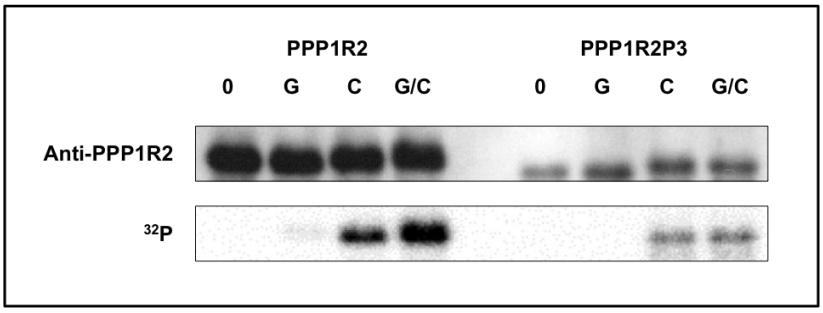


Figure II.A. 4: Comparison of PPP1R2 and PPPR2P3 phosphorylation by GSK3 and CK2. *Recombinant PPP1R2 and PPP1R2P3 were incubated in kinase buffer with reactive ATP (0); or in the presence of GSK3β kinase (G), or CK2 kinase (C); or both kinases (G/C), for 90min at 30°C. Protein phosphorylation was detected by autoradiography (^{32}P) following SDS-PAGE. Immunoreactivity detected with a specific sheep anti-PPP1R2 antibody is shown for comparative purposes.*

PPP1R2 and PPP1R2P3 are present in human ejaculated sperm

Based on biochemical studies it was suggested that sperm PPP1CC2 is regulated by a PPP1R2-like activity [3, 5]. However, definitive evidence for the presence of PPP1R2 in spermatozoa has not been reported. The PPP1R2 mRNA of 1.2kb, 1.4-1.7kb, 2.4-2.7kb and 4kb in testis have previously been described in rat and rabbit [32]. These messages strictly correspond to the polyadenylation signals identified by bioinformatics. Furthermore, a new highly expressed testis-specific and presumably spliced 0.9-1.1kb message from the PPP1R2 gene was also found in rabbit and rat. This message was only detected after 50 days and not before 20 days old rats [32, 43]. Efforts to identify PPP1R2 at the protein level in testis and sperm have proven to be difficult because of the low amount of protein in soluble extracts and the quality of the available antibodies. PPP1R2 is an intrinsically unstructured protein (IUP) with a high proportion of disorder inducing residues versus few hydrophobic residues [38, 44, 45]. Since SDS binds to hydrophobic amino acids, proteins that have few, such as PPP1R2 (predicted, 23kDa), will run in SDS-PAGE at a position corresponding to a higher molecular weight than that predicted by the primary sequence (32kDa, for PPP1R2) [46, 47]. Also, PPP1R2P3 has the same predicted molecular weight of PPP1R2 and in SDS-PAGE migrate at the same position, 32kDa (Figure II.A. 2B). To bypass the low abundance and take advantage of PPP1R2 being heat stable, generally heat stable fractions followed by protein precipitation are used to concentrate the samples [48]. So far only one report showed successful identification of PPP1R2 in heat stable extracts of bull testis and mouse testis and sperm using a PPP1R2 antibody (rabbit anti-PPP1R2, see materials and methods) raised in rabbit against an affinity-purified peptide (^{135}REKKRQFEMKRKLH^{148} from the mouse sequence) [49].

We used a sheep polyclonal anti-PPP1R2 antibody, raised against purified rabbit PPP1R2 as immunogen [46, 50-55] which detected PPP1R2P3, and we observed a band around the expected size (32kDa) in testis and sperm 1% SDS extracts (Figure II.A. 5A). Also, to demonstrate the presence of PPP1CC in the same extracts, an antibody against the C-terminus of PPP1CC was used (CBC3C). We detected in testis two bands that correspond to PPP1CC1 and PPP1CC2, while only one band, PPP1CC2, was detected in sperm, as expected [3]. Moreover, since PPP1R2 and PPP1R2P3 are heat stable PPP1C inhibitors, we prepared heat stable extracts by boiling RIPA supernatant extracts. A band corresponding to PPP1R2/PPP1R2P3 was detected in the heat stable extract at the expected size, 32kDa (Figure II.A. 5B, second lane).

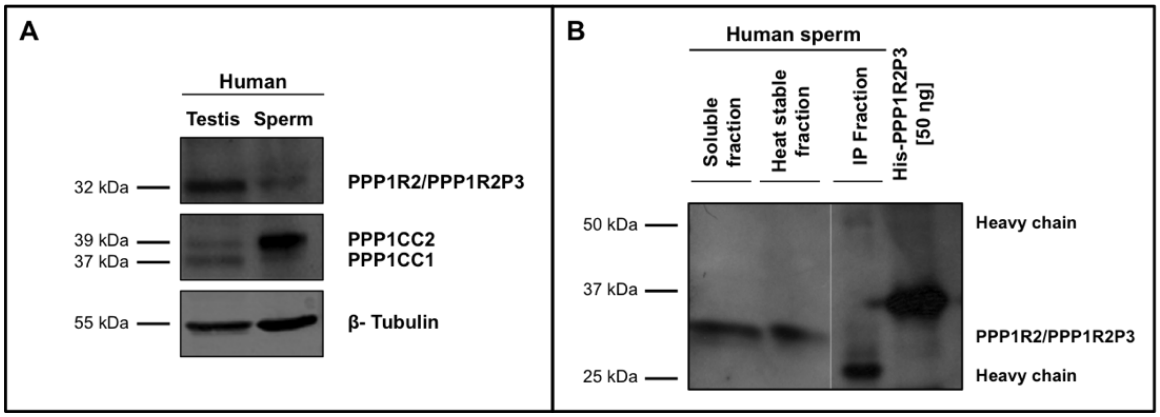


Figure II.A. 5: PPP1R2/PPP1R2P3 is present in human testis and ejaculated sperm. A. *A human testis and sperm screening (50µg) of PPP1R2/PPP1R2P3 and PPP1CC isoforms was performed using sheep anti-PPP1R2 and CBC3C antibodies. Extracts were prepared in 1%SDS. β-Tubulin was used as loading control.* **B.** *PPP1R2/PPP1R2P3 was immunoprecipitated from a human sperm heat stable extract with sheep anti-PPP1R2 antibody followed by Western blotting with the same antibody. Recombinant His-PPP1R2P3 (50ηg) and sperm lysate before boiling (soluble fraction) were used as controls. PPP1R2 was immunoprecipitated using dynabeads protein G. IP, immunoprecipitated.*

To unequivocally confirm the existence of PPP1R2 and the new PPP1R2P3 isoform in human sperm we adopted a innovative strategy, trying to bypass the problem of the aberrant way PPP1R2/PPP1R2P3 migrate in SDS-PAGE and the difficulty of past work to detect this protein in sperm. We performed an immunoprecipitation of PPP1R2/PPP1R2P3 using the PPP1R2 antibodies from five independent human ejaculated sperm samples (four using the sheep anti-PPP1R2 and one using the rabbit anti-PPP1R2) followed by mass spectrometry analysis. Immunoprecipitation was performed using human ejaculate heat stable extracts to further concentrate the protein. We were able to immunoprecipitate sufficient amount of PPP1R2/PPP1R2P3 for detection of protein peptides by mass spectrometry (Table II.A. 1). Using the Orbitrap Velos mass

spectrometer we were able to identify 41 MSMS spectra corresponding to 11 different peptides of PPP1R2/PPP1R2P3. From these, 5 MSMS spectra corresponding to 2 peptides match unequivocally with PPP1R2 and 3 MSMS spectra corresponding to 1 peptide match with PPP1R2P3, with the rest of them matching to both (Table II.A. 1 and Figure II.A. 6). Taking into consideration peptides matching to both and the specific ones for each protein, the sequence coverage obtained was 63.9% for PPP1R2 and 56.6% for PPP1R2P3. Also, the mascot score levels for each protein were near 600 (in addition, spectra were manually evaluated). For the first time both PPP1R2 and PPP1R2P3 were recovered from the same extract of human sperm, using two different antibodies.

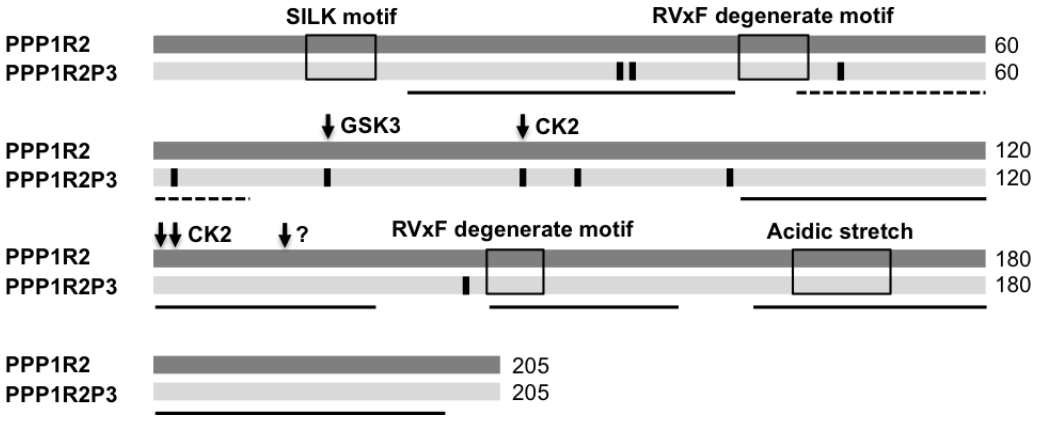


Figure II.A. 6: Schematic representation of the alignment between PPP1R2 and PPP1R2P3 where the peptides obtained in the mass spectrometry are shown. *The sequences for PPP1R2 in dark grey and PPP1R2P3 in light grey are shown in rectangles (see ClustalW2 alignment in figure 1). Boxes represent the important PPP1 binding regions. Black arrows show phosphorylation sites and the respective known kinase is indicated. Black vertical bars represent the residues in PPP1R2P3 that are different from PPP1R2. Black bars below the protein sequences show the coverage of the peptides obtained by mass spectrometry. Black trace bar indicate the peptide that allowed us to distinguish PPP1R2P3 from PPP1R2.*

Table II.A. 1: *Peptides identified by Orbitrap Velos mass spectrometry for PPP1R2, PPP1R2P3 or for both and for PPP1CC2, after immunoprecipitation of human sperm samples with sheep or rabbit anti-PPP1R2 antibodies. aa, amino acids; pI, isoelectric point.*

Protein Name	**Uniprot ID**	**MW (Da)**	**pI**	**Protein size (aa)**	**Coverage**	***Mascot score***
PPP1R2	IPP2_HUMAN	23,015	4.64	205	22.0%	74.12
Peptide		***Range (start-end)***	***Number of spectra***	***m/z meas.***	***z***	***Mascot score***
K.WDEMNILATYHPADKDYGLMK.I		47 - 67	3	848,39	3	47,40
K.TSTTSSMVASAEQPRGNVDEELSK.K		19 - 42	2	847,40	3	26,72
Protein Name	**Uniprot ID**	**MW (Da)**	**pI**	**Protein size (aa)**	**Coverage**	***Mascot score***
PPP1R2P3	IPP2M_HUMAN	23,048	4.82	205	7.3%	48,95
Peptide		***Range (start-end)***	***Number of spectra***	***m/z meas.***	***z***	***Mascot score***
K.WDEINILATYHPADK.G		47 - 61	3	595,96	3	48,95
Protein Name		**MW (Da)**	**pI**	**Protein size (aa)**	**Coverage**	***Mascot score***
Common to both PPP1R2 and PPP1R2P3					49.3%	546.34
Peptide		***Range (start-end)***	***Number of spectra***	***m/z meas.***	***z***	***Mascot score***
R.KLAAAEGLEPK.Y		103 - 113	7	563,83	2	70,38
R.IQEQESSGEEDSDLSPEER.E		116 - 134	5	1082,46	2	117,82
K.LAAAEGLEPK.Y		104 - 113	5	499,78	2	63,21
K.LHYNEGLNIK.L		146 - 155	4	400,88	3	53,17
R.KLHYNEGLNIK.L		145 - 155	4	664,87	2	51,25
K.LHYNEGLNIKLAR.Q		146 - 158	2	514,29	3	23,57
R.IQEQESSGEEDSDLSPEEREK.K		116 - 136	2	807,69	3	59,32
K.DLHDDDEDEEMLETADGESMNTEESNQG STPSDQQQNK.L		164 - 201	1	1434,23	3	30,63
K.YRIQEQESSGEEDSDLSPEER.E		114 - 134	1	828,70	3	11,42
K.LAAAEGLEPKYR.I		104 - 115	1	439,91	3	15,97
K.TSTTSSMVASAEQPR.G		19 - 33	1	784,86	2	49,60
Protein Name	**Uniprot ID**	**MW (Da)**	**pI**	**Protein size (aa)**	**Coverage**	***Mascot score***
PPP1CC2	PP1G_HUMAN	38,518	5.78	337	15%	252,57
Peptide		***Range (start-end)***	***Number of spectra***	***m/z meas.***	***z***	***Mascot score***
R.VASGLNPSIQK.A		315 - 325	5	557,32	2	54,76
R.GVSFTFGAEVVAK.F		222 - 234	2	656,35	2	72,94
K.NVQLQENEIR.G		27 - 36	2	621,83	2	42,11
K.YPENFFLLR.G		114 - 122	1	599,82	2	39,6
K.LNIDSIIQR.L		7 - 15	1	536,31	2	43,16

Mass spectrometry data of the same immunoprecipitates showed that PPP1CC2 was also present. Five peptides out of 10 MSMS spectra were identified that matched

PPP1CC2 with 15% coverage and a protein MASCOT score higher than 200 (manually checked). This indicates that PPP1CC2 binds to PPP1R2/PPP1R2P3 in human spermatozoa (Table Table II.A. 1).

Phosphorylation of Thr73 by GSK3 renders the PPP1R2/PPP1C complex active and this phosphorylation is enhanced when Ser87, 121 and 122 are also phosphorylated by CK2, particularly Ser87 [32, 40]. Based on past biochemical studies, the current hypothesis is that PPP1R2 is phosphorylated at Thr73 leading to activation of the PPP1R2/PPP1CC2 complex and sperm immotility in caput epididymis. In contrast in cauda epididymis and in ejaculated sperm the PPP1 complex is inactive leading to a motile sperm [3, 5].

In extracts from ejaculated sperm, three independent peptides show post-translational modifications, identified as phosphorylations at Ser121, Ser122 and Ser127. However, they could not be assigned to either PPP1R2 or PPP1R2P3 given the fact that both proteins are identical in that region. To our knowledge, this is the first time that a phosphorylation in Ser127 is reported for PPP1R2/PPP1R2P3. No tyrosine or threonine phosphorylations were identified. The peptides obtained with phosphorylations correspond to only 7% of the total number of obtained peptides strongly supporting the fact that in the majority of the PPP1R2/PPP1R2P3 are not phosphorylated, in accordance to the model proposed above.

Peptides have already been misassigned to PPP1R2P3 protein by mass spectrometry in two independent reports [56, 57]. Protein acetylation in Ala2 and two phosphorylations in Ser121 and Ser122 were observed by high throughput MS analysis in human Jurkat T cell leukemia and embryonic kidney (HEK293) cell lines [56, 57]. In Figure II.A. 1 and Figure II.A. 6 it is seen that in these regions (Ala2, Ser121 and Ser122) both PPP1R2 and PPP1R2P3 are similar and so, the peptides assigned to PPP1R2P3 may well be from PPP1R2. So, for the first time we have shown unequivocally that PPP1R2P3 pseudogene is in fact transcribed and translated, being present in human mature sperm.

PPP1R2 subcellular-localization in human spermatozoa

To further pursue the PPP1R2/PPP1R2P3 presence in sperm we studied its subcellular localization and co-localization with PPP1CC2 in human mature spermatozoa. We performed immunocytochemistry experiments using the sheep anti-PPP1R2 and CBC502 antibodies (Figure II.A. 7).

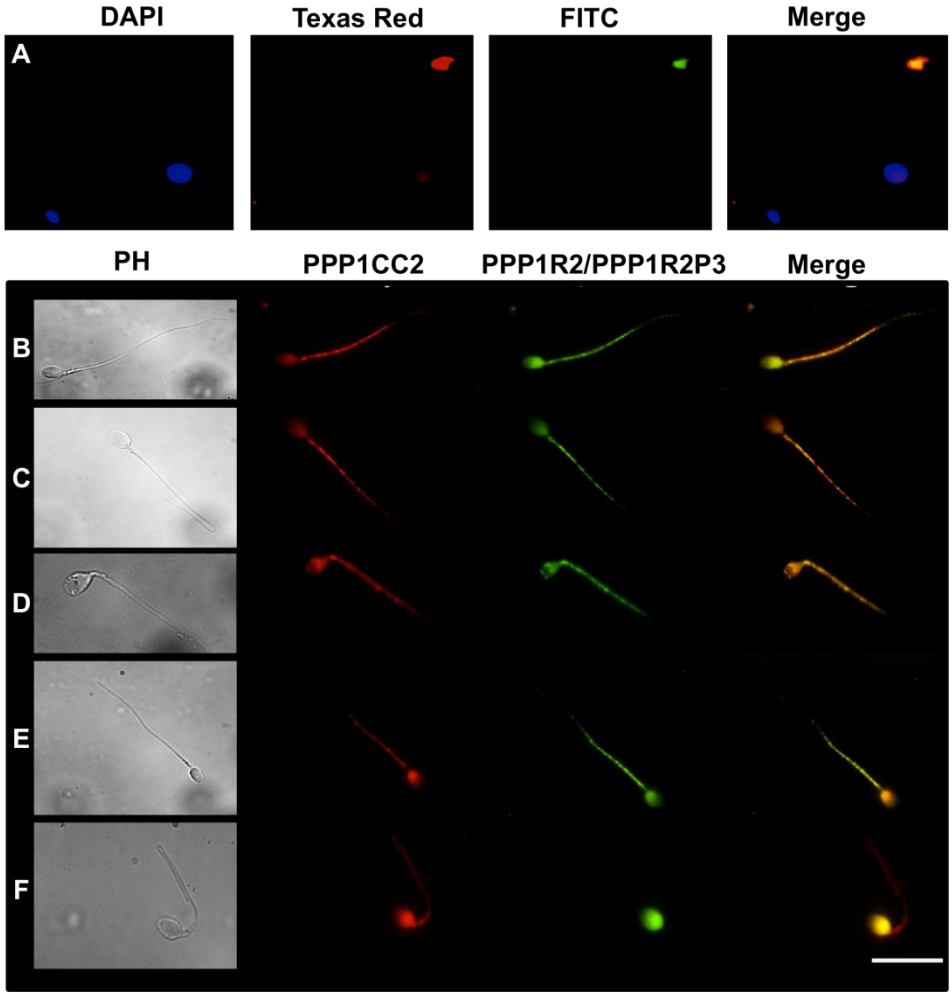


Figure II.A. 7: Co-localization of PPP1CC2 and PPP1R2/PPP1R2P3 in morphologically normal and abnormal spermatozoa. *Human spermatozoa were labeled with rabbit anti-PPP1CC2 and sheep anti-PPP1R2 antibodies, and specific secondary antibodies conjugated with Texas-Red and FITC fluorophores, respectively.* **A.** *Negative control using only the fluorescence labeled secondary antibodies.* **B.** *Normal spermatozoa.* **C.** *Abnormal tail spermatozoa.* **D.** *Abnormal mid-piece spermatozoa.* **E.** *Abnormal head spermatozoa.* **F.** *Multiple abnormalities spermatozoa. Phase contrast (PH). Scale bar=20µm.*

Results showed that PPP1R2/PPP1R2P3 are present along the flagellum, in the mid-piece, principal piece, except the end-piece, and also in the head, more specifically in the equatorial and post-acrosomal regions (Figure II.A. 7B). PPP1CC2 localization is also similar, co-localizing to the same regions as seen with PPP1R2. Furthermore, a negative control using only secondary antibodies showed that this localization pattern is specific (Figure II.A. 7A). The PPP1CC2 localization is also corroborated by a previous report [6] and the localization is consistent with its role in sperm forward and hyperactivated motility and possibly in acrosome reaction [3-5, 58-60]. We also checked if in sperm with abnormalities the sub-localization patterns of both proteins remained the same (Figure II.A. 7 C-F). In sperm with only one type of morphological defect either in the tail, mid-

piece or head, no changes were observed for both proteins [61]. However in sperm with multiple abnormalities it appears that PPP1R2/PPP1R2P3 is relocated to the head, which does not happen with PPP1CC2 expression pattern. Although this observation needs further examination, this differential staining could be the basis for a novel biomarker for sperm with multiple defects (Figure II.A. 7, panel D-G). To further reconfirm the immunolocalization data, RIPA soluble and insoluble extracts were resolved by SDS-PAGE (Figure II.A. 8A). The results showed that PPP1R2/PPP1R2P3 is mainly present in the soluble extracts. Although RIPA lysis buffer is a stringent buffer we still see some PPP1R2 protein in the insoluble fraction. In contrast, PPP1CC2 is present in both fractions but more abundantly in the pellet fraction. These results show that PPP1CC2 is tightly attached to the axoneme.

Furthermore tail and head preparations were also made using a sucrose gradient. In this separation although heads are kept intact, tails are demembranated and only the axoneme structure is maintained, thus all soluble proteins present in the tail are removed. By Western blotting we demonstrate that both PPP1R2/PPP1R2P3 and PPP1CC2 are present in tail and head and in similar amounts (Figure II.A. 8B). This result is important to show that indeed PPP1R2/PPP1R2P3 is associated with the tail, probably with PPP1CC2, but its binding could be abolished using RIPA buffer. That is PPP1CC2 is probably bound to the axoneme along with PPP1R2/PPP1R2P3. A previous report has shown that PPP1 is anchored to *Chlamydomonas* axoneme in the central pair apparatus, associated with the C1 microtubule and at less extent to the outer doublet microtubules, suggesting that PPP1 can control both dynein arms and thereby flagellar motility [62]. This data correlates with what was obtained from the immunocytochemistry and Western blot results (Figure II.A. 7 and Figure II.A. 8), where both proteins were shown to be present in head and tail.

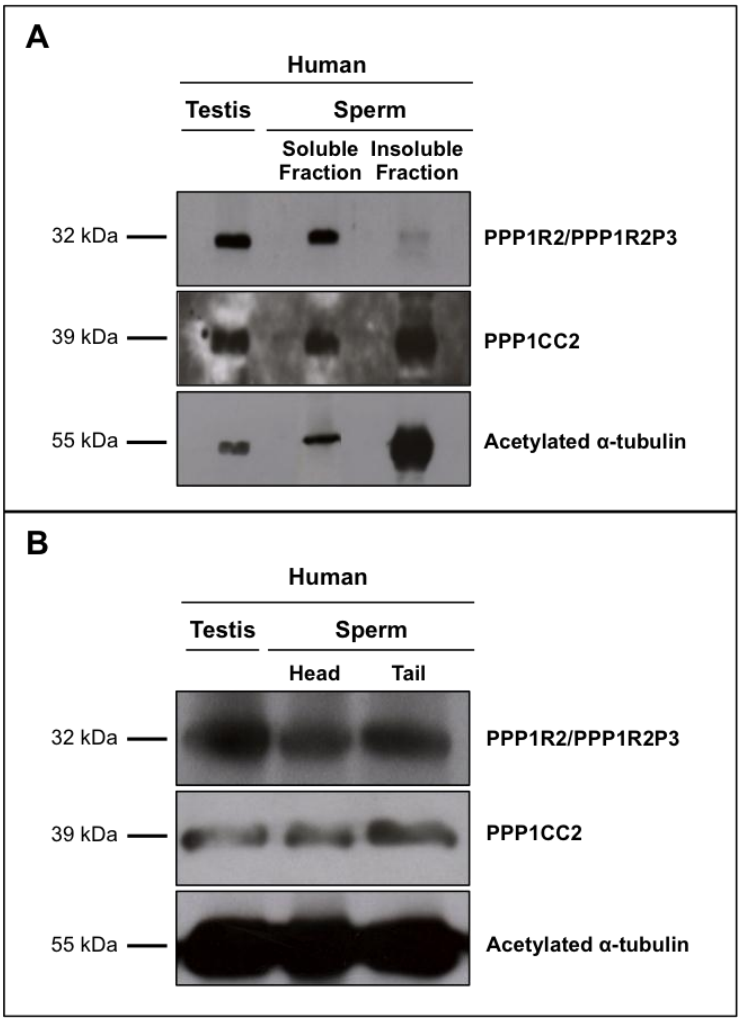


Figure II.A. 8: PPP1R2/PPP1R2P3 are present in human testis and sperm extracts. A. *PPP1R2/ PPP1R2P3 and PPP1CC2 presence in the RIPA supernatant (soluble fraction) and pellet (insoluble fraction) sperm lysates has been performed using sheep anti-PPP1R2 and rabbit anti-PPP1CC2 antibodies. 100µg of both fractions were loaded in a SDS-PAGE gel. Acetylated α-tubulin was used as loading control.* **B.** *A human sperm sample (2x10^8 sperm cells) was sonicated to disrupt head and tail bond, and subjected to a sucrose gradient. Both pools were verified by phase contrast (PH). 100µg of both fractions were loaded in a SDS-PAGE gel. After transfer, blot was probed with sheep anti-PPP1R2 and rabbit anti-PPP1CC2 antibodies. Acetylated α-tubulin was used as loading control.*

Phosphorylation of PPP1R2/PPP1R2P3 in Human sperm

Western blot of human sperm heat stable extract (hsPPP1R2/PPP1R2P3) revealed, depending on the samples, two bands migrating above both purified PPP1R2 and PPP1R2P3 (Figure II.A. 9A).

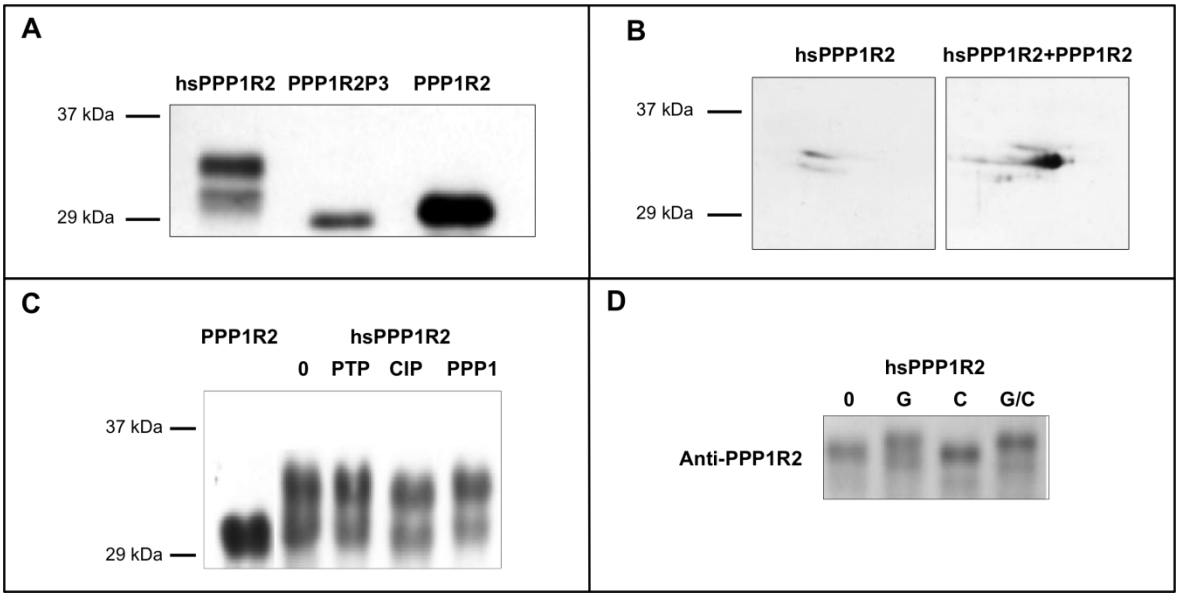


Figure II.A. 9: Western blot analysis of endogenous PPP1R2/PPP1R2P3 from human sperm.

A. *Comparison of endogenous human sperm PPP1R2/PPP1R2P3, with recombinant PPP1R2 and PPP1R2P3, using sheep anti-PPP1R2 antibody. hsPPP1R2, heat stable human sperm extract. Recombinant PPP1R2P3 and PPP1R2 are shown as positive controls.* **B.** *Anti-PPP1R2 Western blot of 2D separation of hsPPP1R2 human sperm extract. Left panel, heat stable human sperm extract; right panel, heat stable human sperm extract supplemented with recombinant PPP1R2.* **C.** *hsPPP1R2 was incubated in the presence of different phosphatases, in the respective buffers at 30ºC for 3hrs. 0, no phosphatase, control; PTP, incubation with Protein Tyrosine Phosphatase 1B; CIP, incubation with Calf Intestinal Phosphatase; PPP1C, incubation with PPP1CC1. Recombinant PPP1R2 is shown as positive control.* **D.** *hsPPP1R2 was incubated in the presence of GSK3 (G) or CK2 (C), or both (G/C), as previously described. Resulting changes in protein migration, probably reflecting alterations in phosphorylation, were detected by Western blot analysis using the sheep anti-PPP1R2 antibody.*

In order to investigate the nature of those bands, human sperm heat stable extract was resolved by 2D-PAGE revealing three isolated spots of different molecular weight and pI, suggesting the existence of different phosphorylated forms of PPP1R2/PPP1R2P3 (Figure II.A. 9B). Besides threonine and serine, PPP1R2 can also be tyrosine phosphorylated [63]. We attempted to dephosphorylate the human sperm heat extract with protein tyrosine phosphatase 1B (PTP), calf intestinal phosphatase (CIP) and PPP1CC1. Under the conditions used, only CIP, a nonspecific phosphatase, could dephosphorylate the hsPPP1R2/PPP1R2P3, which resulted in an increase in electrophoretic mobility. This band did not run at the same molecular weight, as recombinant PPP1R2 (Figure II.A. 9C). When hsPPP1R2/PPP1R2P3 was incubated with GSK3 or CK2 or both, only GSK3 was able to induce a mobility shift (Figure II.A. 9D). Taken together, these results suggest that PPP1R2/PPP1R2P3 in sperm could be phosphorylated at serine residues, probably by CK2 in Ser121, Ser122 and Ser127.

Conclusion

The current hypothesis for sperm motility acquisition states that a PPP1R2-like/PPP1CC2 complex controls the sperm motility and is in turn regulated by a mechanism of reversible phosphorylation by GSK3 and an unknown phosphatase [3, 5] (Figure II.A. 10, left panel). Nevertheless, the PPP1R2-like activity in human sperm was not identified.

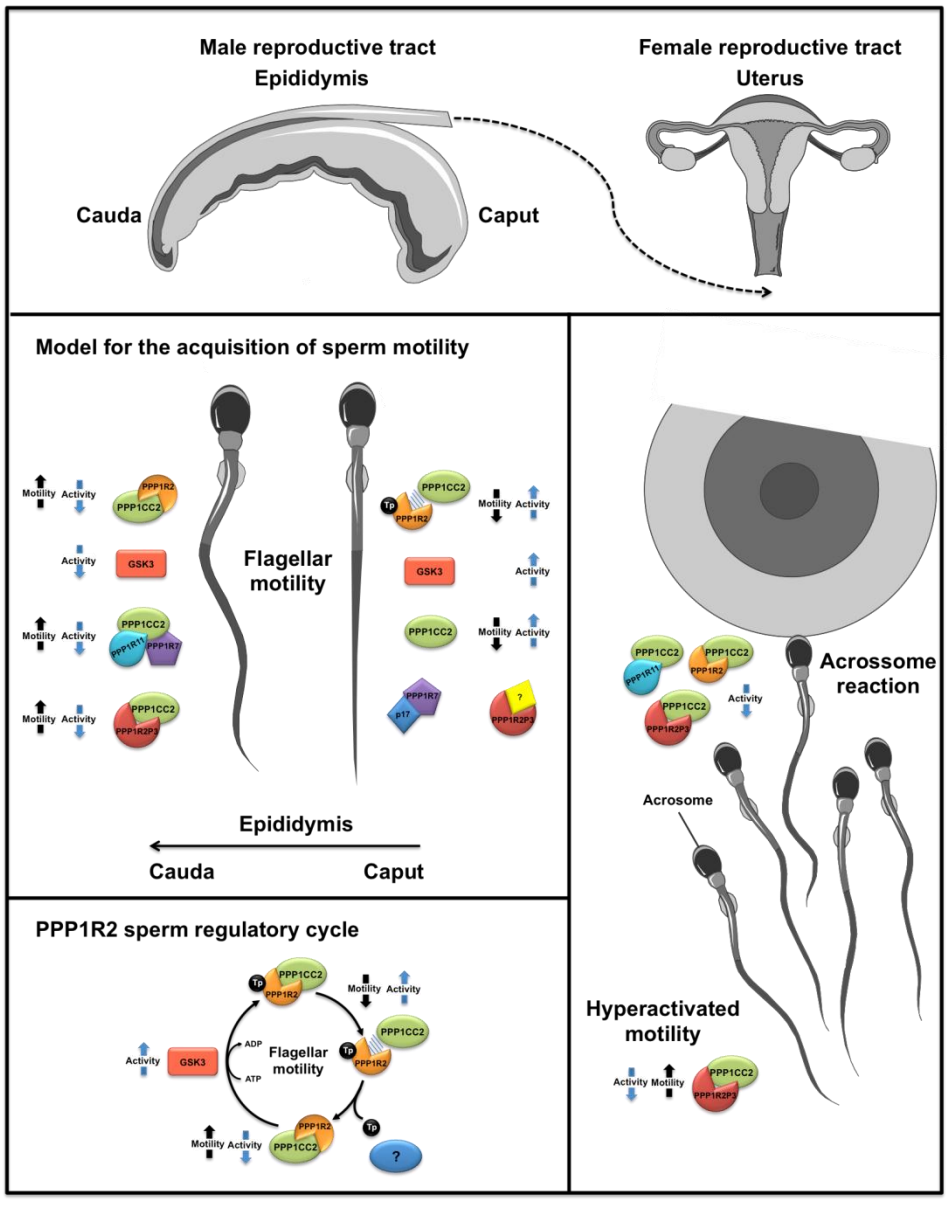


Figure II.A. 10: Novel model for sperm motility acquisition based on PPP1R2 and PPP1R2P3. *Sperm cells mature in the epididymis in order to be able to enter in the female reproductive tract and fertilize the oocyte (upper panel). In the first segment of the epididymis (caput), the sperm is immotile due to the activation of the PPP1R2/PPP1CC2 complex by GSK3, while in the last segment of the epididymis (cauda), PPP1R2P3 substitutes PPP1R2 as an irreversible inhibitor and sperm acquires motility. Since sperm cells are terminally differentiated and essentially devoid of*

transcriptional and translational activity, we assume that in the caput region, PPP1R2P3 is somehow bound to a protein (yellow lozenge) that keeps it from binding to PPP1CC2. PPP1R2/PPP1CC2 complex controls sperm motility and is in turn regulated by a mechanism of reversible phosphorylation by GSK3 and an unknown phosphatase. This complex is operative in axoneme and regulates flagellar motility (left lower panel). PPP1R2P3 may also be involved in the regulation of hyperactivated motility and in the acrosome reaction in the female reproductive tract (right, lower panel). Other PPP1CC2 complexes exist in sperm [2]. Other PPP1C inhibitors exist in sperm forming dimers and even trimers, for example PPP1R11/PPP1R7/PPP1CC2 [70]. PPP1R7 is mainly soluble and is present in the principal piece and head-neck junction of spermatozoa bound to an inhibiting PPP1CC2 [2, 6] and, PPP1R11 exists in tail and head of spermatozoa [2, 71].

Here we show, by mass spectrometry, that PPP1R2 is present in human sperm (Table II.A. 1). Also, peptides for PPP1CC2 were recovered in the same immunoprecipitation extracts, meaning that PPP1R2/PPP1CC2 may in fact form a complex in sperm (Table II.A. 1). We also demonstrated by immunocytochemistry and Western blot that PPP1R2 and PPP1CC2 co-localize in almost all the extent of the spermatozoon and in the same structures (Figure II.A. 5, Figure II.A. 7 and Figure II.A. 8). This strongly supports the hypothesis that the PPP1R2-like activity, identified several years ago might be in fact be assigned to PPP1R2 [3, 5].

Furthermore, we have also identified by mass spectrometry a second PPP1R2-like protein (Table II.A. 1), termed PPP1R2P3 that was consider to be a pseudogene in many databases.

Processed pseudogenes, like PPP1R2P3, are intronless, with polyA traits in the C-terminal, no parental promoter, and are integrated in the genome in a new location. Further, since they were originated from a retrotransposition activity, they have truncated 5'UTR, due to the low processivity of the reverse transcriptase, and direct repeats in both ends [64]. Testis is an organ where many pseudogenes are expressed and the gene products have been shown to actively participate in spermatogenesis or other germ cell functions [65-67]. This happens because the transcription in testis, compared to other somatic tissues, is not as tightly regulated due to hyper transcription rates, which could lead to unbiased activation of otherwise imperfect or weak promoters [67, 68]. We have also identified PPP1R2P3 in a testis cDNA library [24]. Also, the apparent function of pseudogenes in testis germ cells could be a way to facilitate the creation of new genes from the parental ones [65]. We have shown that PPP1R2P3 pseudogene is in fact transcribed and translated, being present in the mature human sperm cell (Table II.A. 1). PP1R2P3 protein binds to PPP1CC isoforms and is a heat stable inhibitor (Fig.2 and 3) that cannot be phosphorylated by GSK3, being therefore a putatively irreversible inhibitor of PPP1C (Fig.4). Also, PPP1R2P3 was phosphorylated by CK2 *in vitro*, probably in

residues 121 and 122, as shown by mass spectrometry (Figure II.A. 4 and Table II.A. 1). Regarding the PPP1R2P3 specific localization in testis/sperm and co-localization with PPP1CC2 we cannot distinguish from that of PPP1R2, since antibodies we have detect both PPP1R2P3 and PPP1R2 (Figure II.A. 5 and Figure II.A. 7 and Figure II.A. 8).

We propose a new model in which in the first segments of the epididymis (caput) the sperm is immotile due to the activation of the PPP1R2/PPP1CC2 complex by GSK3, while in the last segment of the epididymis (cauda), PPP1R2P3 takes place and substitutes PPP1R2 as an irreversible inhibitor of PPP1CC2 triggering sperm motility.

Since sperm are terminally differentiated cells essentially devoid of transcriptional and translational activity, we could not take into consideration the regulation of the expression of both proteins. So, it is possible that PPP1R2P3 is somehow bound another protein that may keep it from binding to PPP1CC2 in immotile caput sperm, in a manner similar to what is suggest to occur with sds22 (PPP1R7) in caput sperm [69] (Figure II.A. 10, left middle panel).

This mechanism of controlling sperm motility by PPP1CC2 and GSK3 seems to have evolved in mammals, since other vertebrate classes do not have this alternatively spliced phosphatase [49]. We suggest that the occurrence of PPP1R2P3 protein, which is only present in primates, seems to lead to a novel regulatory mechanism of PPP1CC2 activity that might have evolved in this order.

In addition, cauda sperm motility is further stimulated by incubation with the same phosphatase inhibitors, okadaic acid and calyculin A, mimicking hyperactivated motility [3, 5]. This raises another interesting possibility in which the initial acquisition of motility along the epididymis is due to the inhibition of a PPP1CC2 pool by PPP1R2, whereas, a second PPP1CC2 pool may be inhibited by PPP1R2P3 leading to hyperactivated motility in the female tract (Figure II.A. 10, right bottom panel).

The mechanisms described above may be occurring simultaneously, since peptides specific for both proteins have been identified in ejaculated sperm. What determine the levels of motility regulated by those processes may be the relative amounts available and the activity of kinases and phosphatases involved in the regulatory process. The final spermatozoon motion and velocity is the result of several complex events, but many involve PPP1CC2 and its PIPs [2].

Acknowledgements

This work was supported by the Centre for Cell Biology of the University of Aveiro, by grants from FCT of the Portuguese Ministry of Science and Higher Education to LKMG (SFRH/BD/42334/2007), WW (SFRH/BD/6879/2001) and EFCS (POCI/SAU-OBS/57394/2004; PPCDT/SAU-OBS/57394/2004), by Re-equipment Grant (REEQ/1023/BIO/2005) and by CRUP (E-92/08;B-32/09).

M Fardilha dedicates this paper to the memory of her mentor ECS (Edgar da Cruz e Silva, deceased) who believed that PPP1R2 and PPP1R2P3 proteins were the key to explain sperm motility initiation, but unfortunately could not prove his theory during his lifetime. Collaborators (D. L. Brautigan and S. Vijayaraghavan) worked with ECS on PPP1 and PPP1R2 for many years.

References

1. Urner F, Sakkas D. Protein phosphorylation in mammalian spermatozoa. Reproduction 2003; 125: 17-26.
2. Fardilha M, Esteves SLC, Korrodi-Gregório L, Pelech S, da Cruz e Silva OAB, da Cruz e Silva E. Protein phosphatase 1 complexes modulate sperm motility and present novel targets for male infertility. Molecular Human Reproduction 2011; 17: 466-477.
3. Vijayaraghavan S, Stephens DT, Trautman K, Smith GD, Khatra B, da Cruz e Silva EF, Greengard P. Sperm motility development in the epididymis is associated with decreased glycogen synthase kinase-3 and protein phosphatase 1 activity. Biol Reprod 1996; 54: 709-718.
4. Smith GD, Wolf DP, Trautman KC, Vijayaraghavan S. Motility potential of macaque epididymal sperm: the role of protein phosphatase and glycogen synthase kinase-3 activities. J Androl 1999; 20: 47-53.
5. Smith GD, Wolf DP, Trautman KC, da Cruz e Silva EF, Greengard P, Vijayaraghavan S. Primate sperm contain protein phosphatase 1, a biochemical mediator of motility. Biol Reprod 1996; 54: 719-727.
6. Huang Z, Khatra B, Bollen M, Carr DW, Vijayaraghavan S. Sperm PP1gamma2 is regulated by a homologue of the yeast protein phosphatase binding protein sds22. Biol Reprod 2002; 67: 1936-1942.
7. Varmuza S, Jurisicova A, Okano K, Hudson J, Boekelheide K, Shipp EB. Spermiogenesis is impaired in mice bearing a targeted mutation in the protein phosphatase 1cgamma gene. Dev Biol 1999; 205: 98-110.
8. Chakrabarti R, Kline D, Lu J, Orth J, Pilder S, Vijayaraghavan S. Analysis of Ppp1cc-Null Mice Suggests a Role for PP1gamma2 in Sperm Morphogenesis. Biology of Reproduction 2007; 76: 992-1001.
9. Ashizawa K, Wishart GJ, Tomonaga H, Nishinakama K, Tsuzuki Y. Presence of protein phosphatase type 1 and its involvement in temperature-dependent flagellar movement of fowl spermatozoa. FEBS Letters 1994; 350: 130-134.
10. Ashizawa K, Magome A, Tsuzuki Y. Stimulation of motility and respiration of intact fowl spermatozoa by calyculin A, a specific inhibitor of protein phosphatase-1 and -2A, via a Ca2+-dependent mechanism. Journal of Reproduction and Fertility 1995; 105: 109-114.
11. Klumpp S, Cohen P, Schultz JE. Okadaic acid, an inhibitor of protein phosphatase 1 in Paramecium, causes sustained Ca2(+)-dependent backward swimming in response to depolarizing stimuli. EMBO J 1990; 9: 685-689.
12. Klumpp S, Schultz JE. Identification of a 42 kDa protein as a substrate of protein phosphatase 1 in cilia from Paramecium. FEBS Letters 1991; 288: 60-64.
13. Habermacher G, Sale WS. Regulation of dynein-driven microtubule sliding by an axonemal kinase and phosphatase in Chlamydomonas flagella. Cell Motil Cytoskeleton 1995; 32: 106-109.
14. Habermacher G, Sale WS. Regulation of flagellar dynein by an axonemal type-1 phosphatase in Chlamydomonas. Journal of Cell Science 1996; 109: 1899-1907.
15. Habermacher G, Sale WS. Regulation of flagellar dynein by phosphorylation of a 138-kD inner arm dynein intermediate chain. J Cell Biol 1997; 136: 167-176.
16. Bollen M, Peti W, Ragusa MJ, Beullens M. The extended PP1 toolkit: designed to create specificity. Trends in biochemical sciences 2010; 35: 450-458.
17. Fardilha M, Esteves SLC, Korrodi-Gregorio L, Silva OABdCe, Silva EFdCe. The Physiological Relevance of Protein Phosphatase 1 and its Interacting Proteins to Health and Disease. Current Medicinal Chemistry 2010; 17.

18. Hemmings BA, Resink TJ, Cohen P. Reconstitution of a Mg-ATP-dependent protein phosphatase and its activation through a phosphorylation mechanism. FEBS Letters 1982; 150: 319-324.
19. Somanath PR, Jack SL, Vijayaraghavan S. Changes in sperm glycogen synthase kinase-3 serine phosphorylation and activity accompany motility initiation and stimulation. J Androl 2004; 25: 605-617.
20. Vijayaraghavan S, Mohan J, Gray H, Khatra B, Carr DW. A Role for Phosphorylation of Glycogen Synthase Kinase-3α in Bovine Sperm Motility Regulation. Biology of Reproduction 2000; 62: 1647-1654.
21. Sanseau P, Jackson A, Alderton RP, Beck S, Senger G, Sheer D, Kelly A, Trowsdale J. Cloning and characterization of human phosphatase inhibitor-2 (IPP-2) sequences. Mamm Genome 1994; 5: 490-496.
22. Browne GJ, Fardilha M, Oxenham SK, Wu W, Helps NR, da Cruz ESOA, Cohen PT, da Cruz ESEF. SARP, a new alternatively spliced protein phosphatase 1 and DNA interacting protein. Biochem J 2007; 402: 187-196.
23. Fardilha M, Wu W, Sa R, Fidalgo S, Sousa C, Mota C, da Cruz e Silva OA, da Cruz e Silva EF. Alternatively spliced protein variants as potential therapeutic targets for male infertility and contraception. Ann N Y Acad Sci 2004; 1030: 468-478.
24. Fardilha M, Esteves SLC, Korrodi-Gregório L, Vintém AP, Domingues SC, Rebelo S, Morrice N, Cohen PTW, Silva OABdCe, Silva EFdCe. Identification of the human testis protein phosphatase 1 interactome. Biochemical Pharmacology 2011; 82: 1403-1415.
25. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, et al. Clustal W and Clustal X version 2.0. Bioinformatics 2007; 23: 2947-2948.
26. Zhang ZJ, Bai G, Shima M, Zhao SM, Nagao M, Lee EYC. Expression and Characterization of Rat Protein Phosphatases-1α, -1γ1, -1γ2, and -1δ. Archives of Biochemistry and Biophysics 1993; 303: 402-406.
27. Helps NR, Street AJ, Elledge SJ, Cohen PTW. Cloning of the complete coding region for human protein phosphatase inhibitor 2 using the two hybrid system and expression of inhibitor 2 in E. coli. FEBS Letters 1994; 340: 93-98.
28. Mota C. Identification of PP1γ2 interacting proteins from human testis using the YTH system. Universidade de Aveiro 2003.
29. Watanabe T, da Cruz e Silva EF, Huang HB, Starkova N, Kwon YG, Horiuchi A, Greengard P, Nairn AC. Preparation and characterization of recombinant protein phosphatase 1. Methods Enzymol 2003; 366: 321-338.
30. da Cruz e Silva EF, Fox CA, Ouimet CC, Gustafson E, Watson SJ, Greengard P. Differential expression of protein phosphatase 1 isoforms in mammalian brain. J Neurosci 1995; 15: 3375-3389.
31. Kim Y-M, Watanabe T, Allen PB, Kim Y-M, Lee S-J, Greengard P, Nairn AC, Kwon Y-G. PNUTS, a Protein Phosphatase 1 (PP1) Nuclear Targeting Subunit. Journal of Biological Chemistry 2003; 278: 13819-13828.
32. Park IK, Roach P, Bondor J, Fox SP, DePaoli-Roach AA. Molecular mechanism of the synergistic phosphorylation of phosphatase inhibitor-2. Cloning, expression, and site-directed mutagenesis of inhibitor-2. Journal of Biological Chemistry 1994; 269: 944-954.
33. WorldHealthOrganization. WHO laboratory manual for the examination and processing of human semen, Organization WH, Editor 2010, WHO press: Geneva, Switzerland.
34. Thiele H, Glandorf J, Hufnagel P. Bioinformatics strategies in life sciences: from data processing and data warehousing to biological knowledge extraction. J Integr Bioinform 2010; 7: 141.

35. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. ELECTROPHORESIS 1999; 20: 3551-3567.
36. Gerhard DS, Wagner L, Feingold EA, Shenmen CM, Grouse LH, Schuler G, Klein SL, Old S, Rasooly R, Good P, Guyer M, Peck AM, et al. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). Genome Res 2004; 14: 2121-2127.
37. Yang J, Hurley TD, DePaoli-Roach AA. Interaction of Inhibitor-2 with the Catalytic Subunit of Type 1 Protein Phosphatase. Journal of Biological Chemistry 2000; 275: 22635-22644.
38. Hurley TD, Yang J, Zhang L, Goodwin KD, Zou Q, Cortese M, Dunker AK, DePaoli-Roach AA. Structural basis for regulation of protein phosphatase 1 by inhibitor-2. J Biol Chem 2007; 282: 28874-28883.
39. Bollen M, Stalmans W. The structure, role, and regulation of type 1 protein phosphatases. Crit Rev Biochem Mol Biol 1992; 27: 227-281.
40. DePaoli-Roach AA. Synergistic phosphorylation and activation of ATP-Mg-dependent phosphoprotein phosphatase by F A/GSK-3 and casein kinase II (PC0.7). J Biol Chem 1984; 259: 12144-12152.
41. Takizawa N, Mizuno Y, Ito Y, Kikuchi K. Tissue distribution of isoforms of type-1 protein phosphatase PP1 in mouse tissues and its diabetic alterations. J Biochem 1994; 116: 411-415.
42. Andreassen PR, Lacroix FB, Villa-Moruzzi E, Margolis RL. Differential Subcellular Localization of Protein Phosphatase-1 α, γ1, and δ Isoforms during Both Interphase and Mitosis in Mammalian Cells. The Journal of Cell Biology 1998; 141: 1207-1215.
43. Osawa Y, Nakagama H, Shima H, Sugimura T, Nagao M. Identification and Characterization of three Isotypes of Protein Phosphatase Inhibitor-2 and their Expression Profiles during Testis Maturation in Rats. European Journal of Biochemistry 1996; 242: 793-798.
44. Dancheck B, Nairn AC, Peti W. Detailed Structural Characterization of Unbound Protein Phosphatase 1 Inhibitors†. Biochemistry 2008; 47: 12346-12356.
45. Li M, Satinover DL, Brautigan DL. Phosphorylation and functions of inhibitor-2 family of proteins. Biochemistry 2007; 46: 2380-2389.
46. Gruppuso PA, Johnson GL, Constantinides M, Brautigan DL. Phosphorylase phosphatase regulatory subunit. "Western" blotting with immunoglobulins against inhibitor-2 reveals a protein of Mr = 60,000. Journal of Biological Chemistry 1985; 260: 4288-4294.
47. Roach P, Roach PJ, DePaoli-Roach AA. Phosphoprotein phosphatase inhibitor-2. Identification as a species of molecular weight 31,000 in rabbit muscle, liver, and other tissues. Journal of Biological Chemistry 1985; 260: 6314-6317.
48. Reddy P, Ernst VG. Partial purification and characterization of heat stable protein phosphatase inhibitor-2 from rabbit reticulocytes. Biochemical and Biophysical Research Communications 1983; 114: 1089-1096.
49. Chakrabarti R, Cheng L, Puri P, Soler D, Vijayaraghavan S. Protein phosphatase PP1[gamma]2 in sperm morphogenesis and epididymal initiation of sperm motility. Asian J Androl 2007; 9: 445-452.
50. Eto M, Bock R, Brautigan DL, Linden DJ. Cerebellar Long-Term Synaptic Depression Requires PKC-Mediated Activation of CPI-17, a Myosin/Moesin Phosphatase Inhibitor. Neuron 2002; 36: 1145-1158.
51. Terry-Lorenzo RT, Carmody LC, Voltz JW, Connor JH, Li S, Smith FD, Milgram SL, Colbran RJ, Shenolikar S. The neuronal actin-binding proteins, neurabin I and neurabin II, recruit specific isoforms of protein phosphatase-1 catalytic subunits. J Biol Chem 2002; 277: 27716-27724.

52. Satinover DL, Leach CA, Stukenberg PT, Brautigan DL. Activation of Aurora-A kinase by protein phosphatase inhibitor-2, a bifunctional signaling protein. Proceedings of the National Academy of Sciences of the United States of America 2004; 101: 8625-8630.
53. Swain JE, Ding J, Brautigan DL, Villa-Moruzzi E, Smith GD. Proper Chromatin Condensation and Maintenance of Histone H3 Phosphorylation During Mouse Oocyte Meiosis Requires Protein Phosphatase Activity. Biology of Reproduction 2007; 76: 628-638.
54. Wang W, Stukenberg PT, Brautigan DL. Phosphatase Inhibitor-2 Balances Protein Phosphatase 1 and Aurora B Kinase for Chromosome Segregation and Cytokinesis in Human Retinal Epithelial Cells. Molecular Biology of the Cell 2008; 19: 4852-4862.
55. Li M, Stukenberg PT, Brautigan DL. Binding of Phosphatase Inhibitor-2 to Prolyl Isomerase Pin1 Modifies Specificity for Mitotic Phosphoproteins†. Biochemistry 2007; 47: 292-300.
56. Mayya V, Han DK. Phosphoproteomics by mass spectrometry: insights, implications, applications and limitations. Expert Review of Proteomics 2009; 6: 605-618.
57. Gauci S, Helbig AO, Slijper M, Krijgsveld J, Heck AJR, Mohammed S. Lys-N and Trypsin Cover Complementary Parts of the Phosphoproteome in a Refined SCX-Based Approach. Analytical Chemistry 2009; 81: 4493-4501.
58. Si Y. Hyperactivation of Hamster Sperm Motility by Temperature-Dependent Tyrosine Phosphorylation of an 80-kDa Protein. Biology of Reproduction 1999; 61: 247-252.
59. Si Y, Okuno M. Role of Tyrosine Phosphorylation of Flagellar Proteins in Hamster Sperm Hyperactivation. Biology of Reproduction 1999; 61: 240-246.
60. Ashizawa K, Wishart GJ, Katayama S, Takano D, Ranasinghe AR, Narumi K, Tsuzuki Y. Regulation of acrosome reaction of fowl spermatozoa: evidence for the involvement of protein kinase C and protein phosphatase-type 1 and/or -type 2A. Reproduction 2006; 131: 1017-1024.
61. WorldHealthOrganization. In: WHO Laboratory for examination of human semen and sperm-cervical mucus interaction, in *Collection and examination of human semen*1999, Cambridge University Press: Cambridge.4-30.
62. Yang P, Fox L, Colbran R, Sale W. Protein phosphatases PP1 and PP2A are located in distinct positions in the Chlamydomonas flagellar axoneme. J Cell Sci 2000; 113: 91-102.
63. Williams JP, Jo H, Hunnicutt RE, Brautigan DL, McDonald JM. Tyrosine phosphorylation of phosphatase inhibitor 2. J Cell Biochem 1995; 57: 415-422.
64. Rouchka EC, Cha IElizabeth. Current Trends in Pseudogene Detection and Characterization. Current Bioinformatics; 4: 112-119.
65. Kleene KC, Mulligan E, Steiger D, Donohue K, Mastrangelo M-A. The Mouse Gene Encoding the Testis-Specific Isoform of Poly(A) Binding Protein (<i>Pabp2</i>) Is an Expressed Retroposon: Intimations That Gene Expression in Spermatogenic Cells Facilitates the Creation of New Genes. Journal of Molecular Evolution 1998; 47: 275-281.
66. Marques AC, Dupanloup I, Vinckenbosch N, Reymond A, Kaessmann H. Emergence of Young Human Genes after a Burst of Retroposition in Primates. PLoS Biol 2005; 3: e357.
67. Huang C-J, Lin W-Y, Chang C-M, Choo K-B. Transcription of the rat testis-specific Rtdpoz-T1 and -T2 retrogenes during embryo development: co-transcription and frequent exonisation of transposable element sequences. BMC Molecular Biology 2009; 10: 74.

68. Vinckenbosch N, Dupanloup I, Kaessmann H. Evolutionary fate of retroposed gene copies in the human genome. Proceedings of the National Academy of Sciences of the United States of America 2006; 103: 3220-3225.
69. Mishra S, Somanath PR, Huang Z, Vijayaraghavan S. Binding and inactivation of the germ cell-specific protein phosphatase PP1gamma2 by sds22 during epididymal sperm maturation. Biol Reprod 2003; 69: 1572-1579.
70. Cheng L, Pilder S, Nairn AC, Ramdas S, Vijayaraghavan S. PP1gamma2 and PPP1R11 are parts of a multimeric complex in developing testicular germ cells in which their steady state levels are reciprocally related. PLoS One 2009; 4: e4861.
71. Pilder SH, Lu J, Han Y, Hui L, Samant SA, Olugbemiga OO, Meyers KW, Cheng L, Vijayaraghavan S. The molecular basis of "curlicue": a sperm motility abnormality linked to the sterility of t haplotype homozygous male mice. Soc Reprod Fertil Suppl 2007; 63: 123-133.

Chapter II.B

The evolution of PPP1R2 related pseudogenes

Luís Korrodi-Gregório[1], Joana Abrantes[2], Thorsten Muller[3], Katrin Marcus[3], Odete A. B. da Cruz e Silva[4], Margarida Fardilha[5] and Pedro José Esteves[2,6*]

[1]*Signal Transduction Laboratory, Centre for Cell Biology, University of Aveiro, Aveiro, Portugal*

[2]*CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Vairão, Portugal*

[3]*Functional Proteomics, Medizinisches Proteom-Center, Ruhr-University Bochum, Bochum, Germany*

[4]*Neuroscience Laboratory, Centre for Cell Biology and Health Science Department, University of Aveiro, Aveiro, Portugal*

[5]*Signal Transduction Laboratory, Centre for Cell Biology and Health Science Department, University of Aveiro, Aveiro, Portugal*

[6]*CITS, Centro de Investigação em Tecnologias da Saúde, IPSN, CESPU, Portugal*

Running Title: Evolution of PPP1R2 pseudogenes

Keywords: PP1, PPP1R2, IPP2, inhibitor-2, pseudogenes, evolution

To be submitted to: Genome Biology (6.885 IF, A)

***Corresponding author:** Pedro José Esteves, CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Rua Padre Armando Quintas, nr. 7, 4485-661 Vairão, Portugal, Tel.: +351 252 660 400; Fax: +351 252 661 780; E-mail: pjesteves@mail.icav.up.pt

Abstract

Protein phosphatase 1 regulatory subunit 2 (PPP1R2), also known as inhibitor-2 (I2), is an inhibitor and binding partner of the phosphoprotein phosphatase 1 (PPP1), a major Ser/Thr protein phosphatase involved in many cellular processes. In the human genome, several sequences collectively named PPP1R2 pseudogenes (PPP1R2P) have been identified that are highly similar to PPP1R2. The evolutionary analysis in mammals here presented shows evidences that two retroposons appeared prior to the great radiation of the mammals being seven primate-specific and retroposed at different times during the evolution of this group. From all ten pseudogenes present in humans only four appear to be transcribed and from these only three translated. Due to PPP1R2 role in sperm motility, mass spectrometry analysis was performed trying to detect new PPP1R2-related proteins. A new PPP1R2-related protein, PPP1R2P9, was found in human sperm. Besides the physiological roles that these proteins seems to have, some appear to be also associated with disease and pathological conditions. Our results show that this evolution process might be in part related with the formation of new genes and the gain of new specific functions and that these pseudogenes are not just silent regions as initially believed.

Background

Protein phosphatase 1 regulatory subunit 2 (PPP1R2), also known as inhibitor-2 (I2), is one of the first regulatory subunits identified as an inhibitor and binding partner of the phosphoprotein phosphatase 1 (PPP1), a major Ser/Thr protein phosphatase involved in many cellular processes from cell cycle to protein synthesis, muscle contraction, glycogen metabolism, cytokinesis, and neuronal signaling [1, 2]. PPP1R2 forms a stable and high affinity complex with PPP1 catalytic subunit (PPP1C) by blocking the active site. The reactivation of the complex is triggered by phosphorylation of PPP1R2 at Thr72 by several kinases, including extracellular signal-regulated kinases (ERKs), cyclin-dependent kinases (CDKs) and glycogen synthase kinase 3 (GSK3) [3-8]. PPP1R2 is also phosphorylated at the residues Ser86, Ser120 and Ser121 by casein kinase 2 (CK2), but these phosphorylations do not alter the inhibitory activity of PPP1R2, with the exception of Ser86 that accelerates the subsequent phosphorylation at Thr72 by GSK3 [3]. PPP1R2 is a potent heat and acid-stable inhibitor (1-10nM) that at low concentrations has been demonstrated to act as chaperone by helping the proper folding of newly synthesized PPP1C [9, 10]. Most PPP1C regulators have well conserved PPP1C binding motifs: the RVxF motif binds to the hydrophobic groove of PPP1C and the SILK motif, when present, is usually N-terminal to the former, among others [11]. PPP1R2 has two conserved degenerated RVxF motifs, the $_{145}$KLHY$_{148}$ and the $_{43}$KSQKW$_{47}$; the later already described to bind to the hydrophobic groove. Besides, a SILK motif ($_{12}$KGILK$_{16}$, in humans) that possibly initiates the binding, and the C-terminal acidic stretch that is required for GSK3 activation, are also present [11-13]. PPP1C/PPP1R2 complex has already been implied in several processes such as cardiac function [14-16], mitosis and meiosis [17-22], tubulin acetylation and neuronal cell survival [23]. Also, it has been shown that a PPP1CC2/PPP1R2-like complex is important in the acquisition of sperm motility [24, 25].

The PPP1R2 gene is conserved throughout all eukaryotes, from yeast to humans [26, 27]. In humans, it is located in chromosome 3 and spans 30kb containing 6 exons, the last with 2521bp, and a mRNA of 3474bp with the codifying sequence (CDS) encompassing nucleotide positions 377 to 994. In the human genome, as observed for other ancient PPP1 inhibitors such as PPP1R8 (NIPP1) and PPP1R11 (I3), several sequences have been identified that are highly similar to PPP1R2. [26]. For PPP1R2, nine loci were found that present hallmark features of processed pseudogenes, such as being intronless, presenting frameshifts disruptions, no parental promoter, truncated 5'UTR and polyA traits at the 3'UTR [28]. These related sequences were collectively named PPP1R2

pseudogenes (PPP1R2P) and were numbered from 1 to 9 [26]. These pseudogenes are scattered in the genome due to random retrotransposition phenomena [29]. These phenomena consist on the reverse transcription of cellular RNAs and insertion into the nuclear genome [30]. For instance, in human, the parental PPP1R2 is present in the chromosome 3 while PPP1R2P9 gene is located in the X chromosome and the PPP1R2P3 in the chromosome 5 [26]. Pseudogenes were often considered to evolve under neutrality, in which mutations that accumulate were out of scope of selection [31]. This assumed that pseudogenes are functionally inert. Several studies indicate that from the nine PPP1R2 pseudogenes, four have been detected at the message level using high throughput techniques. PPP1R2P1 and P2 were discovered in human [32, 33], PPP1R2P3 in human and crab-eating macaque (*Macaca fascicularis*) and PPP1R2P9 (also called I4) was found in human and mouse (*Mus musculus*) [34-36]. This clearly indicates that some of these pseudogenes might in fact have a biological role further supported by recent evidences that suggest that pseudogenes may indeed be functionally active [37-42]. Thus, the study of their evolution and conservation might provide support for a functional role and give insights into their potential mechanism of action. Therefore, in this work we did an exhaustive search for PPP1R2 pseudogenes in publicly available mammalian genomes in order to infer their evolutionary history and to suggest their potential functions. The collected pseudogenes were further characterized and detection at the protein level pursued.

Methods

Sequences retrieval

The human PPP1R2 mRNA sequence (GenBank accession number NM_006241.4) was used to detect orthologs and pseudogene-related sequences by performing a BLAST on Genbank database, National Center for Biotechnology Information (NCBI, http://BLAST.ncbi.nlm.nih.gov/) against all available mammalian genomes. Only sequences with more than 60% of sequence similarity and with query coverage of more than 35% were recovered.

Evolutionary tree reconstruction

The retrieved sequences were aligned using ClustalW implemented in Bioedit 7.0.9.0 [43] (Table II.B. 1). For phylogenetic reconstruction, and to improve accuracy, only sequences encompassing >85% coverage of the human PPP1R2 CDS (nucleotide positions 377-994 of the mRNA sequence) and with >60% of sequence similarity were included in the alignment. In order to infer the phylogenetic relationships between the PPP1R2 gene and related pseudogenes, a Neighbour-Joining (NJ) tree was constructed using MEGA software version 5 [44]. The evolutionary distances were computed using the p-distance method and alignment gaps, missing data and ambiguous bases were treated with the partial deletion option (site coverage cutoff 95%). Reliability of clustering of the branches was calculated using the bootstrap test with 1000 replicates.

Table II.B. 1: *Nucleotide sequences retrieved from NCBI Genbank database that were used for subsequent alignments and evolutionary analysis.*

Order	Species	Gene Name	Alias	ReferenceID	Chr	ChrID
Primates	Human	PPP1R2	I2 IPP2	NM_006241.4	3q29	NC_000003.11
	Homo sapiens	PPP1R2P1		NG_027882.1	6p21.3	NC_000006.11
		PPP1R2P2		NG_000913.3	21q22.13	NC_000021.8
		PPP1R2P3	I2L	NR_038443.1	5q33.3	NC_000005.9
		PPP1R2P4		NG_011458.1	13q14.13	NC_000013.10
		PPP1R2P10		NG_029109.1	13q21.31	NC_000013.10
		PPP1R2P5		NG_021516.1	2q12.3	NC_000002.11
		PPP1R2P6		NG_022575.1	7q34	NC_000007.13
		PPP1R2P8		n.a.	5p14.3	NC_000005.9
		PPP1R2P9	I4 IPP4	NR_002191.2	Xp11.3	NC_000023.10
	Chimpanzee	PPP1R2	I2	XM_516963.3	3	NC_006490.3
	Pan troglodytes	PPP1R2P1		XR_127820.1	6	NC_006473.3

Order	Species	Gene Name	Alias	ReferenceID	Chr	ChrID
		PPP1R2P2		BS000199.1	21	NC_006488.2
		PPP1R2P3	I2L	XM_003310938.1	5	NC_006472.3
		PPP1R2P10		XM_003314161.1	13	NC_006480.3
		PPP1R2P5		XR_127157.1	2A	NC_006469.3
		PPP1R2P6		AC144780.1	7	NC_006474.3
		PPP1R2P8		n.a.	5	NC_006472.3
	Orangutan	PPP1R2	I2 IPP2	XM_002814436.1	3	NC_012594.1
	Pongo abelii	PPP1R2P1		AC206576.3	6	NW_002874546.1
		PPP1R2P2		n.a.	21	NC_012612.1
		PPP1R2P3	I2L	XM_002816121.1	5	NC_012596.1
		PPP1R2P10		XM_002824322.1	13	NC_012604.1
		PPP1R2P10-Like		n.a.	13	NC_012604.1
		PPP1R2P5		n.a.	2A	NC_012592.1
		PPP1R2P6		n.a.	7	NC_012598.1
		PPP1R2P9	I4 IPP4	XM_002831555.1	X	NC_012614.1
	Gibbon	PPP1R2	I2 IPP2	XM_003280422.1	scaf_234 (11)	NW_003501604.1
	Nomascus leucogenys	PPP1R2P1		n.a.	scaf_81 (1b)	NW_003501451.1
		PPP1R2P2		n.a.	scaf_32 (25)	NW_003501402.1
		PPP1R2P3		n.a.	scaf_57 (2)	NW_003501427.1
		PPP1R2P5		XM_003277451.1	scaf_152 (14)	NW_003501522.1
		PPP1R2P6		n.a.	scaf_71 (13)	NW_003501441.1
		PPP1R2P9	I4 IPP4	XM_003271019.1	sca_72 (X)	NW_003501442.1
	Rhesus macaque	PPP1R2	I2 IPP2	XM_001097826.2	2	NC_007859.1
	Macaca mulatta	PPP1R2P1		n.a.	4	NC_007861.1
		PPP1R2P2		n.a.	3	NC_007860.1
		PPP1R2P10		n.a.	17	NC_007874.1
		PPP1R2P5		AC187497	13	NC_007870.1
		PPP1R2P9	I4 IPP4	XM_001088324.2	X	NC_007878.1
	Grivet	PPP1R2P1		AC241599.3	n.a.	n.a.
	Chlorocebus aethiops					
	Marmoset	PPP1R2	I2 IPP2	XM_002758211.1	15	NC_013910.1
	Callithrix jacchus	PPP1R2P5		n.a.	14	NC_013909.1
		PPP1R2P1		AC242643.3	4	NC_013899.1
		PPP1R2P6		n.a.	8	NC_013903.1
		PPP1R2P9-Like		n.a.	X	NC_013918.1
		PPP1R2P9	I4 IPP4	XM_002762795.1	X	NC_013918.1
Artiodactyla	Pig	PPP1R2P9-Like	I4 IPP4	XR_131237.1	X	NC_010461.3

Order	Species	Gene Name	Alias	ReferenceID	Chr	ChrID
	Sus scrofa					
	Cow	PPP1R2	I2 IPP2	NM_001035392.1	1	NC_007299.4
	Bos taurus	PPP1R2P9	I4 IPP4	NM_001079599.1	X	NC_007331.3
	Horse	PPP1R2	I2 IPP2	XM_001500822.3	19	NC_009162.2
	Equus caballus	PPP1R2P9	I4 IPP4	XM_001491737.1	X	NC_009175.2
Proboscidea	Elephant	PPP1R2P9-Like		XM_003418018.1	n.a.	n.a.
	Loxodonta africana					
Carnivora	Giant panda	PPP1R2	I2 IPP2	XM_002921452.1	scaf_1297	NW_003217779.1
	Ailuropoda melanoleuca	PPP1R2-Like		n.a.	scaf_639	NW_003217930.1
		PPP1R2P9	I4 IPP4	XM_002922902.1	scaf_438	NW_003217919.1
	Dog	PPP1R2		XM_535781.2	33	NC_006615.2
	Canis lupus familiaris	PPP1R2P9		XM_548958.2	X	NC_006621.2
Lagomorpha	Rabbit	PPP1R2		NM_001082744.2	14	NC_013682.1
	Oryctolagus cuniculus	PPP1R2P9		XM_002719841.1	X	NC_013690.1
Rodentia	Rat	PPP1R2	I2-α1 I2-α2	NM_138823.2	11	NC_005110.2
	Rattus norvegicus	PPP1R2P9-Like	I2-β	XM_002730182.1 XM_002727659.1	X	NC_005120.2
		PPP1R2P9		XM_002730188.1 XM_002727510.1 NM_001106953.1	X	NC_005120.2
	Mouse	PPP1R2		NM_025800.3	16	NC_000082.5
		PPP1R2P9		NR_033171.1	X	NC_000086.6
		PPP1R2P9-Like		NR_033731.1	X	NC_000086.6
Marsupials	Opossum	PPP1R2-Like		XM_001371978.2	n.a.	NW_001583292.1
	Monodelphis domestica					

Divergence times

Divergence times from the other species in relation to Homo sapiens in millions of years ago (Mya) were obtained from the website http://www.timetree.org/.

Pseudogene classification and conserved linkage

Sequences obtained from the BLAST queries were analyzed in terms of intronic regions presence, polyA traits, truncation of the 5'UTR and chromosomal location.

Chromosomal locations from all mammals were obtained from the Genbank database with the exception of gibbon (*Nomascus leucogenys*) that was obtained from the website http://www.biologia.uniba.it/gibbon/ [45] (Table II.B. 1). Pseudogenes located in the same chromosome and nearby and/or with intronic regions were classified as duplicated pseudogenes. Pseudogenes that were located in different chromosomes and have polyA traits, truncation of the 5'UTR and no introns were classified as processed pseudogenes. Furthermore, genes that are flanking each human PPP1R2 pseudogene and are well conserved among mammals were selected. Conserved linkage, meaning conservation of synteny and also conservation of the gene order, was then searched in order to provide insight regarding the human PPP1R2 pseudogenes orthologous in the other species analyzed.

Distance to closest and repeated regions

The distance of each pseudogene to the closest neighboring gene, not taking into account other pseudogenes present nearby or the transcript orientations, was calculated. Repeated sequences were detected by submitting each pseudogene sequence to the program RepeatMask from Institute for Systems Biology, Seattle, Washington, USA (http://www.repeatmasker.org/).

GC-content and polyA signals

Isobase Isochore Database (http://www.geneinfo.eu:8080/isobase/index.jsp) was searched to assign each PPP1R2 pseudogene to the respective isochore [46]. The GC-content of each pseudogene CDS was calculated and the difference between each pseudogene and the respective flanking regions was performed. PolyA signals were obtained by submitting the sequences to the program PolyApred (http://www.imtech.res.in/raghava/polyapred/).

Sperm extracts and immunoprecipitation.

Testis is one of the organs where most pseudogenes are expressed and it was shown that the gene products have a function in spermatogenesis or other germ cell functions [47-49]. Therefore, since sperm are produced in testis and are easily obtained, the presence of some of the studied pseudogenes was checked in human sperm. Thus, ejaculated sperm was collected from healthy donors and spermograms were performed as described before [50]. In brief, sperm was lysed in 1xRIPA buffer

(radioimmunoprecipitation buffer, Millipore Iberica S.A.U., Madrid, Spain) supplemented with protease inhibitors (10mM benzamidine, 1.5μM aprotinin, 5μM pepstatin A, 2μM leupeptin, 1mM PMSF), sonicated 3x10sec and centrifuged at 16000g for 20min, at 4ºC. RIPA supernatant sperm extract was immunoprecipitated using Dynabeads® Protein G (Life Technologies S.A., Madrid, Spain) and 1μg of rabbit anti-PPP1R2 (against a mouse PPP1R2 peptide, amino acids 134–147) with standard direct immunoprecipitation procedure [50]. Also, an independent RIPA supernatant sperm extract was prepared, boiled in a water bath for 30min, chilled on ice for 2min and centrifuged at 16000g for 20min, 4ºC to obtain a heat stable extract.

Mass spectrometry

For mass spectrometry analysis, the immunoprecipitate and the heat stable extract were resolved by 10% SDS-PAGE along with purified positive controls. Gels were stained with Coomassie blue colloidal (Sigma-Aldrich Química, S.A., Sintra, Portugal) using standard procedures [50].

Bands were then excised from the gel using commercial PPP1R2 band as control and destained. An overnight digestion with trypsin (Promega, Madison, Wisconsin, USA) was performed and resulting peptides were extracted and prepared for mass spectrometry analysis using an Orbitrap Velos mass spectrometer as described elsewhere [50]. Subsequent generated data were imported to ProteinScapeTM (Bruker Daltonik GmbH, Bremen, Germany, [34]) and analyzed using MASCOT (version 2.2.0, Matrix Science, London, UK, [35]) search algorithm. Proteins were considered to be identified if the Mascot score (ProteinScapeTM) was higher than 65.

Results and Discussion

Phylogenetic analysis

A total of 119 sequences were retrieved from the NCBI database using the PPP1R2 nucleotide sequence. Ten pseudogenes were obtained from human sequences increasing by one the previous stated number in the literature [26]. In order to increase the reliability of the alignment for phylogenetic reconstruction, we selected sequences with >85% coverage and >60% similarity with the human PPP1R2 CDS. By doing this, we were able to include in the tree 67 sequences that represented all the pseudogenes with the exception of PPP1R2P7 (Table II.B. 1). The sequences not used encompass pseudogenic fragments and sequences without CDS (PPP1R2P7) or truncated CDS (some PPP1R2P8 and PPP1R2P9). The NJ tree obtained showed two major clusters (Figure II.B. 1). The basal cluster includes all PPP1R2P9 sequences (with the exception of Rodentia and Callithrix PPP1R2P9-Like) and the elephant (*Loxodonta africana*) PPP1R2P9-like sequence. The upper cluster is composed by all PPP1R2 sequences and the other pseudogene groups, with the gray short-tailed opossum (*Monodelphis domestica*) PPP1R2-like sequence appearing as an outgroup for this cluster. This is consistent with the presence of PPP1R2 in eukaryotes, being indeed an ancient and well conserved gene [26]. Concerning the retroposition of the different pseudogenes in the NJ tree two major events are clearly seen. The first event was the retroposition of PPP1R2P9 prior to the separation of Metatheria (marsupial mammals) and Eutheria (placental mammals), evidencing this pseudogene as the most ancient one still present in humans. The second event was the retroposition wave that occurred in the beginning and during the primate radiation with the appearance of the other groups. This is in accordance with the NJ tree that shows these groups after the non-primate PPP1R2 sequences cluster (Figure II.B. 1).

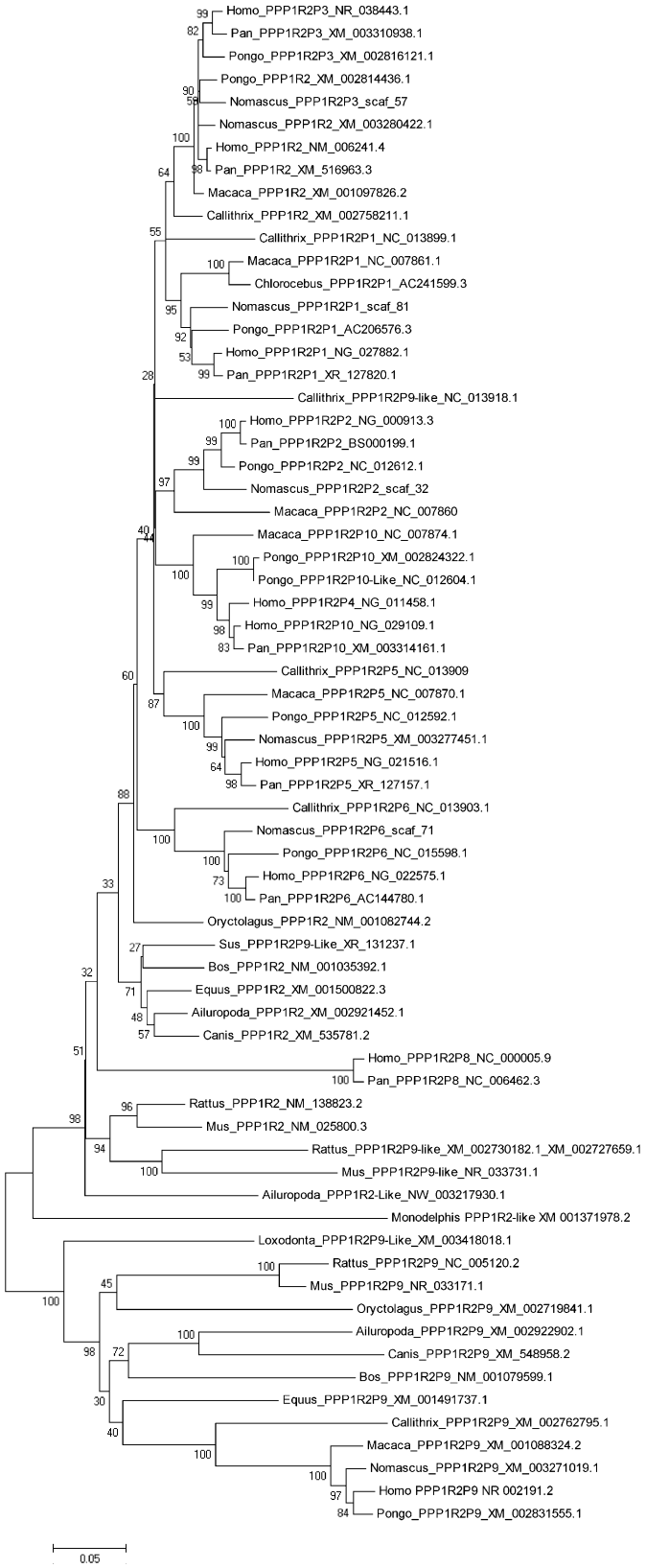


Figure II.B. 1: Evolutionary tree of PPP1R2 and related pseudogenes. *The evolutionary history was inferred using MEGA5 with the Neighbor-Joining method and the partial deletion option (site coverage cutoff 95%). Reliability of the tree was assessed by bootstrap with 1000 replicates and is indicated in the nodes.*

Furthermore, PPP1P2P3 grouping with the primate PPP1R2 sequences is in agreement with it being one of the youngest pseudogenes that have been retroposed, 20.6-29.2 million years ago (Mya) (Figure II.B. 2).

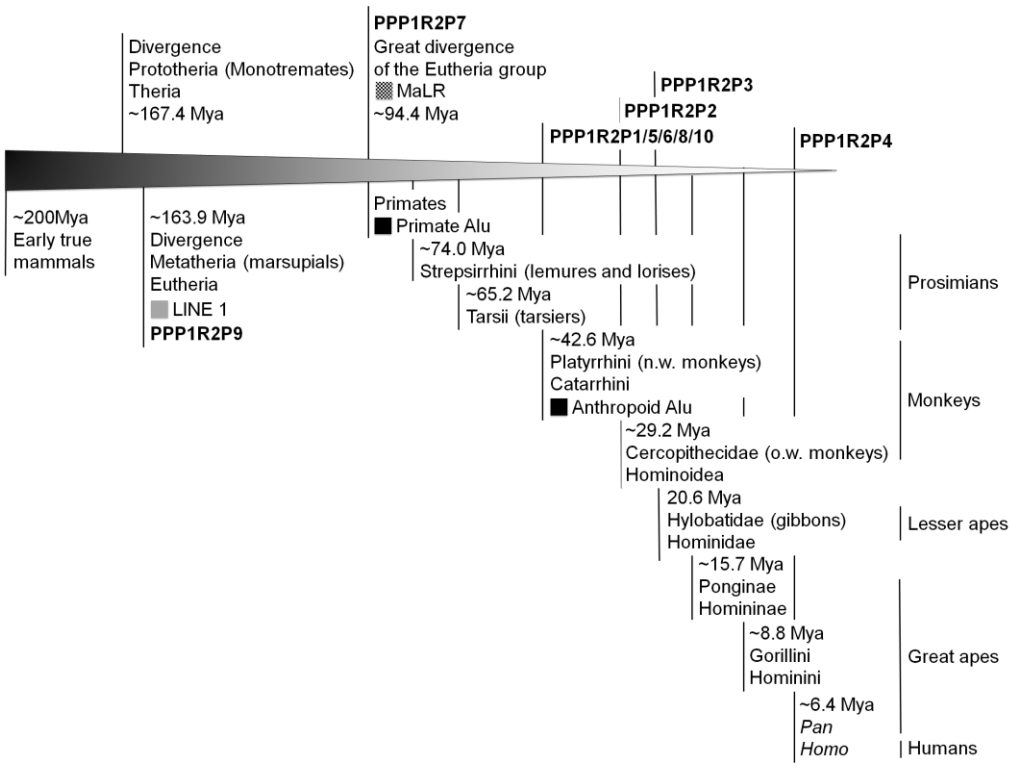


Figure II.B. 2: Evolution of pseudogenes diagram. *Time scale from the early mammals evolution till humans is shown with emphasis in the primate class. The time in million years ago (Mya) indicates the splits between groups. Pseudogenes estimated emergence is shown, as well as, important retrotransposable elements.*

PPP1R2P1

By analyzing the NJ tree, we observed that the PPP1R2P1 orthologs genes cluster together in a highly supported branch (bootstrap value of 90) and are a sister group of the cluster composed by the parental PPP1R2 gene and the primates PPP1R2P3 (Figure II.B. 1). This study demonstrates that PPP1R2P1 is primate-specific and was originated before the division of New World monkeys (Platyrrhini) and Catarrhini that occurred 43.4-65.2 Mya (Figure II.B. 2). Figure II.B. 3 shows that PPP1R2P1s are assigned to the negative strand orientation of chromosome 4 in marmoset (*Callithrix jacchus*) and rhesus monkey (*Macaca mulatta*), chromosome 1b in gibbon and to chromosome 6 in the rest of the primates. The human PPP1R2P1 gene was first mapped to the short arm of chromosome 6 (6p21.31) in the class II major histocompatibility complex (MHC) region by in situ hybridization [51, 52]. The TAP1 (transporter 1, ATP-binding cassette, sub-family B) and

HLA-DMB genes flanking the PPP1R2P1 in *Homo sapiens* are also present in the other primates which reinforces the tree results and shows a conserved linkage in primates.

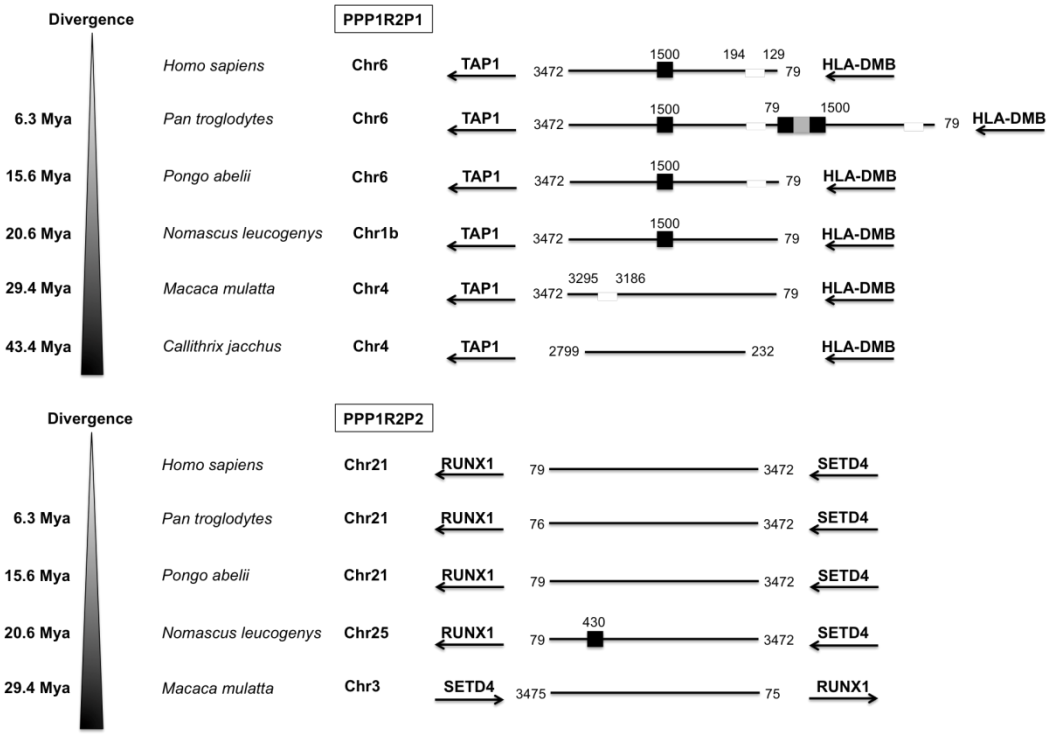


Figure II.B. 3: Conserved linkage of PPP1R2P1 and PPP1R2P2. *PPP1R2P1 and PPP1R2P2 location in terms of chromosome and flanking genes is presented concerning each species where they were found, showing the conserved linkage. Divergence time is shown on the left. The numbers flanking the pseudogenes are related to the parental PPP1R2 message. Black boxes refer to the short interspersed elements (SINEs) Alu repeats that are primate-specific. Grey boxes refer to the long interspersed elements (LINEs) in this case a LINE1 element. Number above the boxes states the location where the repeat interrupted the sequence. In the case of chimpanzee PPP1R2P1 a duplicated pseudogene was originated and the repeats are located in the middle of both, and so, the numbers refer to the final of one pseudogene and the initial of the other. Also, a deletion is shown (129 to 194) that is common to all pseudogenes with the exception of gibbon and marmoset.*

Concerning the genomic sequence, PPP1R2P1 is intronless and has conserved the polyA traits that were present at the 3'UTR of the parental mRNA, meaning that it is a processed pseudogene, with 88% similarity to the parental gene and with 97% sequence coverage. It has a small 70bp deletion that seems to have appeared in Hominidae after the divergence from Hylobatidae, ~20.6 Mya. Also, an Alu repeat was inserted after the radiation of the Hominoidea (that include Hominidae and Hylobatidae), ~29.4 Mya, in the middle of the sequence (~1500bp) disrupting it, but without affecting the CDS (Figure II.B. 2). Interestingly, in chimpanzee (*Pan troglodytes*), PPP1R2P1 suffered a duplication event that gave rise to a second locus that is 100% identical to the parental PPP1R2P1 from which is separated by two Alu repeats flanking a LINE1 element (Figure II.B. 2, not

included in the NJ tree). The high similarity between these two loci suggests that the duplication event occurred recently.

Transcribed processed pseudogenes tend to be closer to other functional genes (median, 22.9kb) than the transcriptionally silent ones (median, 38.7kb) [53].

Chromosomal location of this gene is favorable for the acquisition of a biological role since is in the MHC region and therefore many genes occur nearby, with only 16.6Kb distance to the closest one. MHC region of chromosome 6 has many genes that encode for proteins with important role in immune system and autoimmunity, like the human leucocyte antigen (HLA) [54]. Studies in humans identified in this region a frameshift mutation in PPP1R2P1, stating that it was probably an intronless processed pseudogene [55]. However this and other early studies have used a mixture of multiple haplotypes. Afterwards, eight HLA-homozygous haplotypes were established and changes in these regions were characterized [56]. Subsequent studies identified that the PPP1R2P1 frameshift mutation was not present in all individuals [57-59]. In fact, only haplotypes from SSTO, APD and MCF seem to have this mutation, whereas in PGF, COX, QBL, DBB and MANN it was absent [54]. Furthermore, in a random screening of 18 genomic samples, two were homozygous for the frameshift, seven were heterozygous and nine were homozygous for the continuous open reading frame (ORF), which revealed a frameshift frequency of ~0.31 [57]. Also, as expected by our phylogenetic analysis, studies in the marsupial mammal gray short-tailed opossum (*Monodelphis domestica*), dog (*Canis familiaris*), cat (*Felis catus*) and mouse MHC region showed that PPP1R2P1 gene was absent [60, 61].

The presence of promoters, enhancers and other regulatory elements could be the explanation for why some ESTs have already been found linked to this pseudogene, although basal transcription should not be set aside. Gene expression of PPP1R2P1 was shown to be upregulated comparing to normal tissue by RNA differential display and qPCR cDNA in vascularized human breast carcinoma and was deposited as a mRNA entry in Unigene (NCBI: AF275684.1), NCBI [32]. More recently, human PPP1R2P1 was chosen has a potential gene to be involved in initiation and formation of ganglioglioma although showing chromosomal alterations common to both the benign and malignant components of the tumor in humans [62]. In Gene Expression Omnibus (GEO) from NCBI and in ArrayExpress from Ensembl, high-throughput gene expression databases, 19 GEO profiles and 2 experiments were found, respectively, in humans. Altogether, these results suggest that in human, PPP1R2P1 might be functionally relevant.

Concerning the mRNA stability, only part of the 5'UTR (238bp), due to the low processivity of the reverse transcriptase and the 70bp deletion already described, and part of the 3'UTR (506bp) are present. Therefore, the stability might be compromised although a polyA signal ATTAAA is present near the extreme of the 3'UTR (position 1361 according to Figure II.B. 3). Regarding the human CDS and possible translation, the Kozak sequence, important for translation initiation, is present in the parental gene and is conserved.

Furthermore, in other primates the ORF has frameshift disruptions that introduce premature stop codons, meaning that this pseudogene in these species does not produce a putative functional protein or the protein is truncated.

PPP1R2P2

PPP1R2P2 was originated in Catarrhini after its separation from the Platyrrhini 29.2-42.6 Mya (Figure II.B. 2). Indeed, our phylogenetic analysis demonstrates that PPP1R2P2s form a robust cluster between PPP1R2P5 and P4/P10 genes clusters (Figure II.B. 1). The PPP1R2P2 gene is present in the positive strand orientation of human chromosome 21q22.13 and covers 99% of the parental gene ORF with 86% similarity, being only truncated at the 5'UTR. In the other primates PPP1R2P2 is also present in chromosome 21 and always in the positive strand orientation, with the exception of gibbon where it is localized in the positive strand orientation of chromosome 25 and rhesus monkey where it is localized to chromosome 3 and in the negative strand orientation. Runt-related transcription factor 1 (RUNX1) and SET domain containing 4 (SETD4, a variant of PPP2 inhibitor 2) genes flank PPP1R2P2 in all species analyzed, confirming the conserved synteny and the NJ tree results (Figure II.B. 3). In humans, chromosome 21 has only ~232 genes, a low density only surpass by the Y chromosome (130) and as expected the processed pseudogene density is also low, 34 [63]. Concerning the distance to other genes, and as expected, PPP1R2P2 was inserted in a region without genes. The SETD4 gene is the closest, but still ~145Kb distant.

The only interruption to the ORF occurred in gibbon by an Alu repeat that also disrupts the CDS (~430bp) (Figure II.B. 3). Considering the CDS, many mutations in all species analyzed introduced frameshift mutations leading to premature stop codons. Moreover, all the four polyadenylation signals present in the parental PPP1R2 mRNA are conserved in PPP1R2P2. Although protein expression is unlikely, message was found by qPCR in human testis but not in peripheral blood leukocytes because it is methylated in these cells [33]. Also, two experiments from ArrayExpress Ensembl, report the up/down

regulation of this pseudogene in prostate adenocarcinoma and in a prostate transcriptomic study in Caucasian population [64]. These results might in fact be artifacts or could be assigned to other PPP1R2 pseudogenes/parental gene since this pseudogene is promoterless and no other promoters are nearby.

PPP1R2P3

Evolutionary analyzes indicate that PPP1R2P3 is one of the youngest processed pseudogenes, being originated after the separation of Hominoidea from Cercopithecoidea (old world monkeys), 20.6-29.2 Mya, since rhesus monkey and marmoset do not have a PPP1R2P3 copy (Figure II.B. 2 and Figure II.B. 4).

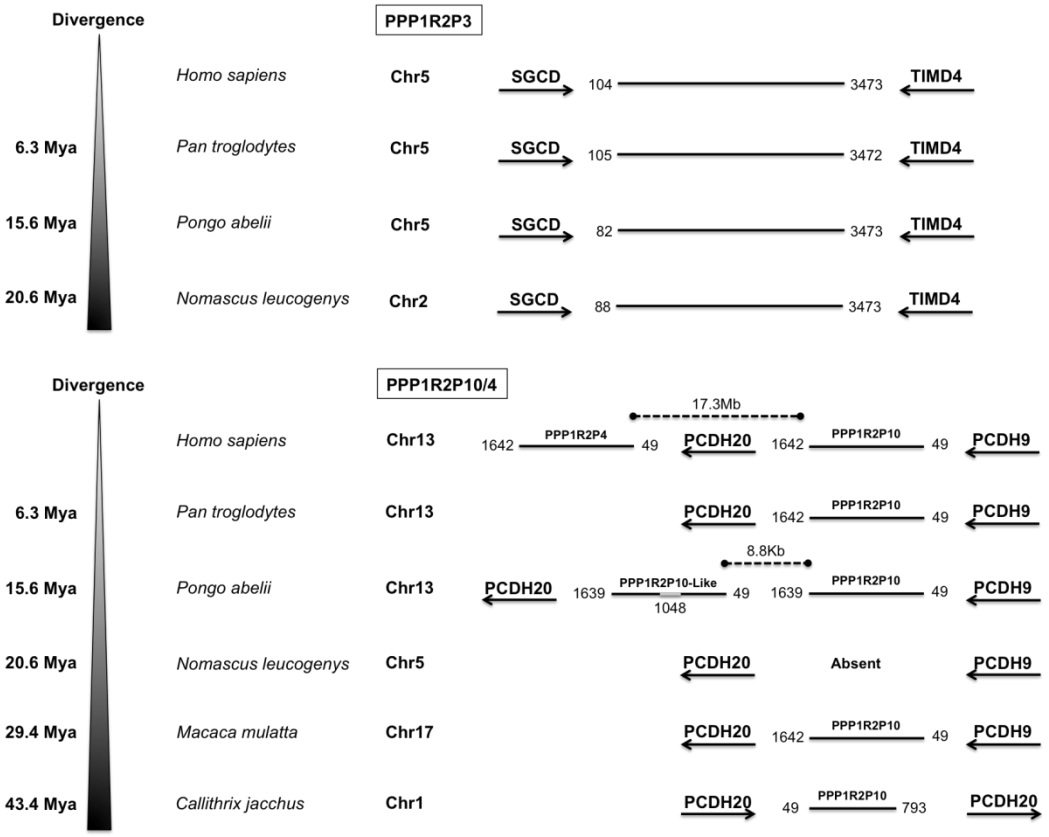


Figure II.B. 4: Conserved linkage of PPP1R2P3 and PPP1R2P10/4. *PPP1R2P3 and PPP1R2P10/4 location in terms of chromosome and flanking genes is presented concerning each species where they were found, to show the conserved linkage in these pseudogenes. Divergence time is shown on the left. The numbers flanking the pseudogenes are related to the parental PPP1R2 message. In orangutan an unknown sequence according to the current genome assembly was inserted in PPP1R10-like and is shown with a number referring to the location. The distances in dashed lines of the duplicated forms in human and orangutan are also indicated.*

Interestingly in crab-eating macaque (*Macaca fascicularis*) with caloric restrictions, PPP1R2P3 gene has shown to be downregulated in skeletal muscle by microarrays and

RT-qPCR techniques when compared to normal feeding animals [65]. Since PPP1R2P3 is not present in rhesus monkey, from the same group, and is highly similar to the parental gene, this result was probably misinterpreted as being PPP1R2P3 instead of parental PPP1R2, or other pseudogenes present in Cercopithecoidea. Phylogenetic analysis shows that PPP1R2P3 nucleotide sequences from human, chimpanzee, orangutan and gibbon are clustered together and are closer to the parental PPP1R2 primate cluster than the rest of human pseudogenes (Figure II.B. 1).

BLAST search using the mRNA of parental PPP1R2 gene indicate 98% coverage and 95% similarity. Human PPP1R2P3 is present in chromosome 5q33.3 and is also located in the same chromosome in the other primates and always in the positive strand orientation, except in gibbon that is in chromosome 2. Furthermore, PPP1R2P3 is flanked by sarcoglycan delta (SGCD) and T-cell immunoglobulin and mucin domain containing 4 (TIMD4) genes in all primates, showing a shared synteny (Figure II.B. 4). The TIMD4 gene is the closest one, ~66Kb apart, meaning that PPP1R2P3 is located in a region with low transcription rate. Past work using in situ hybridization has already mapped hPPP1R2P3 to chromosome 5 (5q33) [52].

PPP1R2P3 shows low divergence from the parental gene and the ORF is complete without any frameshift or element repeats disruptions inside the CDS. The ORF is truncated at the 5'UTR as expected, due to the low processivity of the reverse transcriptase, and in the 3'UTR it lost two of the four polyadenylation signals, that may lead to a short ~1500-1600nt message. We have previously found, by a yeast two hybrid screening of human testis cDNA using as bait PPP1CC1, one clone assigned to PPP1R2P3 (GenBank:JF438008.1) [66, 67]. PPP1R2P3 mRNA has also shown to be present in human testis tissue cDNA libraries by the Mammalian Gene Collection (MGC) program [35]. Search for PPP1R2P3 ESTs in databases, revealed that after the parental gene, this is one of the pseudogenes more represented. Unigene from NCBI has seven ESTs, three of them from testis, two from brain and one from lung, being the other one not assigned. In high-throughput databases GEO profiles and ArrayExpress, 11 Geo profiles and 114 experiments are respectively present, that state up and down-regulation of this pseudogene. Together with our previous data, this strongly suggests that this pseudogene is in fact transcribed. Moreover, in humans, PPP1R2P3 has only 16 nucleotide changes (92.2%), which correspond to 9 amino acid changes in the translated sequence, including the Thr73 to Pro and the Ser87 to Arg that are important phosphorylation sites for the GKS3 and CK2 kinases.

Two independent reports using mass spectrometry have assigned peptides to PPP1R2P3 [68, 69]. However, these peptides share the sequence with both parental gene and PPP1R2P3 being most probably misassigned. Nevertheless, in a recent paper we were capable to recover from human ejaculated sperm a peptide that was specifically assigned to PPP1R2P3, demonstrating for the first time that this previous stated processed pseudogene has evolved to the gene category gaining important functions in sperm physiology [50].

PPP1R2P4/10

PPP1R2P4 and PPP1R2P10 are both located in the long arm of human chromosome 13 and are 17.4Mb apart. Ceulemans et al. first described the presence of PPP1R2P4 in 2002, but no information was at that time provided for PPP1R2P10. In our BLAST search both were detected and retrieved and we find that, in fact, PPP1R2P10 (as stated in NCBI) is the ancestral and PPP1R2P4 a duplication that occurred only in humans, being therefore a duplicated pseudogene (Figure II.B. 1 and Figure II.B. 4). Moreover, PPP1R2P10 is a processed pseudogene and was originated before the division of Platyrrhini and Catarrhini (Figure II.B. 2).

BLAST results show a 41% coverage and 88% similarity with the parental gene. Figure II.B. 1 demonstrates that PPP1R2P4/10s form a robust cluster more distant from the parental gene than the previous described pseudogenes. Concerning the position, PPP1R2P10 (and P4 in humans) is present in the negative strand orientation in all primates except for marmoset, which is present in the positive strand orientation of chromosome 1 (Figure II.B. 2). PPP1R2P10 is flanked in all primates by the protocadherin 20 (PCDH20) and protocadherin 9 (PCDH9) genes showing a conserved linkage and further confirming its chromosomal location. Also, in orangutan, a duplication occurred very close (~8.8kb) to PPP1R2P10 that is not related with human PPP1R2P4, and was hence here named PPP1R2P10-Like, being therefore a duplicated pseudogene (Figure II.B. 1 and Figure II.B. 2).

Surprisingly, PPP1R2P10 is not present in gibbon, which might suggest that it has been deleted. Lesser apes family (Hylobatidae), in contrary to the great apes and humans, have major chromosome translocations [70]. Although these translocations do not explain this particular deletion, the higher rate of rearrangements and possibly errors in crossing over during meiosis might explain the deletion occurred. In fact, when we

looked to orangutan (*Pongo abelii*) chromosomal region between PCDH20 and PCDH9 genes, 5.25Mb separates them with 11 putative pseudogenes in between, but in gibbon the distance between the two genes is only 4.73Mb with 3 putative pseudogenes maintained (legumin-like and steroidogenic acute regulatory protein, mitochondrial-like). This means that 0.52Mb were deleted and probably along with it PPP1R2P10 and other pseudogenes. Another explanation might be the fact that since gibbon genome annotation is still in the first assembly, the pseudogene could be masked.

PPP1R2P10 is surrounded nearby only by pseudogenes, showing that this region is probably less transcribed. Moreover, PPP1R2P4/10/10-Like ORFs are truncated at the 3'UTR (648bp) near the first polyadenylation signal and since no expression pattern was found in ESTs and high-throughput databases, this may indicate that this pseudogene is indeed transcriptionally silent. In terms of CDS, many frameshift disruptions cause premature stop codons in all ORFs, which could indicate that this pseudogene is not translated.

PPP1R2P5

The PPP1R2P5 has a high coverage (98%) with the parental gene although the similarity is lower, 86%. Furthermore, like PPP1R2P1 and P10, PPP1R2P5 was originated before the division of Platyrrhini and Catarrhini (Figure II.B. 2). By phylogenetic analysis, all PPP1R2P5s form a robust cluster (bootstrap value 87) close to PPP1R2P2 and P4/10 clusters (Figure II.B. 1). These four pseudogenes form also a group clearly separated from the PPP1R2P3/P1 and parental gene group (Figure II.B. 1). PPP1R2P5 is located in chromosome 2 in all great apes (2A in chimpanzee and orangutan), chromosome 14 in gibbon and marmoset and chromosome 13 in rhesus monkey, always in negative strand orientation (Figure II.B. 5).

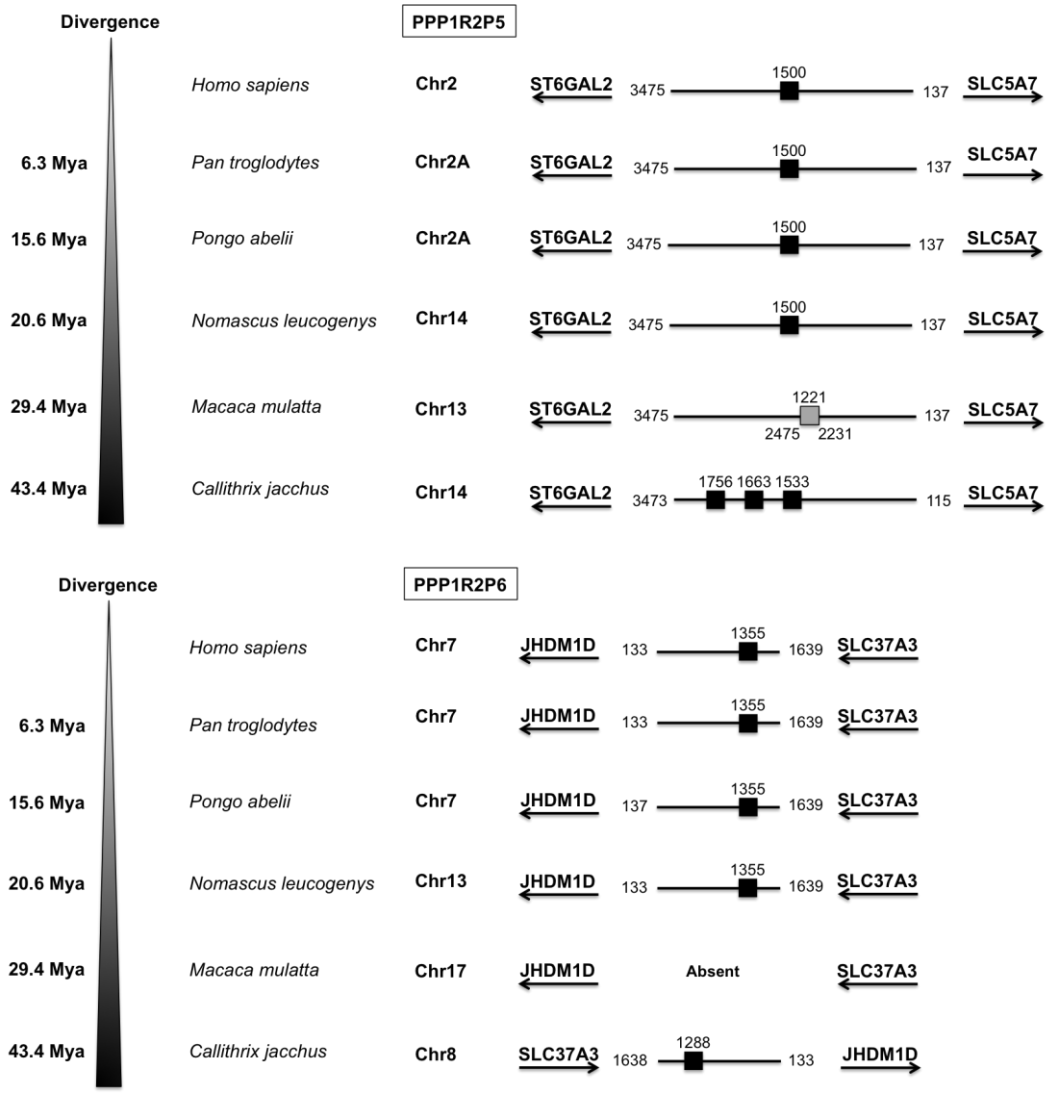


Figure II.B. 5: Conserved linkage of PPP1R2P5 and PPP1R2P6. *PPP1R2P5 and PPP1R2P6 location in terms of chromosome and flanking genes is presented concerning each species where they were found, showing the conserved linkage. Divergence time is shown on the left. The numbers flanking the pseudogenes are related to the parental PPP1R2 message. Black boxes refer to the short interspersed elements (SINEs) Alu repeats that are primate-specific. Number above the boxes states the location where the repeat interrupted the sequence. Grey box delimited with a black line in rhesus monkey PPP1R2P5 refer to a parental PPP1R2 insertion. Number on the top indicates where the insertion took place in the pseudogene, while numbers at the bottom show which region of the parental PPP1R2 was inserted.*

PPP1R2P5 is flanked by the genes, ST6 beta-galactosamide alpha-2,6-sialyltranferase 2 (ST6GAL2) and solute carrier family 5 choline transporter, member 7 (SLC5A7) in all primates analyzed which shows a conserved linkage (Figure II.B. 5). The only interruption to the ORF occurred before the divergence of Hominidae and Hylobatidae, since it is only present in rhesus monkey, by an Alu repeat at the position 1500, but without disrupting the CDS. Other repeats are also present in marmoset but not related with the Hominoidea species Alu repeat (Figure II.B. 5). In rhesus monkey, an

insertion of a parental mRNA portion (2231-2475) in the same orientation and at the position 1221 has occurred but without disrupting the CDS.

Considering the CDS, many mutations in all the species analyzed introduced frameshift mutations leading to premature stop codons. PPP1R2P5 is not surrounded by any genes or pseudogenes, being the closest one, the ST6GAL2 gene, 53.8Kb far. This region is probably less transcribed which is also corroborated by the fact that no ESTs or expression in high-throughput databases were found. Altogether, this suggests that PPP1R2P5 has no relevant biological role and might be a true processed pseudogene.

PPP1R2P6

PPP1R2P6 has been originated before the division of the Platyrrhini and Catarrhini, 42.6-65.2 Mya (Figure II.B. 2). Phylogenetic analysis shows that PPP1R2P6s form a robust cluster (bootstrap value 100) more distant from the parental gene cluster than the previous pseudogenes but closer to the Lagomorpha PPP1R2 (Figure II.B. 1). BLAST results show a 44% coverage (1506/3475) and an 85% similarity with the parental gene. PPP1R2P6 is present in chromosome 7 of great apes and humans and in the chromosome 13 of gibbon always in the positive strand orientation, with the exception being the marmoset in which it is present in the negative strand orientation of chromosome 8 (Figure II.B. 5).

This pseudogene is flanked by the jumonji C domain containing histone demethylase 1 homolog D (JHDM1D) and solute carrier family 37 (glycerol-3-phosphate transporter), member 3 (SLC37A3) genes showing a conserved linkage, being the JHDM1D gene the closest one, 40.4Kb apart (Figure II.B. 5).

Surprisingly, PPP1R2P6 is absent in rhesus monkey what might suggest that it has been deleted. In fact, when looking to the gibbon chromosomal region between JHDM1D and SLC37A3 genes, 194.8Kb distance separates them, but in rhesus monkey the distance between the two genes is only 160.2Kb. This means that 34.6Kb were deleted and probably along with it PPP1R2P6, although caution should be taken because rhesus monkey genome assembly might be incomplete in this region since at least 10.1Kb are still unknown and so PPP1R2P6 could also be masked and therefore not retrieved by our BLAST searches.

Before the divergence of Hominidae and Hylobatidae, 20.6-29.2 Mya, an Alu repeat was inserted at nucleotide position 1355 without disrupting the CDS but disrupting the ORF. An Alu repeat was also inserted in marmoset also disrupting the ORF, but with no

relation with the Alu repeat of Hominoidea (Figure II.B. 5). Also, no common consensus polyadenylation signals were found. This pseudogene is located near other genes, but in accordance to the arguments presented and since no ESTs and no expression in high-throughput databases were found, this pseudogene should be transcriptionally silent. Also, many mutations in all the species analyzed have introduced frameshifts leading to premature stop codons.

PPP1R2P7

PPP1R2P7 is the first described in this work that is not primate specific. Evolutionary analysis suggests that it was originated 163.9-94.4 Mya since it is present in all the mammal orders, with the exception of the Glires clade (Lagomorpha and Rodentia). However, this pseudogene was not included in the phylogenetic analysis since it only has the 3'UTR of the parental message.

PPP1R2P7 is present in chromosome 4 in Homininae (humans, chimpanzee and gorilla) in the negative strand orientation and is flanked by the pyruvate dehydrogenase (lipoamide) alpha 2 (PDHA2) and RAP1, GTP-GDP dissociation stimulator 1 (RAP1GDS1) genes showing a conserved linkage in all species (Figure II.B. 6). BLAST results show a 46% coverage and a 67% similarity with the parental PPP1R2 message.

Since it is one of the most ancient pseudogenes, many mutations and LINE1 repeats have occurred. In particular, a substitution in the final of the sequence has occurred in the primates, or even earlier since it is present in the pig (*Sus scrofa*) sequence (Figure II.B. 6) switching the parental sequence by an unknown sequence. This sequence resembles a downstream region of the parental PPP1R2 gene in Artiodactyla, which indicate that a recombination event, or another retroposition might have occurred. Also, two LINE1 repeats have occurred in Artiodactyla and Carnivora and one LINE1 repeat has occurred before the divergence of the genera *Pan* and *Homo* (Figure II.B. 5).

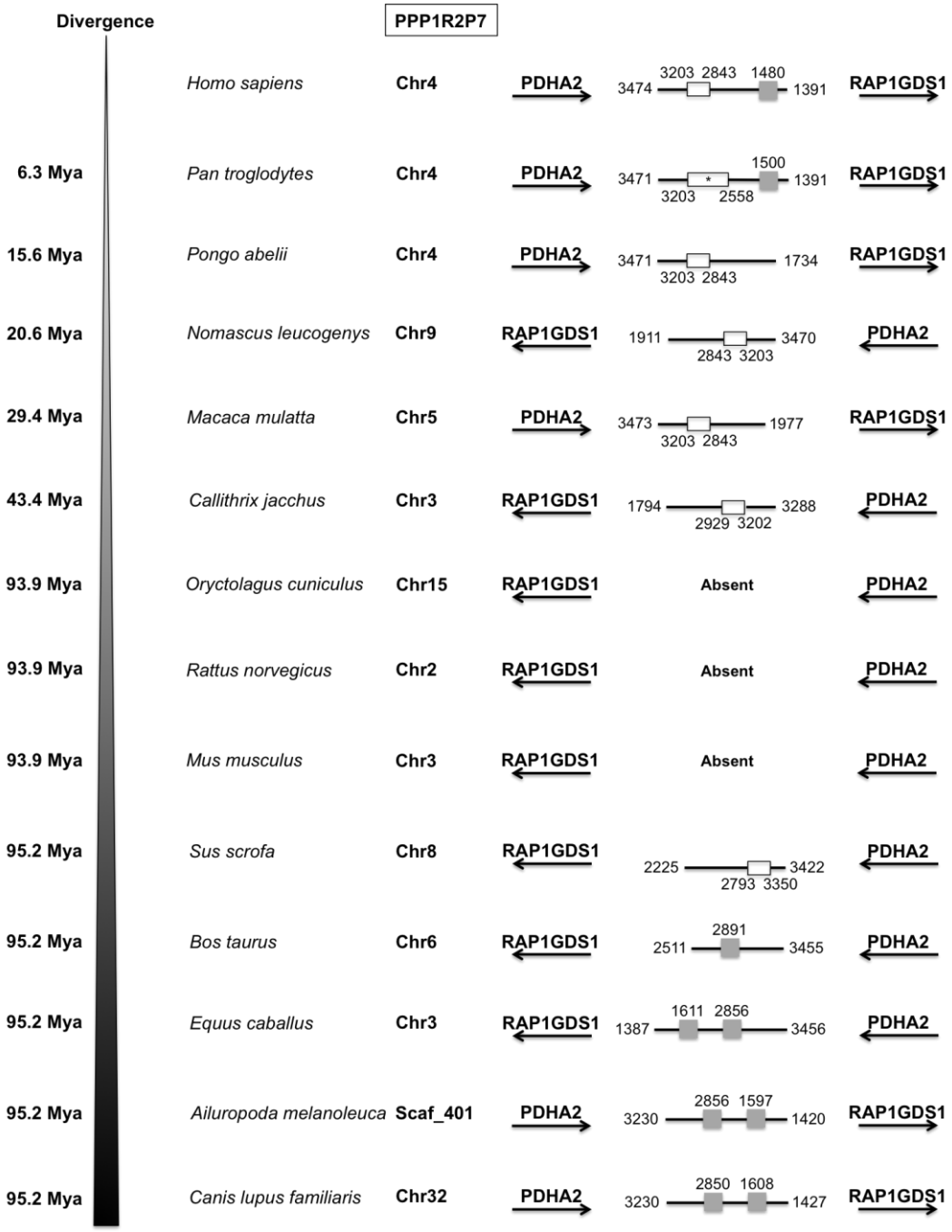


Figure II.B. 6: Conserved linkage of PPP1R2P7. *PPP1R2P7 location in terms of chromosome and flanking genes is presented concerning each species where it was found, showing the conserved linkage. Divergence time is shown on the left. The numbers flanking the pseudogenes are related to the parental PPP1R2 message. Grey boxes refer to the long interspersed elements (LINEs) LINE1 elements. White boxes delimited with a black line show a region that is absent and substituted by other unknown region. Numbers below the boxes show the region that is absent. * part of this sequence has unknown nucleotides and so the range (2558-3203bp) might be similar to the other species (2843-3203bp).*

PPP1R2P8

Evolutionary analysis of PPP1R2P8 shows that it was originated before the division of the Platyrrhini and Catarrhini (Figure II.B. 2). Phylogenetic analysis was performed only in the *Homo* and *Pan* genera since in the other species only the 3'UTR was obtained. In the NJ tree, it forms a robust cluster (bootstrap value 100) distant from the parental primate PPP1R2 group but more closely related to the Rodentia PPP1R2 group (Figure II.B. 1).

BLAST results show a 39% coverage and a 72% similarity with the parental human PPP1R2 message. PPP1R2P8 is located in the negative strand orientation of chromosome 5 of Hominidae and is flanked by the cadherin 18 (CDH18) and cadherin 12 (CDH12) genes showing a conserved linkage in all the primates analyzed (Figure II.B. 7). Furthermore, it is absent in gibbon, although the flanking genes are present and in orangutan only 146bp of the sequence still remain (Figure II.B. 5). Many frameshift disruptions and a large number of sequence repeats were inserted disrupting totally the sequence. These sequence repeats include LINEs (L1, L2 and CR1 families), SINEs (Alu/B1 and MIR families), DNA (hAT-Charlie and TcMar-Tigger families) and LTR (ERV1, ERVL and MaLR families) repeats and other DNA repeat elements (Figure II.B. 7). This pseudogene should be therefore transcriptionally silent.

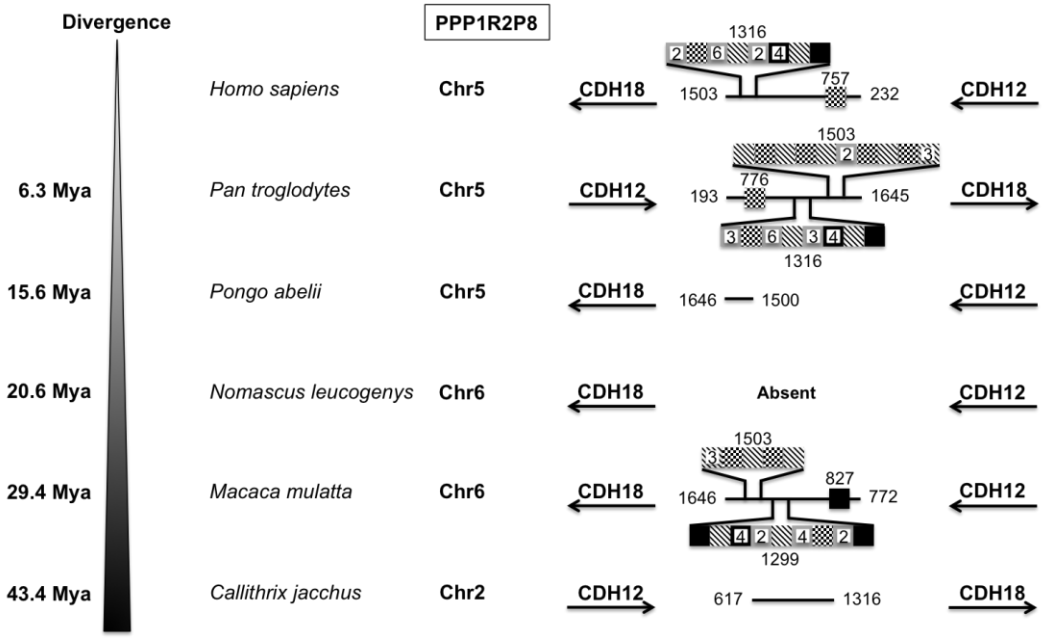


Figure II.B. 7 (previous page): Conserved linkage of PPP1R2P8. *PPP1R2P8 location in terms of chromosome and flanking genes is presented concerning each species where it was found, showing the conserved linkage. Divergence time is shown on the left. The numbers flanking the pseudogenes are related to the parental PPP1R2 message. Grey boxes refer to the long interspersed elements (LINEs) most LINE1 elements and one LINE2 element. Black boxes indicate the short interspersed elements (SINEs) most Alu repeats that are primate-specific but also others. Boxes with squares indicate long terminal repeat (LTR) endogenous retroviral-related (ERVL). Black traced white boxes indicate DNA-related repeats. Number above the boxes states the location where the repeat interrupted the sequence. Numbers inside the boxes indicate if there is more than one in line.*

PPP1R2P9

PPP1R2P9 is probably the first PPP1R2 pseudogene that has been retroposed. Indeed, evolutionary analysis shows that it was originated 163.9-167.4 Mya (Figure II.B. 2) and that it is widespread in mammals, being inclusively present in the marsupial gray short-tailed opossum. Phylogenetic analysis was performed and clearly shows a robust basal cluster independent of the upper cluster that includes all the other PPP1R2-related sequences in which the closest group is the PPP1R2P8 (Figure II.B. 1). This is the most divergent group of pseudogenes when comparing with the parental PPP1R2 message. BLAST results show a 39% coverage and 63% similarity when compared with the human PPP1R2 message. PPP1R2P9 is present always in chromosome X in the negative strand orientation in primates, although the strand orientation varies in the other order representatives (Figure II.B. 8).

This pseudogene was also found duplicated in marmoset (two copy), rat (three copies), mouse (three copies) and pig (two copy), none of them related to each other (Figure II.B. 5). This suggests that the parental PPP1R2 has been retroposed in this species more than once to chromosome X.

PPP1R2P9 is flanked by the calcium/calmodulin-dependent serine protein kinase (CASK) and the monoamine oxidase A (MAOA) genes, showing a conserved linkage (Figure II.B. 6). The CASK gene is 0.85Mb upstream and the MAOA gene is 0.88Mb downstream, and so, their promoters and regulatory sequences are very distant to provide the help to PPP1R2P9 transcription. The rest of the neighborhood is constituted by pseudogenes with very low transcription rates.

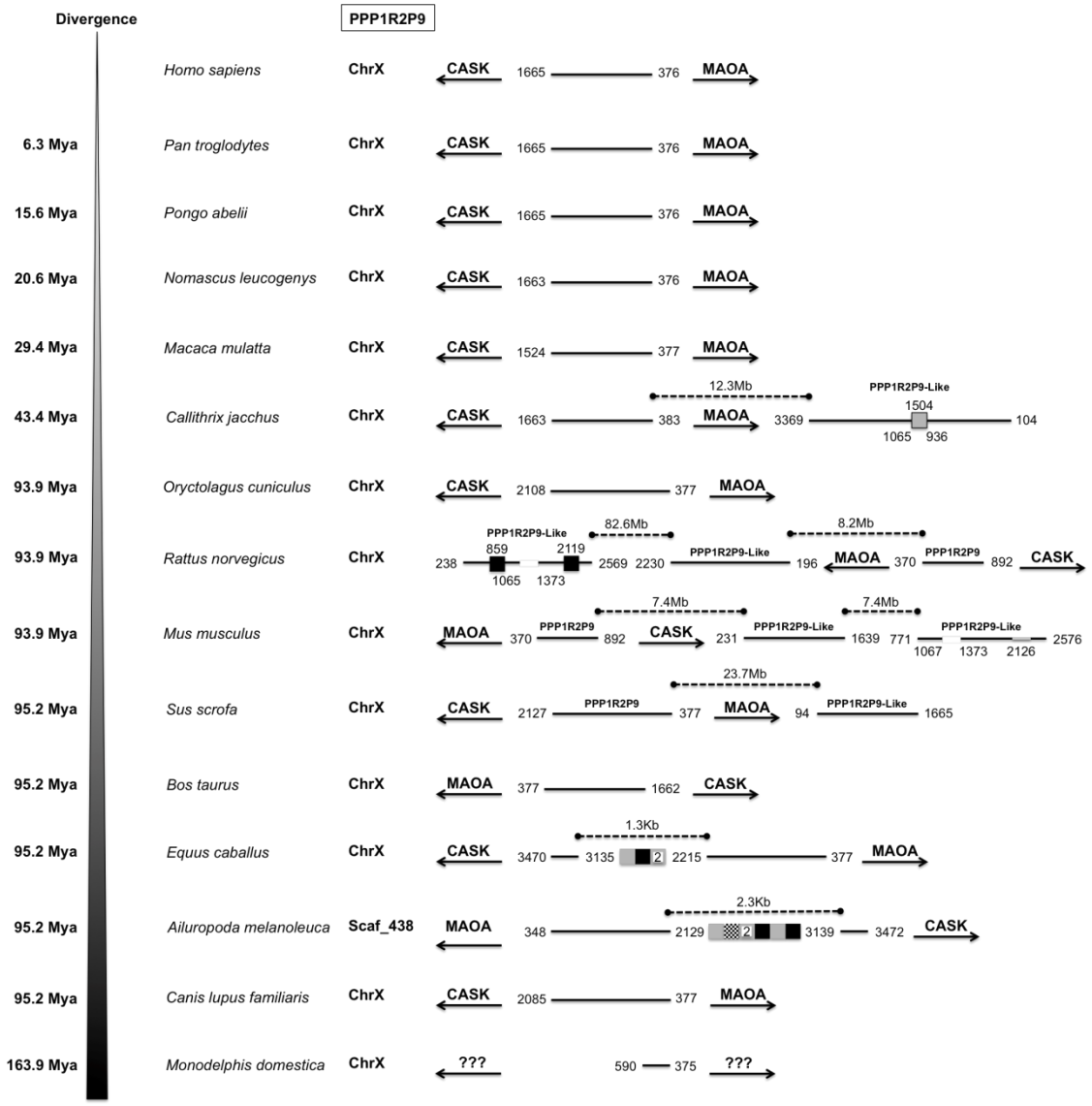


Figure II.B. 8: Conserved linkage of PPP1R2P9. *PPP1R2P9 location in terms of chromosome and flanking genes is presented concerning each species where it was found, showing the conserved linkage. Divergence time is shown on the left. The numbers flanking the pseudogenes are related to the parental PPP1R2 message. Grey boxes refer to the long interspersed elements (LINEs) LINE1 elements. Black boxes refer to the short interspersed elements (SINEs) B2 repeats that are rodent-specific and to the tRNAs present in the horse and in the giant panda sequences. Box with squares refers to long terminal repeat (LTR) in the giant panda sequence, which is an endogenous retroviral-related (ERVL). Number above the boxes states the location where the repeat interrupted the sequence. Numbers inside the boxes indicate if there is more than one in line. Grey box delimited with a black line in marmoset PPP1R2P9-like refer to a parental PPP1R2 insertion. Number on the top refers where the insertion took place in the pseudogene, while numbers at the bottom show which region of the parental PPP1R2 was inserted. In mouse an unknown sequence according to the current genome assembly was inserted in PPP1R9-like and is shown with a number referring to its location. Also, a deletion is shown (1067 to 1373) in mouse PPP1R2P9-like. The distances in dashed lines of the duplicated forms in human and orangutan are also indicated.*

Considering the ORF, Figure II.B. 6 shows that at least in primates it is not disrupted. However, the 5'UTR of the parental gene is completely absent and the 3'UTR is

truncated (671bp in humans). At the 3'UTR there is a single polyA signal in the position 1088 according with the human sequence, what might suggest that a shorter message is made. Sequence repeats, deletions, unknown and known sequence insertions were only found in the duplicated copies. The only exceptions are in mouse and rat (*Rattus norvegicus*) where the 3'UTR was completely deleted in the parental PPP1R2P9 (Figure II.B. 8).

This pseudogene is the one with more transcriptional data related. In Unigene, NCBI, two mRNAs and 9 ESTs, six from testis and one from brain were found. Also 704 Geo profiles and 102 experiments were found in NCBI and ArrayExpress (Ensembl), respectively. This suggests that silent regulatory areas were present in the region were PPP1R2P9 was retroposed and during the evolution PPP1R2P9 might have gain the ability to be transcribed.

PPP1R2P9 was originally found in cDNA libraries of human germ cell tumors. It was found to be 61.5% and 43.7% similar to PPP1R2 at the nucleotide and protein levels respectively, and to maintain all the important phosphorylation sites, PPP1C binding motifs and the nuclear polyadenylation signal (residues 133 to 142) [34]. The nucleotide similarity score is in accordance with our BLAST results and when comparing the human PPP1R2P9 protein with the parental gene product in ClustalW2 we found 113 amino acids changes. This means a 41% score, also very close to that obtained by Shirato and co-workers [34].

Furthermore it binds to PPP1C directly and in heat stable extracts inhibits potently with an IC50 of 0.2nM [34]. More recently it was found in human testis tissue by the MGC project team [35]. Also, PPP1R2P9 gene expression has shown to be affected in transgenic mice of STAT5-induced tumors [36].

Considering the CDS, all species show a continuous CDS with no or small truncations at the C-terminus (e.g. in mouse and rat), with the exception of pig where no protein translation was obtained from the CDS. Furthermore, no continuous CDS was obtained for the PPP1R2P9 duplications with the exception of rat PPP1R2P9-Like (the one found 8.2Mb upstream of the original rat PPP1R2P9, Figure II.B. 8).

Other pseudogenes and pseudogenic fragments

Besides the above-described pseudogenes, several pseudogenic fragments (100-600bp) were identified in all species analyzed. The larger ones (200-600bp) were localized to the 3'UTR when comparing to the parental PPP1R2 message, while the smaller ones (100bp) were localized to the CDS. In primates, only one pseudogenic

fragment is present and is mapped to the 3'UTR. This pseudogenic fragment appeared before the division of Platyrrhini and Catarrhini and is localized to chromosome 12 in humans (Figure II.B. 2).

Other pseudogenes that could not be assigned to the human ones and are species-specific, were found in giant panda (*Ailuropoda melanoleuca*), dog, rabbit (*Oryctolagus cuniculus*) and mouse. These were considered pseudogenes, since are larger than 600bp or, are larger than 400bp but contain part of the CDS. Mouse genome presented the largest number, six, and curiously, three of them are present in chromosome 17 and closely located, suggesting that they are duplications. All of the three have the CDS intact but lost the normal start codon, having however an upstream start codon that might produce, if transcribed, a longer product. This suggests that other retropositions were disseminated in the mammalian genomes.

GC-content

In the human genome, the average GC-content is ~41%, but the CDS of genes and the Alu repeats have normally higher GC percentage, ~53% and ~57%, respectively [71]. Also, GC-rich regions are more densely populated on functional genes [71].

Isochores are long stretches of DNA with a uniform GC-content. There is a clear tendency of LINE1 (long interspersed nuclear elements, family L1) to be present in GC-poor regions (isochores L1 and L2, <41%), Alu elements (SINEs, short interspersed nuclear elements) in GC-rich regions (isochores H2 and H3, >46%) and processed pseudogenes in intermediate GC-content regions (isochore H1, 41-46%) [71, 72]. Conserved transcribed processed pseudogenes tend to have a GC-content higher than the respective flanking regions [73].

Concerning the PPP1R2 gene, the CDS has a GC-content of 46% and is localized in a L2 isochore, which does not follow the normal patterns and could indicate that the pseudogenes will also have a lower GC-content, at least the closest ones.

Therefore, we searched if the human pseudogenes also have this in common and if they have higher probability of being transcribed. PPP1R2P1 has a putative CDS with a GC-content of 43.4%, being localized in a L2 isochore (37-42%) but surrounded by two H2 isochores (47-52%) where the TAP1 and HLA-DMB genes are present (Figure II.B. 3). This pseudogene is below the average of transcribed processed pseudogenes and the GC-content of PPP1R2P1 is not considerably higher than the surrounding flanking region. Taking together, the GC-content, chromosomal location, surrounding genes, conservation of the ORF and past gene expression data, we hypothesize that in humans this

pseudogene might be in fact expressed and may be functional at least in some individuals.

PPP1R2P3 is present in a region of isochore L2, which is also in accordance with the absence of genes in this region. Furthermore PPP1R2P3 GC-content of 46.3% is higher than the flanking region, 37-42% that is common in transcribed processed pseudogenes. This is in agreement with it being in fact transcribed and translated in testis and sperm [50].

PPP1R2P9 is present in a region that has a low GC-content, being localized in a L2 isochore and only represented by pseudogenes. This means that this is supposedly a low transcription region. Moreover, the GC-content of the CDS is 49.8%, one of the highest among the PPP1R2 pseudogenes and considerably higher than the surrounding flanking region (37-42%, L2). This is also in common with transcribed processed pseudogenes.

Concerning the other pseudogenes, most of them are also present in L2 isochore regions, with the exception of PPP1R2P6 that is present in a H1 isochore region. Also, since no ESTs were identified, except for PPP1R2P2, and many mutations in all the species analyzed have introduced frameshifts resulting in premature stop codons, we consider them to be transcriptionally silent, and so true pseudogenes.

PPP1R2-related proteins: protein detection

PPP1CC2, a sperm-specific protein phosphatase, is involved in sperm motility, since inhibition of its activity by okadaic acid and calyculin A causes both initiation and stimulation of motility in caput and cauda sperm, respectively [24, 25]. In fact, PPP1CC2 has two-fold higher activity in immotile bovine caput epididymal sperm compared to mature motile caudal sperm [24, 25]. Moreover, *in vivo*, this inhibition was associated with a PPP1R2-like activity since GSK3 was able to reverse the process [24]. We have previously shown that PPP1R2 and PPP1R2P3 are present in testis and sperm were they may account for this PPP1R2-like activity [50]. Testis is one of the organs where most pseudogenes are expressed and its gene products were shown to have important roles in spermatogenesis and other germ cell related functions [47-49]. This might be due, in part, to the transcription not being as tightly regulated, which could lead to activation of otherwise imperfect or weak promoters [49, 53]. PPP1R2 has been shown to be one of the PPP1C regulators with more pseudogenization [26]. Given that from all pseudogenes, only PPP1R2P1 and PPP1R2P9, besides PPP1R2P3, are putatively translated, we checked if we could obtain peptides representative of the pseudogenes from human ejaculated sperm by mass spectrometry [50]. The molecular weight of these pseudogenes

should be similar to the parental one, as it happens with PPP1R2P3, being therefore present in the same region where the band was extracted to mass spectrometry analysis. Also, since PPP1R2 is a heat stable inhibitor of PPP1C we tested if peptides related to those pseudogenes would appear in heat stable extracts. The antibody that we used to immunoprecipitate PPP1R2-related proteins was raised against a peptide containing amino acid residues 134-147 from the mouse PPP1R2 sequence. In this region of 14 residues, PPP1R2P1 is similar and PPP1R2P9 has three substitutions when comparing to PPP1R2 sequence. This antibody was used previously to detect PPP1R2 and PPP1R2P3 in human sperm samples [50]. Results from human sperm immunoprecipitate and heat stable extract were analyzed using an Orbitrap Velos mass spectrometer and we were able to identify 23 MSMS spectra corresponding to 8 different peptides matching unequivocally with PPP1R2P9 (Figure II.B. 9 and Table II.B. 2)

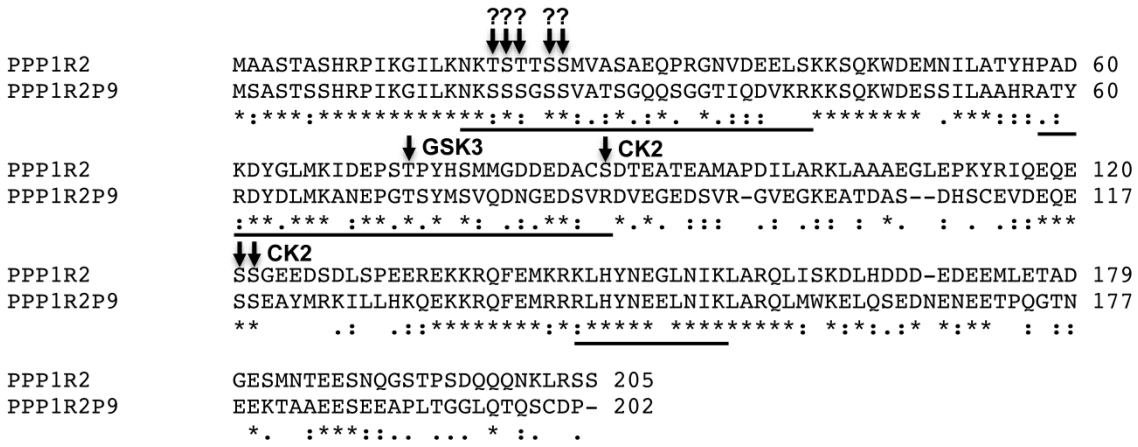


Figure II.B. 9: Protein alignment of PPP1R2P9 and PPP1R2 showing peptide obtained form mass spectrometry. A ClustalW2 alignment was performed using the protein sequence of PPP1R2P9 and PPP1R2. *Black arrows indicate the important phosphorylation sites in PPP1R2 and the respective known kinase. Phosphorylation sites detected by mass spectrometry and respective unknown kinase (?) are also depicted. Black bars at the bottom of each row of alignment show the region covered by the peptides obtained. * represent high conservation, : and . represent low conservation in which the substituted residue has respectively more and less similar properties.*

Table II.B. 2: *Peptides identified by Orbitrap Velos mass spectrometry for PPP1R2P9, from human sperm heat stable extracts and immunoprecipitates using rabbit anti-PPP1R2 antibodies. aa, amino acids; pI, isoelectric point.*

Protein Name	***Uniprot ID***	***MW (Da)***	***pI***	***Protein size (aa)***	***Coverage***	***Mascot score***
PPP1R2P9	IPP4_HUMAN	22,660	5.04	202	36,5%	381,09
Peptide		***Range (start-end)***	***Number of spectra***	***m/z meas.***	***z***	***Mascot score***
K.NKSSSGSSVATSGQQSGGTIQDVK.R		*17 - 40*	*6*	*770,71*	*3*	*21,95*
K.SSSGSSVATSGQQSGGTIQDVK.R		*19 - 40*	*6*	*1034,49*	*2*	*99,44*
K.SSSGSSVATSGQQSGGTIQDVKR.K		*19 - 41*	*4*	*1112,54*	*2*	*38,09*
R.LHYNEELNIK.L		*143 - 152*	*2*	*424,89*	*3*	*31,40*
K.ANEPGTSYMSVQDNGEDSVRDVEGEDSVR.G		*68 - 96*	*2*	*1053,45*	*3*	*63,05*
R.RLHYNEELNIK.L		*142 - 152*	*1*	*476,92*	*3*	*29,73*
R.ATYRDYDLMK.A		*58 - 67*	*1*	*431,20*	*3*	*28,17*
K.ANEPGTSYMSVQDNGEDSVR.D		*68 - 87*	*1*	*1086,46*	*2*	*69,26*

The sequence coverage obtained for PPP1R2P9 was 36.5% and the mascot score levels were 381,09 (in addition, spectra were manually evaluated). This is the first time that PPP1R2P9 protein is detected being clearly recovered from human ejaculated sperm. Furthermore, these results also indicate that native PPP1R2P9 is indeed a heat stable protein and that it migrates at the same position as the parental PPP1R2 and PPP1R2P3. Finally, we were also able to obtain PPP1R2P9 peptides with phosphorylations in serines 19, 20, 21, 23 and 24. Previous mass spectrometry studies also identified Ser20 of parental human PPP1R2 and Ser23 in the parental mouse PPP1R2 phosphorylated [74-77]. The relevance of these phosphorylations in parental PPP1R2 and in PPP1R2P9 still needs to be assessed, however no phosphorylations were obtained for PPP1R2P9 regarding the important functional residues, Thr72 and Ser120/121. Furthermore, no peptides were obtained for PPP1R2P1, which suggests that at least in sperm the protein is not present, although this possibility cannot be excluded.

The human nucleotide sequences of PPP1R2P1/3 and 9 were translated and the corresponding amino acid sequence aligned with the human parental PPP1R2 using ClustalW tool (Figure II.B. 10).

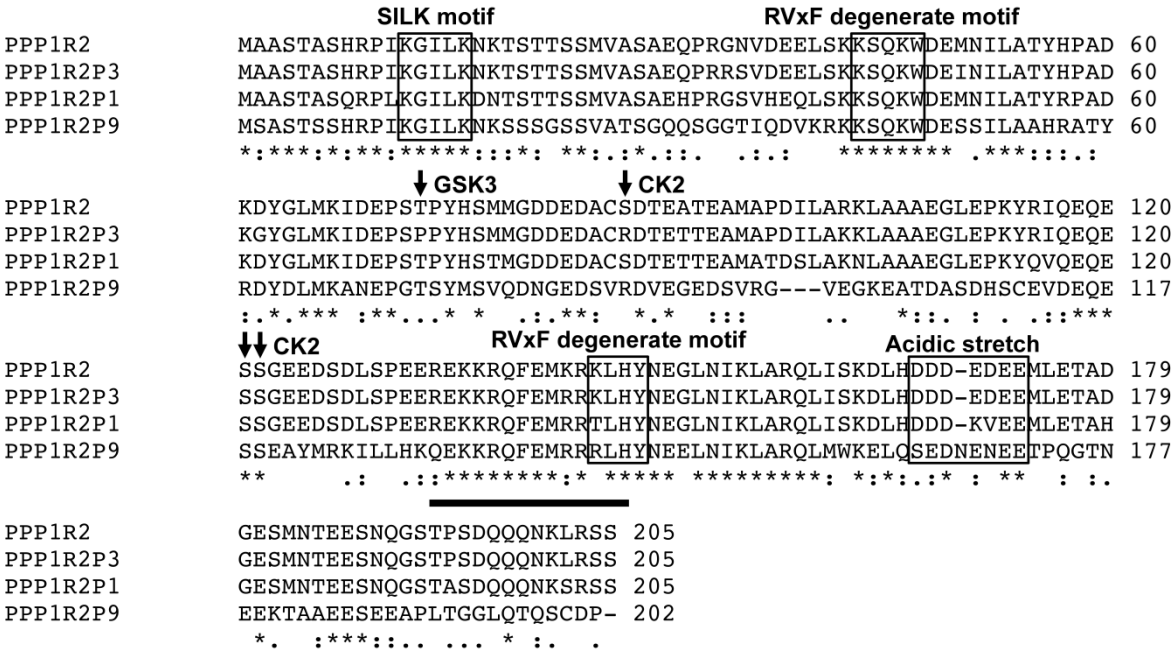


Figure II.B. 10: Alignment of PPP1R2-related proteins reveals high conservation. *A ClustalW2 alignment was performed using the protein sequences of PPP1R2P1, PPP1R2P3, PPP1R2P9 and PPP1R2. Black arrows indicate the important phosphorylation sites in PPP1R2 and the respective known kinase. Black boxes encage each PPP1 binding motif known for PPP1R2 and the acidic stretch. Black bar indicate the peptide for which the the antibody used for immunoprecipitation was raised. * represent high conservation, : and . represent low conservation in which the substituted residue has respectively more and less similar properties.*

Alignment shows that PPP1R2P9 is the most divergent (41%) and PPP1R2P3 is the most similar (95%) when comparing with the parental PPP1R2 (Figure II.B. 10). Concerning the important regions, we observe that Ser120 and 121, SILK and KSQKW motifs and the nuclear localization signal (residues 133 to 142) are conserved in all PPP1R2-related proteins. The other PPP1 binding motif KLHY is conserved in PPP1R2P3 but a substitution in the first residue to Thr or Arg is observed for PPP1R2P1 or PPP1R2P9, respectively. The acidic stretch (DDDEDEE) required fo

r GSK3 phosphorylation is maintained in PPP1R2P3 although the GSK3 phosphorylation site is substituted to Pro. The other two PPP1R2-related proteins maintain the GSK3 phosphorylation site but the acidic stretch has several changes particularly in PPP1R2P9. Finally, concerning the CK2 phosphorylation site important to enhance GSK3 phosphorylation, it is conserved in PPP1R2P1 but substituted by Arg in PPP1R2P3 and PPP1R2P9. Overall, the results show that these three PPP1R2-related proteins should maintain the ability to bind to PPP1C as was already demonstrated for PPP1R2P3 and PPP1R2P9 [34, 50]. Also, the ability to regulate the holoenzyme activity

by GSK3 phosphorylation appears to be compromised in PPP1R2P3, as we have shown previously [50], and at less extent in PPP1R2P9, due to the Ser87 change to Arg.

Conclusions

Evidences of retroposons from the gene PPP1R2 is ancient, prior to the great radiation of the mammals, supported by the presence of PPP1R2P9 and PPP1R2P7 in the different groups of mammals. All the other pseudogenes found are primate-specific and were retroposed at different times during the evolution of this group. For instance, PPP1R2P3 exists only in the members of the Hominoidea family, whereas PPP1R2P8, the most distinct, is present in all groups and was retroposed 42.6-65.2 Mya. This reveals that retropositions have occurred in waves and in a unique way similar to the Alu repeats explosion that occurred 40-50 Mya after the divergence of simian ancestors from the prosimians (lemurs and lorises) [63]. The recent pseudogene duplication in humans, PPP1R2P4, and in chimpanzee, PPP1R2P1, suggests that evolution of pseudogenes is still an active process.

In this study we clearly confirm that PPP1R2P9 is also present in human sperm and is a heat stable protein in its native form. The importance of these PPP1R2-related proteins in physiological conditions, such as spermatogenesis and sperm physiology, is relevant for further studies. Besides this, PPP1R2P1 and PPP1R2P9 were also found associated with pathological conditions [32, 34, 36, 62].

Furthermore, it has been shown that pseudogenes can regulate their parental counterparts at the message level either by siRNA generation and consequent gene silencing or by competition for positive and negative stabilizing factors and miRNAs, leading to an alteration of the parental mRNA levels [78]. Although PPP1R2P2 translation is very unlikely, its expression is stated and so, it is feasible that PPP1R2P2 or even other pseudogene messages could regulate the parental PPP1R2 message levels and therefore its function.

These observations indicate that PPP1R2 pseudogenes have possible biological functions and could not be non-functional relics as was initially believed and that there evolution process might be in part related with the formation of new genes and the gain of new specific functions. Therefore, their denomination as pseudogenes must be rethought.

Acknowledgements and Funding

We thank Srinivasan Vijayaraghavan (Anatomy and Cell Biology Department, Kent State University, Kent, Ohio, USA) for kindly providing the rabbit anti-PPP1R2 antibody. FCT supported the doctoral fellowship of Luis Korrodi-Gregório (SFRH/BD/41751/2007) and the post-doctoral fellowship of Joana Abrantes (SFRH/BPD/73512/2010) and Pedro J. Esteves (SPRH/BPD/27021/2006).

References

1. Cohen PT. Protein phosphatase 1--targeted in many directions. J Cell Sci 2002; 115: 241-256.
2. Ceulemans H, Bollen M. Functional diversity of protein phosphatase-1, a cellular economizer and reset button. Physiol Rev 2004; 84: 1-39.
3. DePaoli-Roach AA. Synergistic phosphorylation and activation of ATP-Mg-dependent phosphoprotein phosphatase by F A/GSK-3 and casein kinase II (PC0.7). J Biol Chem 1984; 259: 12144-12152.
4. Aitken A, Holmes CFB, Campbell DG, Resink TJ, Cohen P, Leung CTW, Williams DH. Amino acid sequence at the site on protein phosphatase inhibitor-2, phosphorylated by glycogen synthase kinase-3. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 1984; 790: 288-291.
5. Holmes CF, Kuret J, Chisholm AA, Cohen P. Identification of the sites on rabbit skeletal muscle protein phosphatase inhibitor-2 phosphorylated by casein kinase-II. Biochim Biophys Acta 1986; 870: 408-416.
6. Wang QM, Guan K-L, Roach PJ, DePaoli-Roach AA. Phosphorylation and Activation of the ATP-Mg-dependent Protein Phosphatase by the Mitogen-activated Protein Kinase. Journal of Biological Chemistry 1995; 270: 18352-18358.
7. Puntoni F, Villamoruzzi E. Phosphorylation of the Inhibitor-2 of Protein Phosphatase-1 by cdc2-Cyclin B and GSK3. Biochemical and Biophysical Research Communications 1995; 207: 732-739.
8. Agarwal-Mawal A, Paudel HK. Neuronal Cdc2-like Protein Kinase (Cdk5/p25) Is Associated with Protein Phosphatase 1 and Phosphorylates Inhibitor-2. Journal of Biological Chemistry 2001; 276: 23712-23718.
9. Alessi DR, Street AJ, Cohen P, Cohen PTW. Inhibitor-2 functions like a chaperone to fold three expressed isoforms of mammalian protein phosphatase-1 into a conformation with the specificity and regulatory properties of the native enzyme. European Journal of Biochemistry 1993; 213: 1055-1066.
10. Bennett D, Szöőr B, Alphey L. The Chaperone-like Properties of Mammalian Inhibitor-2 Are Conserved in a Drosophila Homologue†,‡. Biochemistry 1999; 38: 16276-16282.
11. Hurley TD, Yang J, Zhang L, Goodwin KD, Zou Q, Cortese M, Dunker AK, DePaoli-Roach AA. Structural basis for regulation of protein phosphatase 1 by inhibitor-2. J Biol Chem 2007; 282: 28874-28883.
12. Park IK, Roach P, Bondor J, Fox SP, DePaoli-Roach AA. Molecular mechanism of the synergistic phosphorylation of phosphatase inhibitor-2. Cloning, expression, and site-directed mutagenesis of inhibitor-2. Journal of Biological Chemistry 1994; 269: 944-954.
13. Yang J, Hurley TD, DePaoli-Roach AA. Interaction of Inhibitor-2 with the Catalytic Subunit of Type 1 Protein Phosphatase. Journal of Biological Chemistry 2000; 275: 22635-22644.
14. Kirchhefer U, Baba HA, Bokník P, Breeden KM, Mavila N, Brüchert N, Justus I, Matus M, Schmitz W, DePaoli-Roach AA, Neumann J. Enhanced cardiac function in mice overexpressing protein phosphatase Inhibitor-2. Cardiovascular Research 2005; 68: 98-108.
15. Yamada M, Ikeda Y, Yano M, Yoshimura K, Nishino S, Aoyama H, Wang L, Aoki H, Matsuzaki M. Inhibition of protein phosphatase 1 by inhibitor-2 gene delivery ameliorates heart failure progression in genetic cardiomyopathy. The FASEB Journal 2006; 20: 1197-1199.
16. Brüchert N, Mavila N, Boknik P, Baba HA, Fabritz L, Gergs U, Kirchhefer U, Kirchhof P, Matus M, Schmitz W, DePaoli-Roach AA, Neumann J. Inhibitor-2 prevents protein phosphatase 1-induced cardiac hypertrophy and mortality.

American Journal of Physiology - Heart and Circulatory Physiology 2008; 295: H1539-H1546.
17. Eto M, Bock R, Brautigan DL, Linden DJ. Cerebellar Long-Term Synaptic Depression Requires PKC-Mediated Activation of CPI-17, a Myosin/Moesin Phosphatase Inhibitor. Neuron 2002; 36: 1145-1158.
18. Leach C, Shenolikar S, Brautigan DL. Phosphorylation of Phosphatase Inhibitor-2 at Centrosomes during Mitosis. Journal of Biological Chemistry 2003; 278: 26015-26020.
19. Satinover DL, Leach CA, Stukenberg PT, Brautigan DL. Activation of Aurora-A kinase by protein phosphatase inhibitor-2, a bifunctional signaling protein. Proceedings of the National Academy of Sciences of the United States of America 2004; 101: 8625-8630.
20. Li M, Stukenberg PT, Brautigan DL. Binding of Phosphatase Inhibitor-2 to Prolyl Isomerase Pin1 Modifies Specificity for Mitotic Phosphoproteins†. Biochemistry 2007; 47: 292-300.
21. Wang W, Stukenberg PT, Brautigan DL. Phosphatase Inhibitor-2 Balances Protein Phosphatase 1 and Aurora B Kinase for Chromosome Segregation and Cytokinesis in Human Retinal Epithelial Cells. Molecular Biology of the Cell 2008; 19: 4852-4862.
22. Khandani A, Mohtashami M, Camirand A. Inhibitor-2 induced M-phase arrest in <i>Xenopus</i> cycling egg extracts is dependent on MAPK activation. Cellular & Molecular Biology Letters 2011; 16: 669-688.
23. Zambrano CA, Egaña JT, Núñez MT, Maccioni RB, González-Billault C. Oxidative stress promotes τ dephosphorylation in neuronal cells: the roles of cdk5 and PP1. Free Radical Biology and Medicine 2004; 36: 1393-1402.
24. Vijayaraghavan S, Stephens DT, Trautman K, Smith GD, Khatra B, da Cruz e Silva EF, Greengard P. Sperm motility development in the epididymis is associated with decreased glycogen synthase kinase-3 and protein phosphatase 1 activity. Biol Reprod 1996; 54: 709-718.
25. Smith GD, Wolf DP, Trautman KC, da Cruz e Silva EF, Greengard P, Vijayaraghavan S. Primate sperm contain protein phosphatase 1, a biochemical mediator of motility. Biol Reprod 1996; 54: 719-727.
26. Ceulemans H, Stalmans W, Bollen M. Regulator-driven functional diversification of protein phosphatase-1 in eukaryotic evolution. BioEssays 2002; 24: 371-381.
27. Li M, Satinover DL, Brautigan DL. Phosphorylation and functions of inhibitor-2 family of proteins. Biochemistry 2007; 46: 2380-2389.
28. Rouchka EC, Cha IElizabeth. Current Trends in Pseudogene Detection and Characterization. Current Bioinformatics; 4: 112-119.
29. Maestre J, Tchenio T, Dhellin O, Heidmann T. mRNA retroposition in human cells: processed pseudogene formation. EMBO J 1995; 14: 6333-6338.
30. Weiner AM, Deininger PL, Efstratiadis A. Nonviral Retroposons: Genes, Pseudogenes, and Transposable Elements Generated by the Reverse Flow of Genetic Information. Annual Review of Biochemistry 1986; 55: 631-661.
31. Balakirev ES, Ayala FJ. PSEUDOGENES: Are They “Junk” or Functional DNA? Annual Review of Genetics 2003; 37: 123-151.
32. Wu I, Moses MA. Cloning of a cDNA encoding an isoform of human protein phosphatase inhibitor 2 from vascularized breast tumor. DNA Seq 2001; 11: 515-518.
33. Yamada Y, Watanabe H, Miura F, Soejima H, Uchiyama M, Iwasaka T, Mukai T, Sakaki Y, Ito T. A Comprehensive Analysis of Allelic Methylation Status of CpG Islands on Human Chromosome 21q. Genome Research 2004; 14: 247-266.

34. Shirato H, Shima H, Sakashita G, Nakano T, Ito M, Lee EYC, Kikuchi K. Identification and Characterization of a Novel Protein Inhibitor of Type 1 Protein Phosphatase†,‡. Biochemistry 2000; 39: 13848-13855.
35. Gerhard DS, Wagner L, Feingold EA, Shenmen CM, Grouse LH, Schuler G, Klein SL, Old S, Rasooly R, Good P, Guyer M, Peck AM, et al. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). Genome Res 2004; 14: 2121-2127.
36. Eilon T, Barash I. Distinct gene-expression profiles characterize mammary tumors developed in transgenic mice expressing constitutively active and C-terminally truncated variants of STAT5. BMC Genomics 2009; 10: 231.
37. Hirotsune S, Yoshida N, Chen A, Garrett L, Sugiyama F, Takahashi S, Yagami K-i, Wynshaw-Boris A, Yoshiki A. An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. Nature 2003; 423: 91-96.
38. Piehler A, Hellum M, Wenzel J, Kaminski E, Haug K, Kierulf P, Kaminski W. The human ABC transporter pseudogene family: Evidence for transcription and gene-pseudogene interference. BMC Genomics 2008; 9: 165.
39. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, Surani MA, Sakaki Y, et al. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 2008; 453: 539-543.
40. Hawkins PG, Morris KV. Transcriptional regulation of Oct4 by a long non-coding RNA antisense to Oct4-pseudogene 5. Transcription 2010; 1: 165-175.
41. Chiefari E, Iiritano S, Paonessa F, Le Pera I, Arcidiacono B, Filocamo M, Foti D, Liebhaber SA, Brunetti A. Pseudogene-mediated posttranscriptional silencing of HMGA1 can result in insulin resistance and type 2 diabetes. Nat Commun 2010; 1: 40.
42. Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature 2010; 465: 1033-1038.
43. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 1999; 41: 95-98.
44. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011; 28: 2731-2739.
45. Roberto R, Capozzi O, Wilson RK, Mardis ER, Lomiento M, Tuzun E, Cheng Z, Mootnick AR, Archidiacono N, Rocchi M, Eichler EE. Molecular refinement of gibbon genome rearrangements. Genome Research 2007; 17: 249-257.
46. Schmidt T, Frishman D. Assignment of isochores for all completely sequenced vertebrate genomes using a consensus. Genome Biology 2008; 9: R104.
47. Kleene KC, Mulligan E, Steiger D, Donohue K, Mastrangelo M-A. The Mouse Gene Encoding the Testis-Specific Isoform of Poly(A) Binding Protein (<i>Pabp2</i>) Is an Expressed Retroposon: Intimations That Gene Expression in Spermatogenic Cells Facilitates the Creation of New Genes. Journal of Molecular Evolution 1998; 47: 275-281.
48. Marques AC, Dupanloup I, Vinckenbosch N, Reymond A, Kaessmann H. Emergence of Young Human Genes after a Burst of Retroposition in Primates. PLoS Biol 2005; 3: e357.
49. Huang C-J, Lin W-Y, Chang C-M, Choo K-B. Transcription of the rat testis-specific Rtdpoz-T1 and -T2 retrogenes during embryo development: co-transcription and frequent exonisation of transposable element sequences. BMC Molecular Biology 2009; 10: 74.
50. Korrodi-Gregório L, Ferreira M, Vintém AP, Wu W, Muller T, Marcus K, Vijayaraghavan S, Brautigan DL, Silva OABdCe, Fardilha M, Silva EFdCe.

Discovery and characterization of human sperm protein phosphatase 1 inhibitor 2 proteins. (Submited to Biochem J 2011).
51. Helps NR, Street AJ, Elledge SJ, Cohen PTW. Cloning of the complete coding region for human protein phosphatase inhibitor 2 using the two hybrid system and expression of inhibitor 2 in E. coli. FEBS Letters 1994; 340: 93-98.
52. Sanseau P, Jackson A, Alderton RP, Beck S, Senger G, Sheer D, Kelly A, Trowsdale J. Cloning and characterization of human phosphatase inhibitor-2 (IPP-2) sequences. Mamm Genome 1994; 5: 490-496.
53. Vinckenbosch N, Dupanloup I, Kaessmann H. Evolutionary fate of retroposed gene copies in the human genome. Proceedings of the National Academy of Sciences of the United States of America 2006; 103: 3220-3225.
54. Horton R, Gibson R, Coggill P, Miretti M, Allcock R, Almeida J, Forbes S, Gilbert J, Halls K, Harrow J, Hart E, Howe K, et al. Variation analysis and gene annotation of eight MHC haplotypes: The MHC Haplotype Project. Immunogenetics 2008; 60: 1-18.
55. Mungall AJ, Palmer SA, Sims SK, Edwards CA, Ashurst JL, Wilming L, Jones MC, Horton R, Hunt SE, Scott CE, Gilbert JG, Clamp ME, et al. The DNA sequence and analysis of human chromosome 6. Nature 2003; 425: 805-811.
56. Allcock RJN, Atrazhev AM, Beck S, De Jong PJ, Elliott JF, Forbes S, Halls K, Horton R, Osoegawa K, Rogers J, Sawcer S, Todd JA, et al. The MHC haplotype project: A resource for HLA–linked association studies. Tissue Antigens 2002; 59: 520-521.
57. Stewart CA, Horton R, Allcock RJN, Ashurst JL, Atrazhev AM, Coggill P, Dunham I, Forbes S, Halls K, Howson JMM, Humphray SJ, Hunt S, et al. Complete MHC Haplotype Sequencing for Common Disease Gene Mapping. Genome Research 2004; 14: 1176-1187.
58. Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, Lush MJ, Povey S, Talbot CC, Wright MW, Wain HM, Trowsdale J, et al. Gene map of the extended human MHC. Nat Rev Genet 2004; 5: 889-899.
59. Traherne JA, Horton R, Roberts AN, Miretti MM, Hurles ME, Stewart CA, Ashurst JL, Atrazhev AM, Coggill P, Palmer S, Almeida J, Sims S, et al. Genetic Analysis of Completely Sequenced Disease-Associated MHC Haplotypes Identifies Shuffling of Segments in Recent Human History. PLoS Genet 2006; 2: e9.
60. Belov K, Deakin JE, Papenfuss AT, Baker ML, Melman SD, Siddle HV, Gouin N, Goode DL, Sargeant TJ, Robinson MD, Wakefield MJ, Mahony S, et al. Reconstructing an Ancestral Mammalian Immune Supercomplex from a Marsupial Major Histocompatibility Complex. PLoS Biol 2006; 4: e46.
61. Debenham SL, Hart EA, Ashurst JL, Howe KL, Quail MA, Ollier WER, Binns MM. Genomic sequence of the class II region of the canine MHC: comparison with the MHC of other mammalian species. Genomics 2005; 85: 48-59.
62. Pandita A, Balasubramaniam A, Perrin R, Shannon P, Guha A. Malignant and benign ganglioglioma: a pathological and molecular study. Neuro Oncol 2007; 9: 124-134.
63. Ohshima K, Hattori M, Yada T, Gojobori T, Sakaki Y, Okada N. Whole-genome screening indicates a possible burst of formation of processed pseudogenes and Alu repeats by particular L1 subfamilies in ancestral primates. Genome Biology 2003; 4: R74.
64. Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J, Guigo R, Dermitzakis ET. Transcriptome genetics using second generation sequencing in a Caucasian population. Nature 2010; 464: 773-777.
65. Wang ZQ, Floyd ZE, Qin J, Liu X, Yu Y, Zhang XH, Wagner JD, Cefalu WT. Modulation of Skeletal Muscle Insulin Signaling With Chronic Caloric Restriction in Cynomolgus Monkeys. Diabetes 2009; 58: 1488-1498.

66. Fardilha M, Esteves SLC, Korrodi-Gregório L, Vintém AP, Domingues SC, Rebelo S, Morrice N, Cohen PTW, Silva OABdCe, Silva EFdCe. Identification of the human testis protein phosphatase 1 interactome. Biochemical Pharmacology 2011; 82: 1403-1415.
67. Fardilha M, Esteves SLC, Korrodi-Gregório L, Pelech S, da Cruz e Silva OAB, da Cruz e Silva E. Protein phosphatase 1 complexes modulate sperm motility and present novel targets for male infertility. Molecular Human Reproduction 2011; 17: 466-477.
68. Mayya V, Han DK. Phosphoproteomics by mass spectrometry: insights, implications, applications and limitations. Expert Review of Proteomics 2009; 6: 605-618.
69. Gauci S, Helbig AO, Slijper M, Krijgsveld J, Heck AJR, Mohammed S. Lys-N and Trypsin Cover Complementary Parts of the Phosphoproteome in a Refined SCX-Based Approach. Analytical Chemistry 2009; 81: 4493-4501.
70. Misceo D, Capozzi O, Roberto R, Dell'Oglio MP, Rocchi M, Stanyon R, Archidiacono N. Tracking the complex flow of chromosome rearrangements from the Hominoidea Ancestor to extant Hylobates and Nomascus Gibbons by high-resolution synteny mapping. Genome Research 2008; 18: 1530-1537.
71. Zhang Z, Harrison PM, Liu Y, Gerstein M. Millions of Years of Evolution Preserved: A Comprehensive Catalog of the Processed Pseudogenes in the Human Genome. Genome Research 2003; 13: 2541-2558.
72. Pavlicek A, Gentles AJ, Pačes J, Pačes V, Jurka J. Retroposition of processed pseudogenes: the impact of RNA stability and translational control. Trends in Genetics 2006; 22: 69-73.
73. Khachane A, Harrison P. Assessing the genomic evidence for conserved transcribed pseudogenes under selection. BMC Genomics 2009; 10: 435.
74. Molina H, Horn DM, Tang N, Mathivanan S, Pandey A. Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. Proc Natl Acad Sci U S A 2007; 104: 2199-2204.
75. Yu Y, Yoon SO, Poulogiannis G, Yang Q, Ma XM, Villen J, Kubica N, Hoffman GR, Cantley LC, Gygi SP, Blenis J. Phosphoproteomic analysis identifies Grb10 as an mTORC1 substrate that negatively regulates insulin signaling. Science 2011; 332: 1322-1326.
76. Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, Beausoleil SA, Villen J, Haas W, Sowa ME, Gygi SP. A tissue-specific atlas of mouse protein phosphorylation and expression. Cell 2010; 143: 1174-1189.
77. Wisniewski JR, Nagaraj N, Zougman A, Gnad F, Mann M. Brain phosphoproteome obtained by a FASP-based method reveals plasma membrane protein topology. J Proteome Res 2010; 9: 3280-3289.
78. Pink RC, Wicks K, Caley DP, Punch EK, Jacobs L, Carter DR. Pseudogenes: pseudo-functional or key regulators in health and disease? RNA 2011; 17: 792-798.

Chapter II.C

Overexpression of PPP1R2 and PPP1R2P3 in mouse – a transgenic approach

In the prior chapters the presence and localization of PPP1R2 in spermatozoa as well as the relevance of related pseudogenes were discussed. In this chapter PPP1R2 and PPP1R2P3 will be overexpressed in testis using the mouse model to try to unravel the role of these proteins in mammalian sperm motility.

Introduction

A transgenic animal carries a foreign gene that has been deliberately inserted into its genome. In opposition, in a knock-out animal a target gene has been deliberately inactivated or "knocked-out". Nowadays both terms, transgenic and knock-out, are commonly used as transgenic, being the first referred as standard transgenic when we just want to emphasize this technique. Mice are the favorite model to use the standard transgenic and knock-out techniques because they are mammals, have a relatively low cost of maintenance and a generation time of only nine weeks. Besides this, mice naturally develop conditions that mimic human disease, such as cardiovascular disease, cancer and diabetes, and so are widely used to study inherited human diseases. These animals are also very useful for delineating the function of newly discovered genes as well as for producing useful proteins in large animals [1].

Transgenic mice production

There are three alternative methods to produce a transgenic animal: (i) DNA microinjection (ii) Embryonic stem (ES) cell-mediated gene transfer and (iii) Retrovirus-mediated gene transfer.

The first method involves the direct microinjection of a chosen gene construct into the pronucleus of a fertilized egg. It is one of the first methods that proved to be effective in mammals [2]. To minimize interference from the bacterial vector sequences, the transgene is usually excised from the plasmid vector before the microinjection. The insertion of the DNA is, however, a random process, and there is a high probability that the introduced gene will be inserted in silenced regions of the chromosome. Also multiple transgene copies can be inserted at a single locus and expression levels do not correlate

well with the copy number. The manipulated fertilized egg is then transferred into the oviduct of a pseudopregnant foster female that has been induced to act as a recipient by mating with a vasectomized male. A major advantage of this method is its applicability to a wide variety of species and is generally used in standard transgenic mice techniques.

The second method involves prior insertion of the desired DNA sequence by homologous recombination into an *in vitro* culture of ES cells [3]. These cells are then incorporated into an embryo at the blastocyst stage of development. The result is a chimeric animal. ES cell-mediated gene transfer is the method of choice for the knock-out technique. This technique works particularly well in mice and has the advantage of allowing precise targeting of defined mutations in the gene via homologous recombination.

The third method is used to increase the probability of gene expression [4]. The gene transfer is mediated with the help of a carrier or vector, generally a virus or a plasmid. Offspring derived from this method are chimeric and transmission of the transgene is only possible if the retrovirus integrates into some of the germ cells.

The success rate in terms of live birth of animals containing the transgene or in which the gene is disabled is extremely low. Providing that the genetic manipulation does not lead to abortion, the result is a first generation (F1) of animals that need to be tested for the expression of the transgene. Depending on the technique used, F1 generation will result in heterozygous mice, in the case of DNA microinjection, or in chimeras, in the case of the other methods. Subsequent crossing of the quimeras with wild type mice will result in heterozygous offspring. To obtain a homozygous mouse, a female and a male heterozygous need to be crossed. In all these steps confirmation of heterozygosity by PCR or Southern blot need to be performed. Consequent crossing of homozygous will guarantee an offspring of homozygous without the need of further confirmation.

Knock-out technique

The knock-out technique can be performed globally, by disabling the gene in all cells of the animal, or locally in specific cells, using the Cre recombinase system [5] or, more recently, the yeast-derived Flip (FLP) recombinase [6]. The first one is generally applied, as a first approach to get fast conclusions about the importance of a specific gene, but has the disadvantage that some genes can be embryonically or neonatal lethal. The second (Cre/FLP) is applied when the gene needs to be studied in a particular tissue or there is the need to overcome the embryonical lethality by inducing the knock-out at latter stage in a time-dependent manner, as a switch. Both use the homologous recombination to switch-off the relevant gene. The Cre-recombinase system, also called, Cre-*lox*P

system involves the action of Cre (causes recombination), a type I topoisomerase from P1 bacteriophage that catalyzes site-specific recombination of DNA between *lox*P (locus of X over in PI) sites, a 34 base pair (bp) sequence composed of two 13bp inverted repeats flanking an 8bp spacer region which confers directionality sites. Similarly, the system FLP-recombinase catalyzes the recombination targeted to short (34 bp) DNA recognition sequences termed FRT sites. Two mice strains are needed, one in which the region of the gene targeted for deletion is flanked by the *lox*P sites or FRT sites in the same orientation (floxed or flipped mice), and a second line in which the expression of Cre/FLP recombinase is under the control of a time- and/or tissue-specific promoter.

Besides the already stated advantages, Cre/FLP-recombinase systems can also be used to generate global knock-outs. So this flexibility allows that if global or different tissue-/time-dependent lines of Cre/FLP mice are available the role of the target gene can be studied at different places. The main difference between Cre and FLP recombination systems is that FLP system is very useful for the removal of the selection gene from the targeted gene at the ES cell stage.

The expression of Cre/FLP-recombinases in a time dependent manner can be achieved by placing the Cre/FLP expression under the control of a minimal promoter and by directing the expression of the inducer protein (transactivator) to certain cell types. Transcription of Cre/FLP in mice can be activated by administration of the small molecules, such as tetracycline [7], Tamoxifen [8], RU486 [9] or ecdysone [10]. The major disadvantage of these regulatory systems is the separate insertion of the gene encoding the transactivator and that of the controllable Cre/FLP into the genome leading to large numbers of transgenic mice, which have to be intercrossed and tested for the desired combination of the transgenes.

The knock-in technique uses the same system of the knock-out and consists of an endogenous gene that is replaced with a mutant variant in order to address the role of specific functional domains, amino acid residues, or signaling pathways *in vivo*.

Standard transgenic technique

The first widely used approach to study gene function *in vivo* was to produce transgenic mice that overexpressed target genes. This requires the full-length coding sequence (cDNA) of a gene to be cloned downstream of a promoter that will provide ubiquitous or tissue-specific expression. The major advantages of the transgenic approach are that it is relatively straightforward and inexpensive, high levels of target gene expression can be achieved, and transgenic overexpressing mice often demonstrate

an obvious phenotype. One disadvantage is that the site of integration of the transgene into the genome can seriously affect tissue specificity and levels of transgene expression, like for instance in regions of heterochromatin, so a high number of founder lines may need to be screened. Also the transgene can be incorporated in another gene disabling it, and leading to a wrong analysis of the phenotype. However, even well characterized promoters are often expressed at lower levels in non-target tissues, so rigorous analysis should include examination of transgene expression in a range of tissues. Although overexpression models by their nature produce nonphysiological levels of gene and protein expression, they can still provide valuable insights into normal gene function in vivo [11].

Vector construction

The vector construction must take into account the final purpose of the experiment. In the standard transgenic technique, a transgene typically contains a promoter, a cDNA, an intron, and a polyadenylation signal. The cDNA should include at least the CDS and part of the 3'UTR and the intron should be added to stimulate the transport of mRNA out of the nucleus, since this process is coupled to the splicing process [12-14]. The artificial intron should be placed at the 3-prime end of the cDNA. The SV40 polyadenylation sequence is often used because of its strong polyadenylation signal that helps to stabilize the message. Also restriction sites should be added for excision of the entire transgene before microinjection. Tags are also often used in order to distinguish the transgene from the wild type form (Table II.C. 1).

Table II.C. 1: *Common tags used in transgenic mice systems.*

Tag	Epitope	Description	Reference
GFP		green fluorescent protein and its variants, yellow, cyan and red fluorescent proteins (YFP, CFP and RFP), enables the visualization and localization without any further staining. It was isolated from jellyfish *Aequorea aequorea* and is a 238 amino acid protein with an apparent molecular weight of about 27-30 kDa	[15]

Tag	Epitope	Description	Reference
AU1	DTYRYI	derived from the major capsid protein of bovine papillomavirus-1	[16]
HA	YPYDVPDYA	derived from the human influenza hemagglutinin protein	[17]
Myc	EQKLISEEDL	derived from the human protooncogene myc	[18]
V5	GKPIPNPLLGLDST	derived from a small epitope (Pk) present on the P and V proteins of the paramyxovirus of simian virus 5	[18]
FLAG	DYKDDDDK	The first one to be described, it's a peptide highly hydrophilic	[19]

Multiple tags, double or triple, using 1 or 2 amino acid linkers between tags are now starting to be used (ex. 3xmyc, [20]; 3xHA [21], 3xFlag [22], and apparently help to increase the signal [23].

The choice of the appropriate promoter is extremely important. This is particularly evident in meiotic spermatogenic cells where a failure to mimic the correct expression of the transgene could lead to a decreased expression. To mimic the expression of the selected transgene, the same promoter, a variation or another well-studied promoter should be chosen. The promoters frequently used can be divided in viral and eukaryotic promoters [24]. The viral promoters were the first to be used, however, eukaryotic cells have evolved mechanisms to detect and silence viral transgene expression. Viral promoters have shown a frequent inability to sustain transgene expression *in vivo*, but despite this a large proportion of gene therapy applications continue to use them to drive transgene expression. Examples include the ubiquitous cytomegalovirus immediate early (CMV-IE) promoter and enhancer, the simian virus 40 (SV40), the Rous sarcoma virus long terminal repeat (RSV-LTR) and the Moloney murine leukemia virus (MoMLV) LTR promoters [24].

The eukaryotic promoters emerging prove to be highly advantageous in achieving long-term expression *in vivo* and can also drive the expression in a tissue-/cell type-dependent manner. They are classified as ubiquitous/non-tissue specific or tissue-/cell type-/disease-specific promoters. Examples of non-tissue specific promoters are the elongation factor 1α (EF1α) and the human phosphoglycerate kinase 1 (PGK1) [25]. The list of tissue-/disease-specific promoters has expanded considerably over the past few

years. Examples of tissue-specific promoters are various, the human α1-antitrypsin (hAAT) in liver [26], the adipocyte P2 in fat cells [27], the myosin light-chain in muscle [28], the amylase in acinar pancreas [29] and insulin in islets of Langerhans β cells [30]. In spermatogenic cells, the human PGK2 human drives the transgene expression at the stages of meiotic preleptotene spermatocytes (transcription) and pachytene spermatocytes (translation) (-445/+70, 515bp [31, 32], whereas mouse protamine promoter (mPrm) 1 and 2 drives the expression in the late stages of spermatogenesis, in round (transcription) and elongated spermatids (translation) (mPrm1 -240/-37, 203bp; mPrm2 -170/-82, 88bp, [33-36].

In the case of knock-out and knock-in techniques, to design the gene targeting constructs one should also take into account the homology arms in order to the homologous recombination to occur. The homology arms consist of 2 long segments of genomic DNA (gDNA) that flank a selection cassette. The gDNA homology arms can undergo a double-reciprocal recombination event with their matching sequences on one chromosome, carrying the selection cassette with them. The selection cassette thereby replaces the gDNA between the regions of homology on the chromosome. Also positive and negative selection cassettes are generally introduced to select positive ES cells (ES cells that have suffer homologous recombination). One of the most commonly used positive selection marker is the *neo*^{r}, a gene that encodes an enzyme that inactivates the antibiotic neomycin and its relatives, like the drug G418, which is lethal to mammalian cells. Therefore only positive ES will survive. However, the presence of *neo*^{r} in an intron can result in an alteration of gene function and therefore produce an unwanted or even lethal phenotype [37]. This problem can be avoided by the use both the Cre and FLP recombination systems.

A targeting vector containing both a FLP-flanked neo^r marker and a *lox*P-flanked exon can be introduced into ES cells. After selection, the neo^r can be removed with FLP recombinase before the ES cells be injected into host blastocysts. With this system, the chimeric offspring contain only a minimal genetic modification (the addition of two *lox*P sites and one Frt site) in the gene of interest, limiting the probability of a complicating phenotype.

One commonly used negative selection marker contains the herpes simplex virus gene that encodes for thymidine kinase (tk). Thymidine kinase is an enzyme that phosphorylates the nucleoside analogue ganciclovir. Viral ganciclovir is a non-toxic drug for the cells but when mono-phosphorylated by tk and subsequently bi- and tri-phosphorylated by cellular kinases is incorporated into the DNA of replicating cells by

DNA polymerase, blocking the cell cycle and inducing apoptosis [38]. Cells that are transfected and exposed to ganciclovir can also kill adjacent, untransfected cells by the so-called bystander effect, in which transfer of phosphorylated ganciclovir molecules occur between neighbor cells via gap junctions [39]. The tk negative selection cassette is cloned into the targeting construct outside of the homology arms, and so it will not be incorporated during homologous recombination. But it will be incorporated during most random integrations helping to select against those clones [40].

Another negative selection marker is the diphtheria toxin A (DTA) gene. The A subunit inhibits protein synthesis via covalent modification of elongation factor 2 and has the advantage over tk/ganciclovir because is not taken up by other cells and works without the need of a second drug [41].

Objective

The aim of this work was to use the DNA microinjection technique to achieve PPP1R2 and PPP1R2P3 overexpression in mice testis in order to address the functions of those two PPP1 inhibitors in spermatogenesis and sperm physiology.

Material and Methods

Plasmids construction

For the generation of transgenic mice two constructs were produced using the pBlueScriptII(SK-) vector (Stratagene now Agilent Technologies UK Ltd, Edinburgh, UK). Both constructs were designed in a similar way, having a human PGK2 promoter (hPGK2p), 3x HA tag, the cDNA and the SV40 polyadenylation signal. The cDNAs coded for hPPP1R2 (inhibitor-2) and hPPP1R2P3 (also called inhibitor-2-like). The parental pBlueScriptII(SK-) plasmid containing the hPGK2p, PPP1CC and the SV40 polyadenylation signal was kindly provided by our collaborator Prof. Srinivasan Vijayaraghavan [42]. Parental plasmid was cut with EcoRI and XhoI (all restriction enzymes were bought from NEB, New England Biolabs (UK), Herts, UK) in order to remove the PPP1CC cDNA, run in an extraction 1% agarose gel and the corresponding band purified using QIAquick spin purple columns (QIAGEN, Dusseldorf, Germany). ClaI and SalI sites were taken out from the original vector. hPPP1R2 and hPPP1R2P3 cDNAs were excised from pET28a-hPPP1R2 and pET28a-hPPP1R2P3 (MERCK, Darmstadt, Germany) plasmids already described, using the same enzymes. Both cDNAs corresponding to the CDS with start and stop codons were prepared using the same methodology has for the vector backbone. Ligations were performed using T4 DNA ligase (NEB (UK), Herts, UK) according to manufacturer's protocol thereby producing the final plasmids pBlueScriptII(SK-)-hPGK2-PPP1R2/PPP1R2P3-SV40.

The 3x HA tag fragment with restriction sites overhangs for EcoRI was obtained by hybridization of the two primers 3HA-FW (108nt, 5´-attcatgtacccatacgatgttccagattacgctaccggatacccatacgatgttccagattacgctaccggatacccatacgatgttccagattacgctaccggag-3´) and 3HA-RV (108nt, 5´-aattctccggtagcgtaatctggaacatcgtatgggtatccggtagcgtaatctggaacatcgtatgggtatccggtagcgtaatctggaacatcgtatgggtacatg-3´) (all primers were synthesized by Eurofins MWG Operon Ebersberg, Germany). In brief both primers were resuspended in a volume of Tris annealing buffer (10mM Tris, pH 8.0), at the same molar concentration (100mM), then a dilution of 10mM was obtained and the primers mixed in equal volumes in a 1.5mL eppendorf tube. The tube was placed in a heatblock at 95 °C for 5min., and subsequently cooled to room temperature. Then, the 3x HA fragment was ligated as explained before producing the final constructs pBlueScriptII(SK-)-hPGK2-3xHA-hPPP1R2/hPPP1R2P3-SV40. In all steps the positive clones were sequenced in the ABI PRISM 310 Genetic

Analyser (Portugal Applied Biosystems, Porto, Portugal) to ensure the right orientation and sequence of the product.

DNA microinjection testing

In order for the constructs to be sent to the Case Western Reserve University (CWRU, Cleveland, Ohio, USA), Case Transgenic and Targeting Facility the following was done: 200ug of pure plasmid DNA kit-purified at a concentration of 1 µg/µl or greater were prepared, information was given on the restriction enzymes which will allow to cut and gel-purify the insert from the backbone, a photo of a gel with ladder showing the insert released from the plasmid backbone by restriction digest with the same enzymes that will be used to gel-purify the insert and a demonstration that 1 copy or less of the transgene could be detected by Southern blot or PCR methods in a background of mouse genomic DNA, using an assay on 1, 0.1, 0.01, 0.001 transgene copy equivalents until reach the point of no detection, in what is called, spike the plasmid, was also performed.

In detail a bulk production of the plasmid was performed using the PureYield™ Plasmid Maxiprep System (Promega, Southampton, UK). Then a restriction digest using BamHI and KpnI restriction enzymes was performed in order to release the insert, PGK2-3xHA-hPPP1R2/hPPP1R2P3-SV40 from the backbone, pBlueScriptII(SK-). The restriction cut was run in a 1% agarose gel along with 1kb plus DNA ladder (**Life Technologies S.A.,** Madrid, Spain) and pictures were taken using the AlphaImager HP system (Fisher Scientific, Loures, Portugal). Finally in order to spike the plasmid, first an ear punch from a wild type SJL/BL6 mouse was obtained and the tissue was digested in a heat block for 1hr at 95º in 50µL of alkaline lysis buffer (25mM NaOH and 2mM EDTA, pH12.0 in ddH_2O) and neutralized by mixing in 50µL neutralization buffer (40mM Tris-HCl, pH5.0 in ddH_2O). A centrifugation was performed at maximum speed to get read of the fur. The supernatant containing the genomic DNA was used (1µL) to mask the plasmid during the spike. A PCR was subsequently done using the primers PGK2-FW (20nt, 5'-gcgcacacctcaggactatt-3') and SV40-RV (23nt, 5'-cttggcgtaatcatggtggtacc-3'), and increasing amounts of plasmidic DNA (0,1; 0,5; 1; 2; 5 and 10pg) to obtain a final product of 1310nt. The PCR conditions were done has follows: initial denaturation (92ºC for 2min.), 30 cycles (92ºC for 30sec., 59.5ºC for 30sec. and 72ºC for 1min.) and final extension (72ºC for 5min.).

Transgenic mice generation

The two constructs, pBlueScriptII(SK-)-hPGK2-3xHA-hPPP1R2/hPPP1R2P3-SV40, were sent to the Case Western Reserve University (CWRU, Cleveland, Ohio, USA), Case

Transgenic and Targeting Facility. In CWRU the insert was first extracted using the same enzymes of the subcloning and gel-purified it before microinjection. For microinjection of the DNA into the pronucleus, F2 (generation 2, SJL/BL6) zygotes were first obtained from a cross between F1 hybrids of SJL and C57BL/6J. Microinjection was performed in these zygotes and they were implanted in a foster mother. After the gestation of almost 3 weeks, the mice were maintained in their facility for a postnatal growth of 6 weeks and after they were weaned. Then the mice were shipped to the animal facility in the Department of Biological Sciences at Kent State University, Kent, Ohio, USA.

Identification and genotyping

Transgenic mouse production and use at Kent State University follows approved Institutional Animal Care and Use Committee (IACUC) protocols adapted from the National Research Council publication, Guide for the Care and Use of Laboratory Animals. The animals were received in the animal facility, separated in cages, maximum 5 animals per cage, and divided by sex. Mice came in a rainbow of coat colors, white, brown, black, and yellow, because they are hybrids of two inbred strains (Table II.C. 2). They were kept in a quarantine room for two months, and were subjected to blood analysis, identification and genotyping.

Table II.C. 2: Number of mice received in orders hPPP1R2 and hPPP1R2P3. *Sex and colors are also shown for each order.*

	hPPP1R2		hPPP1R2P3		
	Male	Female	Male	Female	Total
Brown	14	15	19	22	70
Black	6	4	7	9	26
White	14	2	10	14	40
Yellow	0	1	2	2	5
Total	34	22	38	47	141

The marking was done by earpunch using an ear puncher and according with a pattern described in Figure II.C.1.

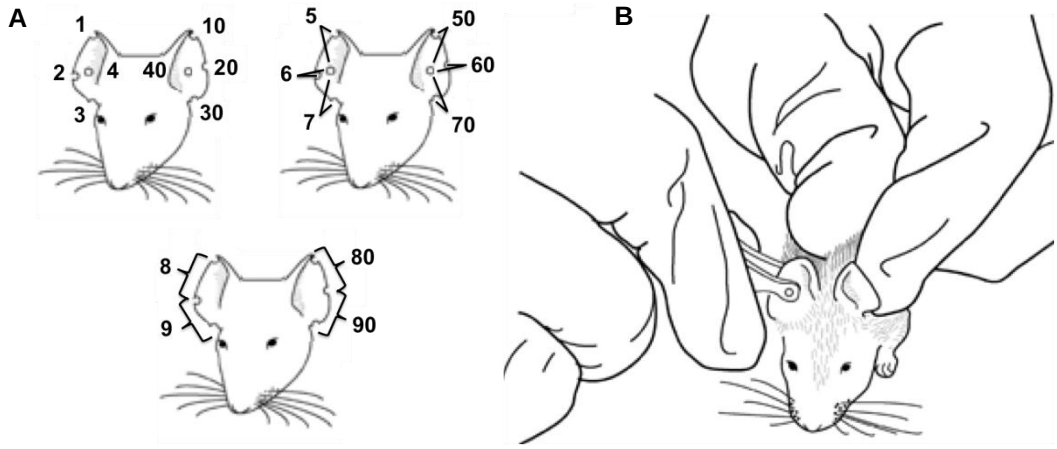


Figure II.C.1: Ear punching of mice. ***A.*** *Punch(es) according to the illustrated code.* ***B.*** *Restraining the animal manually, place the ear puncher in the desired position and firmly quickly punch a hole (figure adapted from [43]).*

The ear punch was also used to genotype in the same conditions as explained in the section Plasmids construction, with the difference that a second set of primers, PPP1R2-Int-FW (26nt, 5'-actgcagatggagaaagcatgaatac-3') and SV40-RV, was also selected for reconfirmation, giving a fragment of 302nt. Also tail snips of all mice were cut and the genomic DNA extracted with DNeasy® Blood & Tissue kit (QIAGEN, Dusseldorf, Germany), in order to check for false negatives. All the inserts were also subcloned in pGEM-T Easy (Promega, Madison, Wisconsin, USA) and sequenced (Cleveland Clinic Genomic Core sequencing service, Cleveland, Ohio, USA). After that, all positives (founders) were transported to a separate room and the lines established by breeding the founders with the wild type CD1 mice. Generally female mice reach sexual maturity around 6 weeks and males around 8 weeks.

The new generations were allowed to growth for 3 weeks in the presence of the mother and after that separated to new cages, marked and genotyped. The negatives were euthanized by carbon dioxide according to the rules. Positives were kept for posterior experiments and for continuing the lineage. All the cages were identified with cards, and the mating pairs, founders and generations recorded.

For fertility test a single male positive for the transgene and two female CD1 mice in reproductive age, were put together in a separate cage. Regularly and in the morning females were checked for plug presence. Plugs are useful for obtaining timed mating. A plug is hardened semen, blocking the vagina, and remains in place for about 12 hours after mating. Plugs are detected by visual inspection or by probing gently with a sterile toothpick on an immobilized female. Mating is assumed to occur at the midpoint of the

dark cycle. For example if the 12hr on/off cycle starts at 6 pm the midnight will be the midpoint, and thus noon of the next day is 0.5 days of gestation. Generally the gestation occurs for 18 to 20 days, and when the litter is born the date and number is recorded.

Also fertility should be checked in at least 2 males of each founder line and caution should be taken if the females don't deliver pups. In this case females in the box should be switched by new ones.

Fertility test

Transgenic male mice from the positive lines were mated with wild-type CD1 females over a period of 5 weeks, and the number of offspring in each litter was recorded. CD1 females that failed to become pregnant were subsequently tested for fertility by mating to wild-type CD1 males.

Samples preparation

For all the experiments, mice were sacrificed using carbon dioxide and dissected under the fume hood. When possible a littermate negative or a wild type CD1 of approximate age, was used as negative control. The mice were open laterally with small incisions. The testis, epididymis and vas deferens were cut from each side. Each testis was weight with tunica in an analytical balance. Then, for protein extracts the tunica was removed and the testis homogenized in 1% SDS using a Model Pro 200 tissue homogenizer (Pro Scientific Inc., Oxford, Connecticut, USA).

For immunohistochemistry, the testis with tunica was immersed in 4% paraformaldehyde/1x PBS and kept at 4ºC overnight. Cauda of epididymis and vas deferens were separated from the rest of the epididymis and immersed in 1x PBS solution warmed to 37ºC. The cauda of epididymis was punctuated with a fine-tip needle, cut as a fillet and vas deferens was squeezed. Sperm motility in caput and cauda of epididymis was also manually accessed under the stereomicroscope. Then epididymal sperm was allowed to exit, collected and washed two times with 1x PBS by centrifugation at 600g, room temperature. Sperm was subsequently counted in a haemocytometer chamber, and before preparing the sperm extracts an aliquot was mixed with 4% paraformaldehyde to visualize the morphology in differential interference contrast (DIC). The remaining sperm was diluted to $2x10^5$ sperm/µL in 1% SDS and sonicated three times 10sec. For the preparation of supernatant and pellet extract fractions, homogenization buffer (10mM Tris pH7.0, 1mM EDTA, 1mM EGTA) with protease inhibitors (10mM Benzamidine, 1mM PMSF and 0.1mM TPCK), HB+, was used instead of 1% SDS. After centrifugation at

16000g, 20min., 4ºC, the supernatant fraction was collected to a new tube and the pellet was dissolved in 1% SDS (pellet fraction).

Testis heat stable extracts were prepared from the supernatant fraction of homogenization buffer by boiling for 10min. at 95ºC, 2min. in ice and centrifugation at 16000g, for 15min. at 4ºC. Supernatant fraction was then recovered to a new tube.

SDS-PAGE

In SDS polyacrylamide gel electrophoresis (SDS-PAGE) separations were carried out using established methods [44]. Small 12.5% gels were used for the separation of hPPP1R2 and hPPP1R2P3 proteins. In brief, the resolving gel was pipetted down the spacer into the gel apparatus, leaving some space for the stacking gel. Then, isopropanol was added to cover the top of the gel and the gel was allowed to polymerize for 1hr. The isopropanol was then poured out, the stacking gel was added to the apparatus, a comb inserted and the gel was allowed to polymerize for 30min.

After the gel polymerization, the combs were removed, and the wells filled and washed with running buffer. The samples were prepared by adding to the protein sample 4x loading buffer (LB). Samples were loading according with sperm number or total testis weight. Total testis weight of wild type mice is around 130-160mg, which corresponds roughly to 4-4.2µg/µL if homogenized in 1mL (calculated before by Bradford standard protein quantification method). Also brain, liver, lung and spleen tissues were sometimes used to check for the correct expression of the transgene. The gel was run for 50min., at 200V using a power supply.

Western blotting

For electroblotting, the tank transfer system was used as follows: 3MM blotter paper (Whatman, GEHealthcare, Waukesha, Wisconsin, USA) was cut to fit the transfer cassette and was blotted to polyvinylidene fluoride membranes (PVDF, Immobilon-P, Millipore, Bedford, Massachusetts, USA) to fit the gel size. The gel was removed from the electrophoresis apparatus (Biorad, Hercules, California, USA) and the stacking gel removed and discarded. The transfer sandwich was assembled immersed in the transfer buffer to avoid trapping air bubbles. The cassette was placed in the transfer apparatus and filled with transfer buffer.

Transfer was allowed to proceed for 1hr at 100V in ice. Afterwards, the transfer cassettes were disassembled, the membrane carefully removed and allowed to air dry prior to further manipulations. The membranes were incubated with specific antibodies

diluted in 3% low fat milk/1x TBS-T: anti-HA hybridome (raised in mouse, 1:1000), anti-PPP1R2 (raised in rabbit against the peptide ^{134}EKKRQFEMKRKLH^{147}, of mPPP1R2, 1:2000) or with anti-actin (raised in mouse, 1:5000, GenScript, Piscataway, New Jersey, USA), the first two overnight and the third for 2hrs. Subsequently the membranes were washed once in 1x TBS-T, incubated with horseradish-peroxidase conjugated secondary antibodies (mouse and rabbit, 1:5000, GenScript, Piscataway, New Jersey, USA) for 1hr and then washed 2 times for 10min. in 1x TBS-T.

The immunodetection was performed using a homemade enhanced chemiluminescence (ECL) and the pictures were taken in image acquisition system LAS-3000 imager (Fujifilm, Tokyo, Japan).

Morphology preparations

For morphology, an aliquot of epididymal sperm collected and washed was diluted 100x in 4% paraformaldehyde/1x PBS and put at 4ºC, during 2hrs. Then 10µL were applied in a slide, dried for 20min., a coverslip was put on top and then sealed using nail polish. Slides were then visualized under Olympus IX70 inverted microscope with a color camera and a cooled grayscale CCD camera (Olympus, Center Valley, Pennsylvania, USA) using DIC, an optical microscopy illumination technique used to enhance the contrast in unstained, transparent samples.

Immunohistochemistry

Whole testis was kept in 4% paraformaldehyde/1x PBS overnight, for fixation, as stated before in samples preparation section. Prior to embedded in paraffin, testis was washed in the following order: 75% ethanol for 1hr, 95% ethanol for 1hr, 100% ethanol for 40min. twice and finally citrosol for 1hr. Then the tissue was put inside a plastic capsule and waxed for 1hr in Shandon Citadel 2000 Tissue Processor (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Finally the tissue was embedded with wax inside a plastic mold and allowed to solidify on the cold plate cover of Shandon Histocentre 2 Embedding Center (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

Tissues were kept in the plastic mold at 4ºC till processing. For immunohistochemistry the tissue was first sectioned using a microtome (Leica Microsystems Inc., Buffalo Grove, Illinois, USA) in 8µm slices and then slices were put in a electrothermal paraffin section mounting bath (Thermo Fisher Scientific, Waltham, Massachusetts, USA) for easily mounting the slides. Paraffin was removed from sections by sequential washings in citrosol 2x for 5min., 100% ethanol 2x for 5min., 95% ethanol

for 5min., 80% ethanol for 5min., 70% ethanol for 5min., 50% ethanol for 3min., and finally in ddH_2O for 2min. The slides were put in a boiled citrate buffer bath and microwaved until boiling and then rest for 1min.

After repeat the process twice, the slides were allowed to cool down to room temperature in the citrate buffer and then washed twice in 1x PBS-T for 2min. and 1xPBS for 5min. Slides were wiped carefully and the border of the section marked using a liquid blocker super pap pen (EM Sciences, Hatfield, Pennsylvania, USA). Sections were blocked by incubation during 1hr in 5% goat serum/1% bovine serum albumin/1x PBS in a humid chamber at room temperature. After wash twice in 1x PBS for 5min., sections were permeabilized for 15min. with 0.2% Triton X-100/1x PBS and then washed again in 1x PBS for 5min. An overnight incubation was performed with the primary antibodies, anti-HA (1:250) or anti-I2 (1:250) at 4ºC in the humid chamber. After washing twice in 1x PBS during 10min., sections were incubated with FITC (1:500, mouse, Santa Cruz Biotechnology Inc., Heidelberg, Germany) or Cy3-conjugated (1:250, rabbit, Jackson ImmunoResearch Laboratories, Inc., West Grove, Pennsylvania, USA) secondary antibodies for 1hr at room temperature, in a humid dark chamber. After washing 3x with 1x PBS for 10min., the sections were mounted with mounting media vectashield, sealed with nail polish and viewed under an Olympus Fluoview 300 confocal microscope (Olympus, Center Valley, Pennsylvania, USA).

Results and Discussion

Plasmids construction

Plasmids were constructed as depicted in material and methods section. Afterwards PCR was performed to spike the plasmid till reach the point of no detection (Figure II.C. 2A). A restriction digestion of both plasmids was also executed to show the separation between the insert (hPGK2-3x HA-hPPP1R2/hPPP1R2P3-SV40) and the plasmid backbone (pBlueScriptII(SK-)) using the same enzymes of the subcloning (Figure II.C. 2B). A final map of both plasmids can be seen in Figure II.C. 2C. Subsequently the plasmids and all the data were sent to CWRU for mouse microinjection.

Genotyping and establishing the lines

A total of 141 animals were received from CWRU in the animal facility, 56 from the hPPP1R2 order and 85 from the hPPP1R2P3 order (Table II.C. 3).

Table II.C. 3: *Number of positives and the respective percentage.*

Order	Mice	Male	Positives	Female	Positives	% (+)
PPP1R2	56	34	M13, M14, M5	22	F10	7,1
PPP1R2P3	85	38		47	F29, F45	2,4
Total	141	72	3	69	3	

After screening all the mice using the earpunches only 2 positives (M13 and M14) for the hPPP1R2 order and 2 positives (F29 and F45) for the hPPP1R2P3 order were found, although an uncertain M34 (hPPP1R2 order) also gave positive but the insert obtained in the PCR had a lower size than the expected 1310bp band (Figure II.C. 3). A re-confirmation was made using a second pair of primers, which resulted in a 302bp band. As expected all the initial positives were obtained, but also M5 and the F10 (hPPP1R2 order) and the M34 in a correct position. Moreover no more positives were obtained for the hPPP1R2P3 order (Figure II.C. 3).

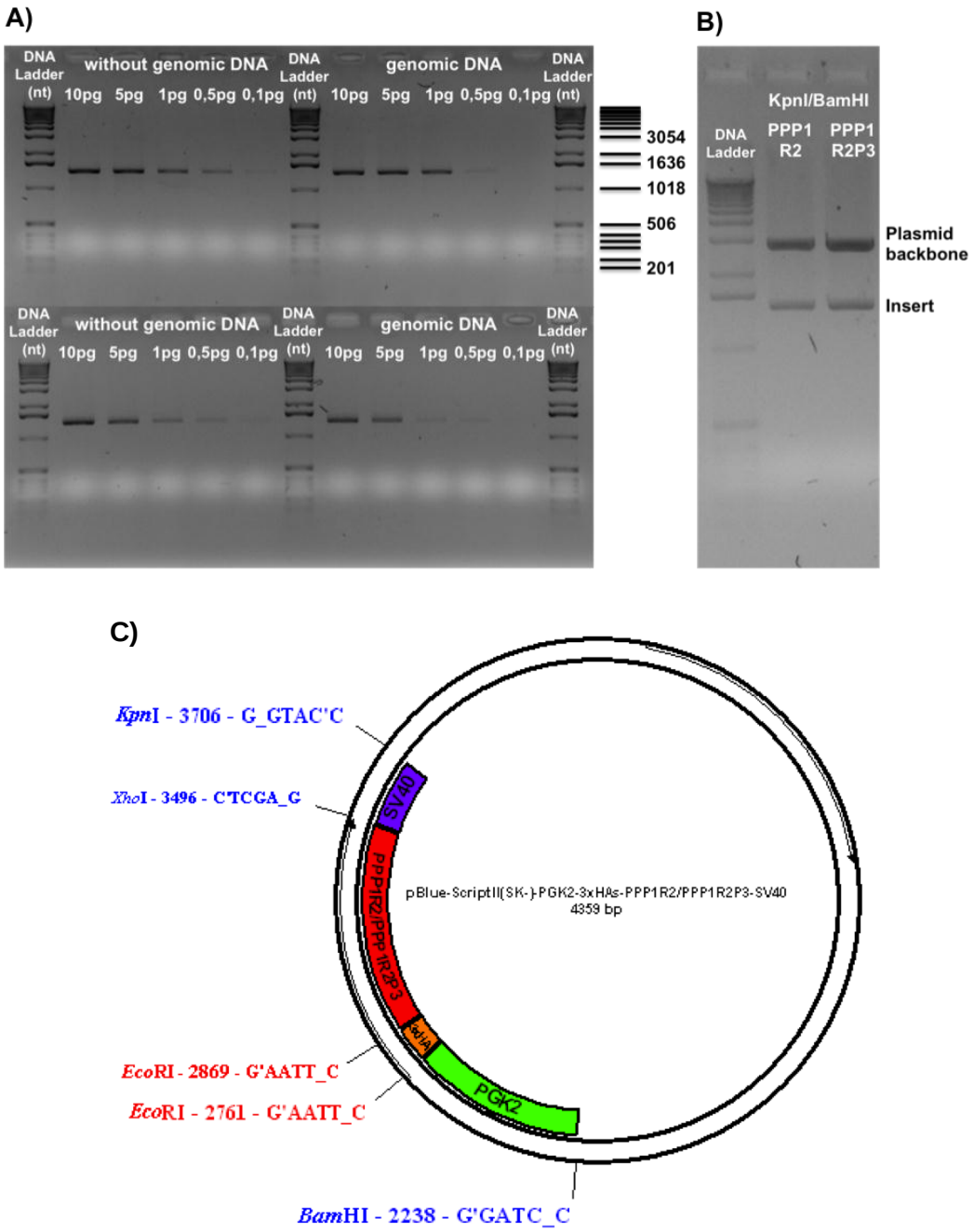


Figure II.C. 2: Data requested by CWRU to proceed for microinjection. A. *Agarose gel showing the spike of the plasmid till reaches the point of no detection. The fragment obtained (1394bp) was amplified using primers that flank PGK2 and SV40 regions. Mouse ear genomic DNA (1µL) was used to mask the plasmid and mimic the future conditions positive mice for both sets of primers.* **B.** *Agarose gel showing the restriction digestion of 500ηg of both constructs. Plasmid backbone (2895bp) and insert (1464bp).* C. Plasmid map. Arrows show the ORFs (insert and ampicillin).

These new positives that were obtained with the second set of primers and not with the first could mean that PGK2-FW primer (forward of the first set) could not anneal because of the high GC content of the locus near the PGK2 promoter where it was integrated, whereas the SV40-RV annealed alone. CWRU state that although they don't

provide a guarantee, the goal is to deliver at least 3 founders and the average stated in the site is 9.1 founders. The percentage of frequency of transgenic founders per construct is 6% for 2 founders and 9% for 3 founders. This clearly means that by genotyping all mice lines the number of founders obtained was very low (7.1% for PPP1R2 and 2.4% for PPP1R2P3).

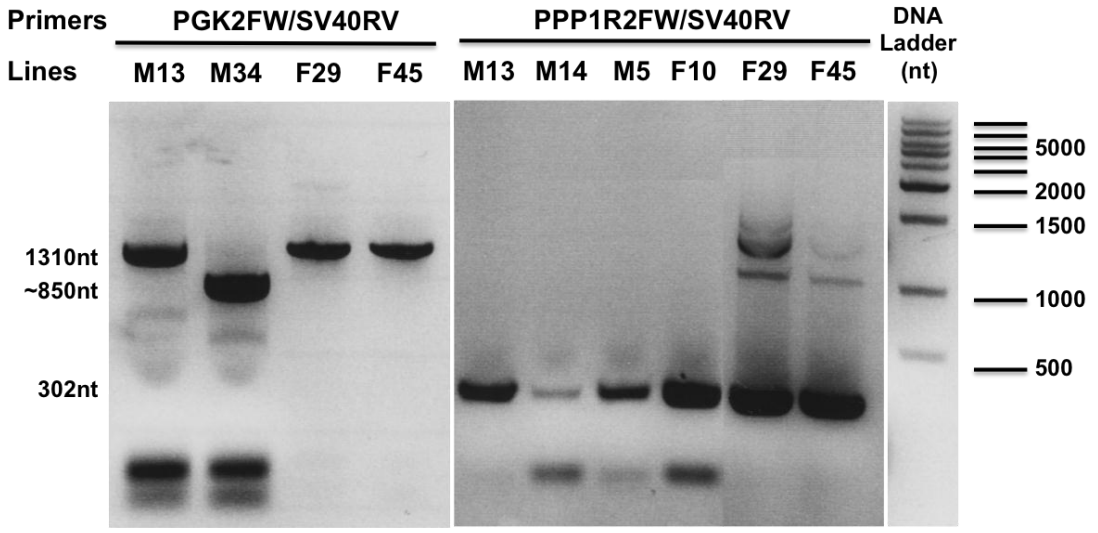


Figure II.C. 3: Agarose gel showing positive mice for both sets of primers. *The first set (PGK2FW/SV40) was used to select M13, F29 and F45 lines as positives. M34 line is positive but with a lower band, due to a deletion. M14 positive line is not shown. The second set of primers was used to re-confirm the positives and also to increase the possibility of getting more positives. As expected M13, M14, F29 and F45 lines showed up again. Two more lines, M5 and F10, were selected. M34 line is not shown, but the band size is the same as the others.* **Nt, nucleotide; FW, forward; RV, reverse; M, male; F, female**

To further re-confirm this outcome, tails snips of all animals and column purification of genomic DNA was performed and all the mice were genotyped again. Once again, the same results were obtained. Additionally, all the inserts, except for M5 and F10 mice, were sequenced, and as suspected M34 line which gave a short band size with the first set of primers and the normal band with the second set, showed a deletion in the last part of hPGK2, the entire HA-tag and the first part of hPPP1R2 cDNA. This deletion is in accordance with the size obtained and also with the restriction digestion, because EcoRI did not cut (HA-tag is flanked by EcoRI) and BstXI showed the same difference in size (Figure II.C. 4). This deletion might have occurred in CWRU facility during the restriction enzyme cut to release the insert for microinjection.

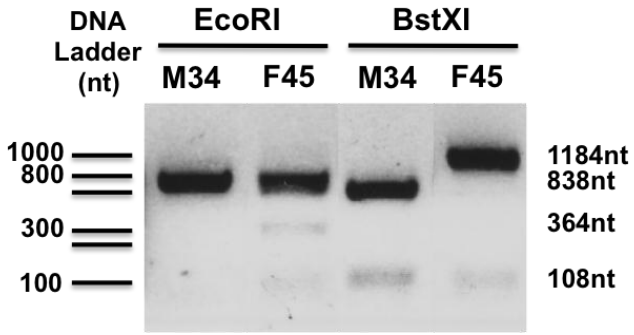


Figure II.C. 4: **Agarose gel showing M34 and F45 cut with EcoRI and BstXI restriction enzymes.** *The restriction cut was performed in the fragment obtained from the PCR with the PGK2FW/SV40 set of primers. The PCR using this pair of primers produces a band of 1310nt when positive. These results show that EcoRI does not cut M34 line and that when M34 is cut with BstXI (middle of mPGK2 promoter) it produces a common band (108nt) and a lower band size (~800nt) than expected (1184nt). Nt, nucleotide.*

So the founders M5, M13, M14, F10, F29 and F45 were subsequently put to breed with wild type CD1 mice in order to see if the transgene was carried over to the next generation, to establish the lines and finally to check if the mice were fertile, specially the males since the transgene was under the expression of the testis-specific promoter hPGK2. All lines were established. Minor complications for the F29 founder occurred, since it was killing the pups after birth. The average litter size and percentage of positives for each line is stated in Table II.C. 4. The lines M5, M13 and M14 have shown to be fertile in the first cross since the founders were males, but for the rest of the lines a positive male from the second generation was chosen to test the fertility. These mice were also fertile.

Table II.C. 4: Average litter size and percentage of positives in the founder lines.

Transgene	Founder Line	Number of litters	Average nº of pups (F1)	Average nº of pups (F2)	Positives % (F1)	Positives % (F2)
PPP1R2	M5	3	8	n.d.	34	n.d.
	M13	8	11,2	13,5	5	56
	M14	5	14,3	8	42	42
	F10	4	10,5	n.d.	17	n.d.
PPP1R2P3	F29	13	6,8	13,8	32	31
	F45	9	9,5	12,3	68	43

In the case of the pronuclear microinjection technique, the insertion of a single transgene will indicate that the germ line transmission will follow the Mendelian inheritance. But generally a given founder will not necessarily transmit the transgene in a Mendelian fashion. Founders can be mosaic for the transgene, if integration occurs after the first cell division. Mosaicism occurrence account for 30% and can result in a frequency of inheritance lower than 50% in the first generation offspring [45]. Also, in some founders, the transgene might be integrated into more than one locus, resulting in a frequency of inheritance of more than 50%. In this case, the expression levels among the first generation offspring may vary depending on which integration site they inherit. Although the identification of the transgene insertion site is often not essential for functional analysis of the transgene, identifying the site or sites can have practical benefit.

Furthermore, identification of the insertion site might be necessary to analyze unexpected phenotypes that might arise by insertional inactivation of an endogenous gene. Liang Z., and co-workers have shown an easy and straightforward way to do this. It consists in digesting the genomic DNA from a tail snip with a restriction enzyme present in distal 3' of the transgene. After the digestion T4 DNA ligase is subsequently added to circularize the extracted fragments. Then a PCR using primers specific for the terminal transgene fragment will generate a amplicon that includes the flanking chromosomal sequence, that could then be sequenced [46].

According to Jackson Lab (The Jackson Laboratory, Bar Harbor, Maine, USA) for the C57BL/6J the average litter size is 5 to 6, SJL is around 6 to 7 and the CD1 is around 11 to 12. This in accordance with our data in that the male founders (SJL/BL6, hybrids of the first two) paired with wild type CD1 females delivered between 8 (M5) to 14,3 (M14), and the female founders (SJL/BL6) paired with wild type CD1 males delivered 6,8 (F29) to 9,5 (F45). Also the F1 generation of F29 and F45 delivered more pups than the founders because this time the males were paired with wild type CD1 females.

The only exception is the M14 line in which the fertility was substantially lower in F1, although the number of positives obtained remained the same. Interestingly in the M13 line the percentage of positives that were obtained in F2 was higher and all of them were males. This result means that the transgene has been incorporated in the Y chromosome, but unfortunately also means that the probability of the transgene being expressed is low. Y chromosome consists predominantly of highly repeated DNA sequences with no obvious function [47]. Additionally, because of its extremely late replication during mitosis, characteristic of heterochromatic segments, an optimal expression of transgenes on the mouse Y chromosome may not be expected [48]. Moreover in F1 only some males were

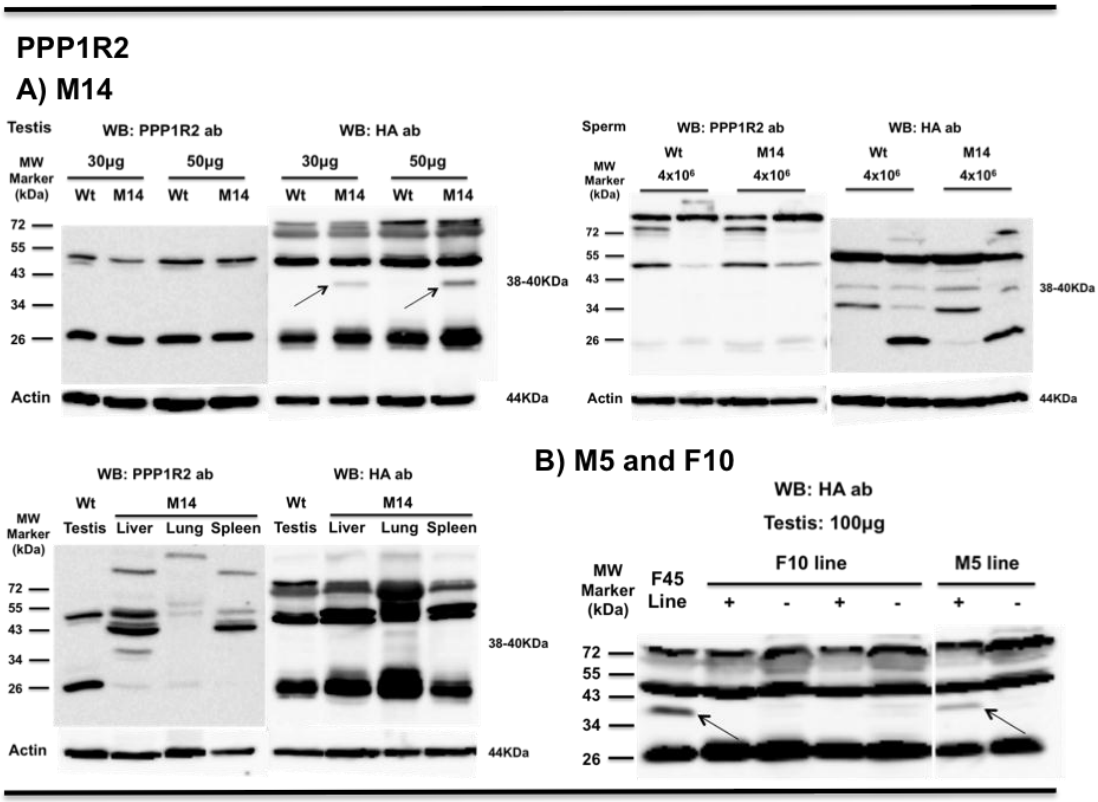
transgene positive. The only explanation could be a mosaic of spermatozoa genotypes in this founder line, with ones carrying the transgene in the Y chromosome whereas others not. This will happen when the transgene integration occurred after the first cell division.

The same case could be explained regarding the F29 founder line where the percentage of transgene positives was also lower than the Mendelian pattern. In this case some positives of the second generation should recover to the Mendelian pattern. If we check each pair of breeders of F1, we can see that two were low (33% and 12%) like the founder but one was closed (42%) to the normal Mendelian pattern. In the F45 founder line of hPPP1R2P3, the percentage of transgenic positives in F1 was higher than the Mendelian pattern (68%), which is explained by the multiple locus integration of the transgene. In F2 the percentage fall down to a value closer to the Mendelian pattern (42%).

Protein expression and phenotyping

Protein expression of the transgene was accessed by Western blot in testis and sperm extracts. Also lung, liver and spleen tissues were used, to check if the transgene expression under the control of the hPGK2 promoter was restricted to testis as expected. All these tissues have expression of the wild type mPPP1R2, with exception of spleen. Since the transgenic proteins are carrying a HA-tag, two antibodies were used to fulfill this task. Results show that only HA-tag hybridome antibody, and not PPP1R2 antibody, was able to detect a band that could be the transgenic hPPP1R2/hPPP1R2P3 in testis (Figure II.C. 5, A and D first blot).

Other commercial available HA-tag antibodies were also tested with no positive results (data not shown) even using 100µg testis extracts. Unfortunately no protein was detected in sperm extracts of lines M14, F29 and F45, using 4-5x10^6 sperm (Figure II.C. 5, A second blot, C and D, first blot). These results show that protein was probably kicked out of the sperm in the final steps of spermiation where the cytoplasmic droplet containing all the unnecessary material is released [49].


PPP1R2
A) M14
Testis
WB: PPP1R2 ab
WB: HA ab
MW Marker (kDa)
30µg
50µg
Wt M14 Wt M14
72
55
43
34
26
Actin
38-40KDa
44KDa
Sperm
4x10⁶
B) M5 and F10
Testis: 100µg
F45 Line
F10 line
M5 line
Liver Lung Spleen

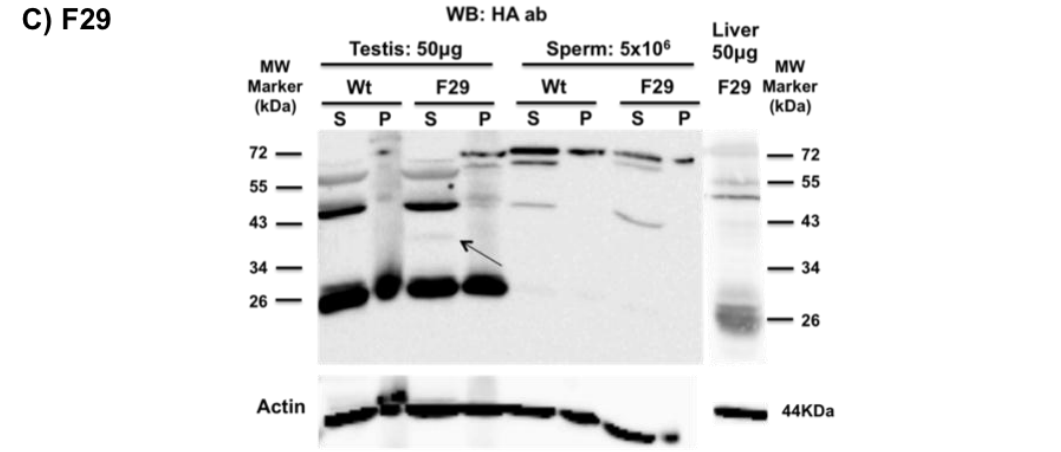

PPP1R2P3
C) F29
WB: HA ab
Testis: 50µg
Sperm: 5x10⁶
Liver 50µg
Wt F29 Wt F29 F29
S P S P S P S P
MW Marker (kDa)
72
55
43
34
26
Actin
44KDa
D) F45
Sperm
WB: PPP1R2 ab
WB: HA ab
Wt F45 Wt F45
Testis F45
Liver Spleen
38-40KDa

Figure II.C. 5: (previous page): Western blotting data of all founder lines. A. *A male from the M14 F1 generation was sacrificed along with a wild type CD1 of the same age. Testis, liver, lung, spleen and sperm extracts were prepared in HB+ as described in materials and methods. For the Western blot, HA and PPP1R2 antibodies were used in separate membranes. The arrow points to a band (38-40kDa) that only appears in testis extracts from M14 line. All other tissues are negative for this band.* **B.** *Two males from F10 and one from M5 founder lines from the F1 generation were sacrificed along with negative littermates. Only testis extracts were prepared in HB+. A male from F45 line was used has positive control. Results show that genotype positives of F10 line do not produce the transgenic protein hPPP1R2 in detectable amounts. By contrary the male from M5 line produced a band with the same molecular weight as the positive control, F45 (arrow).* **C.** *A male from the F29 F1 generation was sacrificed along with a wild type CD1 of the same age. Testis, liver and sperm extracts were prepared in HB+. Pellet fraction was also loaded. HA antibody was used to detect the transgenic protein hPPP1R2P3. Results show the specific 38-40kDa band only in the testis of the F29 male (arrow).* **D.** *A male from the F45 F1 generation was sacrificed along with a wild type CD1 of the same age. Testis, liver, spleen and sperm extracts were prepared in 1%SDS. Pellet fraction was also loaded for sperm. HA antibody was used to detect the transgenic protein hPPP1R2P3. Results show the specific 38-40kDa band only in the testis of the F45 male (arrow). Anti-actin was used as loading control. WB, Western blot; ab, antibody; Wt, wild type; S, supernatant; P, pellet; +, positive; -, negative; HB+, homogenization buffer with protease inhibitors; MW, molecular weight.*

The M5 line was the only line that presence of hPPP1R2 transgene was not accessed in sperm extracts.

Transgenic positive lines expressed a band around 38-40kDa that is not present in the wild type or littermate controls. This band is only seen in 30-100µg soluble extracts and only with HA hybridome antibody (Figure II.C. 5, A first blot, B, C and D second blot, Fig.IV.6). M13 and F10 lines of hPPP1R2 did not produce sufficient amount of protein to be detected even with 100µg testis extracts (Figure II.C. 5, B and

Figure II.C. 6). M13 line, was the one in which the transgene was integrated in the Y chromosome. Both founder lines, M13 and F10 were subsequently discarded.

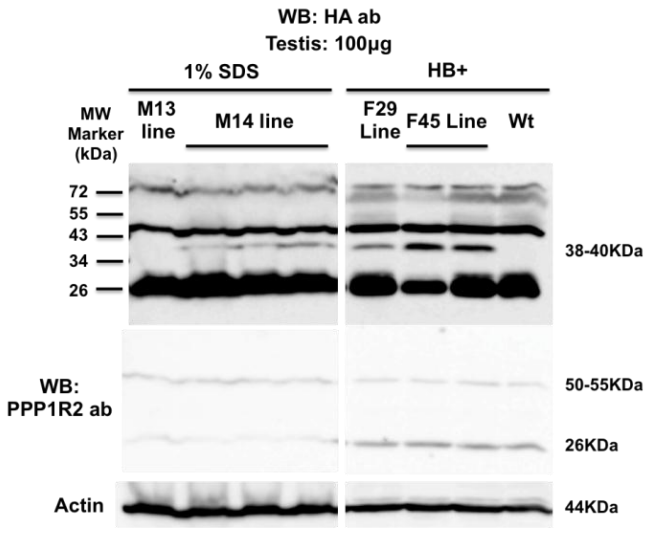


Figure II.C. 6: Western blot showing the positives of both mice orders that produced a band with the first set of primers. *Testis extracts of several positives were prepared in HB+ or 1%SDS. Wild type CD1 mouse was used as negative control. For M14 line a positive from F1 and two positives from F2 were used. For F45 both positives are from F1 generation. Both HA and PPP1R2 antibodies were used in separate blots. Results show that all positives have a band at the expected size except for M13. Anti-actin was used as loading control.*

Also other tissues, such as liver, lung and spleen, were used to show that expression was only driven in testis (Figure II.C. 5, A third blot, C and D second blot).

Since PPP1R2 is a heat stable protein, the testis supernatant extracts were boiled to check if the 38-40kDa band remained unaltered. Results for the M5 line using 2 positives show that protein was still present in these extracts further confirming that the 38-40kDa should be indeed the transgenic hPPP1R2 (Figure II.C. 7).

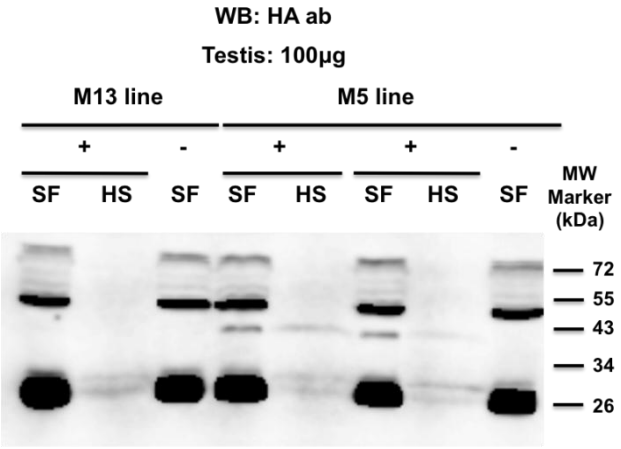


Figure II.C. 7: Western blotting of heat stable extracts. Two positive males from M5 F2 and one from M13 F1 were sacrificed along with negative littermates. *Testis heat stable extracts were prepared from the supernatant of HB+ as described before. HA antibody was used in the Western blot. Results show that the 38-40kDa protein is present in the M5 positives as expected and is heat stable. SF, soluble fraction; HS, heat stable fraction.*

The phenotyping of each founder line was also performed during the establishment of the lines and the analysis of protein expression. Results for testis weight, sperm number and motility are stated in Table II.C. 5 for all mice opened. In each sacrifice a littermate negative or wild type CD1 mice were also opened and testis weight, sperm number and motility were recorded (these results are not present in the Table II.C. 5). Regarding testis weight all mice were normal (see wild type CD1 inTable II.C. 5). Values lower than 120mg are explained due to the younger age of the mice opened (51 to 71days). The only exception was observed for the 86days older mouse from M14 line F1 generation, where the testis showed a different size and weight.

The sperm number observed was in the limits established as normal, and is very similar to the normal sperm number of wild type CD1 males. Motility was accessed in caput and cauda sperm by visual observation as described in material and methods. Other methodology more accurate is to use a light microscope with a camera coupled to a computer that runs CASA (computer assisted sperm analyzer) software. Although visual observation is a roughly measure, it permits a first approach in accessing sperm motility. All observed spermatozoa seemed to be normal with no motility in caput epididymis and forward motility in cauda of epididymis and vas deferens.

Table II.C. 5: *Data showing the age, testis weight and sperm number, motility and morphology of each mouse sacrificed for Western blotting.*

Animal	Age (days)	Testis Weight (mg)	Sperm Number/mL	Motility	Morphology (DIC)
M5 F1	58	115.5/122	-	-	-
	116	130.2/129.8	-	-	-
	116	134.6/143	-	-	-
M13 F1	71	101.8/110.3	-	Cauda sperm more hyperactive	-
	145	122.6/123.2	-	-	-
M14 F1	51	115/112.5	$1.7x10^7$	Normal	Normal
	86	104.2/134.5	$2.5x10^7$	Normal	Normal
	86	121.5/115.7	$1.7x10^7$	Normal	Normal
	106	120.9/122.1	$2.5x10^7$	Normal	Normal
F10 F1	59	127.5/123.1	-	-	-
	107	140.7/145.1	-	-	-
F29 F1	122	148.5/160.5	$4.4x10^7$	Normal	Normal
F45 F1	60	109/117.9	$1.9x10^7$	Cauda sperm more hyperactive. Caput sperm 30% vigorous	Normal
	176	134.6/131.3	$3.5x10^7$	Normal	Normal
F45 F2	54	113.2/112.9	$2.3x10^7$	Normal	Normal
Wt CD1		~120-160	$\sim 2.4x10^7$	-	-

The only exceptions were for the M13 and F45 lines. These differences observed in testis weight and sperm motility were not present in the rest of the mice from the same lines, which indicates that they were isolated cases. Also M13 line did not show expression of the transgenic protein so the results cannot be correlated.

Morphology of sperm was also accessed by differential interference contrast (DIC). Images representative of each line are presented in Figure II.C. 8. As expected no major abnormalities were found and the morphology seemed to be similar to littermate negatives and wild type CD1 of the same age.

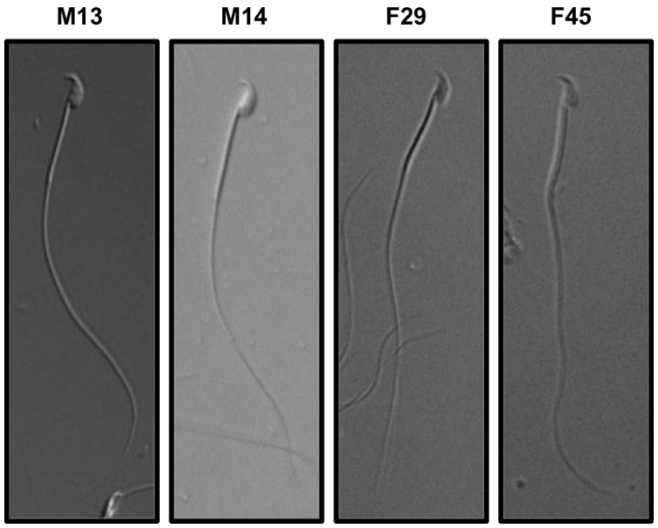


Figure II.C. 8: DIC images of representative cauda sperm from each founder. *Cauda sperm slides were prepared for each founder line. Slides were then observed under the microscope and DIC images were obtained. Results show that the vast majority of the total cauda sperm observed in all positives were morphologically normal.*

In relation to testis histology and if normal spermatogenesis is occurring, generally a haematoxylin staining is used as a first approach. Haematoxylin stains the nuclei of the cells and permits to see if normal mitosis and meiosis is occurring. The haematoxylin staining was not accessed in this study, although paraffin blocks of all positive opened were stored for posterior analysis. In spite of that, sperm morphology and number, and also fertility suggest that normal spermatogenesis seems to be occurring in all founder lines. Also in a first approach, HA and PPP1R2 antibodies were used in immunohistochemistry of a male from M14 line, with no conclusive results because of the high background of the secondary antibodies and the reactivity of HA antibody in wild type mice (Figure II.C. 9). However the testes of this male seemed to be normal compared with the wild type CD1. Further, testicular sperm was observed in the lumen also supporting that normal spermatogenesis is occurring.

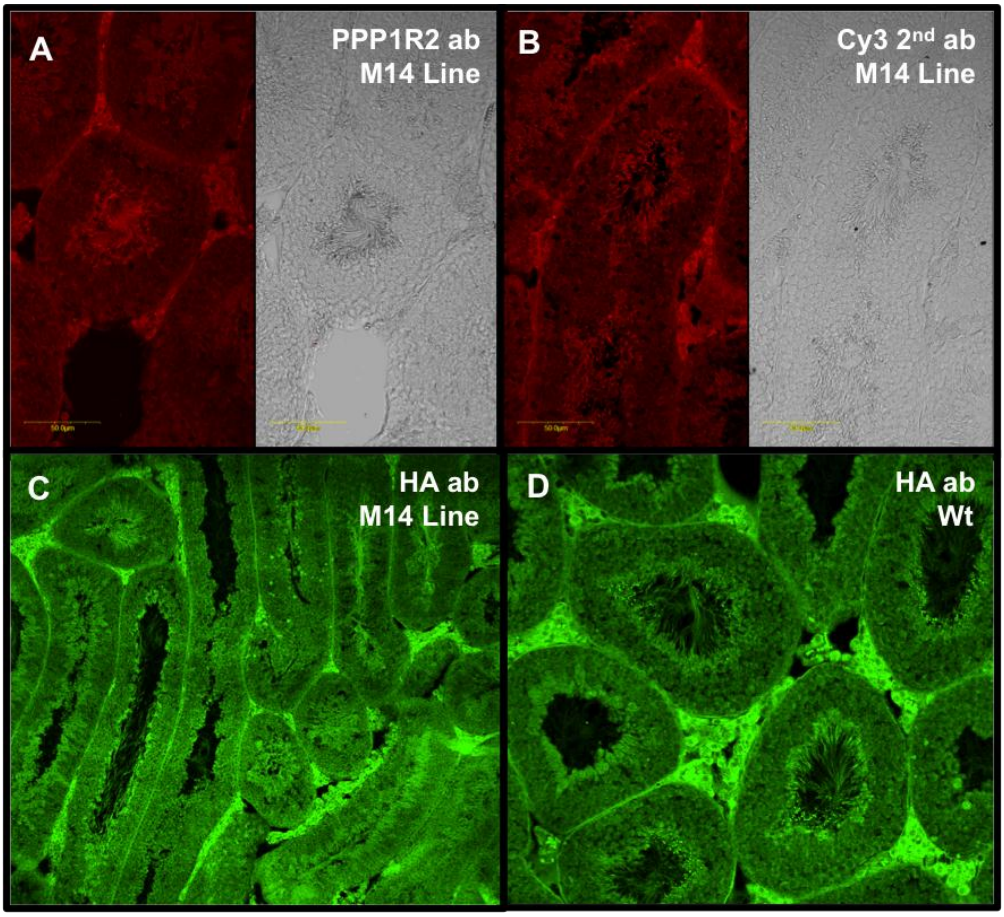


Figure II.C. 9: Testes immunohistochemistry of M14 line using PPP1R2 and HA antibodies. *Testis slices for immunohistochemistry were prepared as described in material and methods section. Pictures were taken in a confocal microscope. Panels* **A.** *and* **B.** *Unspecific PPP1R2 antibody staining is due to the secondary antibody, Cy3. Panels* **C***. and* **D.** *Same staining is shown in M14 Line and wild type CD1 mice using HA antibody. Seminiferous tubules are normal and testicular sperm is produced.*

Message expression data, either by Northern blot or by RT-PCR was not performed since the first goal and final aim were always to detect the protein and achieve a good expression. However since this objective was not accomplished, the message pattern can be accessed in order correlate message versus protein. Also by gaining insight on the correct temporal transcription of the normal wild type PPP1R2 message we can choose a better promoter to mimic the natural protein. If the message is being made in high amounts and we do not see enough protein this may suggest a tight control from the testis. However, if low amount of message is being transcribed we can interpret this as a problem in the integration of the transgene or the stability of the message. This can be overcome with more positives, using site-direct integration of the transgene, correct spatial and temporal transcription of the message, adding an internal intron or finally with a different polyadenylation signal.

Conclusion

Standard transgenic technique was used to overexpress hPPP1R2 and hPPP1R2P3 in mouse testis. PPP1R2P3 differs between PPP1R2 mainly in the Thr72 and Ser85 amino acids that are extremely important for the PPP1R2/PPP1C holoenzyme function. Since Thr72 is absent from PPP1R2P3 it is possible that phosphorylation by GSK3β is prevented in PPP1R2P3/PPP1C complex and PPP1C remains permanently inactive. This could mean that PPP1R2P3 is an irreversible inhibitor of PPP1C. Previous results have shown that PPP1CC2 is extremely important for sperm motility [50, 51]. When it is active the sperm is immotile and when it is inactivated by specific PPP1C inhibitors, sperm becomes motile. The increase (overexpression) of hPPP1R2 in mouse testis could lead to a concomitant increase in PPP1CC2 inhibition. The expectation was that this could be overcame for instance by increasing the amount or activity of GSK3β. However, by overexpressing hPPP1R2P3 the phenotype would be different, since GSK3β would not reverse the phenotype. Without a way to surpass the inhibition of PPP1CC2, the sperm would start to acquire motility as soon as it reached caput or even earlier. Whether the ATP production factories, mitochondria, will follow this increased early demand and will be able to sustain it is a question that remains unsolved.

Thus this work was extremely relevant in the way it provided the molecular tools to initiate the characterization of the mechanisms behind PPP1R2/PPP1CC2 and PPP1R2P3/PPP1CC2 role in spermatozoa.

From the F0 generation received from CWRU only 4 and 2 founder mice were selected by genotyping for hPPP1R2 and hPPP1R2P3, respectively. From these only M14 and M5 for hPPP1R2 and F29 and F45 for hPPP1R2P3 showed protein expression, but only in testis and at low levels. No phenotype was observed in testis weight, sperm number or in morphology. These results are discouraging and clearly indicate that: more founders are needed to surpass the random integration of the transgene and achieve a founder line in which the expression of the protein is in good levels or use a different method to integrate the transgene (ex. Knock-in), albeit more costly; a correct spatio/temporal transcription of the transgene is needed by using a promoter more closely related and the use of an internal intron that will force the message to get out of the nucleus and to be spliced increasing therefore its stability. Parallel results from our collaborator (Prof. Srinivasan Vijayaraghavan) [52] suggest that mPPP1R2 message starts to be made in later spermatogenic cells, more specifically in round spermatids. This means that maybe the correct promoter to use is the mProtamine promoter instead of hPGK2 promoter. Thus we have already started making novel constructs with this

promoter for microinjection hoping this time to overcome these issues and unravel the role of these proteins in mammalian sperm motility.

References

1. Jaenisch, R., *Transgenic animals.* Science, 1988. **240**(4858): p. 1468-74.
2. Gordon, J.W. and F.H. Ruddle, Integration and stable germ line transmission of genes injected into mouse pronuclei. Science, 1981. **214**(4526): p. 1244-6.
3. Gossler, A., et al., *Transgenesis by means of blastocyst-derived embryonic stem cell lines.* Proceedings of the National Academy of Sciences, 1986. **83**(23): p. 9065-9069.
4. Jaenisch, R., Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. Proc Natl Acad Sci U S A, 1976. **73**(4): p. 1260-4.
5. Rajewsky, K., et al., *Conditional gene targeting.* J Clin Invest, 1996. **98**(3): p. 600-3.
6. Chen, Y. and P.A. Rice, New insight into site-specific recombination from Flp recombinase-DNA structures. Annu Rev Biophys Biomol Struct, 2003. **32**: p. 135-59.
7. Gossen, M. and H. Bujard, *Tight control of gene expression in mammalian cells by tetracycline-responsive promoters.* Proceedings of the National Academy of Sciences, 1992. **89**(12): p. 5547-5551.
8. Metzger, D., et al., Conditional site-specific recombination in mammalian cells using a ligand-dependent chimeric Cre recombinase. Proceedings of the National Academy of Sciences, 1995. **92**(15): p. 6991-6995.
9. Wang, Y., et al., Positive and negative regulation of gene expression in eukaryotic cells with an inducible transcriptional regulator. Gene Ther, 1997. **4**(5): p. 432-41.
10. No, D., T.P. Yao, and R.M. Evans, *Ecdysone-inducible gene expression in mammalian cells and transgenic mice.* Proceedings of the National Academy of Sciences, 1996. **93**(8): p. 3346-3351.
11. Brown, G.A.J. and T.J. Corbin, *Transgenesis in the Mouse: Oocyte injection*, in *Transgenesis Techniques. Methods in Molecular Biology*, A.R. Clarke, Editor 2002, Humana Press: Totowa, New Jersey. p. 39-70.
12. Clark, A.J., et al., *Enhancing the Efficiency of Transgene Expression.* Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 1993. **339**(1288): p. 225-232.
13. Choi, T., et al., *A generic intron increases gene expression in transgenic mice.* Mol. Cell. Biol., 1991. **11**(6): p. 3070-3074.
14. Duncker, B.P., P.L. Davies, and V.K. Walker, *Introns boost transgene expression in Drosophila melanogaster.* Mol Gen Genet, 1997. **254**(3): p. 291-6.
15. Feng, G., et al., Imaging Neuronal Subsets in Transgenic Mice Expressing Multiple Spectral Variants of GFP. Neuron, 2000. **28**(1): p. 41-51.
16. Shen, W.-H., et al., Cardiac Restricted Overexpression of Kinase-dead Mammalian Target of Rapamycin (mTOR) Mutant Impairs the mTOR-mediated Signaling and Cardiac Function. Journal of Biological Chemistry, 2008. **283**(20): p. 13842-13849.
17. Pillai-Nair, N., et al., Neural cell adhesion molecule-secreting transgenic mice display abnormalities in GABAergic interneurons and alterations in behavior. J Neurosci, 2005. **25**(18): p. 4659-71.
18. Martin, A.F., et al., *Expression and function of COOH-terminal myosin heavy chain isoforms in mouse smooth muscle.* American Journal of Physiology - Cell Physiology, 2007. **293**(1): p. C238-C245.
19. Newton, K., et al., Myodegeneration in EDA-A2 Transgenic Mice Is Prevented by XEDAR Deficiency. Mol. Cell. Biol., 2004. **24**(4): p. 1608-1613.
20. Greggio, E., et al., The Parkinson Disease-associated Leucine-rich Repeat Kinase 2 (LRRK2) Is a Dimer That Undergoes Intramolecular Autophosphorylation. Journal of Biological Chemistry, 2008. **283**(24): p. 16906-16914.

21. Johnson, R.A., G.I. McFadden, and C.D. Goodman, Characterization of Two Malaria Parasite Organelle Translation Elongation Factor G Proteins: The Likely Targets of the Anti-Malarial Fusidic Acid. PLoS One, 2011. **6**(6): p. e20633.
22. Rozeboom, A.M., H. Akil, and A.F. Seasholtz, *Mineralocorticoid receptor overexpression in forebrain decreases anxiety-like behavior and alters the stress response in mice.* Proceedings of the National Academy of Sciences, 2007. **104**(11): p. 4688-4693.
23. Zhang, L., R. Hernan, and B. Brizzard, Multiple tandem epitope tagging for enhanced detection of protein expressed in mammalian cells. Molecular Biotechnology, 2001. **19**(3): p. 313-321.
24. Papadakis, E.D., et al., Promoters and control elements: designing expression cassettes for gene therapy. Curr Gene Ther, 2004. **4**(1): p. 89-113.
25. Ramezani, A., T.S. Hawley, and R.G. Hawley, *Lentiviral Vectors for Enhanced Gene Expression in Human Hematopoietic Cells.* Mol Ther, 2000. **2**(5): p. 458-469.
26. Schiedner, G., et al., Genomic DNA transfer with a high-capacity adenovirus vector results in improved in vivo gene expression and decreased toxicity. Nat Genet, 1998. **18**(2): p. 180-183.
27. Graves, R.A., et al., Identification of a fat cell enhancer: analysis of requirements for adipose tissue-specific gene expression. J Cell Biochem, 1992. **49**(3): p. 219-24.
28. Franz, W.-M., et al., Analysis of tissue-specific gene delivery by recombinant adenoviruses containing cardiac-specific promoters. Cardiovascular Research, 1997. **35**(3): p. 560-566.
29. Osborn, L., et al., Tissue-specific and insulin-dependent expression of a pancreatic amylase gene in transgenic mice. Mol. Cell. Biol., 1987. **7**(1): p. 326-334.
30. Sarvetnick, N., et al., Insulin-dependent diabetes mellitus induced in transgenic mice by ectopic expression of class II MHC and interferon-gamma. Cell, 1988. **52**(5): p. 773-782.
31. Robinson, M.O., J.R. McCarrey, and M.I. Simon, *Transcriptional regulatory regions of testis-specific PGK2 defined in transgenic mice.* Proceedings of the National Academy of Sciences, 1989. **86**(21): p. 8437-8441.
32. Zhang, L.P., et al., Multiple Elements Influence Transcriptional Regulation from the Human Testis-Specific PGK2 Promoter in Transgenic Mice. Biol Reprod, 1999. **60**(6): p. 1329-1337.
33. Peschon, J.J., et al., *Spermatid-specific expression of protamine 1 in transgenic mice.* Proceedings of the National Academy of Sciences, 1987. **84**(15): p. 5316-5319.
34. Zambrowicz, B.P., et al., *Analysis of the mouse protamine 1 promoter in transgenic mice.* Proceedings of the National Academy of Sciences, 1993. **90**(11): p. 5071-5075.
35. Bunick, D., et al., Transcription of the testis-specific mouse protamine 2 gene in a homologous in vitro transcription system. Proceedings of the National Academy of Sciences, 1990. **87**(3): p. 891-895.
36. Steger, K., et al., *Expression of protamine-1 and -2 mRNA during human spermiogenesis.* Molecular Human Reproduction, 2000. **6**(3): p. 219-225.
37. Scacheri, P.C., et al., Bidirectional transcriptional activity of PGK-neomycin and unexpected embryonic lethality in heterozygote chimeric knockout mice. genesis, 2001. **30**(4): p. 259-263.
38. Robe, P., et al., Dexamethasone inhibits the HSV-tk/ ganciclovir bystander effect in malignant glioma cells. BMC Cancer, 2005. **5**(1): p. 32.
39. Robe, P.A., et al., Modulation of the HSV-TK/ganciclovir bystander effect by n-butyrate in glioblastoma: correlation with gap-junction intercellular communication. Int J Oncol, 2004. **25**(1): p. 187-92.

40. Mansour, S.L., K.R. Thomas, and M.R. Capecchi, Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature, 1988. **336**(6197): p. 348-352.
41. Pappenheimer, A.M., *Diphtheria Toxin.* Annual Review of Biochemistry, 1977. **46**(1): p. 69-94.
42. Soler, D.C., et al., Expression of Transgenic PPP1CC2 in the Testis of Ppp1cc-Null Mice Rescues Spermatid Viability and Spermiation but Does Not Restore Normal Sperm Tail Ultrastructure, Sperm Motility, or Fertility. Biol Reprod, 2009. **81**(2): p. 343-352.
43. Donovan, J. and P. Brown, *Animal Identification*, in *Current Protocols in Immunology*2006, John Wiley & Sons, Inc.
44. Sambrook J, F.E.a.M.T., *Molecular cloning: a laboratory manual*1989, New York: Cold Spring Harbor Laboratory Press.
45. Wilkie, T.M., R.L. Brinster, and R.D. Palmiter, *Germline and somatic mosaicism in transgenic mice.* Developmental Biology, 1986. **118**(1): p. 9-18.
46. Liang, Z., et al., *Identifying and genotyping transgene integration loci.* Transgenic Research, 2008. **17**(5): p. 979-983.
47. Bishop, C.E., Encyclopedia of the mouse genome III. October 1993. Mouse Y chromosome. Mamm Genome, 1993. **4 Spec No**: p. S282-3.
48. Pravtcheva, D.D., et al., Mosaic expression of an Hprt transgene integrated in a region of Y heterochromatin. J Exp Zool, 1994. **268**(6): p. 452-68.
49. Cooper, T.G., et al., Cytoplasmic droplets are normal structures of human sperm but are not well preserved by routine procedures for assessing sperm morphology. Human Reproduction, 2004. **19**(10): p. 2283-2288.
50. Smith, G.D., et al., Primate sperm contain protein phosphatase 1, a biochemical mediator of motility. Biol Reprod, 1996. **54**(3): p. 719-727.
51. Vijayaraghavan, S., et al., Sperm motility development in the epididymis is associated with decreased glycogen synthase kinase-3 and protein phosphatase 1 activity. Biol Reprod, 1996. **54**(3): p. 709-718.
52. Korrodi-Gregório L, et al. Identification and characterization of a testis specific isoform of inhibitor-2 highly expressed in haploid germ cells during late stages of sperm development. in Europhosphatases. 2011. Vienna, Austria.

Chapter III

Chapter III

Introduction

Phosphoprotein phosphatase 1 (PPP1), one of the major eukaryotic Ser/Thr-PPs, has exquisite specificities *in vivo*, both in terms of substrates and cellular localization. Over the past two decades, it has become apparent that PPP1 versatility is achieved by its ability to interact with multiple targeting/regulatory subunits known as PPP1 interacting proteins - PIPs [1, 2]. Several PPP1-PIP complexes have been involved in cytoskeleton functions [3-5], placing PPP1 in the regulation of cytoskeleton dynamics not only at the actin plane but also at the tubulin level [6, 7].

In this Chapter we confirmed a novel PIP of PPP1CC2, the t-complex testis expressed protein 1 domain containing 4, Tctex1d4, using a yeast two-hybrid screen from a human testis cDNA library [7]. In Chapter III.A is described the binding of Tctex1d4 to PPP1C and the co-localization of the holoenzyme in culture cells. Further, we addressed Tctex1d4 tissue expression and localization in testis and sperm. We also showed the importance of the PPP1 binding motif (PPP1BM) in the complex formation. We pursued the PPP1BM motif relevance by analyzing Pika (*Ochotona* genus) aberrant motif (Chapter III.B). Mutation screening was followed to analyze in detail the PPP1BM and surrounding region. These findings were applied to understand the evolutionary mechanisms that were behind these dramatic amino acid changes in Pika.

The results described in Chapter III open new avenues to the possible roles of this dynein light chain, together with PPP1, in microtubule dynamics, sperm function, acrosome reaction and in the regulation of the blood testis barrier.

References

1. Cohen, P.T., *Protein phosphatase 1--targeted in many directions.* J Cell Sci, 2002. **115**(Pt 2): p. 241-56.
2. Virshup, D.M. and S. Shenolikar, *From promiscuity to precision: protein phosphatases get a makeover.* Mol Cell, 2009. **33**(5): p. 537-45.
3. Kao, S.C., et al., *Identification of phostensin, a PP1 F-actin cytoskeleton targeting subunit.* Biochem Biophys Res Commun, 2007. **356**(3): p. 594-8.
4. Lai, N.S., et al., *Phostensin caps to the pointed end of actin filaments and modulates actin dynamics.* Biochem Biophys Res Commun, 2009. **387**(4): p. 676-81.
5. Allen, P.B., et al., *Phactrs 1-4: A family of protein phosphatase 1 and actin regulatory proteins.* Proc Natl Acad Sci U S A, 2004. **101**(18): p. 7187-92.
6. Yang, P., et al., *Protein phosphatases PP1 and PP2A are located in distinct positions in the Chlamydomonas flagellar axoneme.* J Cell Sci, 2000. **113 (Pt 1)**: p. 91-102.
7. Fardilha, M., et al., *Identification of the human testis protein phosphatase 1 interactome.* Biochemical Pharmacology, 2011. **82**(10): p. 1403-1415.

Chapter III.A

TCTEX1D4, a novel Protein Phosphatase 1 interacting protein involved in tubulin dynamics in testis and sperm

Luís Korrodi-Gregório[1], Sandra I. Vieira[2], Sara L. C. Esteves[1], Georg Luers[3], Odete A. B. da Cruz e Silva[2], Margarida Fardilha[4*] and Edgar F. da Cruz e Silva[1†]

[1]*Laboratory of Signal Transduction, Centre for Cell Biology, Biology Department, University of Aveiro, 3810-193 Aveiro, Portugal*

[2]*Laboratory of Neurosciences, Centre for Cell Biology, Health Sciences Department and Biology Department, University of Aveiro, 3810-193 Aveiro, Portugal*

[3]*Institut für Anatomie und Experimentelle Morphologie, Zentrum für Experimentelle Medizin, Universitätsklinikum Hamburg-Eppendorf, D-20246 Hamburg, Germany*

[4]*Laboratory of Signal Transduction, Centre for Cell Biology, Biology Department; Health Sciences Department, University of Aveiro, 3810-193 Aveiro, Portugal*

[†]*deceased on 2nd of March, 2010*

Short-title: Novel PP1 holoenzyme in testis and sperm

Summary sentence: This manuscript describes a physiologically relevant PPP1 interacting protein in testis, Tctex1d4 that is involved in microtubule dynamics, sperm motility and in the regulation of the blood testis barrier.

Keywords: Tctex2beta, dynein, PP1, microtubules, PP1 interacting protein, TGFbeta signaling

Submitted to: Biol Reprod (3.870 IF, A)

* Address all correspondence and requests for reprints to: Margarida Fardilha, Centro de Biologia Celular, Universidade de Aveiro, 3810-193 Aveiro, Portugal, Tel.: 00351 91 8143947; Fax: 00351 234 37039; E-mail: mfardilha@ua.pt

Footnote: This work was supported by the Centre for Cell Biology of the University of Aveiro, by grants from FCT of the Portuguese Ministry of Science and Higher Education to LKMG (SFRH/BD/42334/2007), SV (SFRH/BPD/19515/2004), SLCE (SFRH/BD/41751/2007) and EFCS (POCI/SAU-OBS/57394/2004; PPCDT/SAU-OBS/57394/2004), by Re-equipment Grant, REEQ/1023/BIO/2005, by CRUP (E-92/08;B-32/09) and by the German Academic Exchange Service (DAAD) with grant to GL (DAAD/50112116).

Abstract

Reversible phosphorylation plays an important role as a mechanism of intracellular control in eukaryotes. PPP1, a major eukaryotic Ser/Thr-protein phosphatase, acquires its specificity by interacting with different protein regulators, also known as PPP1 interacting proteins (PIPs). In this work we characterized a physiologically relevant PIP in testis. Using a yeast two-hybrid screen from a human testis cDNA library, we identified a novel PIP of PPP1CC2, the T-complex testis expressed protein 1 domain containing 4 (TCTEX1D4) that has recently been described as a Tctex1 dynein light chain family member. The results conclusively confirm that TCTEX1D4 interacts with the different spliced isoforms of PPP1CC. Also, the binding domain occurs in the N-terminal, where a consensus PPP1 binding motif (PPP1BM) RVSF is present. TCTEX1D4 distribution in testis is consistent with it being involved in distinct functions, from a potentially important player in the TGFbeta signaling at the blood testis barrier to a role in microtubule dynamics in sperm. Immunofluorescence in sperm shows that TCTEX1D4 is present in the tail and in the acrosome region of the head. Moreover, TCTEX1D4 and PPP1 co-localize in the microtubule organizing center (MTOC) and microtubules in cultured cells. Importantly, the TCTEX1D4 PPP1BM seems to be relevant for complex formation, for PPP1 retention in the MTOC and movement along microtubules.

These novel results open new avenues to possible roles of this dynein, together with PPP1. In essence TCTEX1D4/PPP1C complex appears to be involved in microtubule dynamics, sperm motility, acrosome reaction and in the regulation of the blood testis barrier.

Introduction

Reversible protein phosphorylation is a post-transcriptional event, regulated by both protein kinases and phosphatases (PPs), which play an important role as a mechanism of intracellular control in eukaryotes, being involved in almost all cellular functions, from metabolism to signal transduction and cell division [1]. Phosphoprotein phosphatase 1 (PPP1), one of the major eukaryotic Ser/Thr-PPs, has exquisite specificities *in vivo*, both in terms of substrates and cellular localization. Over the past two decades, it has become apparent that PPP1 versatility is achieved by its ability to interact with multiple targeting/regulatory subunits known as PPP1 interacting proteins - PIPs [1, 2]. PPP1 catalytic subunit (PPP1C) is encoded by three different genes giving rise to α/A, β/B, and γ/C isoforms. After transcription, *PPP1CC* undergoes tissue-specific splicing, originating a ubiquitously expressed isoform, PPP1CC1 and a testis-enriched and sperm-specific isoform, PPPCC2 [3].

To date, more than 200 PIPs have been identified, most of them having the consensus PPP1 binding motif (PPP1BM) RVxF, that binds to the catalytic subunit of PPP1, determining its function and specific cellular location [4, 5]. Several PPP1-PIP complexes are involved in cytoskeleton functions. For instance, PPP1-Phostensin holoenzyme has been implicated in actin rearrangements [6]. Phostensin targets PPP1 to F-actin, being an actin filament pointed end-capping protein that is capable of modulating actin dynamics [7]. The protein family PHACTR (all members 1-4) is involved in synaptic activity through the control of the actin cytoskeleton and regulate PPP1 and actin [8]. The above mentioned PIPs bind actin through the amino acids – RPEL - and may direct PPP1 to a panoply of actin-associated substrates. Thus, several lines of evidence place PPP1 in the regulation of cytoskeleton dynamics, together with various PIPs. This occurs not only at the actin plane but also at the tubulin level. PPP1 has been shown to be anchored to *Chlamydomonas* central pair apparatus axoneme, associated with the C1 microtubule, and to a lesser extent to the outer doublet microtubules, suggesting that PPP1 can control both dynein arms and thereby flagellar motility [9]. Also, recent data from our laboratory showed that PPP1 co-immunoprecipitates with tubulin from human sperm [10].

Clearly the key to characterizing the diverse roles of PPP1 relies on the identification of novel PIPs, and in understanding the specific functions of PPP1/PIP complexes. Therefore, we focused on identifying novel PIPs, through yeast two-hybrid screens where PPP1CC isoforms were used as baits [10-14]. In this study, we present a novel partner of PPP1, the T-complex testis expressed protein 1 domain containing 4 (TCTEX1D4/Tctex2β), which has recently been described as a novel Tctex1 dynein light

chain family member [15]. Further the TCTEX1D4/PPP1CC subcellular co-localization and its dependence on TCTEX1D4-PPP1CC binding, support functions for the complex in microtubules dynamics. Simultaneously, the data also contributes to our understanding of the molecular basis of sperm motility as well as the dynamic and varied functional nature of PPP1.

Materials and Methods

Plasmid constructs

Construction of plasmids was carried out as previously described [14]. The following plasmids were prepared:

pAS2-PPP1CC1 – *PPP1CC1* cDNA was directionally subcloned into *Sal*I/*Sma*I digested pAS2-1 (Clontech, Saint Germain-en-Laye, France) to produce pAS2-PPP1CC1. pAS2-PPP1CC2/pAS2-PPP1CC2end - The 200bp *PPP1CC2*-specific C-terminal-containing *Pst*I fragment was transferred from pTacTac-PPP1CC2 into *Pst*I digested pAS-PPP1CC1 to produce pAS-PPP1CC2 or into *Pst*I digested pAS2-1 to produce pAS-PPP1CC2end [12]. The pAS2-PPP1CC constructs were used in the yeast two-hybrid screening. pET-TCTEX1D4 - *TCTEX1D4* cDNA was PCR-amplified forward (5'-gcgaattcatggccagcaggcctc-3') and reverse (5'-ccgctcggtcactcgcagtagagc-3') from clone IMAGE 30531412 and inserted into pET-28a vector (Novagen, Madison, Wisconsin, USA). pET-TCTEX1D4-NT/CT – The N-terminal portion of the *TCTEX1D4* cDNA was PCR amplified with pET forward primer and an internal reverse primer (5'-ccgctcgagtcaggcgggcgccaggg-3'). The pET-TCTEX1D4-NT construct comprises amino acids 1 to 101. Likewise, the C-terminal portion of the *TCTEX1D4* cDNA was PCR amplified with an internal forward primer (5'-gcgaattccgttgggtggcgcc-3') and pET reverse primer. The pET-TCTEX1D4-CT construct comprises amino acids 102 to 221. The pET-TCTEX1D4 and NT and CT constructs were used in the overlay assay for PPP1C binding. Myc-TCTEX1D4 - *TCTEX1D4* cDNA was PCR-amplified from pET-TCTEX1D4 with primers forward (5'-gcgaattccgatggccagcaggcctc-3') and reverse (5'-ccgctcggtcactcgcagtagagc-3'), and inserted into *EcoRI*/*XhoI* sites of pCMV-Myc vector (Clontech, Saint Germain-en-Laye, France). This construct was used in immunoprecipitation and in immunofluorescence. pACT-TCTEX1D4 - *TCTEX1D4* cDNA was digested from Myc-TCTEX1D4 with *EcoRI*/*XhoI*, and inserted into *EcoRI*/*XhoI* digested pACT-2 (Clontech, Saint Germain-en-Laye, France), using standard molecular biology procedures. The pACT-TCTEX1D4 and pAS2-PPP1CC constructs were used in the co-transformation assay. The pAS2-1/pACT-2 and the pVA3/pTD1 vectors (Clontech, Saint Germain-en-Laye, France) were used as co-transformation controls. pET-TCTEX1D4-AAAA/Myc-TCTEX1D4-RVSA - Mutagenesis of PPP1BM RVSF in *TCTEX1D4* cDNA was performed using the QuikChange Site-Directed Mutagenesis Kit (Stratagene now Agilent Technologies UK Ltd, Edinburgh, UK) by mutating just the last or all the four amino acids to alanine in order to disrupt the PPP1BM. The pET-TCTEX1D4-

AAAA was used in the overlay assay, whereas, Myc-TCTEX1D4-RVSA construct was used in immunofluorescence studies.

Antibodies

The mouse anti-Myc·tag (Cell Signaling, Danvers, Massachusetts, USA), the mouse acetylated anti-α-tubulin (Zymed Laboratories Inc., Cambridge, UK), the mouse anti-β-tubulin (Zymed Laboratories Inc., Cambridge, UK) and the rabbit anti-TCTEX1D4N (N-terminal, RP11269F19.9, Sigma-Aldrich Química, S.A., Sintra, Portugal) antibodies were purchased from the respective companies. The antibodies CBC3C (against the C-terminal of PPP1CC, detects both isoforms [3]), CBC502 (specific for the C-terminal of PPP1CC2) and CBC8C (against the C-terminal of TCTEX1D4) were raised in rabbit.

Yeast two-hybrid screening

The methods for yeast two-hybrid screening of a human testis cDNA library using human PPP1 have been described previously [10, 12, 13].

Bioinformatics analysis

Full protein sequences of intermediate chains (ICs), light intermediate chains (LICs) and light chains (LCs) were obtained from Ensembl database to check for PPP1 binding motifs. The ICs human homologues are: the axonemal inner arm IC1 (ENSP00000242317) and IC2 (ENSP00000308312), axonemal outer arm IC138 (ENSP00000360065) and IC140 (ENSP00000294664), and cytoplasmic DYNC1I1 (ENSP00000320130), DYNC1I2 (ENSP00000380308). The LICs human homologues are the cytoplasmic DYNC1LI1 (ENSP00000273130), DYNC1LI2 (ENSP00000258198) and cytoplasmic from intraflagellar transport DYNC2LI1 (ENSP00000330752). The LCs human homologues are the DYNLRB1/LC7/Roadblock (NP_054902.1), DYNLRB2/LC7/Roadblock (ENSP00000302936), DYNLL1/LC8 (ENSP00000376297), DYNLL2/LC8 (ENSP00000240343), TXNDC3/LC5 (ENSP00000199447), TXNDC6/LC5 (ENSP00000372667), DNAL1/LC1 (ENSP00000310360), DNAL4/LC6 (ENSP00000216068) DYNLT1/Tctex1, DYNLT3/RP3, TCTEX1D1, TCTEX1D2/Tctex2b, TCTEX1D3/Tctex2/Tcte3, and TCTEX1D4/Tctex2β. The last 6 comprise the Tctex1 family and the respective Ensembl IDs are depicted in the alignment (Figure III.A. 2A). Eukaryotic Linear Motif (ELM) [16], Psipred [17], ScanProsite (EXPASy Proteomics Server), NetPhos, and NetNGlyc1.0 (CBS Prediction Servers) search engines were used

to further characterize relevant motifs and post-translational modifications. For homology and evolutionary purposes, ClustalW2 and MEGA programs were used [18, 19].

Yeast Co-transformation

Yeast competent AH109 cells were co-transformed with pACT-TCTEX1D4 and pAS2-PPP1CC1, pAS2-PPP1CC2 or pAS2-PPP1CC2end, by the lithium acetate method [10, 12, 20]. Afterwards, the transformation mixture was plated on selective media containing X-α-Gal and incubated at 30ºC to check for MEL1 expression as indicated by the appearance of a blue color (Clontech, Saint Germain-en-Laye, France).

Co-Immunoprecipitation and TCTEX1D4 tissue expression screening

COS-7 cells were grown in the appropriate medium (DMEM) and transfected with Myc-TCTEX1D4, harvested and lysed in lysis buffer (50mM Tris-HCl, 120mM NaCl, 4% CHAPS, 0,1mg/ml Pepstatin A, 0,03mM Leupeptin, 145mM Benzamidine, 0,37mg/ml Aprotinin and 4,4mM PMSF in isopropanol). The lysates were pre-cleared with protein A sepharose slurry (Pharmacia, LKB Biotechnology, Bromma, Sweden) during 1hr at 4ºC with shaking. After centrifugation, protein sepharose and CBC3C (2µg) were added to the supernatant, followed by overnight incubation at 4ºC with shaking. Subsequently, the beads were washed three times with 50mM Tris-HCl, 120mM NaCl and resuspended in loading buffer. Samples were loaded in SDS-PAGE gel, and transferred to a nitrocellulose membrane. Immunodetection was performed using anti-Myc antibody (1:5000). For the tissue screening, tissues from rat and human testis were lysed using a homogenizer in 1%SDS. Both CBC8C (1:100) and anti-TCTEX1D4N (1:1000) antibodies were used in separate blots for immunodetection of protein expression. Immunoreactive bands were detected by enhanced chemiluminescence (ECL, GE Healthcare, Amersham Biosciences Europe GmbH, Freiburg**,** Germany) [21].

Overlay assays

A single *Rosetta* (DE3) (Novagen, Madison, Wisconsin, USA) colony expressing His-tagged TCTEX1D4 was selected and grown overnight in 3 ml Luria Bertani medium containing ampicillin (50 µg/mL) at 37ºC. Expression was induced with 0,1mM isopropyl-β-D-thio-galactopyranoside at 37ºC. Samples were then treated as described elsewhere [13]. The same procedure was also performed for TCTEX1D4-AAAA, TCTEX1D4-NT and

TCTEX1D4-CT. Blots were overlaid with purified PPP1CC1 or PPP1CC2 (25pmol/mL) and detected with CBC3C (1:5000) [13].

Immunohistochemistry

For fluorescence microscopy analysis, cryosections were prepared. C57/Bl6 mice testes were fixed by perfusion with 4% paraformaldehyde, 0.1M HEPES, pH 7.4 and subsequently immersed in 10%, 20% and 25% (w/v) sucrose solution, each for about 4hrs. The tissue was frozen in isopentane at -30°C and stored at -80°C. Cryosections of about 10-14µm were cut using a Microm HM5000 (Zeiss, Wetzlar, Germany). For immunohistochemistry, tissue sections were incubated with 10% (v/v) Roti-Block (Roth, Karlsruhe, Germany) in PBS for 1hr to reduce non-specific binding of antibodies. Sections were further incubated with primary antibodies against TCTEX1D4 (CBC8C, 1:100) in 10% (v/v) Roti-Block in PBS for 1hr at room temperature. After extensive washing with PBS, tissue sections were incubated with a secondary antibody (Cy3-conjugated goat anti rabbit IgG, diluted 1:300, Sigma, Taufkirchen, Germany) suspended in PBS, for 1hr, to visualize immune complexes. For DNA labeling, the sections were incubated in PBS supplemented with 1 µg/ml 4',6-Diamidino-2-phenylindole (DAPI, Gibco BRL) for 1hr at room temperature. Sections were washed in PBS and mounted with 50% glycerol in PBS containing 1.5% (w/v) n-propyl gallate. Specimens were analyzed using a Leica DMRE fluorescence microscope with standard filters for detection of Cy3 and DAPI. Digital images were obtained with a Nikon DXM1200F digital camera using the Nikon ACT-1 software.

Immunocytochemistry

GC1-spg and COS-7 cell lines were grown using previously established conditions [22]. Cells were then transfected with Myc-TCTEX1D4 or Myc-TCTEX1D4-RVSA with Lipofectamine 2000 (Invitrogen, Life Technologies S.A., Madrid, Spain), using standard procedures [14, 23]. Preparation of cells for immunocytochemistry was achieved by cold methanol fixation as previously described [21]. Human ejaculated sperm were first washed three times in PBS, diluted and applied to coated coverslips. Subsequent steps were similar to those applied for the COS-7 cells. Anti-Myc (1:5000), CBC3C (1:1000), CBC8C (1:100) and CBC502 (1:1000) primary antibodies were used to detect the respective proteins. Fluorescent secondary antibodies anti-mouse Texas-Red (1:300) and anti-rabbit Alexa488 (1:300) were subsequently used. Nuclei were stained with Hoechst 33258 (1:2000, Polysciences Europe GmbH, Eppelheim, Germany) or DAPI (1:200,

Vectashield, Vector Laboratories Burlingame, California, USA). Fluorescence images were acquired in an Olympus IX-81 inverted epifluorescence microscope (Olympus Portugal - Opto-Digital Tecnologias, S.A., Lisboa, Portugal) using an x100 objective, or on a Zeiss LSM-510 confocal microscope (Carl Zeiss Ltd., Wetzlar, Germany). For confocal microscopy, quantitative correlation analysis of PPP1CC and wild type and mutant TCTEX1D4 was carried out with the Zeiss LSM 510 4.0 software [23], using images of delimited TCTEX1D4 transfected single cells. The co-localization coefficients were determined as the percentage of PPP1CC/TCTEX1D4 co-localizing pixels relatively to the number of pixels in the PPP1CC and in the TCTEX1D4 channels.

Statistical analysis

SigmaPlot statistical package (SigmaPlot v.11, Systat Software Inc.) was used for statistical analysis. The Kolmogorov–Smirnov test was employed to test normality of distribution of the data. Significant differences between co-localization coefficients were evaluated with the unpaired Student's t-test ($p<0.001$, alpha=0.050).

Results

Identification and *in silico* characterization of TCTEX1D4, a novel PPP1CC binding protein

A yeast two-hybrid screen using as bait the C-terminal portion of PPP1CC2 was performed against a human testis cDNA library [20]. Four clones were obtained encoding the T-complex testis expressed protein 1 domain 4 (TCTEX1D4). TCTEX1D4 is a novel member of the dynein LC Tctex1 family, and was recently identified as a binding partner of endoglin, a transmembranar glycoprotein involved in the transforming growth factor beta (TGFβ) signaling [15]. TCTEX1D4 has 221 amino acids, with an expected molecular mass of 23352 Da. The gene maps to chromosome 1p34.1 and has 2 exons [15].

In order to identify physiologically relevant motifs and phosphorylation sites from the signaling perspective, we undertook a bioinformatic search using the human TCTEX1D4 protein sequence in ELM [16], PsiPred [17], ScanProsite, NetPhos and NetNGlyc1.0 servers. The conservation of motifs identified was analyzed by comparing with other mammals, and only results with high scores are shown (score 0,9 in NetPhos). Based in ELM and Psipred, the TCTEX1D4 sequence was first divided in two domains: the disordered domain (residues 1 to 120) and the globular domain (residues 121 to 221) (Figure III.A. 1).

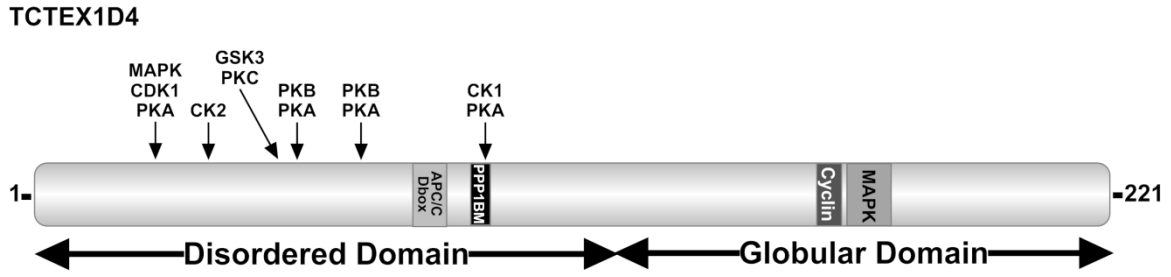


Figure III.A. 1: Schematic representation of the human TCTEX1D4. *The human protein sequence of TCTEX1D4 (ENSP00000361274) was obtained from Ensembl database and submitted to ELM to appoint relevant motifs. TCTEX1D4 disordered (amino acids 1-120) and globular (amino acids 121-221) domains are shown. Putative conserved motifs and phosphorylation sites for specific kinases are indicated.*

Secondary structure prediction of the globular domain shows two alpha helixes, followed by four beta strands resembling that of DYNLT1. Since amino acids from the disordered domain should be more accessible to kinases and phosphatases, only the putatively phosphorylatable motifs of the disordered domain are shown in Figure III.A. 1. The bioinformatic approach revealed that the TCTEX1D4 disordered domain contains

many putative Ser phosphorylation sites, such as consensus sites for protein kinase A (PKA, Ser24/53/66/92), B (PKB, Ser53/66) and C (PKC, Ser49), cyclin dependent kinase 1 (CDK1, Ser24), casein kinase 1 (CK1, Ser92) and 2 (CK2, Ser34), mitogen-activated protein kinase (MAPK, Ser24), and glycogen synthase kinase 3 (GSK3, Ser49). Remarkably, no potential Thr and Tyr phosphorylation sites were found.

The NetNGlyc1.0 server indicated a putative N-glycosylation motif in the extreme C-terminus (residues 205 to 210). Putative binding domains were also identified for the motif "anaphase-promoting complex (APC/C) binding site through the destruction box" (residues 77 to 85), for cyclins (residues 161 to 165), for MAPK (residues 167 to 176) and for PPP1 (residues 90 to 93) (Figure III.A. 1). Additionally and particularly relevant to PIPs, the canonical PPP1BM RVxF was detected in the bioinformatics analysis, ^{90}RVSF^{93}, and reinforces the fact that TCTEX1D4 is a PPP1 binding partner (Figure III.A. 1).

TCTEX1D4 belongs to the Tctex1 family of dynein LCs that share the Tctex1 globular domain, a region of high homology present among the different family members (Figure III.A. 1and Figure III,A. 2)

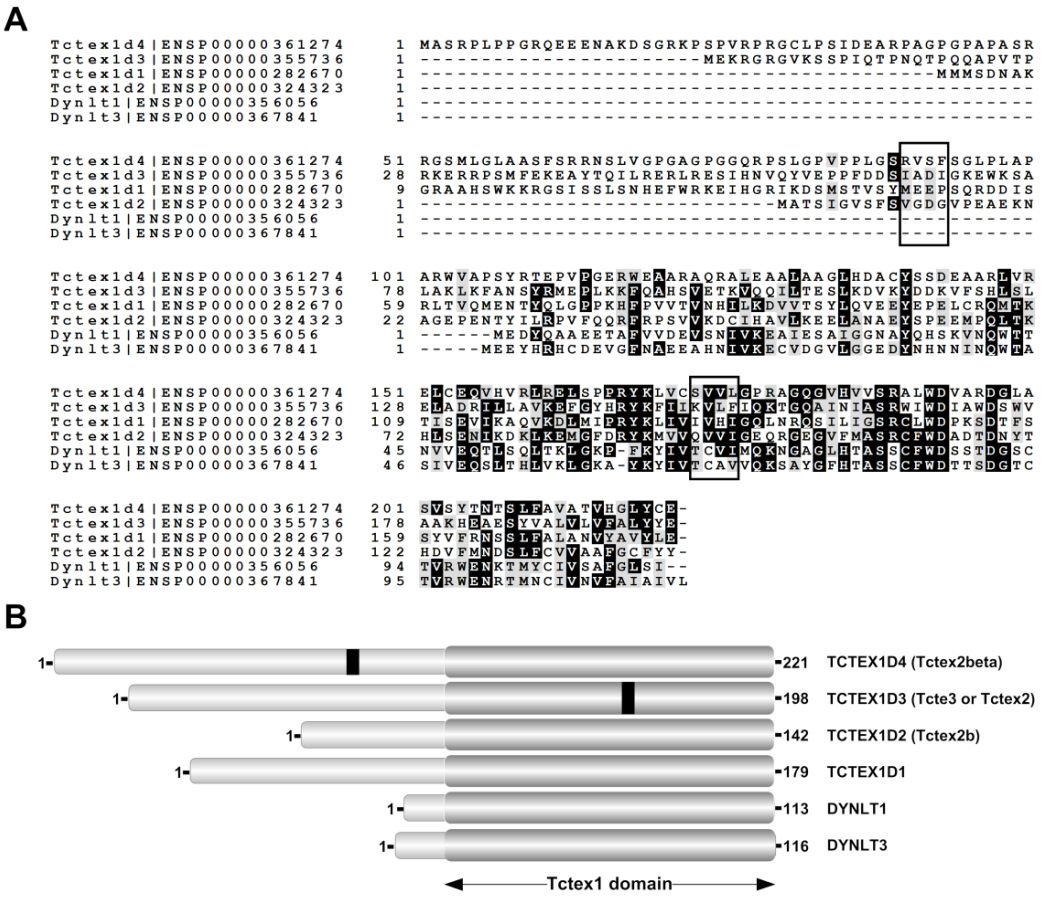


Figure III.A. 2: Alignment and schematic representation of TCTEX1D4 and its family members. *The human sequences of Tctex1 family proteins were obtained from Ensembl database. A. Sequences were submitted to ClustalW2 and the resulting aligning output file shaded*

with BOXSHADE. Open squares indicate the PPP1BM RVSF of TCTEX1D4 and the KVLF motif of TCTEX1D3. B. Schematic representation of human Tctex1 family members with their respective TCTEX1 globular domains (dark grey). The black squares show the position of the PPP1BM of TCTEX1D4 and TCTEX1D3.

The human Tctex1 family comprises 6 members, DYNLT1 (Tctex1), DYNLT3 (RP3), TCTEX1D1, TCTEX1D2 (Tctex2b), TCTEX1D3 (Tctex2/Tcte3), and TCTEX1D4 (Tctex2β). A ClustalW2 protein alignment with the full protein sequences (Figure III.A. 2A) revealed that the Tctex1 domain is common to all the proteins studied. Subsequently for an alignment using only the TCTEX1 globular domain, a score of 57% was obtained between DYNLT1 and DYNLT3, and for the other four family members of the Tctex1 domain containing proteins, the score varied between 25% and 33% (alignment not shown). This implies the formation of two distinct groups, one containing DYNLT1 and DYNLT3 and the other the Tctex1 domain containing proteins. These results are in accordance with the work of DiBella and Meng [15, 24].

Since reversible protein phosphorylation is an important mechanism in the control of dynein function [25, 26], a bioinformatics analyses using ScanProsite to search for all the PPP1 binding motifs known to date in all dynein LCs, light intermediate chains (LICs), and intermediate chains (ICs) was carried out [27]. Interestingly, from all LCs, besides TCTEX1D4, only TCTEX1D3 possesses a consensus PPP1BM (Figure III.A. 2A, B). Between the ICs and LICs, the IC2, DYNC1LI1 and DYNC1LI2 proteins also have a canonical PPP1BM.

In order to evaluate the evolutionary conservation of the phosphorylation sites and motifs described above (Figure III.A. 1), a ClustalW2 alignment was performed using representative orthologs of TCTEX1D4 (Figure III.A. 3A).

Results show that phosphorylation sites Ser34/49/53/66, consensus sites for CK2, GSK3, PKA, PKB and PKC, are conserved in all mammals, while the Ser92 is present in all mammals with the exception of pig (*Sus scrofa*) and kangaroo rat (*Dipodomys ordii*), and Ser24 in all Primates (except *Microcebus murinus* representative of the lower primates Stepsirrhini, lemurs), Artiodactyla, Carnivora (except dog, *Canis familiaris*) and Chiroptera. The N-glycosylation motif is highly conserved in mammals with the exception of Murinae. In addition, the APC/C binding site is conserved in all mammals with the exception of *Cavia porcellus*, the cyclins binding site is present in all mammals with the exception of *Felis catus*, Insectivora and Rodentia (excluding *Dipodomys ordii*), and the MAPK binding site is conserved in all mammals, fish and birds. Finally, the PPP1BM is conserved in all placental mammals, with the exception of *Pteropus vampyrus* from the order Chiroptera (Figure III.A. 3A).

A phylogenetic tree was further constructed using the Neighbour-Joining method in MEGA program, revealing that TCTEX1D4 follows the modern mammal taxonomy (Figure III.A. 3B). Besides being present in placental mammals, TCTEX1D4 is also present in marsupials (*Monodelphis domestica*) and monotremes mammals (*Ornithorhynchus anatinus*), birds (*Gallus gallus*) and fishes (*Danio rerio*, outgroup of the evolutionary tree).

A

```
ENSP00000361274|H.sapiens        MASRPLPPGRQEEE-NAKDSGRKPSPVRPRGCLPSIDEARPAGPGP--APASRRGSMLGLAASFSRRNSLVGPGAGPGGQ [ 80]
ENSPTRP00000001151|P.troglodytes ....S.........-.............................GPG............................. [ 80]
ENSMMUP00000008656|M.mulatta     .............-....Y.P.........R.............--G..........V.T........A........R [ 80]
ENSMICP00000007342|M.murinus     ..G...........-T.....P.LP.M....R......V..T.L----G.......V.............A...V....R [ 80]
ENSMUSP00000052243|M.musculus    ..C.T..SR.....-TT..LAL.LP.GK.G.H......T..I..----G.......LP..HP........A..LV....R [ 80]
ENSOCUP00000010732|O.cuniculus   ..A....T.C....-T...P.P..VA.....R......D..T.V----G.......V..........S..A...T....R [ 80]
ENSSSCP00000004243|S.scrofa      ..G..V.A......-L...P.Q.L.LA..P.R......T.....----G.......V..P.S........A.A......R [ 80]
ENSCAFP00000007008|C.familiaris  ..G.SA......AD-TP..P.P.LPAA..A.H........L...----G.S.....V..P.S........A........R [ 80]
ENSPVAP00000009298|P.vampyrus    ..G....T.C....-A...P.L.L..L.LP.R.S..........----G.......V....S........A.S.VS...R [ 80]
ENSLAFP00000023230|L.africana    ..G...H.RL....KT...P.P..P.L..L.R.........T...L--S...H...V...ST............V....R [ 80]
ENSPCAP00000009744|P.capensis    ..G...HS--....KT....RP.LP.M..L.R.............L--GL......V...GM............V....R [ 80]
ENSETEP00000005024|E.telfairi    ..GK.....I....KI.....P.LP.A..L.R....C....T..----G.......L..Q.................A.R [ 80]
ENSDNOP00000013171|D.novemcinctu ..GK..A..C....-T.....P..LA...-.R.........V..----G.......V..P.................I..R [ 80]

ENSP00000361274|H.sapiens        RPSLGPVPP-LGSRVSFSGLPLAPARWVAPSYRTEPVPGERWEAARAQRALEAALAAGLHDACYSSDEAARLVRELCEQV [160]
ENSPTRP00000001151|P.troglodytes .........-.....................................C.................................. [160]
ENSMMUP00000008656|M.mulatta     .........-.....................Q.............................S...R......A......Q...... [160]
ENSMICP00000007342|M.murinus     .........-.A...................R.....C...A.A.................E.C.T...GA..G..A......I [160]
ENSMUSP00000052243|M.musculus    .........-.....................MLP.RM.....L..A...H....G.........TTQ.NGV..CGS..GK..QA....I [160]
ENSOCUP00000010732|O.cuniculus   ......L..-.....G.....................R....H.R.........V.............R.QNT...GT..G..TK...D.. [160]
ENSSSCP00000004243|S.scrofa      .........-..W.....................G...RP.........A.....................R..R.VGA..G..A....DLA [160]
ENSCAFP00000007008|C.familiaris  .........-.....................L.RA.........A.........V..S.......R.RST...GA..GP.T.....LL [160]
ENSPVAP00000009298|P.vampyrus    .A.......-....A..L....V.S.RM.........A.......SQ......E...S..Q.T...GM..GP.A....KL. [160]
ENSLAFP00000023230|L.africana    .........-.S.....................V...R......M..L.....QT....S...E....K.Y.V...GS..GP.A....D.. [160]
ENSPCAP00000009744|P.capensis    ......M.A-.....................S......R......M..S.....Q..H..NT...V..EA.YNT..AGA..GP.A.KI.... [160]
ENSETEP00000005024|E.telfairi    .........P.....................V.R.........A.....Q.........S....A.RSK..LGA..GP.T..V..R. [160]
ENSDNOP00000013171|D.novemcinctu .........-....I.....V...R.........A.....Q.V....V..........RN....GH..GS.A........ [160]

ENSP00000361274|H.sapiens        HVRLRELSPPRYKLVCSVVLGPRAGQGVHVVSRALWDVARDGLASVSYTNTSLFAVATVHGLYCE [225]
ENSPTRP00000001151|P.troglodytes ..................................................................... [225]
ENSMMUP00000008656|M.mulatta     ..........................I.......................................... [225]
ENSMICP00000007342|M.murinus     ..............................Q.........A.T.....AT...A.............. [225]
ENSMUSP00000052243|M.musculus    .T.V...NL.......N......E.............AVH.....ATF..P.........AV.W. [225]
ENSOCUP00000010732|O.cuniculus   R......C.............................A.......ATF..A............Y. [225]
ENSSSCP00000004243|S.scrofa      R.H............................A......T.C.....ATV..A.............. [225]
ENSCAFP00000007008|C.familiaris  R......C........G.............A......AES.....ATF..A......M....... [225]
ENSPVAP00000009298|P.vampyrus    ......................T.............E.......ATF..A......M....... [225]
ENSLAFP00000023230|L.africana    RL.....................D.R........TT......ATF..A......I....... [225]
ENSPCAP00000009744|P.capensis    .L.............N......T..D.R........T.C.RM..ATF..A......M....... [225]
ENSETEP00000005024|E.telfairi    RL....F................V..D..........A.......ATF..S............Y. [225]
ENSDNOP00000013171|D.novemcinctu RL.....................T....R........E.C.....AAF..S.............. [225]
```

B

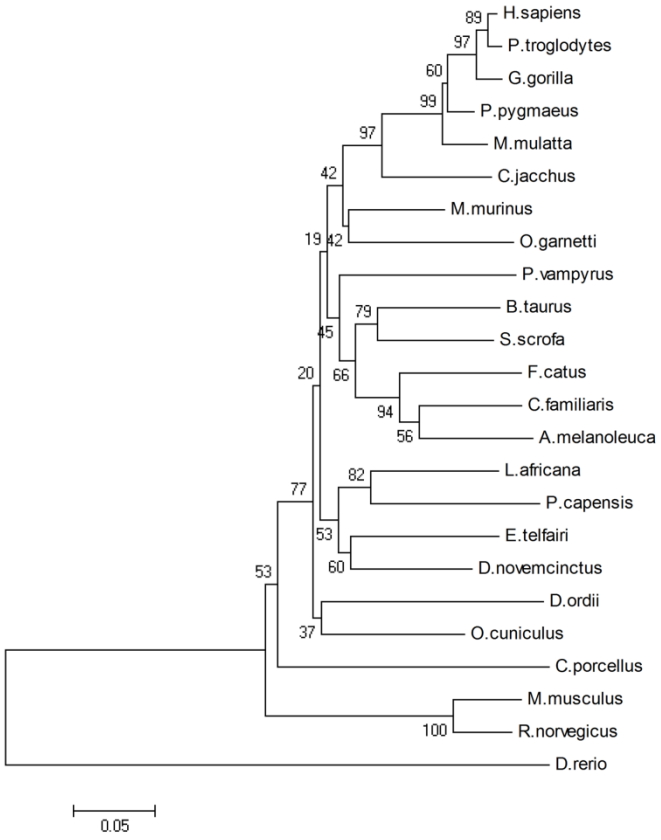


Figure III.A. 3 (previous page): Alignment and phylogenetic relationship between human TCTEX1D4 and its homologues. *The protein sequences of TCTEX1D4 homologues were obtained from the Ensembl database, by performing a Blastp search through the metazoans group using the human TCTEX1D4 sequence.* **A.** *The chosen sequences were submitted to a ClustalW2 alignment. Open box indicate conservation of the RVSF motif across the placental mammals.* **B.** *The resulting phylogenetic tree output was obtained by employing a Neighbor-Joining method with bootstrap test in MEGA program. D. rerio (zebrafish) was chosen as out-group. Scale bar of 0,05 substitutions per residue.*

Validation of the TCTEX1D4 - PPP1 interaction

To demonstrate and validate the interaction between TCTEX1D4 and PPP1C isoforms, AH109 yeast strain was co-transformed with pACT-2-TCTEX1D4 and with *PPP1CA*, *PPP1CC1*, *PPP1CC2* or the *PPP1CC2end* in pAS2-1 [12, 20]. After growing in selective media, colonies were transferred to plates with X-α-Gal, and all colonies turned blue, indicating that TCTEX1D4 interacts with the PPP1C isoforms (Figure III.A. 4A).

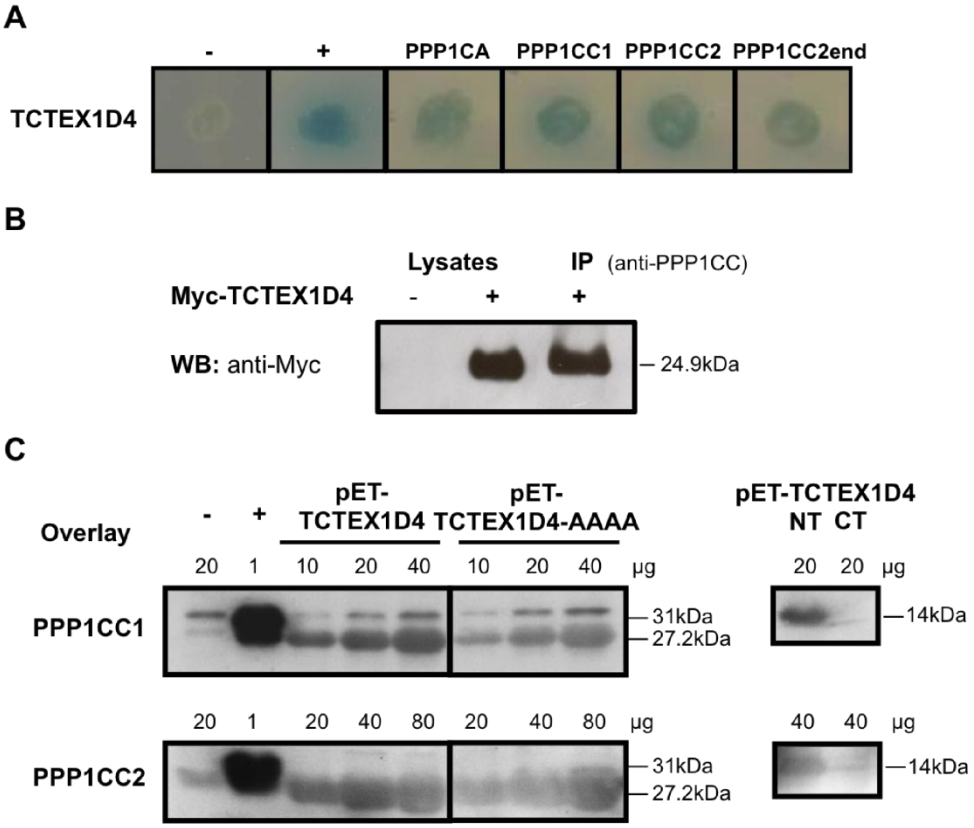


Figure III.A. 4: TCTEX1D4 binds to PPP1. A. *Yeast co-transformation of pACT-2-TCTEX1D4 with PPP1CA, PPP1CC1, PPP1CC2 or PPP1CC2end in pAS2-1 vector, using the Li-Ac method. For negative and positive controls pAS2-1/pACT-2 and pVA3/pTD1 vectors were used, respectively.* **B.** *Western blot showing that TCTEX1D4 binds to and co-immunoprecipitates PPP1CC in COS-7 cells transfected with Myc-TCTEX1D4. Non-transfected and transfected COS-7 cells were used as negative and positive controls, respectively.* **C.** *Bacterial extracts expressing pET-TCTEX1D4, pET-TCTEX1D4-AAAA, pET-TCTEX1D4-NT (NT) and pET-TCTEX1D4-CT (CT) were run in a SDS-*

PAGE gel and overlaid with purified PPP1CC isoforms. pET vector alone was used as negative control and pET-NEK2C, a known PIP, was used as positive control.

The existence of TCTEX1D4-PPP1 complexes *in vivo* was shown by co-immunoprecipitation of the latter from cell lysates of COS-7, previously transfected with Myc-TCTEX1D4 (Figure III.A. 4B). Indeed, when PPP1CC was immunoprecipitated from cells transfected with Myc-TCTEX1D4, this protein was highly co-immunoprecipitated, as denoted by the intense Myc-immunopositive band appearing in the western blot (Figure III.A. 4B).

In order to test if the interaction between TCTEX1D4 and PPP1 was direct, an overlay was performed in blots of bacterial cell lysates expressing recombinant TCTEX1D4 (Figure III.A. 4C). Results show that TCTEX1D4 binds directly to both PPP1CC1 (upper panel) and PPP1CC2 (lower panel) purified proteins.

Several papers demonstrate that, mutation of the PPP1BM either to AAxA [28], RAxA or to RVxA abolish the PPP1/PIP interaction, although some cases exist where interaction still occurs at less extent [29-32]. We also tested by overlay, if mutation of the TCTEX1D4 PPP1BM was important to complex formation. Results show that, mutation of the TCTEX1D4 PPP1BM RVSF to AAAA only partially decreases the binding by 35% (pET-TCTEX1D4-AAAA lanes), with the same being true when RVSF is mutated to RVSA (*data not shown*). Furthermore, when using TCTEX1D4 N-terminal (NT) and C-terminal (CT) truncated forms in the overlay assay (Figure III.A. 4C, right panel), it was clearly shown that the important motifs for PPP1C binding, such as the RVSF, were present in the N-terminal portion of the TCTEX1D4 protein, although it cannot be excluded that the C-terminus might be important for the binding stabilization.

TCTEX1D4 profile in different tissues

An exhaustive NCBI EST database analysis with *TCTEX1D4* mRNA (NM_001013632.2) permitted the identification of a total of 42 hits in *Sus scrofa*, *Rattus norvegicus*, *Mus musculus*, *Homo sapiens* and *Danio rerio*. From the total ESTs identified, 45% corresponded to female reproductive tract related tissues (ovary, oviduct, placenta, uterus and embryonic tissue). Head related tissues (head, hypothalamus, brain, corpus striatum and tongue) corresponded to 21% and lung to 17% of the total ESTs encountered. Moreover, in these public databases no ESTs were found in testis, although previous work has already described the message in this tissue by northern blot [15].

A tissue screen of TCTEX1D4 protein expression was consequently performed on several rat tissues and human testis (Figure III.A. 5).

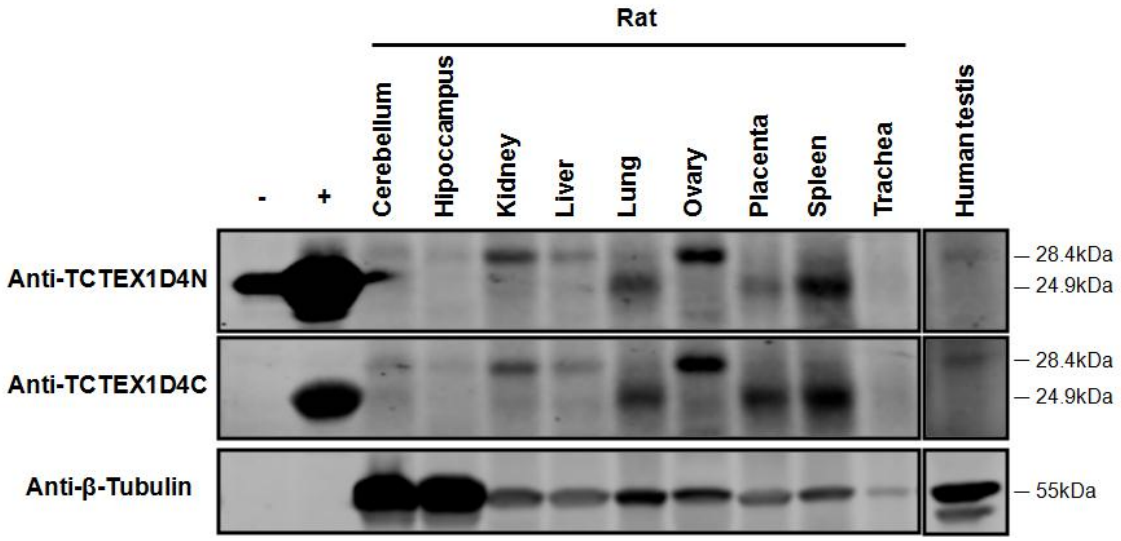


Figure III.A. 5: TCTEX1D4 protein expression profile. *Rat tissues and human testis extracts were prepared as described in material and methods. A tissue protein expression profile (100µg) of TCTEX1D4 was performed by western blot using CBC8C (anti-TCTEX1D4C) and anti-TCTEX1D4N antibodies. The negative control consisted of bacterial extract expressing the pET vector alone, and the positive control of a bacterial extract expressing pET-TCTEX1D4. A loading control, β-tubulin, was also performed.*

The results show that TCTEX1D4 is highly expressed in ovary, spleen, lung, placenta and kidney. It is also expressed in the other tissues tested, but in lower amounts. Interestingly, both antibodies, recognizing the N- or the C-terminus of TCTEX1D4, revealed the same immunoreactive profile, meaning that the two bands detected with each of the antibodies (24.9 and 28.4 kDa) correspond to the full-length TCTEX1D4. While the protein predicted molecular weight is of 23.4 kDa, the band shifts observed in the immunoblots (Figure III.A. 5) most probably result from post-translational modifications, such as phosphorylations and/or glycosylations, sites for which have been predicted (Figure III.A. 1).

Cellular localization of TCTEX1D4 in mouse testis and human spermatozoa

The cellular and subcellular localization of TCTEX1D4 was analyzed by immunohistochemistry in mouse testis (Figure III.A. 6).

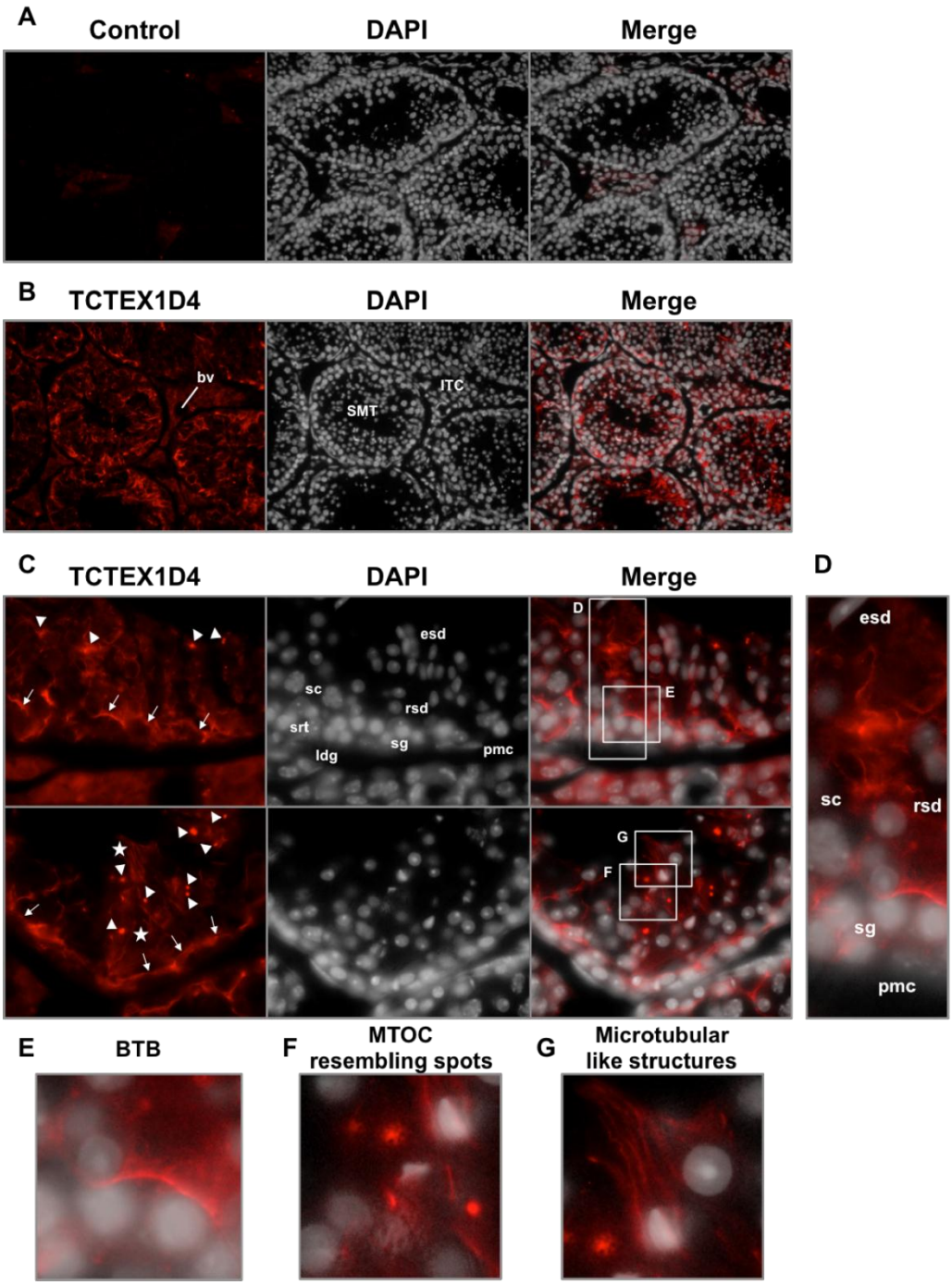


Figure III.A. 6: TCTEX1D4 cellular localization in testis. *Mouse testis sections were stained with DAPI (nucleus, grey in A-G), and an anti-TCTEX1D4 antibody (CBC8C), and visualized by Cy3-labelled secondary antibody (red in A-G).* **A.** *Control staining with only Cy3-labelled secondary antibody (1:300). Image obtained at 20x magnification.* **B.** *Sections stained with CBC8C (20x magnification). bv, blood vessel; SMT, seminiferous tubule; ITC, intertubular compartment.* **C.** *Sections stained with CBC8C (63X magnification). Stars indicate microtubular like structures, arrowheads indicate spots that resemble microtubule organizing center (MTOC) and thin arrows the blood testis barrier (BTB). sp, spermatogonia; sc, spermatocyte; rsd, round spermatid; esd, elongate spermatid; srt, Sertoli cell; ldg, Leydig cell; pmc, peritubular myoid cell* **D.** *ROI (region of interest), showing the principal germ cells.* **E.** *ROI, showing the microtubular structures.* **F.** *ROI, showing the MTOC resembling spots.* **G.** *ROI, showing the BTB.*

TCTEX1D4 was observed in the cytoplasm of most cells of the seminiferous tubules, including germ cells and Sertoli cells. The immunoreactivity was also present in Leydig cells of the interstitial compartment, but could not be detected in the peritubular myoid cells (Figure III.A. 6, panels B and C). Three distinct patterns of subcellular distribution were evident (Figure III.A. 6, panel C). The immunoreactivity was most intense near the cell-cell junctions (Figure III.A. 6, panel C, arrows and panel E) that are consistent with the blood testis barrier (BTB). TCTEX1D4 immunoreactivity was also observed to be highly intense in spots that resemble the microtubule organizing center (MTOC) structure in late stage germ cells (Figure III.A. 6, panel C, arrowheads and panel F). In addition, the subcellular localization of TCTEX1D4 acquires a characteristic pattern of microtubular like structures in late stage spermatids (Figure III.A. 6, panel C, asterisks and panel G).

Given the presence of TCTEX1D4 in late stage germ cells, we further analyzed its subcellular localization in mature human sperm ejaculate (Figure III.A. 7).

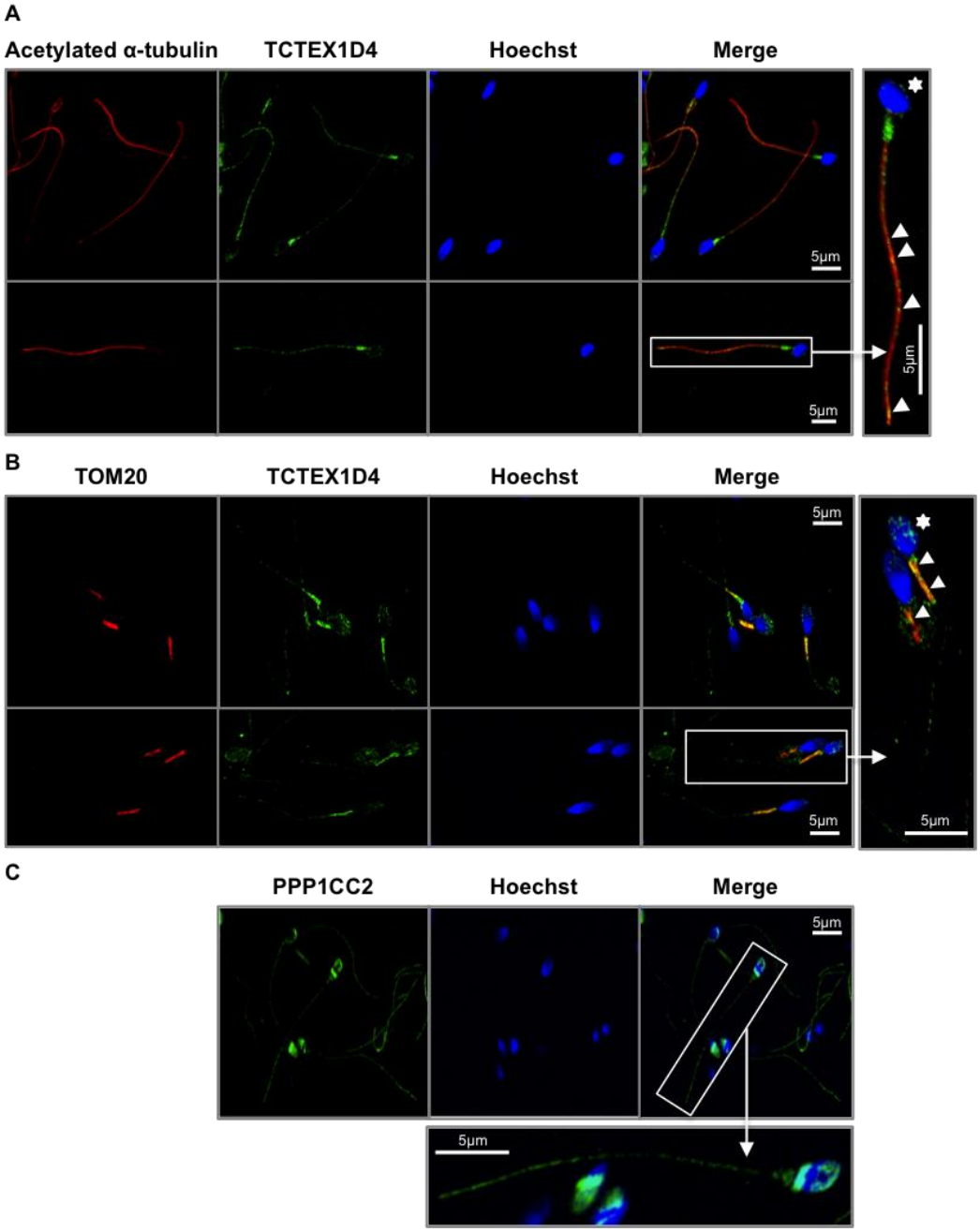


Figure III.A. 7: TCTEX1D4 subcellular localization in human sperm. A. *Human ejaculate sperm was stained with α-tubulin (anti-acetylated-α-tubulin, red), TCTEX1D4 (CBC8C, green) and nucleus (DAPI, blue), and subjected to confocal microscopy analysis. ROI image on the right is shown for an easier visualization of the co-localization in the axoneme (white arrowheads) and in the acrosome (white asterisks)* **B.** *Spermatozoa were stained for mitochondria (TOM20, red), TCTEX1D4 (CBC8C, green) and nucleus (DAPI, blue). ROI image on the right is shown for an easier visualization of the co-localization in the midpiece (white arrowheads) and in the acrosome (white asterisks)* **C.** *Spermatozoa were stained for PPP1CC2 (CBC502, green) and nucleus (DAPI, blue). ROI image on the bottom is shown for a closer visualization of the PPP1CC2 localization. All images were obtained at 100X magnification and the scale bar is shown.*

Results show that TCTEX1D4 is present along the entire length of the spermatozoa tail, including principal and endpiece, and more predominantly in the midpiece region,

where mitochondria are concentrated (Figure III.A. 7A). TCTEX1D4 is also present in head, particularly in the acrosome, but this pattern is only visible in some spermatozoa (Figure III.A. 7A and B, asterisks). Of note, this subcellular localization pattern is similar when using either the N- or the C-terminal anti-TCTEX1D4 antibodies (*data not shown*). To confirm and further characterize TCTEX1D4 subcellular localization, antibodies against the axonemal component tubulin (anti-acetylated-α-tubulin) and mitochondria (TOM20) were used (Figure III.A. 7A and B, respectively). Merged images clearly show that TCTEX1D4 co-localizes with axoneme and mitochondria. The co-localization pattern obtained for α-tubulin supports the previous observations of TCTEX1D4 microtubular like structures in mouse testis (Figure III.A. 6C). Relatively to the potential subcellular sites of PPP1CC2-TCTEX1D4 co-localization in human sperm, PPP1CC2 was observed in the posterior, equatorial and acrosomal regions of the head and along the entire flagellum, including the mid-piece, upon sperm staining with the CBC502 antibody (Figure III.A. 7, panel C). This is the expected PPP1CC2 sperm distribution and, indeed, past work from our laboratory has shown the interaction of PPP1CC2 with β-tubulin by mass spectrometry in human sperm [10].

TCTEX1D4 and PPP1 co-localize at MTOC/microtubules in mammalian cells

A Myc-TCTEX1D4 construct and mouse anti-Myc antibodies were used to confirm TCTEX1D4 subcellular localization and address TCTEX1D4-PPP1 co-localization in mammalian cells. The subcellular localization of transfected TCTEX1D4 was first analyzed in a spermatogonia cell line, the GC1-spg cells, where TCTEX1D4 was observed to be present in the cell nucleus and cytoplasm, being enriched in the MTOC and in the emergent microtubules, as confirmed by its co-localization with specific subcellular markers: centrin, that specifically stains the centrioles present in MTOC, and β-tubulin, to stain microtubules (Figure III.A. 8A).

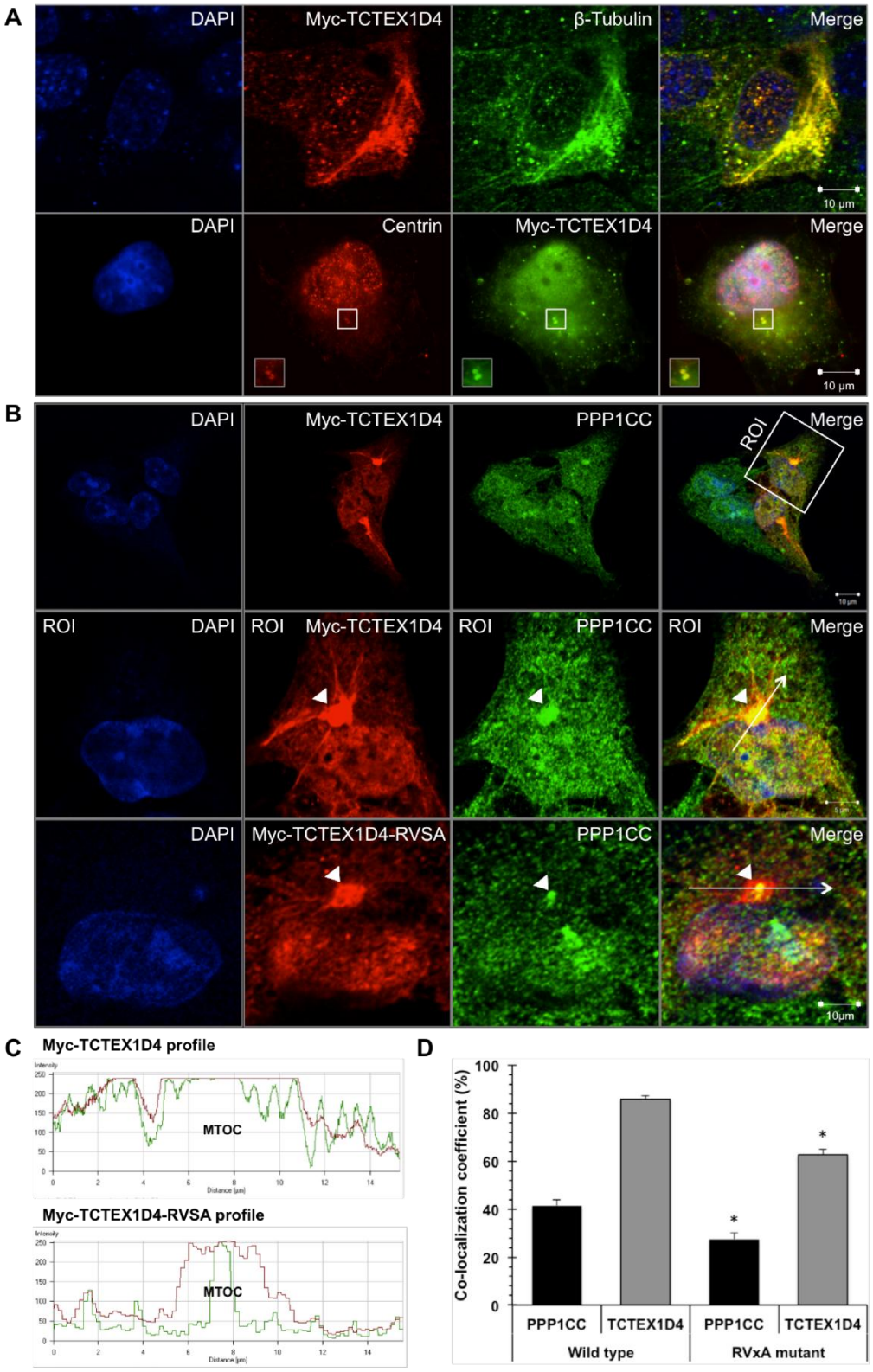


Figure III.A. 8: TCTEX1D4-PPP1 binding regulates the localization of the holoenzyme at microtubules and MTOC. A. *GC1-spg cells (top) and COS-7 cells (bottom) were transfected with Myc-TCTEX1D4, labeled with anti-Myc and Texas-Red anti-mouse antibodies. GC1-spg cells were also immunostained with anti-β-tubulin antibody to visualize microtubules (confocal microscopy analysis), while COS-7 cells were immunostained with anti-centrin antibody to label the microtubule organizing center (MTOC) (epifluorescence microscopy analysis). All images are at 100X*

magnification. Scale bar is shown at bottom left. **B.** *COS-7 cells were transfected with wild type Myc-TCTEX1D4 or the Myc-TCTEX1D4-RVSA mutant, which has reduced ability to bind PPP1C. Arrowhead indicates MTOC. All images are at 100X magnification and were visualized in a confocal microscope. Scale bar is shown at bottom left. White arrows indicate the profiles localization.* **C.** *Fluorescence intensity profiles representing the voxels through the white arrowed lines indicated in B.* **D.** *Co-localization coefficients were determined for the percentage of PPP1CC co-localizing with TCTEX1D4 (black columns), and the percentage of TCTEX1D4 co-localizing with PPP1CC (grey columns), in cells transfected either with wild type (Myc-TCTEX1D4) or mutated (Myc-TCTEX1D4-RVSA) TCTEX1D4. Results were plotted and significant differences were found (*, p<0,001; n=15).*

To further characterize the TCTEX1D4-PPP1 holoenzyme, COS-7 cells were transfected with Myc-TCTEX1D4 and stained with anti-Myc (TCTEX1D4 detection, red) and CBC3C (endogenous PPP1 detection, green) antibodies (Figure III.A. 8B). Again, TCTEX1D4 was observed in the cell nucleus and cytoplasm, where it was present in microtubules and enriched in the MTOC (Figure III.A. 8B). PPP1CC co-localizes with TCTEX1D4 mainly in these regions: inside the nucleus and in the MTOC (Figure III.A. 8B). Interestingly, PPP1CC also accompanies TCTEX1D4 in the microtubules emerging from the MTOC, strongly suggesting that TCTEX1D4-PPP1CC binding might be important for PPP1CC microtubular localization. In order to address this hypothesis we used the Myc-TCTEX1D4-RVSA mutant, where TCTEX1D4-PPP1CC binding decreased by 35% (Figure III.A. 4C). Remarkably, PPP1CC/TCTEX1D4-RVSA co-localized to a much lower extent in the MTOC and along the microtubules emerging from the MTOC (Figure III.A. 8B), as clearly indicated by the fluorescence intensity profiles (Figure III.A. 8C) representing the voxels through the white arrows indicated in Figure III.A. 8B microphotographs. Of note, the PPP1CC signal was particularly decreased in both the MTOC and in microtubules, suggesting that TCTEX1D4 is at least partially responsible for PPP1C MTOC/microtubular localization/transport. Results were further confirmed by quantitative correlation analysis of PPP1CC/TCTEX1D4 co-localization percentages in the cytoplasm of TCTEX1D4 transfected cells (Figure III.A. 8D). While the percentage of cytoplasmic PPP1CC co-localizing with TCTEX1D4 decreased by ~34% (from 41.5±2.5 to 27.5±2.7%), when the RVSF motif was mutated to RVSA the percentage of the TCTEX1D4 cytoplasmic pool co-localizing with PPP1CC decreased by ~27% (85.9±1.3 to 62.7±2.2%). Further, the percentage of transfected cells where PPP1CC/TCTEX1D4 co-localize decreased 3-fold for the TCTEX1D4-RVSA mutant. In essence, both PPP1CC localization at MTOC and its microtubular transport appear to be dependent on PPP1CC binding to the dynein LC TCTEX1D4.

Discussion

The work here described revealed that TCTEX1D4, a new dynein LC present in sperm and testis, forms a complex with PPP1CC2, a phosphatase with an established role in sperm motility [33-36]. The PPP1-TCTEX1D4 interaction was first identified by a yeast two-hybrid approach, being further confirmed by overlay and co-immunoprecipitation techniques (Figure III.A. 4). TCTEX1D4 has a PPP1BM, evolutionarily conserved among species (Figure III.A. 3), and studies with TCTEX1D4 PPP1-binding mutants strengthened the importance of this PPP1 interaction motif for the complex formation (Figure III.A. 4C). TCTEX1D3 is the only dynein light chain, besides TCTEX1D4 that possesses a putative motif (150KVLF153), but the motif is mapped to the globular domain Tctex1, raising the possibility that it might be masked and not as easily accessible for PPP1 binding (Figure III.A. 2). Functionally, while the ICs/LCs form the cargo complex, the LC confer specificity to this binding [37, 38], regulate other molecules or stabilize the assembly of the motor dynein complex [39]. PPP1CC is known to bind to and dephosphorylate the IC [40], and most probably PPP1 binding to TCTEX1D4 would facilitate its access and binding to the IC. Furthermore, PPP1 may additionally dephosphorylate TCTEX1D4 itself, and therefore regulate its function. Of note, TCTEX1D4 has a comparatively longer disordered N-terminus (Figure III.A. 2), giving it a more flexible and exposed arm to bind different cargo and diverse regulatory proteins, as well as being regulated by reversible phosphorylation.

Serine phosphorylation appears to be the most relevant post-translational modification for TCTEX1D4, with several putative target kinases being identified, and supporting the relevance of its binding to the Ser/Thr phosphatase PPP1 (Figure III.A. 1). Also, the results of the putative binding sites for cyclins and MAPK reinforce the Ser24 phosphorylation site predicted for cyclin/CDK complex and MAPK, sustaining a possible role for TCTEX1D4 in the cell cycle, proliferation and differentiation processes, also controlled by the TGFβ signaling pathway [41]. However, the binding sites for cyclins and MAPK are localized in the globular domain, and so the motifs' availability, for instance in loops within the globular domain, still requires further confirmation (Figure III.A. 1). Moreover, TCTEX1D4, in contrast to TCTEX1, TCTEX1D2 and TCTEX1D3 that have orthologs in Chlamydomonas genus [24, 42, 43], emerged only in the vertebrates, maybe by duplication, which may suggest the gain of novel and specialized functions for this protein in this phylum (Figure III.A. 3). One of these functions might have been the binding to PPP1 since the PPP1BM appeared in the placental mammals, interestingly, at the same time as PPP1CC2.

Tctex1d4 binding mutants show a decrease in the binding but not a complete ablation. This fact may be explained by the unusual sequence surrounding the TCTEX1D4 RVSF motif, which is similar to a palindrome with a high percentage of prolines – PSLGPVPPLGSRVSFSGLPLAPARWVAP (bold, prolines, italic, palindrome, underlined, PPP1BM). This sequence is likely to form a structured arm forcing towards the RVSF motif, even when this is mutated to AAAA, to enter the PPP1 hydrophobic pocket to which the RVxF motif is known to bind. It is possible that RVSF ablation leads to the destruction of the arm, and consequently to the ablation of TCTEX1D4-PPP1 binding, but this needs further proof. Also, another possible hypothesis explaining the observed partial reduction in binding is that other important binding sites, not yet described, could also be present.

TCTEX1D4 tissue expression profile indicated that this protein is expressed in several tissues, but to a higher extent in ovary, spleen, lung and placenta, where PPP1 is also present [3] (Figure III.A. 5). These results are in agreement with the EST TCTEX1D4 profile, and also with previous results, describing TCTEX1D4 in a human placenta cDNA library by yeast two-hybrid and by RT-PCR in human testis and placenta [15].

Immunofluorescence studies additionally revealed that, in mouse testis, TCTEX1D4 is present at all germ cells stages, as well as in Sertoli and Leydig cells, being enriched at the cell-cell junctions of BTB and in the late germ cells in a distinctive microtubular/MTOC pattern (Figure III.A. 6). TCTEX1D4 interacts with endoglin and TGFβRII receptors, and inhibits both TGFβ1/3 signaling by increasing the retention time of both receptors at the cell surface and blocking their internalization [15]. TGFβ signaling has also been shown to participate in the regulation of the BTB physiology [44, 45]. BTB, although being one of the tightest blood tissue barriers, has some permeability during the stages VIII to IX of the germ cell cycle to allow for the migration of preleptotene/leptotene spermatocytes towards the adluminal compartment. The current hypothesis is that the internalization (endocytosis) and recycling events are evenly balanced. However, during the stages VIII to IX the internalization event is increased leading to an imbalance and a permissive BTB [44]. Cytokines, TGFβ3 and TNFα, were already shown to be involved in these phenomenon by increasing endocytosis of the membrane receptors [44, 46-48]. Therefore, TCTEX1D4 might have an important role in the regulation of the BTB, either by preventing higher levels of receptor internalization and thereby maintaining the balance before the stages VIII to IX, or by stopping the effect of the cytokines and re-establishing the balance after the preleptotene/leptotene spermatocytes migration. Further, immunocytochemistry analysis confirmed that TCTEX1D4 is present in mature

spermatozoa, particularly in the tail, being enriched in the midpiece region (Figure III.A. 7A and B). This type of expression (testis and sperm) is consistent with a role for TCTEX1D4 as a dynein LC that is present in both cytoplasm and axoneme. Therefore, TCTEX1D4 could be involved in both organelle rearrangements and protein transport (cytoplasmic functions) and in ciliar/flagellar motility (axonemal functions), similarly to its family member DYNLT1, which is both axonemal and cytoplasmic [43, 49]. In sperm tail immunofluorescence microphotographs it is difficult to distinguish different structures, thus it remains unclear whether TCTEX1D4 will function in the axoneme apparatus flagellar motility (for sperm motility) or as a minus-end cytoplasmic dynein motor in the intraflagellar transport (IFT) [50] (Figure III.A. 7A). IFT is a motility event in flagella, unrelated to dynein-based motility, which has been observed as a bidirectional transportation of granule-like particles along the length of flagella [51]. According to both PPP1C and TCTEX1D4 distribution in sperm, the localization of the TCTEX1D4/PPP1CC2 complex appears to be primarily restricted to the flagellum and to a lesser extent in the head (Figure III.A. 7). This suggests a putative role for the TCTEX1D4-PPP1CC2 holoenzyme in sperm motility, where TCTEX1D4 could have its dynein functions altered by PPP1CC2 dephosphorylation and/or would function to transport PPP1CC2 to other possible motility-related PPP1C targets. In the head, the complex might have an important role in the acrosome reaction. Studies in rainbow trout and chum salmon sperm have shown that a TCTEX1D3 homologue (LC2) of the outer dynein arm is phosphorylated when sperm is activated [52] and that dephosphorylation by PPP2 induces immotility in sperm [53]. This model could be similar to what happens in human sperm with TCTEX1D4 and PPP1CC2 (Figure III.A. 9).

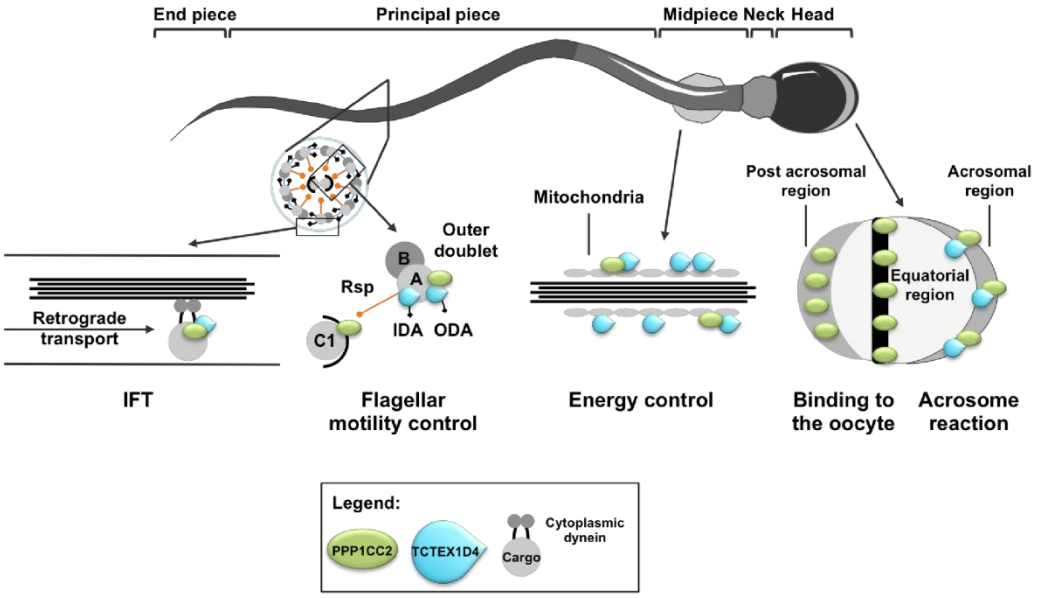


Figure III.A. 9: Schematic representation of TCTEX1D4 and TCTEX1D4-PPP1CC2 localization in sperm. *TCTEX1D4 is present in the head and along the tail of human mature spermatozoa. PPP1CC2 is present in the posterior, equatorial and acrosome regions of the head, as well as, in the entire tail, including midpiece. PPP1 was shown to be present in the central pair apparatus axoneme, associated with the C1 microtubule, and to a lesser extent in the outer doublet microtubules, in Chlamydomonas. The diagram shows the putative roles that could be assigned to TCTEX1D4 (and the TCTEX1D4-PPP1CC2 complex) at these locations. In the head, TCTEX1D4 may have a role in the acrosome reaction, while along the tail it could have a function in sperm flagellar motility and/or in vesicular intraflagellar transport (IFT). In the midpiece, where mitochondria are concentrated, and TCTEX1D4 is highly enriched, TCTEX1D4 may be involved in the energy production necessary for the flagellar motility. Rsp, radial spoke; C1, central pair microtubule 1; B, tubule B; A, tubule A; IDA, inner dynein arm; ODA, outer dynein arm.*

Since mammalian sperm contains a different isoform of PPP1C, the sperm-specific PPP1CC2, and TCTEX1D4 only appeared in the vertebrate branch, this dynein light chain could be the effector of PPP1CC2 activity in mammals, inducing sperm immotility. Figure III.A. 9 is depicts other possible functions of the complex TCTEX1D4/PPP1CC2 in sperm physiology.

Furthermore, PPP1 and TCTEX1D4 were confirmed to co-localize in the MTOC and microtubules, and TCTEX1D4 appears to be at least partially responsible for PPP1 transport along microtubules and PPP1 targeting to MTOC (Figure III.A. 8). This localization is consistent with TCTEX1D4 role as a dynein LC thus linked to microtubules and responsible for PPP1C microtubule-dependent retrograde transport. In this way, TCTEX1D4 may regulate PPP1 functions since PPP1 localization at microtubules is important to regulate microtubule dynamics (e.g. in mitosis [54]) and to regulate cargo transport, by mediating cargo dissociation from the kinesin motor unit [55, 56]. Of note, only part of the TCTEX1D4 cytoplasmic pool is associated with microtubules, as was also reported to occur for TCTEX1, suggesting several distinct roles for these proteins [57].

The presence of TCTEX1D4 in the microtubules and MTOC agrees with the previous immunohistochemistry observations, and strongly suggests a role for TCTEX1D4 in microtubules organization and dynamics. Indeed, dyneins were previously shown to have an important role in MTOC cellular positioning, re-orientation throughout the cell cycle and migration, and in microtubule dynamics [58-60].

Interestingly, the testis protein TLRR (lrrc67), found in the MTOC of germ cells and in culture cells [61], binds to PPP1CC2, β-tubulin, KIF5B (Kinesin-1B), DYNC1I1, and was suggested to have a role in sperm tail formation [61]. The authors further proposed that the TLRR-PPP1CC2 complex regulates the activity of the (plus end) kinesin-1B motor unit in testis. In light of our studies, we suggest that the TCTEX1D4-PPP1CC2 complex may function in the opposite direction, regulating the (minus end) motor unit dynein that comprises TCTEX1D4 as its LC. Additional questions regarding the specific role of this complex in spermatogenesis and sperm physiology will require further work, but will definitely provide new answers to the biology of human reproduction

References

1. Cohen PT. Protein phosphatase 1--targeted in many directions. J Cell Sci 2002; 115: 241-256.
2. Virshup DM, Shenolikar S. From promiscuity to precision: protein phosphatases get a makeover. Mol Cell 2009; 33: 537-545.
3. da Cruz e Silva EF, Fox CA, Ouimet CC, Gustafson E, Watson SJ, Greengard P. Differential expression of protein phosphatase 1 isoforms in mammalian brain. J Neurosci 1995; 15: 3375-3389.
4. Hendrickx A, Beullens M, Ceulemans H, Den Abt T, Van Eynde A, Nicolaescu E, Lesage B, Bollen M. Docking Motif-Guided Mapping of the Interactome of Protein Phosphatase-1. Chemistry & biology 2009; 16: 365-371.
5. Fardilha M, Esteves SL, Korrodi-Gregorio L, da Cruz e Silva OA, da Cruz e Silva FF. The physiological relevance of protein phosphatase 1 and its interacting proteins to health and disease. Curr Med Chem 2010; 17: 3996-4017.
6. Kao SC, Chen CY, Wang SL, Yang JJ, Hung WC, Chen YC, Lai NS, Liu HT, Huang HL, Chen HC, Lin TH, Huang HB. Identification of phostensin, a PP1 F-actin cytoskeleton targeting subunit. Biochem Biophys Res Commun 2007; 356: 594-598.
7. Lai NS, Wang TF, Wang SL, Chen CY, Yen JY, Huang HL, Li C, Huang KY, Liu SQ, Lin TH, Huang HB. Phostensin caps to the pointed end of actin filaments and modulates actin dynamics. Biochem Biophys Res Commun 2009; 387: 676-681.
8. Allen PB, Greenfield AT, Svenningsson P, Haspeslagh DC, Greengard P. Phactrs 1-4: A family of protein phosphatase 1 and actin regulatory proteins. Proc Natl Acad Sci U S A 2004; 101: 7187-7192.
9. Yang P, Fox L, Colbran R, Sale W. Protein phosphatases PP1 and PP2A are located in distinct positions in the Chlamydomonas flagellar axoneme. J Cell Sci 2000; 113: 91-102.
10. Fardilha M, Esteves SLC, Korrodi-Gregório L, Vintém AP, Domingues SC, Rebelo S, Morrice N, Cohen PTW, Silva OABdCe, Silva EFdCe. Identification of the human testis protein phosphatase 1 interactome. Biochemical Pharmacology 2011; 82: 1403-1415.
11. Fardilha M. Characterization of PP1 Interactome from human testis. Universidade de Aveiro 2004.
12. Fardilha M, Wu W, Sa R, Fidalgo S, Sousa C, Mota C, da Cruz e Silva OA, da Cruz e Silva EF. Alternatively spliced protein variants as potential therapeutic targets for male infertility and contraception. Ann N Y Acad Sci 2004; 1030: 468-478.
13. Browne GJ, Fardilha M, Oxenham SK, Wu W, Helps NR, da Cruz ESOA, Cohen PT, da Cruz ESEF. SARP, a new alternatively spliced protein phosphatase 1 and DNA interacting protein. Biochem J 2007; 402: 187-196.
14. Wu W, Baxter JE, Wattam SL, Hayward DG, Fardilha M, Knebel A, Ford EM, da Cruz e Silva EF, Fry AM. Alternative splicing controls nuclear translocation of the cell cycle-regulated Nek2 kinase. J Biol Chem 2007; 282: 26431-26440.
15. Meng Q, Lux A, Holloschi A, Li J, Hughes JM, Foerg T, McCarthy JE, Heagerty AM, Kioschis P, Hafner M, Garland JM. Identification of Tctex2beta, a novel dynein light chain family member that interacts with different transforming growth factor-beta receptors. J Biol Chem 2006; 281: 37069-37080.
16. Puntervoll P, Linding R, Gemund C, Chabanis-Davidson S, Mattingsdal M, Cameron S, Martin DM, Ausiello G, Brannetti B, Costantini A, Ferre F, Maselli V, et al. ELM server: A new resource for investigating short functional sites in modular eukaryotic proteins. Nucleic Acids Res 2003; 31: 3625-3630.

17. David T J. Protein secondary structure prediction based on position-specific scoring matrices. Journal of Molecular Biology 1999; 292: 195-202.
18. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011; 28: 2731-2739.
19. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, et al. Clustal W and Clustal X version 2.0. Bioinformatics 2007; 23: 2947-2948.
20. Mota C. Identification of PP1γ2 interacting proteins from human testis using the YTH system. Universidade de Aveiro 2003.
21. Fardilha M, Vieira SI, Barros A, Sousa M, Da Cruz e Silva OA, Da Cruz e Silva EF. Differential distribution of Alzheimer's amyloid precursor protein family variants in human sperm. Ann N Y Acad Sci 2007; 1096: 196-206.
22. da Cruz ESOA, Vieira SI, Rebelo S, da Cruz e Silva EF. A model system to study intracellular trafficking and processing of the Alzheimer's amyloid precursor protein. Neurodegener Dis 2004; 1: 196-204.
23. Vieira S, Rebelo S, Esselmann H, Wiltfang J, Lah J, Lane R, Small S, Gandy S, da Cruz e Silva E, da Cruz e Silva O. Retrieval of the Alzheimer's amyloid precursor protein from the endosome to the TGN is S655 phosphorylation state-dependent and retromer-mediated. Molecular Neurodegeneration 2010; 5: 40.
24. DiBella LM, Smith EF, Patel-King RS, Wakabayashi K, King SM. A novel Tctex2-related light chain is required for stability of inner dynein arm I1 and motor function in the Chlamydomonas flagellum. J Biol Chem 2004; 279: 21666-21676.
25. Wirschell M, Hendrickson T, Sale WS. Keeping an eye on I1: I1 dynein as a model for flagellar dynein assembly and regulation. Cell Motil Cytoskeleton 2007; 64: 569-579.
26. Vaughan PS, Leszyk JD, Vaughan KT. Cytoplasmic dynein intermediate chain phosphorylation regulates binding to dynactin. J Biol Chem 2001; 276: 26171-26179.
27. Pazour GJ, Agrin N, Walker BL, Witman GB. Identification of predicted human outer dynein arm genes: candidates for primary ciliary dyskinesia genes. J Med Genet 2006; 43: 62-73.
28. Liu D, Vleugel M, Backer CB, Hori T, Fukagawa T, Cheeseman IM, Lampson MA. Regulated targeting of protein phosphatase 1 to the outer kinetochore by KNL1 opposes Aurora B kinase. J Cell Biol 2010; 188: 809-820.
29. Chang JS, Henry K, Wolf BL, Geli M, Lemmon SK. Protein Phosphatase-1 Binding to Scd5p Is Important for Regulation of Actin Organization and Endocytosis in Yeast. Journal of Biological Chemistry 2002; 277: 48002-48008.
30. Traweger A, Wiggin G, Taylor L, Tate SA, Metalnikov P, Pawson T. Protein phosphatase 1 regulates the phosphorylation state of the polarity scaffold Par-3. Proceedings of the National Academy of Sciences 2008; 105: 10402-10407.
31. Bollen M. Combinatorial control of protein phosphatase-1. Trends in Biochemical Sciences 2001; 26: 426-431.
32. Wakula P, Beullens M, van Eynde A, Ceulemans H, Stalmans W, Bollen M. The translation initiation factor eIF2beta is an interactor of protein phosphatase-1. Biochemical Journal 2006; 400: 377-383.
33. Fardilha M, Esteves SLC, Korrodi-Gregório L, Pelech S, da Cruz e Silva OAB, da Cruz e Silva E. Protein phosphatase 1 complexes modulate sperm motility and present novel targets for male infertility. Molecular Human Reproduction 2011; 17: 466-477.
34. Smith GD, Wolf DP, Trautman KC, da Cruz e Silva EF, Greengard P, Vijayaraghavan S. Primate sperm contain protein phosphatase 1, a biochemical mediator of motility. Biol Reprod 1996; 54: 719-727.

35. Smith GD, Wolf DP, Trautman KC, Vijayaraghavan S. Motility potential of macaque epididymal sperm: the role of protein phosphatase and glycogen synthase kinase-3 activities. J Androl 1999; 20: 47-53.
36. Vijayaraghavan S, Stephens DT, Trautman K, Smith GD, Khatra B, da Cruz e Silva EF, Greengard P. Sperm motility development in the epididymis is associated with decreased glycogen synthase kinase-3 and protein phosphatase 1 activity. Biol Reprod 1996; 54: 709-718.
37. Tai AW, Chuang JZ, Bode C, Wolfrum U, Sung CH. Rhodopsin's carboxy-terminal cytoplasmic tail acts as a membrane receptor for cytoplasmic dynein by binding to the dynein light chain Tctex-1. Cell 1999; 97: 877-887.
38. Lo KW, Kogoy JM, Pfister KK. The DYNLT3 light chain directly links cytoplasmic dynein to a spindle checkpoint protein, Bub3. J Biol Chem 2007; 282: 11205-11212.
39. DiBella LM, Gorbatyuk O, Sakato M, Wakabayashi K, Patel-King RS, Pazour GJ, Witman GB, King SM. Differential light chain assembly influences outer arm dynein motor function. Mol Biol Cell 2005; 16: 5661-5674.
40. Whyte J, Bader JR, Tauhata SB, Raycroft M, Hornick J, Pfister KK, Lane WS, Chan GK, Hinchcliffe EH, Vaughan PS, Vaughan KT. Phosphorylation regulates targeting of cytoplasmic dynein to kinetochores during mitosis. J Cell Biol 2008; 183: 819-834.
41. Tian M, Schiemann WP. The TGF-β paradox in human cancer: an update. Future Oncology 2009; 5: 259-271.
42. Pazour GJ, Koutoulis A, Benashski SE, Dickert BL, Sheng H, Patel-King RS, King SM, Witman GB. LC2, the chlamydomonas homologue of the t complex-encoded protein Tctex2, is essential for outer dynein arm assembly. Mol Biol Cell 1999; 10: 3507-3520.
43. Harrison A, Olds-Clarke P, King SM. Identification of the t complex-encoded cytoplasmic dynein light chain tctex1 in inner arm I1 supports the involvement of flagellar dyneins in meiotic drive. J Cell Biol 1998; 140: 1137-1147.
44. Xia W, Wong EWP, Mruk DD, Cheng CY. TGF-β3 and TNFα perturb blood–testis barrier (BTB) dynamics by accelerating the clathrin-mediated endocytosis of integral membrane proteins: A new concept of BTB regulation during spermatogenesis. Developmental Biology 2009; 327: 48-61.
45. Xia W, Mruk DD, Lee WM, Cheng CY. Cytokines and junction restructuring during spermatogenesis—a lesson to learn from the testis. Cytokine & growth factor reviews 2005; 16: 469-493.
46. Xia W, Yan Cheng C. TGF-β3 regulates anchoring junction dynamics in the seminiferous epithelium of the rat testis via the Ras/ERK signaling pathway: An in vivo study. Developmental Biology 2005; 280: 321-343.
47. Xia W, Mruk DD, Lee WM, Cheng CY. Differential Interactions between Transforming Growth Factor-β3/TβR1, TAB1, and CD2AP Disrupt Blood-Testis Barrier and Sertoli-Germ Cell Adhesion. Journal of Biological Chemistry 2006; 281: 16799-16813.
48. Li MWM, Xia W, Mruk DD, Wang CQF, Yan HHN, Siu MKY, Lui W-y, Lee WM, Cheng CY. Tumor necrosis factor α reversibly disrupts the blood–testis barrier and impairs Sertoli–germ cell adhesion in the seminiferous epithelium of adult rat testes. Journal of Endocrinology 2006; 190: 313-329.
49. King SM, Dillman JF, Benashski SE, Lye RJ, Patel-King RS, Pfister KK. The Mouse t-Complex-encoded Protein Tctex-1 Is a Light Chain of Brain Cytoplasmic Dynein. Journal of Biological Chemistry 1996; 271: 32281-32287.
50. Patel-King RS, Benashski SE, Harrison A, King SM. A Chlamydomonas homologue of the putative murine t complex distorter Tctex-2 is an outer arm dynein light chain. J Cell Biol 1997; 137: 1081-1090.

51. Kozminski KG, Johnson KA, Forscher P, Rosenbaum JL. A motility in the eukaryotic flagellum unrelated to flagellar beating. Proceedings of the National Academy of Sciences 1993; 90: 5519-5523.
52. Inaba K, Kagami O, Ogawa K. Tctex2-related outer arm dynein light chain is phosphorylated at activation of sperm motility. Biochem Biophys Res Commun 1999; 256: 177-183.
53. Inaba K. Dephosphorylation of Tctex2-related dynein light chain by type 2A protein phosphatase. Biochem Biophys Res Commun 2002; 297: 800-805.
54. Tournebize R, Andersen SSL, Verde F, Doree M, Karsenti E, Hyman AA. Distinct roles of PP1 and PP2A-like phosphatases in control of microtubule dynamics during mitosis. EMBO J 1997; 16: 5537-5549.
55. Morfini G, Szebenyi G, Brown H, Pant HC, Pigino G, DeBoer S, Beffert U, Brady ST. A novel CDK5-dependent pathway for regulating GSK3 activity and kinesin-driven motility in neurons. EMBO J 2004; 23: 2235-2245.
56. Morfini G, Szebenyi G, Elluru R, Ratner N, Brady ST. Glycogen synthase kinase 3 phosphorylates kinesin light chains and negatively regulates kinesin-based motility. EMBO J 2002; 21: 281-293.
57. Tai AW, Chuang JZ, Sung CH. Localization of Tctex-1, a cytoplasmic dynein light chain, to the Golgi apparatus and evidence for dynein complex heterogeneity. J Biol Chem 1998; 273: 19639-19649.
58. Gonczy P, Pichler S, Kirkham M, Hyman AA. Cytoplasmic dynein is required for distinct aspects of MTOC positioning, including centrosome separation, in the one cell stage Caenorhabditis elegans embryo. J Cell Biol 1999; 147: 135-150.
59. Palazzo AF, Joseph HL, Chen YJ, Dujardin DL, Alberts AS, Pfister KK, Vallee RB, Gundersen GG. Cdc42, dynein, and dynactin regulate MTOC reorientation independent of Rho-regulated microtubule stabilization. Curr Biol 2001; 11: 1536-1541.
60. Schmoranzer J, Fawcett JP, Segura M, Tan S, Vallee RB, Pawson T, Gundersen GG. Par3 and dynein associate to regulate local microtubule dynamics and centrosome orientation during migration. Curr Biol 2009; 19: 1065-1074.
61. Wang R, Kaul A, Sperry AO. TLRR (lrrc67) interacts with PP1 and is associated with a cytoskeletal complex in the testis. Biology of the Cell 2010; 102: 173-189.

Chapter III.B

An intriguing switch in the novel PPP1C binding partner Tctex1d4, between the binding motif and a glycosylation site, in Pikas (Ochotona sp.)

Luís Korrodi-Gregório[1], Ana Margarida Lopes[2], Sara L. C. Esteves[1], Sandra Afonso[2], Ana Matos[2,3], Alexey Lissovsky[4], Odete A. B. da Cruz e Silva[5], Edgar F. da Cruz e Silva[1†], Pedro José Esteves[2,6] and Margarida Fardilha[7*]

[1]*Signal Transduction Laboratory, Centre for Cell Biology, University of Aveiro, Aveiro, Portugal*

[2]*CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Vairão, Portugal*

[3]*Department of Microbiology and Immunology, Stritch School of Medicine, Loyola University Chicago, USA*

[4]*Zoological Museum, Moscow State University, Moscow, Russia*

[5]*Neuroscience Laboratory, Centre for Cell Biology and Health Science Department, University of Aveiro, Aveiro, Portugal*

[6]*CITS, Centro de Investigação em Tecnologias da Saúde, IPSN, CESPU, Portugal*

[7]*Signal Transduction Laboratory, Centre for Cell Biology and Health Science Department, University of Aveiro, Aveiro, Portugal*

Running Title: Tctex1d4 evolution in Pikas revealed positive selection

Keywords: PP1, PP1 binding motif, glycosylation, dynein light chain, Tctex1d4, Pika, positive selection

To be submitted to: The FEBS Journal (3.13 IF, B)

*Corresponding author: Margarida Fardilha, Centro de Biologia Celular, Universidade de Aveiro, 3810-193 Aveiro, Portugal, Tel.: 00351 91 8143947; Fax: 00351 234 370 200 E-mail: mfardilha@ua.pt

Abstract

T-complex testis expressed protein 1 domain containing 4 (Tctex1d4) contains the canonical phosphoprotein phosphatase 1 (PPP1) binding motif (PPP1BM), composed by the amino acid sequence, RVSF [1]. We identified and validate the binding of Tctex1d4 to PPP1 and demonstrated that indeed this protein is a novel PPP1 Interacting Protein [2]. The analyses of twenty-eight mammalian species showed that the PPP1BM $_{90}$RVSF$_{93}$ is present in all and is flanked by a palindromic sequence, PLGS, except in the three *Ochotona* species (*O. princeps*, *O. daurica* and *O. pusilla*). Furthermore, for the *Ochotona* species an extra glycosylation site, motif $_{96}$NLS$_{98}$, and the loss of the palindromic sequence were observed. The comparison with other Lagomorphs suggests that this event happened before the *Ochotona* radiation. The dN/dS for the sequence region comprising the PPP1BM and the flanking palindrome highly supports the hypothesis that for *Ochotona* species this sequence has been evolving under positive selection. Besides this, mutational screening shows that the ability to bind to PPP1 is maintained in Tctex1d4 from Pika, although the PPP1BM being absent, and the N- and C-terminal surrounding residues being also abrogated. These observations converted Pika into an ideal model to study novel PPP1/PIPs regulatory mechanisms.

Introduction

Phosphoprotein phosphatase 1 (PPP1), one of the major eukaryotic serine/threonine protein phosphatase, has exquisite specificities *in vivo*, both in terms of substrates and cellular localization. Over the past two decades, it has become apparent that PPP1 versatility is achieved by its ability to interact with multiple targeting/regulatory subunits known as PPP1 Interacting Proteins (PIPs) [3, 4]. To date, more than 200 PIPs have been identified, most of them having the consensus PPP1 binding motif (PPP1BM), RVxF that binds to the catalytic subunit of PPP1, determining its targeting thus specifying cellular location and ultimately determining its function [5, 6]. Therefore, the key to characterizing the diverse roles of PPP1 relies on the identification of novel PIPs and in understanding the PPP1/PIP complex specific functions. Thus, several novel PIPs were identified, through a yeast two-hybrid system, using PPP1 as bait [2, 7-10].

A novel partner of PPP1 was thus identified, recently described as a novel Tctex1 dynein light chain family member, the t-complex testis expressed protein 1 domain containing 4, Tctex1d4 (Tctex2β) [2, 11]. Our results show that Tctex1d4 is evolutionarily conserved among mammals and ubiquitously expressed, particularly in ovary, spleen, lung and placenta, where PPP1 is also present [2, 12]. Moreover Tctex1d4 interacts directly with PPP1 catalytic subunit (PPP1CC) and possesses a canonical PPP1BM, commonly referred as RVxF [2, 13-16]. We have also demonstrated that Tctex1d4 and PPP1C co-localize in the microtubule organizing center and in microtubules having a probable role in the cytoplasmic transport of the cell [2]. Other groups have also shown that Tctex1d4 interacts with membrane receptors, inhibiting TGFβ signaling [11] and suggested its involvement in peripheral inflammation [17].

Cytoplasmic dyneins are responsible for the retrograde transport, the minus-end directed trafficking in the cytoskeletal microtubules [18]. More specifically, light chains might confer specificity to the cargo binding [19, 20], regulate other molecules [21] or stabilize the assembly of the motor dynein complex [22].

The RVxF motif is present in about 70% of all PIPs [5]. This motif is usually surrounded by basic residues in the N-terminal and by acidic residues in the C-terminal. The binding of this motif to a hydrophobic groove in PPP1C does not alter PPP1C conformation, but anchors the PIP to PPP1C [13, 14, 23, 24]. Nevertheless, the initial binding of this motif to PPP1C is essential to bring PIPs into PPP1C proximity, allowing for secondary interactions that strength holoenzyme binding, and determining substrate specificity, enzyme activity and PPP1 isoform selectivity [25].

It is expected that proteins with relevant biological functions, like dyneins, will be present in all mammals and probably evolved under different pressure levels. Following an extensive search in NCBI and ENSEMBL databases, the Tctex1d4 PPP1 binding motif, $_{90}$RVSF$_{93}$ (amino acid 90 to 93, according to *Homo sapiens* sequence), was shown to be present in all mammals except in the Lagomorpha species, *Ochotona princeps* (subgenus Pika).

The order Lagomorpha is divided into the families Ochotonidae (pikas) and Leporidae (rabbits and hares). Ochotonidae has a single genus (*Ochotona* sp.), which is divided in three subgenera (Conothoa, Ochotona and Pika) [26]. The leporids, encompassing eleven genera, which include the *Oryctolagus*, *Lepus* and *Sylvilagus* [27, 28]. In spite of the studies performed in the past few years, using fossil and molecular data, the divergence time between these two families remains vague. According to Matthee et al. [29], the leporid-ochotonid split was around 31 my ago. On the other hand, McKenna and Bell [30] and Asher et al. [31] suggested that the families' separation occurred around 37 my ago. Three other authors [32-34] suggested 65 my ago as the leporid-ochotonid divergence time.

These different molecular dating models of leporid-ochotonid separation were used by Lanier and Olson [35] to infer a common ancestor for pikas, however the *Ochotona* genus taxonomy still is one of the poorly resolved [35-38]. Lanier and Olson [35] suggested radiation time estimations for the *Ochotona* subgenera Pika (between 6 and 13 my ago), Conothoa (between 7 and 16 my ago) and Ochotona (between 10 and 20 my ago).

In this work we compared different lagomorphs, namely *Ochotona* species belonging to the subgenus Ochotona, in order to validate the observation that Tctex1d4 PPP1BM is absent in *O. princeps*. Also different mutants mimicking Pika PPP1BM and surrounding amino acids were produced and the binding efficiency was determined by the overlay technique. Further, these findings were applied to understand the evolutionary mechanisms that are behind these dramatic amino acid changes.

Materials and Methods

Analyses of Tctex1d4 evolution

Twenty-one different mammal protein sequences from Tctex1d4 were retrieved from NCBI GenBank (http://www.ncbi.nlm.nih.gov/) and from Ensembl (http://www.ensembl.org/index.html/). Table III.B. 1 contains the denominations, GenBank accession numbers and Ensembl scaffolds for all the acquired sequences. Additionally, two other Tctex1d4 protein sequences from *Ochotona* species, three sequences from *Lepus* species and two from *Sylvilagus* species were also included in this study.

Tissue samples from *Ochotona dauurica* (Ocda) and *Ochotona pusilla* (Ocpu), belonging to the subgenus Ochotona, were used. These samples were provided by the Zoological Museum of Moscow State University, Russia. Genomic DNA was extracted using the E.Z.N.A. ® Tissue DNA Kit (Omega Bio-Tek, Norcross, Georgia, USA) according to manufactures' instructions. A pair of primers were designed according to the available sequence for *Ochotona princeps* (subgenus Pika) available in Ensembl (forward 5'-ATGGCTGGCAGGCCTCTGCC-3' and reverse 5'-CTCGCAGTAGAGCCCGTGGA-3') generating a PCR fragment of 657bp. A touchdown PCR was performed and the thermal profile used was the following: initial denaturation (95ºC for 15min.); 5 cycles of denaturation (95ºC for 30sec.), annealing (66ºC for 30sec., 1ºC decrease/cycle) and extension (72ºC for 45sec.); 30 cycles of denaturation (95 ºC for 30sec.), annealing (62ºC for 30sec.) and extension (72ºC for 45sec.); and a final extension (72ºC for 20min). Sequencing was performed on an ABI PRISM 310 Genetic Analyzer (PE Applied Biosystems), where the ABI PRISM BigDye Terminator Cycle sequencing protocols were followed. PCR products were sequenced in both directions.

Genomic DNA samples identified as Jackrabbit (*Lepus* genus; Jack) and Cottontail (*Sylvilagus* genus; Cott) were provided by the Department of Microbiology and Immunology of Loyola University Chicago, USA. *Lepus europaeus* (Leeu) and *Lepus granatensis* (Legr) tissue samples were supplied by CIBIO, Vairão, Portugal and *Sylvilagus bachmani* (Syba) tissue samples were provided by the Blue Oak Ranch Reserve of the University of California, USA. Total RNA isolation from tissues (guanidinium thiocyanate-phenol-chloroform extraction) and cDNA synthesis were performed. A set of primers was designed according to the available sequence for *Oryctolagus cuniculus* existent in Ensembl (forward 5' TGCCAGGAGGAGGAGACTG 3'

and reverse 5' CACGCTGCACACCAGCTTG 3') generating a PCR fragment of approximately 500bp. The PCR thermal profile used was the following: initial denaturation (95ºC for 3min.); 40 cycles of denaturation (95ºC for 45sec.), annealing (57ºC for 1min.) and extension (72ºC for 1min.) and a final extension (72ºC for 10min.).

The nucleotide sequences were translated and aligned using ClustalW [39] and adjusted by visual examination (data not shown). The sequences obtained in this work have been deposited into NCBI GenBank under the accession numbers xxxx,xxxx (7 sequences). Phylogenetic and molecular evolutionary analyses were conducted using the software package MEGA4.1 [40].

Table III.B. 1: *List of the mammalian species in which the coding sequence of Tctex1d4 was retrieved from NCBI or ENSEMBL and used in this study.Tctex1d4 mutant constructs.*

Species name	Common name	Database ID	Database Source
Bos taurus (Bota)	Cow	XM_594476	NCBI GenBank
Callithrix jacchus (Caja)	Marmoset	Chromosome 7: 79,571,655-79,572,320	Ensembl
Canis familiaris (Cafa)	Dog	Chromosome 15: 18,523,787-18,524,446	Ensembl
Cavia porcellus (Capo)	Guinea Pig	scaffold_165: 1,051,684-1,052,355	Ensembl
Dasypus novemcinctus (Dano)	Armadillo	GeneScaffold_5913: 7,015-7,671	Ensembl
Echinops telfairi (Ecte)	Lesser Hedgehog Tenrec	GeneScaffold_7046: 35,805-36,470	Ensembl
Felis catus (Feca)	Cat	GeneScaffold_2486: 79,592-80,251	Ensembl
Gorilla gorilla (Gogo)	Gorilla	Chromosome 1: 46,433,826-46,434,515	Ensembl
Homo sapiens (Hosa)	Human	NM_001013632	NCBI GenBank
Loxodonta africana (Loaf)	Elephant	SuperContig scaffold_34: 24,862,938-24,863,606	Ensembl
Macaca mulatta (Mamu)	Rhesus Monkey	XM_001099595	NCBI GenBank
Microcebus murinus (Mimu)	Mouse Lemur	GeneScaffold_1367: 59,896-60,555	Ensembl
Mus musculus (Mumu)	Mouse	NM_175030	NCBI GenBank
Ochotona princeps (Ocpr)	American Pika	GeneScaffold_4323: 138,892-139,548	Ensembl
Oryctolagus cuniculus (Orcu)	European Rabbit	scaffold_0: 101,766,508-101,767,167	Ensembl
Otolemur garnetti (Otga)	Bushbaby	GeneScaffold_2671: 80,377-81,036	Ensembl
Pan troglodytes (Patr)	Chimpanzee	Chromosome 1: 45,543,578-45,544,952	Ensembl
Pongo pygmaeus (Popy)	Orangutan	Chromosome 1: 185,125,020-185,125,685	Ensembl
Procavia capensis (Prca)	Hyrax	GeneScaffold_6128: 11,550-12,212	Ensembl
Rattus norvegicus (Rano)	Rat	XM_233427	NCBI GenBank
Sus scrofa (Susc)	Pig	NM_001032356	NCBI GenBank

Mutagenic primers were designed according to the sequence of human Tctex1d4 (NCBI: NM_001013632.2) and were used to obtain the desired mutations (Table III.B. 2). Starting with pET-Tctex1d4 plasmid as template, and along with appropriate mutagenic primers, the mutants HA+INL+WS, HA+WS, HA+INL, INL+WS, HA, INL and WS were created using the QuikChange® Site-Directed Mutagenesis Kit (Stratagene, Agilent Technologies UK Ltd, Edinburgh, UK). PCR conditions for site-directed mutagenesis were as followed: initial denaturation (95ºC for 1min.); 18 cycles of denaturation (95ºC for 30sec.), annealing (55ºC for 1min.) and extension (68ºC for 7min.), using KOD polymerase (Novagen, Madison, Wisconsin, USA). DNA was then digested by DpnI restriction enzyme and transformed into *E. coli* XL1-Blue strain (Stratagene Agilent Technologies UK Ltd, Edinburgh, UK). Sequencing was performed on an ABI PRISM 310 Genetic Analyzer (Perkin-Elmer, Applied Biosystems, Barcelona, Spain), where the ABI PRISM BigDye Terminator Cycle sequencing protocols were followed. Positive clones were sequenced in both directions using universal T7 promoter and T7 terminator primers.

Table III.B. 2: Human and Pika PPP1BM and surrounding sequences are shown on the top. *Mutation sites are highlighted in bold and underlined. On the bottom, oligonucleotides used for site direct mutagenesis are shown.*

Homo sapiens (Humans)		*Ochotona princeps* (Pika)
$_{85}$**PP**LGSR**VSF**SG**LP**$_{97}$		$_{83}$**HA**LGSR**INL**SG**WS**$_{95}$
Mutation	**Primer Name**	**Sequence**
$_{85}$PP$_{86}$ to $_{85}$HA$_{86}$	HA-FW	5’-GGGCCCGGTGCACGCTCTGGGCTCAAG-3’
	HA-RV	5’-CTTGAGCCCAGAGCGTGCACCGGGCCC-3’
$_{91}$VSF$_{93}$ to $_{91}$INL$_{93}$	INL-FW	5’-CTCTGGGCTCAAGGATCAACTTATCAGGGTTGCCCC-3’
	INL-RV	5’-GGGGCAACCCTGATAAGTTGATCCTTGAGCCCAGAG-3’
$_{97}$LP$_{98}$ to $_{97}$WS$_{98}$	WS-FW	5’-GCTTCTCAGGGTGGTCCCTGGCGCCCG-3’
	WS-RV	5’-CGGGCGCCAGGGACCACCCTGAGAAGC-3’
$_{97}$LP$_{98}$ to $_{97}$WS$_{98}$	L..WS-FW	5’-CAACTTATCAGGGTGGTCCCTGGCGCCCGCC-3’
	L..WS-RV	5’-GGCGGGCGCCAGGGACCACCCTGATAAGTTG-3’

Protein expression and Overlay assay

Each His-tagged mutant was transformed into *E. coli* Rosetta strain (Novagen, Madison, Wisconsin, USA). A single colony was selected and grown overnight at 37ºC in the appropriate media until the optic density of 0.6-0.7 was reached. Expression was induced using 1M IPTG (isopropyl-β-D-thio-galactopyranoside), at 37ºC with shaking, for 3hrs. Culture cells were recovered by centrifugation and treated as described in (Browne et al., 2007). Lysates were then mass normalized using a BCA® assay (Fisher Scientific, Loures, Portugal) and 10µg of each extract was loaded in a 12% SDS-PAGE gel. The proteins were subsequently transferred to a nitrocellulose membrane and then overlaid with 25pmol/mL of purified PPP1CC1 for one 1hr. Membranes were incubated with either mouse anti-His monoclonal (1:1000, Novagen, Madison, Wisconsin, USA) or rabbit CBC3C (anti-PPP1CC, 1:1000) antibodies, followed by the respective anti-mouse and anti-rabbit infrared secondary antibodies (1:5000, Li-Cor Biosciences UK Ltd, Cambridge, UK) and bands developed in Odyssey infrared-imaging system and quantified using Odyssey v1.2 software (Li-Cor Biosciences UK Ltd, Cambridge, UK). The same procedure was also performed for pET-Tctex1d4 (positive control) and pET vector (negative control).

Statistical analysis

SigmaPlot statistical package (SigmaPlot v.11, Systat Software Inc.) was used for statistical analysis. Data were tested for normal distribution and homogeneity of variances. Student's *t*-test ($p<0.05$, alpha=0.050) was used to detect the differences between each mutation by comparison to the control, pET-Tctex1d4.

Results

Analyses of Tctex1d4 evolution

When comparing the Tctex1d4 PPP1BM, $_{90}$RVSF$_{93}$, in twenty-one mammalian species we observed that it was present in all except in *Ochotona princeps* creating an extra glycosylation site (motif $_{90}$NLS$_{92}$). To confirm if this was not an artifact of the database, *Ochotona dauurica* and *Ochotona pusilla* Tctex1d4 were sequenced. For other five lagomorphs (*Lepus* and *Sylvilagus* genera), the Tctex1d4 coding region was partially sequenced. These sequences were confirmed through BLAST, compared with other mammalian sequences and translated into amino acids (Table III.B. 3). The nucleotide substitution in the *Ochotona* species generated amino acid changes that confirmed the elimination of the consensus PPP1BM, $_{90}$RVSF$_{93}$, and the appearance of a glycosylation site. *Lepus* and *Sylvilagus* genera maintained the canonical PPP1BM.

The alignment between the sequences acquired in this work and the twenty-one sequences available for the different mammals (data not shown) allowed the construction of a Neighbor-Joining phylogenetic tree (Figure III.B. 1a). The topology obtained was in accordance with the mammalian taxonomy proposed and actually accepted [41], suggesting that Tctex1d4 has been evolving under neutral selection. A new Neighbor-Joining tree using amino acid p-distance (Figure III.B. 1b) was constructed using only the twelve amino acids, four upstream and four downstream of the motif $_{90}$RVSF$_{93}$. This choice of amino acids was related with PPP1BM being flanked by an unusual palindromic sequence, $_{86}$PLGS$_{89}$, according to *Homo sapiens* sequence alignment with other mammals. As expected, the obtained tree revealed that the three *Ochotona* species formed an independent cluster, highly supported by a bootstrap value of 99 (see Figure III.B. 1b).

Table III.B. 3: Amino acids and nucleotides sequences, corresponding to the 12 amino acids region (numbered according to human Tctex1d4 CDS), for the 28 mammals used in this study. *Shadowed region: RVSF motif; Underlined region: novel glycosylation site in Ochotona species; Bolded region: non-synonymous substitutions; Italic region: nucleotide sequence corresponding to the novel glycosylation site in Ochotona species. CDS, codifying sequence*

	222	---	222	222	222	222	222	222	222	222	222	222	222
8-88899999999	555	---	566	666	666	667	777	777	777	888	888	888	899
6-78901234567	678	---	901	234	567	890	123	456	789	012	345	678	901
Hosa P-LGSRVSFSGLP	CCT	---	CTG	GGC	TCA	AGG	GTC	AGC	TTC	TCA	GGG	TTG	CCC
Caja .-...Q.....	...	---	...	...	...	**CA**.	...	...	...	..G	...	...	...
Gogo .-.....	...	---	...	...	...	...	...	...	...	...	...	...	...
Mamu .-.....	...	---	...	...	...	C..	...	...	...	...	...	...	..T
Mimu .-.A.....	...	---	..A	**.C.**	..G	C..	...	..T	...	..T	...	...	...
Otga L-.A.....	**.T.**	---	...	**.C.**	..G	..A	...	..T	...	..T	...	...	...
Patr .-.....	...	---	...	...	...	...	...	...	...	...	...	...	...
Popy .-.....	...	---	...	...	...	C..	...	...	...	...	...	...	...
Cafa .-.....	..C	---	..A	...	..G	...	..T	...	...	..G	...	...	...
Feca L-.....	**.T.**	---	T..	...	...	...	..T	...	...	..T	...	...	...
Dano .-....I.....	...	---	...	..A	...	...	**A**..	...	...	..G	...	...	...
Ecte .P.....	...	**CCT**	..T	...	...	C.A	...	...	...	..G	...	...	...
Loaf .-.S.....	...	---	..A	**A.T**	..C	C..	...	...	...	...	...	C..	...
Capo .-.A..I.....	...	---	..A	**.CT**	..T	...	**A**..	..T	...	...	...	...	...
Mumu .-.....	...	---	..A	..T	...	C.A	..T	..T	..T	..T	...	C..	...
Rano .-.....	...	---	..A	..T	...	C.A	..T	..T	...	..T	...	C.A	...
Prca A-.....S.	**G**.C	---	...	..T	..C	C..	...	..T	...	...	...	.C.	...
Ocda A-....LNL..WS	**G**..	---	...	...	...	C..	**C.T**	**.A.**	**..A**	*..C*	...	**.G.**	**T.T**
Ocpr A-....INL..WS	**G**..	---	...	...	...	C..	**A.T**	***.AT***	**..A**	*..C*	..A	**.G.**	**T.T**
Ocpu A-....INL..WS	**G**..	---	...	...	...	C..	**A.T**	**.A.**	**..A**	*..T*	..A	**.G.**	**T.T**
Jack .-.....G.....	...	---	..C	...	...	C..	..T	**G**..	...	..T	...	...	..T
Leeu .-.....G.....	...	---	..C	...	...	C..	..T	**G**..	...	..T	...	...	..T
Legr .-.....G.....	...	---	..C	...	...	C..	..T	**G**..	...	..T	...	...	..T
Orcu .-.....G.....	...	---	..C	...	...	C..	..T	**G**..	...	..C	...	...	..T
Cott .-.....G.....	...	---	..C	...	...	C..	..T	**G**..	...	..C	...	...	..T
Syba .-.....G.....	...	---	..C	...	...	C..	..T	**G**..	...	..C	...	...	..T
Bota .-.....	...	---	..A	...	..G	C.A	...	...	...	..T	..C	...	...
Susc .-..W.....	..G	---	...	...	**.GG**	C..	...	...	...	..G	..C	C..	...

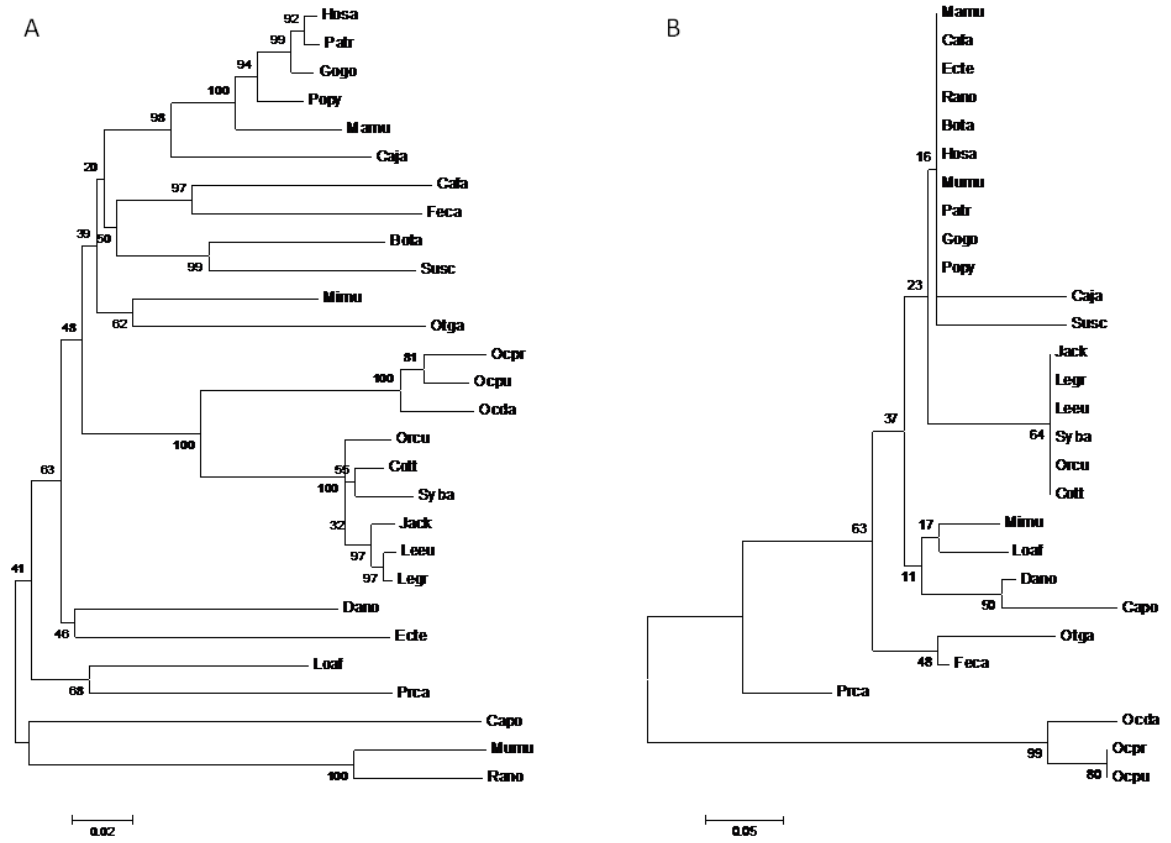


Figure III.B. 1: *Neighbor-Joining tree* **A.** *resulting from the alignment of the 28 mammalians species coding region of Tctex1d4, and* **B.** *using only 12 amino acids, four upstream and four downstream of the motif* $_{90}$*RVSF*$_{93}$*.*

Tctex1d4 RVSF-palindrome studies

To further study the significance of the Bioinformatic studies, mutants based on the Pika sequence ($_{83}$*HALGS*RINL*SGWS*$_{95}$) corresponding to the human Tctex1d4 PPP1BM and flanking regions ($_{85}$*PPLGS*RVSF*SGLP*$_{97}$) were generated by directed mutagenesis followed by bacterial expression of those mutants and PPP1C binding screening by overlay (Fig. 2). The band intensities indicate the amount of PPP1CC that is bound to the bacterial expressed Tctex1d4 recombinant mutant proteins. Since all proteins were expressed with a N-terminal His-tag, anti-His antibody was used to normalize the amount of recombinant protein loaded in each lane. Subsequently band intensities were normalized according to the amount of recombinant protein loaded and compared to the pET-Tctex1d4 control. Negative controls include Rosetta cell extract and Rosetta cell extract expressing pET vector alone. Controls have not shown any binding to PPP1CC.

Results show that HA+INL+WS mutant has a binding profile similar to the wild type human Tctex1d4 since there was no statistical difference in the binding capacity. The

results of the other mutations, single or double, also show no statistical difference when comparing to the control, pET-Tctex1d4 (Figure III.B. 2).

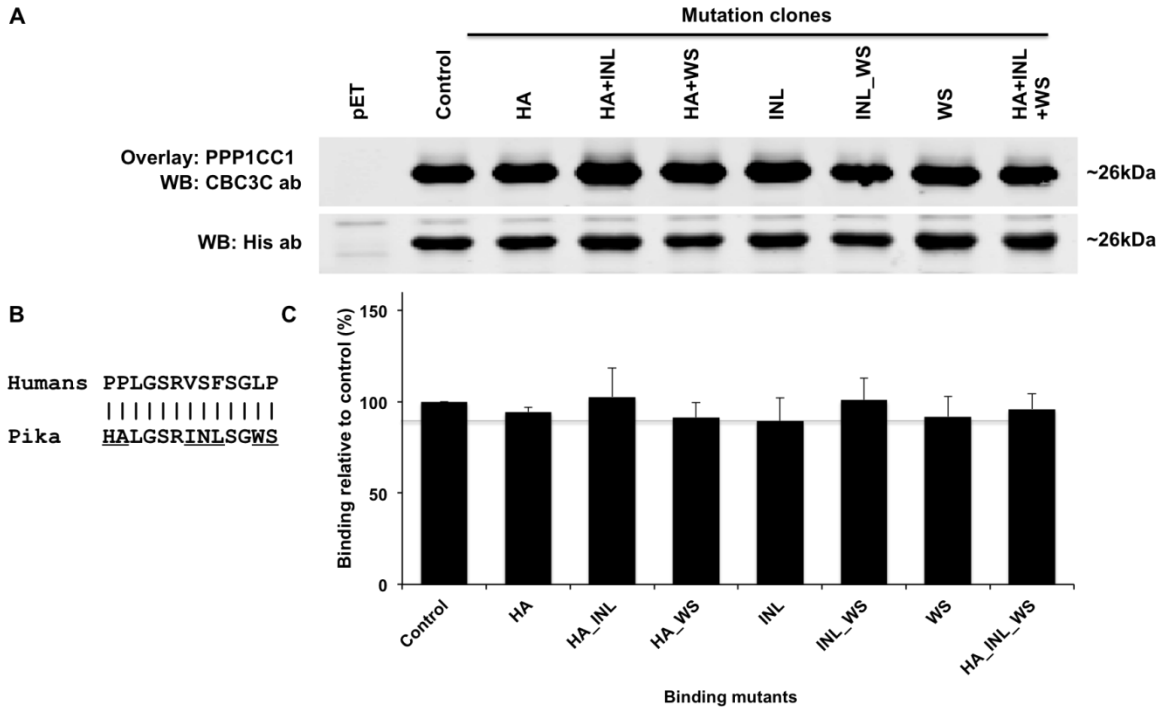


Figure III.B. 2: Tctex1d4 PPP1BM mutants and PPP1C binding analysis. A. *Bacterial cell cultures expressing each construct were loaded in a SDS-PAGE gel (10µg). Membranes were then overlaid with purified PPP1CC and detected using CBC3C antibody. Total amount of recombinant protein was accessed using an anti-His antibody.* **B.** *Alignment of Human and Pika Tctex1d4 PPP1BM and surrounding palindromic sequences. Mutations are underlined* **C.** *Subsequent analysis of band intensities versus total amount of recombinant protein was performed and the results plotted in a graph by comparison with the pET-Tctex1d4 control. Results show that Pika Tctex1d4 aberrant RVxF motif, RINL, and respective non-palindromic surrounding region were sufficient for the PPP1CC binding. Negative controls include Rosetta cell extract and Rosetta cell extract expressing pET vector alone. Error bars represent the standard error of the mean of triplicates. WB, Western blot.*

Discussion

The identification of Tctex1d4 as a new PIP has already been described [2]. This new interaction was supported by the yeast two-hybrid approach, co-immunoprecipitation and overlay techniques. Results from our laboratory showed that the N-terminal domain, where the PPP1BM is present, is essential for the binding. Furthermore, *in vitro* studies with Tctex1d4 PPP1 binding mutants strengths the importance of the PPP1BM to Tctex1d4/PPP1C interaction. Indeed, the mutation of the motif RVSF to AAAA decreases the binding by 35%. This was not in accordance with our expectations, since the mutation of the PPP1BM either to AAxA [42], RAxA [14, 43] or to RVxA [44] usually abrogates PPP1/PIP interaction. Nevertheless, some cases exist where interaction still occurs but at less extent [45-48]. Also, there are some PIPs that still bind PPP1C in the presence of an excess of a synthetic RVxF peptide [47], that usually disrupts the PPP1/PIP complex [9, 49]. The explanation might be that other motifs besides RVxF are present in these proteins and are also important for the binding.

Tctex1d4 possesses an unusual bizarre sequence surrounding the RVSF motif. The twelve amino acid sequences contain a palindrome – *PLGS*RVSF*SGLP*. The PPP1BM binds to PPP1 in a hydrophobic pocket [50]. This palindrome may form a structured arm forcing the RVSF motif, even when it is mutated into AAAA, to enter the PPP1 pocket, since the palindrome contains several rigid prolines. Perhaps, if the RVSF is completely removed, and the arm destroyed, Tctex1d4 will no longer bind PPP1C.

When the twelve amino acids Neighbor-Joining tree was constructed the three *Ochotona* species formed an independent cluster and this observation raised two explanatory hypotheses: or the new motif present in the *Ochotona* species resulted from gene conversion with adjacent genes or a pattern of nucleotide substitution in this specific motif happened. Furthermore, the palindrome is also highly conserved among mammals but is completely lost in Pika sequences (Figure III.B. 1 and Table III.B. 2).

Gene conversion has been reported in other mammalian genes. For example, in leporids a gene conversion event was observed between the two chromosomally adjacent genes CCR2 and CCR5, where the sequence motif $_{194}$QTLKMT$_{199}$ of the CCR5 protein was replaced by the HTIMRN motif, which is characteristic of CCR2 [51, 52]. In the present study, none of the genes chromosomally adjacent showed a clear evidence of gene conversion with Tctex1d4, being this event an unlikely hypothesis. Furthermore, no significant BLAST was obtained when compared with mammalian NCBI database.

Under neutrality, the expected ratio of non-synonymous (dN) to synonymous (dS) substitutions in a gene is one and significant deviations from this value can be interpreted

as evidence of either positive selection (ω>1) or purifying selection (ω<1). To consider a specific pattern of nucleotide substitution, synonymous and non-synonymous substitution rates were estimated using the Nei-Gojobori method [53] and a non-synonymous to synonymous substitution ratio was calculated for the previously referred twelve amino acidic positions. Comparing dN/dS between all analysed mammals, but excluding the three *Ochotona* species, the presented values were on average lower than 0.3, suggesting a strong purifying selection. However, when comparing dN/dS between the three *Ochotona* species and each of the mammalian sequences, on average, the obtained value was 1.6, suggesting that for *Ochotona* species this fragment lost the mutations constrains imposed by purifying selection or/and evolved under Darwinian or positive selection. When focusing the analysis on the Superorder Glires (Order Rodentia and Order Lagomorpha), the main representatives of rodents (mouse and rat) showed a ratio of zero, meaning that Tctex1d4 was under purifying selection for this group. While comparing the rodents sequences nucleotide corresponding to the twelve amino acids region with the one from human, a total of ten substitutions caused no amino acid changes (Table III.B. 2). On the other hand, for the *Ochotona* species, a total of twelve substitutions caused six amino acid alterations. Furthermore, when comparing all species from the three Lagomorpha genera, *Lepus*, *Oryctolagus* and *Sylvilagus*, with the three *Ochotona* species, the dN/dS ratio ranged between 1.7 and 7.0. This obtained dN/dS ratio, clearly higher than 1, and the fact that amino acid alterations created a new putative glycosylation site, highly support the hypothesis that for *Ochotona* sp. this sequence fragment has been evolving under positive selection. The occurrence of this nucleotide pattern in the three *Ochotona* species studied in this work and its absence in the other lagomorphs, suggests that this evolutionary event happened before the radiation of the *Ochotona* genus (between 6 and 20 Million years) [35] and after the split of Ochotonidae and Leporidae families (between 31 and 65 Million years) [32-34].

The creation of a novel putative glycosylation site in the three *Ochotona* species by positive selection clearly indicates a physiological important function. The amino acidic sequence, $_{96}$NLS$_{98}$, suggests a N-linked glycosylation attached to the nitrogen of the asparagine [54, 55]. The probability of Pika Tctex1d4 being glycosylated is increased by the fact of this motif to be located more than sixty amino acids upstream of the C-terminal [56]. The remaining unsolved question is the acquired function of Tctex1d4 in *Ochotona* sp. This new putative glycosylation site in the protein may increase the half-life of the protein, which in turn will stay longer in the membrane attached to endoglin [11] being a stronger inhibitor of TGFβ in *Ochotona* sp. than in other mammals. Also, Tctex1d4 in

Ochotona sp. lost the PPP1BM. Further, it also lost the palindromic sequence, PLGS, probably important for the binding of Tctex1d4 to PPP1. Thus, it would be expected that it would no longer bind to PPP1 directly. Evolutionarily, it is not clear what happened first, the lost of the palindrome with subsequent mutation of the PPP1BM to a glycosylation site or the acquisition of a glycosylation site by positive selection followed by loss of the palindrome.

These bioinformatic results suggest that an alternative mechanism may exist in *Ochotona* sp. for Tctex1d4 binding to PPP1. Thus, we employed an overlay screening with different binding mutants to test this hypothesis. Results show that Pika Tctex1d4 aberrant RVxF motif and respective non-palindromic surrounding region, $_{83}$*HALGS*RINL*SGWS*$_{95}$, mutant were sufficient to sustain the binding of the Tctex1d4 mutant to PPP1CC, at the same levels by comparison to the wild type human Tctex1d4 (Figure III.B. 2). Moreover, in single and double mutants, the binding capacity was also maintained, which clearly shows that although substantial differences were found in Pika RVxF and surrounding regions, these differences alone and together do not contribute to a binding disruption. Our earlier results have shown that a mutation of the RVSF motif to AAAA it only decreases the overall binding efficiency in 35% [2]. Furthermore, we have also shown that important regions for this binding are concentrated in the N-terminal, where the RVxF is also present. This means that, or the RVxF motif is not the only point of contact, or the RVxF surrounding region is also important for this binding. Here, using Pika aberrant motif we clearly show that the second hypothesis does not explain why the binding is not abolished when we mutate the RVxF motif. However, we cannot exclude the possibility that both surrounding regions, the Pika and the rest of the mammals, support the binding and are important for it.

PPP1BM RVxF motif is usually surrounded by basic residues (arginine, lysine and histidine) in the N-terminal and by acidic residues (aspartate and glutamate) in the C-terminal [14, 16]. Analysis of 143 RVxF motifs in known and novel PIPs revealed that five to six of these flanking basic and acidic residues are relatively common among PIPs [5]. Human Tctex1d4 RVSF motif is a strong motif according to this analysis but the palindromic region that surrounds it does not follow this pattern, since no basic or acidic amino acids are present. Even so, all the flaking residues are present at some extent in other PIPs. By comparing the above results with ours, the PP to HA mutation would not lead to any difference because some PIPs also have these amino acids in these positions (P 11%, P 4% comparing to H 4%, A 10%). In what concerns to the VSF to INL mutation (V 94%, S 21%, F 83% comparing to I 6%, N 5% and L 0%) we can infer that the binding

would be potentially abrogated, but our results show that it is maintained. Finally, relatively to the LP to WS mutation (L 3%, P 7% comparing to W 0%, S 7%) the outcome is quit unknown but our results show that this mutation does not alter the binding. So, we have shown that the strange palindromic sequence, evolutionarily conserved, appears to be irrelevant for the binding, since the HA and WS mutations, single or double, resulted in the same binding capacity, and that the strange RVxF motif, RINL, seems to sustain the binding. This is surprising because looking evolutionarily one might say that in Pika Tcted1d4 had lost the ability to bind to PPP1C, but our results show undoubtedly that even with the motif and flanking regions evolving under positive selection, both regions seem to still sustain the binding. The hypothesis of another N-terminal region being important for the binding arises and might explain why the RVxF motif seems to be another point of contact that helps to stabilize the complex.

In conclusion, Tctex1d4 evolutionary analysis revealed that in Pika the PPP1BM was lost and replaced by a new putative glycosylation site. Furthermore, we also observed, in *Ochotona* sp., the lost of a highly conserved palindrome present among mammals. Also, the presence of the HA, INL and WS substitutions in *Ochotona* sp, does not alter the binding capacity. The combination of these factors in Pika species makes it a perfect model to study the biology of PPP1/Tctex1d4 complex and can be expanded to understand PPP1/PIPs interactions, increasing the number of PIPs previously expected to exist based on the consensus RVxF motif.

Acknowledgements

FCT supported the doctoral fellowship of Ana Matos (SFRH/BD/48566/2008), Luis Korrodi-Gregório (SFRH/BD/41751/2007) and Sara L. C. Esteves (SFRH/BD/42334/2007) and post-doctoral fellowship of Pedro Esteves (SPRH/BPD/27021/2006). We would like to thank Jeff Wilcox and Dr. Michael Hamilton from Blue Oak Ranch Reserve, University of California, USA, for providing *Sylvilagus bachmani* tissue samples. Also, we would like to thank Dr. Dennis Lanning from the Department of Microbiology and Immunology of Loyola University Chicago, USA, for supplying Jackrabbit and Cottontail genomic DNA samples.

References

1. Fardilha, M., et al., *Identification of the human testis protein phosphatase 1 interactome.* Biochemical Pharmacology, 2011. **82**(10): p. 1403-1415.
2. Korrodi-Gregório, L., et al., *Characterization of a novel Protein Phosphatase 1 interacting Protein, Tctex1d4, in male gonads and sperm, involved in tubulin dynamics.* 2011.
3. Cohen, P.T., *Protein phosphatase 1--targeted in many directions.* J Cell Sci, 2002. **115**(Pt 2): p. 241-56.
4. Virshup, D.M. and S. Shenolikar, *From promiscuity to precision: protein phosphatases get a makeover.* Mol Cell, 2009. **33**(5): p. 537-45.
5. Hendrickx, A., et al., *Docking Motif-Guided Mapping of the Interactome of Protein Phosphatase-1.* Chemistry & biology, 2009. **16**(4): p. 365-371.
6. Fardilha, M., et al., *The Physiological Relevance of Protein Phosphatase 1 and its Interacting Proteins to Health and Disease.* Current Medicinal Chemistry, 2010. **17**(33).
7. Fardilha, M., *Characterization of PP1 Interactome from human testis*, 2004, Universidade de Aveiro.
8. Fardilha, M., et al., *Alternatively spliced protein variants as potential therapeutic targets for male infertility and contraception.* Ann N Y Acad Sci, 2004. **1030**: p. 468-78.
9. Browne, G.J., et al., *SARP, a new alternatively spliced protein phosphatase 1 and DNA interacting protein.* Biochem J, 2007. **402**(1): p. 187-96.
10. Wu, W., et al., *Alternative splicing controls nuclear translocation of the cell cycle-regulated Nek2 kinase.* J Biol Chem, 2007. **282**(36): p. 26431-40.
11. Meng, Q., et al., *Identification of Tctex2beta, a novel dynein light chain family member that interacts with different transforming growth factor-beta receptors.* J Biol Chem, 2006. **281**(48): p. 37069-80.
12. da Cruz e Silva, E.F., et al., *Differential expression of protein phosphatase 1 isoforms in mammalian brain.* J Neurosci, 1995. **15**(5 Pt 1): p. 3375-89.
13. Egloff, M.P., et al., *Structural basis for the recognition of regulatory subunits by the catalytic subunit of protein phosphatase 1.* EMBO Journal, 1997. **16**(8): p. 1876-87.
14. Wakula, P., et al., *Degeneracy and Function of the Ubiquitous RVXF Motif That Mediates Binding to Protein Phosphatase-1.* Journal of Biological Chemistry, 2003. **278**(21): p. 18817-18823.
15. Ceulemans, H. and M. Bollen, *Functional diversity of protein phosphatase-1, a cellular economizer and reset button.* Physiol Rev, 2004. **84**(1): p. 1-39.
16. Meiselbach, H., H. Sticht, and R. Enz, *Structural Analysis of the Protein Phosphatase 1 Docking Motif: Molecular Description of Binding Specificities Identifies Interacting Proteins.* Chemistry & Biology, 2006. **13**(1): p. 49-59.
17. Marques, F., et al., *The choroid plexus response to a repeated peripheral inflammatory stimulus.* BMC Neurosci, 2009. **10**: p. 135.
18. Vallee, R.B., et al., *Dynein: An ancient motor protein involved in multiple modes of transport.* J Neurobiol, 2004. **58**(2): p. 189-200.
19. Tai, A.W., et al., *Rhodopsin's carboxy-terminal cytoplasmic tail acts as a membrane receptor for cytoplasmic dynein by binding to the dynein light chain Tctex-1.* Cell, 1999. **97**(7): p. 877-87.
20. Lo, K.W., J.M. Kogoy, and K.K. Pfister, *The DYNLT3 light chain directly links cytoplasmic dynein to a spindle checkpoint protein, Bub3.* J Biol Chem, 2007. **282**(15): p. 11205-12.
21. Benashski, S.E., et al., *Dimerization of the highly conserved light chain shared by dynein and myosin V.* J Biol Chem, 1997. **272**(33): p. 20929-35.

22. DiBella, L.M., et al., *Differential light chain assembly influences outer arm dynein motor function.* Mol Biol Cell, 2005. **16**(12): p. 5661-74.
23. Hurley, T.D., et al., *Structural basis for regulation of protein phosphatase 1 by inhibitor-2.* J Biol Chem, 2007. **282**(39): p. 28874-83.
24. Terrak, M., et al., *Structural basis of protein phosphatase 1 regulation.* Nature, 2004. **429**(6993): p. 780-4.
25. Bollen, M., et al., *The extended PP1 toolkit: designed to create specificity.* Trends in biochemical sciences, 2010. **35**(8): p. 450-458.
26. Hoffmann, R.S. and A.T. Smith, *Order Lagomorpha*, in *Mammal Species of the World: a taxonomic and geographic reference*, D.E. Wilson and D.M. Reeder, Editors. 2005, The Johns Hopkins University Press: Baltimore, Maryland. p. 185-201.
27. Dawson, M., *Evolution of modern leporids*, in *Proceedings of the world lagomorph conference*, K. Myers and C. MacInnes, Editors. 1981, University of Guelph Press: Ontario.
28. Corbet, G., *A review of classification in the family Leporidae.* Acta Zoolog Fennica, 1983(174): p. 1-15.
29. Matthee, C.A., et al., *A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene.* Syst Biol, 2004. **53**(3): p. 433-47.
30. Mckenna, M.C. and S.K. Bell, *Classification of mammals above the species level*1997, New York: Columbia University Press.
31. Asher, R.J., et al., *Stem Lagomorpha and the antiquity of Glires.* Science, 2005. **307**(5712): p. 1091-4.
32. Springer, M.S., et al., *Placental mammal diversification and the Cretaceous-Tertiary boundary.* Proc Natl Acad Sci U S A, 2003. **100**(3): p. 1056-61.
33. Bininda-Emonds, O.R., et al., *The delayed rise of present-day mammals.* Nature, 2007. **446**(7135): p. 507-12.
34. Rose, K.D., et al., *Early Eocene lagomorph (Mammalia) from Western India and the early diversification of Lagomorpha.* Proc Biol Sci, 2008. **275**(1639): p. 1203-8.
35. Lanier, H.C. and L.E. Olson, *Inferring divergence times within pikas (Ochotona spp.) using mtDNA and relaxed molecular dating techniques.* Mol Phylogenet Evol, 2009. **53**(1): p. 1-12.
36. Yu, N., et al., *Molecular systematics of pikas (genus Ochotona) inferred from mitochondrial DNA sequences.* Mol Phylogenet Evol, 2000. **16**(1): p. 85-95.
37. Niu, Y., et al., *Phylogeny of pikas (Lagomorpha, Ochotona) inferred from mitochondrial cytochrome b sequences.* Folia Zoologica, 2004. **53**(2): p. 141-156.
38. Lissovsky, A.A., N.V. Ivanova, and A.V. Borisenko, *Molecular Phylogenetics and Taxonomy of the =Subgenus pika (Ochotona, Lagomorpha).* Journal of Mammalogy, 2007. **88**(5): p. 1195-1204.
39. Thompson, J.D., D.G. Higgins, and T.J. Gibson, *CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.* Nucleic Acids Res, 1994. **22**(22): p. 4673-80.
40. Tamura, K., et al., *MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0.* Mol Biol Evol, 2007. **24**(8): p. 1596-9.
41. Springer, M.S., et al., *The adequacy of morphology for reconstructing the early history of placental mammals.* Syst Biol, 2007. **56**(4): p. 673-84.
42. Liu, D., et al., *Regulated targeting of protein phosphatase 1 to the outer kinetochore by KNL1 opposes Aurora B kinase.* J Cell Biol, 2010. **188**(6): p. 809-20.

43. Novoyatleva, T., et al., *Protein phosphatase 1 binds to the RNA recognition motif of several splicing factors and regulates alternative pre-mRNA processing.* Human Molecular Genetics, 2008. **17**(1): p. 52-70.
44. Takemiya, A., C. Ariyoshi, and K.-i. Shimazaki, *Identification and Functional Characterization of Inhibitor-3, a Regulatory Subunit of Protein Phosphatase 1 in Plants.* Plant Physiology, 2009. **150**(1): p. 144-156.
45. Chang, J.S., et al., *Protein Phosphatase-1 Binding to Scd5p Is Important for Regulation of Actin Organization and Endocytosis in Yeast.* Journal of Biological Chemistry, 2002. **277**(50): p. 48002-48008.
46. Traweger, A., et al., *Protein phosphatase 1 regulates the phosphorylation state of the polarity scaffold Par-3.* Proceedings of the National Academy of Sciences, 2008. **105**(30): p. 10402-10407.
47. Bollen, M., *Combinatorial control of protein phosphatase-1.* Trends in Biochemical Sciences, 2001. **26**(7): p. 426-31.
48. Wakula, P., et al., *The translation initiation factor eIF2beta is an interactor of protein phosphatase-1.* Biochemical Journal, 2006. **400**(2): p. 377-83.
49. Cohen, P.T., et al., *Assay of protein phosphatase 1 complexes.* Methods in Enzymology, 2003. **366**: p. 135-44.
50. Gibbons, J.A., D.C. Weiser, and S. Shenolikar, *Importance of a surface hydrophobic pocket on protein phosphatase-1 catalytic subunit in recognizing cellular regulators.* Journal of Biological Chemistry, 2005. **280**(16): p. 15903-11.
51. Carmo, C.R., et al., *Genetic variation at chemokine receptor CCR5 in leporids: alteration at the 2nd extracellular domain by gene conversion with CCR2 in Oryctolagus, but not in Sylvilagus and Lepus species.* Immunogenetics, 2006. **58**(5-6): p. 494-501.
52. Abrantes, J., et al., *A shared unusual genetic change at the chemokine receptor type 5 between Oryctolagus, Bunolagus and Pentalagus.* Conservation Genetics, 2009: p. 1-6.
53. Nei, M. and T. Gojobori, *Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions.* Molecular Biology and Evolution, 1986. **3**(5): p. 418-26.
54. Spiro, R.G., *Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds.* Glycobiology, 2002. **12**(4): p. 43R-56R.
55. Vagin, O., J.A. Kraut, and G. Sachs, *Role of N-glycosylation in trafficking of apical membrane proteins in epithelia.* Am J Physiol Renal Physiol, 2009. **296**(3): p. F459-69.
56. Nilsson, I. and G. von Heijne, *Glycosylation efficiency of Asn-Xaa-Thr sequons depends both on the distance from the C terminus and on the presence of a downstream transmembrane segment.* J Biol Chem, 2000. **275**(23): p. 17338-43.

Chapter V

Conclusion

Chapter IV.

Conclusion

The main objective of this thesis was to study novel PPP1 interacting proteins (PIPs) in testis and sperm, in order to further characterize the complexes these proteins make with PPP1 in mammalian reproduction.

To fulfill this aim we addressed the presence, localization and putative roles of PPP1R2, a previously well known PIP, in testis and sperm, and two novel PPP1CC2 testis/sperm-specific PIPs, PPP1R2P3 and Tctex1d4, that were first identified in a Yeast Two Hybrid (YTH) screen using a human testis cDNA library [1].

In Chapter II.A we showed for the first time and unequivocally the presence of PPP1R2 and PPP1R2P3 in human ejaculated sperm by mass spectrometry. PPP1R2P3 is a heat stable protein that binds to PPP1CC *in vitro*. Moreover, PPP1R2P3 is an inhibitor of PPP1CC that cannot be phosphorylated by GSK3. Also, we showed that PPP1R2P3 was phosphorylated by CK2 *in vitro*, probably in residues 121 and 122, as shown by mass spectrometry. Finally, PPP1R2 and/or PPP1R2P3 are serine phosphorylated, *in vivo*, in human sperm. Co-localization of PPP1R2/PPP1R2P3-PPP1CC2 holoenzymes in the head and tail of sperm and in both soluble and insoluble fractions was demonstrated by immunocytochemistry and Western blot. Our current hypothesis is that the holoenzymes localized in the head may have an important role in the acrosome reaction while the axoneme bound holoenzymes are clearly important for the control of flagellar motility. This results help to unravel one more piece of the epididymal sperm motility initiation puzzle involving the PPP1R2/PPP1R2P3-PPP1CC2 complexes.

To further address the PPP1R2 significance, we described in Chapter II.B, for the first time, and in great detail its pseudogenes. Recent evidence suggests that pseudogenes are functionally active, and therefore studying their evolution and conservation might support a functional role and give insight into their potential mechanism of action [2-7]. PPP1R2 pseudogenes were analyzed in terms of evolutionary history and putative functions. In humans, PPP1R2 has nine processed pseudogenes and one duplicated. By phylogenetic analysis the most ancient one is PPP1R2P9 (163.9-167.4 Mya) that was originated before the great radiation of mammals. The other processed pseudogenes are primate specific, with the exception of PPP1R2P7, and were originated in waves similarly to the Alu repeats explosion that occurred 40-50 Mya after the divergence of simian ancestors from the prosimians [8]. Evidences for pseudogene

duplication in humans, PPP1R2P4, and chimpanzee, PPP1R2P1, show that pseudogenization is still an active process.

Bioinformatic studies and database mining showed that PPP1R2P1, PPP1R2P2, PPP1R2P3 and PPP1R2P9 are the pseudogenes with more probability of being transcribed. Our studies indicate that indeed PPP1R2P3 and PPP1R2P9 are present in human sperm, being therefore translated and having a putative function in sperm maturation [1, 9, 10]. Also, pseudogenes altered expression seems to be involved in cancer, like for instance PPP1R2P1 in breast carcinoma [11] and in ganglioglioma [12] and PPP1R2P9 in germ cell tumors [13] and in STAT5-induced tumors [14].

Pseudogenes can also regulate their parental counterparts at the message level, leading to an alteration of the parental mRNA levels [15]. This could be what happens with PPP1R2P2, since its translation is very unlikely and it was already found to be deregulated in prostate adenocarcinomas [16].

The fact that several pseudogenes are associated with physiological and pathological states, indicate that this evolution process may be in part related with the formation of new genes or in the control of their parental PPP1R2 message. This work shows the importance of pseudogene studies in unraveling their possible biological functions to reverse the current thought that they are non-functional relics. Therefore, their names should also be revised.

In Chapter II.C a standard transgenic technique was used to overexpress hPPP1R2 and hPPP1R2P3 in mouse testis under the control of a spermatocyte specific promoter (PGK2). This work was extremely relevant in the way that it provided the molecular tools to initiate the characterization of the mechanisms behind PPP1R2/PPP1CC2 and PPP1R2P3/PPP1CC2 role in spermatozoa. The hypothesis was to test if an increase (overexpression) of hPPP1R2 in mouse testis leads to a concomitant increase in PPP1CC2 inhibition and initiation of sperm motility. Also, if this inhibition could be overcame by increasing the amounts or activity of GSK3β. Concerning hPPP1R2P3 the phenotype would be different, since GSK3β would not reverse the inhibition given the fact that Thr72 is absent. However, very few founders were obtained and no relevant phenotype was observed in the testis weight, sperm number or in morphology of these lines. Thus, no feedback was obtained from these lines and new lines expressing these transgenes under the control of a different promoter (Protamine, spermatid specific) are underway since parallel results from our collaborator (Prof. Srinivasan Vijayaraghavan) [17] suggest that mPPP1R2 message appears in round spermatids.

In Chapter III.A a novel PIP of PPP1CC2 isoform was identified, termed t-complex testis expressed protein 1 domain containing 4 (Tctex1d4), a dynein light chain. The binding results supported by YTH [1], overlay and co-immunoprecipitation approaches showed that Tctex1d4 interacts with the different spliced isoforms of PPP1CC. It was also shown that Tctex1d4 is localized in Sertoli, Leydig and germ cells of mouse testis and showed to be highly enriched in the cell-cell junctions of blood testis barrier and in the microtubules and MTOC of late stage germ cells. Further, Tctex1d4 is present in the entire tail and in the acrosome of human mature sperm. Using cell culture it was shown that Tctex1d4 and PPP1 co-localize in the MTOC structure and microtubules and that the Tctex1d4 PPP1BM seems to be important for the complex formation. Tctex1d4 PPP1BM is essential to retain PPP1 in the MTOC, and also to disrupt or delay its movement along microtubules. Furthermore, these results open new avenues to the possible roles of this dynein light chain, together with PPP1 in microtubule dynamics, sperm motility, acrosome reaction and in the regulation of the blood testis barrier possibly via TGFß signaling [18, 19].

Continuing the Tctex1d4 work, PPP1BM and its surrounding palindromic region were analyzed. Chapter III.B describes the conservation of the PPP1BM and respective palindromic region in all mammals with the exception of *Ochotona* species (*O. princeps*, *O. daurica* and *O. pusilla*). This work suggests that this event happened before the *Ochotona* radiation between 6 and 20 Million years ago. The dN/dS for this region highly supports the hypothesis that for *Ochotona* species it has been evolving under positive selection. Also, mutational screening showed that the ability of Tctex1d4 to bind to PPP1 is maintained in Pika, although both the PPP1BM and palindromic region were absent.

Concluding, the work described in this thesis characterizes four PIPs, PPP1R2, PPP1R2P3, PPP1R2P9 and Tctex1d4, in view of their importance in mammalian male reproduction. Also, by describing four new complexes in testis and sperm, this thesis opens doors to new findings in this area and to unravel the role of PPP1CC2 in spermatogenesis and sperm physiology (Figure IV. 1).

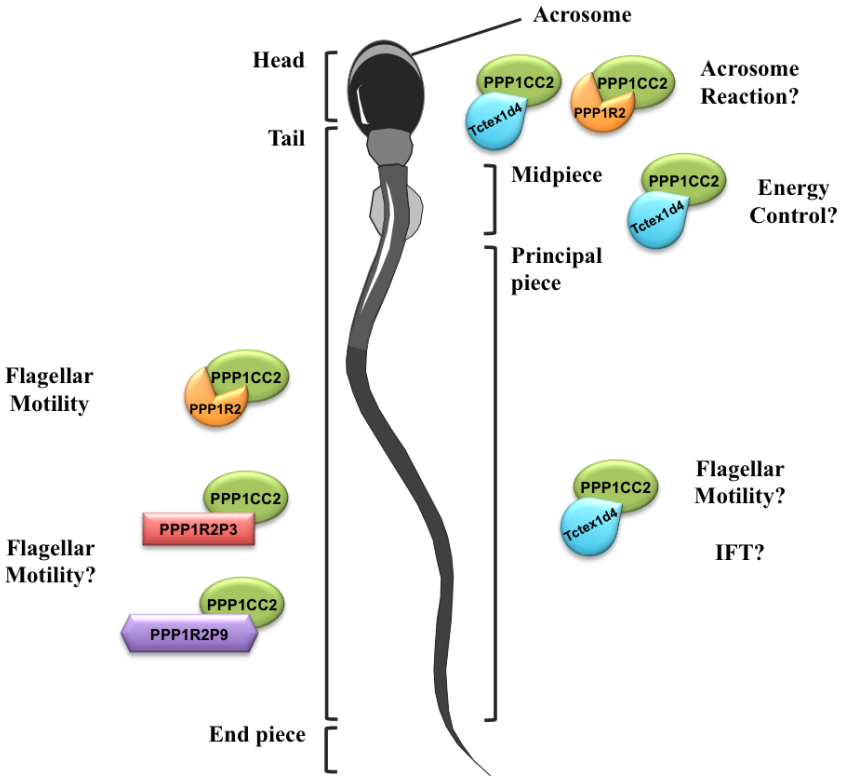


Figure IV. 1: Diagram representing the PIP-PPP1C holoenzymes characterized in the work presented in this thesis. *Respective holoenzymes localizations in spermatozoa were gathered from the immunofluorescence studies, except for PPP1R2P9. Putative functions discussed in each chapter are also presented.*

References

1. Fardilha M, Esteves SLC, Korrodi-Gregório L, Vintém AP, Domingues SC, Rebelo S, Morrice N, Cohen PTW, Silva OABdCe, Silva EFdCe. Identification of the human testis protein phosphatase 1 interactome. Biochemical Pharmacology 2011; 82: 1403-1415.
2. Hirotsune S, Yoshida N, Chen A, Garrett L, Sugiyama F, Takahashi S, Yagami K-i, Wynshaw-Boris A, Yoshiki A. An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. Nature 2003; 423: 91-96.
3. Piehler A, Hellum M, Wenzel J, Kaminski E, Haug K, Kierulf P, Kaminski W. The human ABC transporter pseudogene family: Evidence for transcription and gene-pseudogene interference. BMC Genomics 2008; 9: 165.
4. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, Surani MA, Sakaki Y, et al. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 2008; 453: 539-543.
5. Hawkins PG, Morris KV. Transcriptional regulation of Oct4 by a long non-coding RNA antisense to Oct4-pseudogene 5. Transcription 2010; 1: 165-175.
6. Chiefari E, Iiritano S, Paonessa F, Le Pera I, Arcidiacono B, Filocamo M, Foti D, Liebhaber SA, Brunetti A. Pseudogene-mediated posttranscriptional silencing of HMGA1 can result in insulin resistance and type 2 diabetes. Nat Commun 2010; 1: 40.
7. Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature 2010; 465: 1033-1038.
8. Ohshima K, Hattori M, Yada T, Gojobori T, Sakaki Y, Okada N. Whole-genome screening indicates a possible burst of formation of processed pseudogenes and Alu repeats by particular L1 subfamilies in ancestral primates. Genome Biology 2003; 4: R74.
9. Fardilha M, Esteves SLC, Korrodi-Gregório L, Pelech S, da Cruz e Silva OAB, da Cruz e Silva E. Protein phosphatase 1 complexes modulate sperm motility and present novel targets for male infertility. Molecular Human Reproduction 2011; 17: 466-477.
10. Korrodi-Gregório L, Ferreira M, Vintém AP, Wu W, Muller T, Marcus K, Vijayaraghavan S, Brautigan DL, Silva OABdCe, Fardilha M, Silva EFdCe. Discovery and characterization of human sperm protein phosphatase 1 inhibitor 2 proteins. (Submited to Biochem J 2011).
11. Wu I, Moses MA. Cloning of a cDNA encoding an isoform of human protein phosphatase inhibitor 2 from vascularized breast tumor. DNA Seq 2001; 11: 515-518.
12. Pandita A, Balasubramaniam A, Perrin R, Shannon P, Guha A. Malignant and benign ganglioglioma: a pathological and molecular study. Neuro Oncol 2007; 9: 124-134.
13. Shirato H, Shima H, Sakashita G, Nakano T, Ito M, Lee EYC, Kikuchi K. Identification and Characterization of a Novel Protein Inhibitor of Type 1 Protein Phosphatase†,‡. Biochemistry 2000; 39: 13848-13855.
14. Eilon T, Barash I. Distinct gene-expression profiles characterize mammary tumors developed in transgenic mice expressing constitutively active and C-terminally truncated variants of STAT5. BMC Genomics 2009; 10: 231.
15. Pink RC, Wicks K, Caley DP, Punch EK, Jacobs L, Carter DR. Pseudogenes: pseudo-functional or key regulators in health and disease? RNA 2011; 17: 792-798.

16. Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J, Guigo R, Dermitzakis ET. Transcriptome genetics using second generation sequencing in a Caucasian population. Nature 2010; 464: 773-777.
17. Korrodi-Gregório L, Shandilya R, Nilam S, Fardilha M, EF dCeS, S. V. *Identification and characterization of a testis specific isoform of inhibitor-2 highly expressed in haploid germ cells during late stages of sperm development*. in *Europhosphatases*. 2011. Vienna, Austria.
18. Xia W, Mruk DD, Lee WM, Cheng CY. Cytokines and junction restructuring during spermatogenesis—a lesson to learn from the testis. Cytokine & growth factor reviews 2005; 16: 469-493.
19. Xia W, Wong EWP, Mruk DD, Cheng CY. TGF-β3 and TNFα perturb blood–testis barrier (BTB) dynamics by accelerating the clathrin-mediated endocytosis of integral membrane proteins: A new concept of BTB regulation during spermatogenesis. Developmental Biology 2009; 327: 48-61.

Appendix

Appendix

Culture medium and solutions

LB (Luria-Bertani) Medium
To 950 mL of deionised H_2O add:
LB 25 g
Agar 20 g (for plates only)
Shake until the solutes have dissolved. Adjust the volume of the solution to 1 liter with deionised H_2O. Sterilize by autoclaving.

SOB Medium
To 950 mL of deionised H_2O add:
25,5 g SOB Broth
Shake until the solutes have dissolved. Add 10mL of a 250mM KCl (prepared by dissolving 1.86g of KCl in 100 mL of deionised H_2O). Adjust the pH to 7.0 with 5N NaOH. Adjust the volume of the solution to 1 liter with deionised H_2O. Sterilize by autoclaving. Just prior to use add 5 mL of a sterile solution of 2M $MgCl_2$ (prepared by dissolving 19 g of $MgCl_2$ in 90 mL of deionised H_2O; adjust the volume of the solution to 1000 mL with deionised H_2O and sterilize by autoclaving).

SOC Medium
SOC is identical to SOB except that it contains 20 mM glucose. After the SOB medium has been autoclaved, allow it to cool to 60ºC and add 20mL of a sterile 1M glucose (this solution is made by dissolving 18 g of glucose in 90 mL of deionised H_2O; after the sugar has dissolved, adjust the volume of the solution to 1 L with deionised H_2O and sterilize by filtration through a 0.22-micron filter).

Yeast Media

YPD medium
To 950mL of deionised H_2O add:
50 g YPD
20 g Agar (for plates only)
Shake until the solutes have dissolved. Adjust the volume to 1 L with deionised H_2O and sterilize by autoclaving. Allow medium to cool to 60ºC and add glucose to 2% (50mL of a sterile 40% stock solution).

SD synthetic medium
To 800mL of deionised H_2O add:
6.7g Yeast nitrogen base without amino acids (DIFCO)
20g Agar (for plates only)
Shake until the solutes have dissolved. Adjust the volume to 850mL with deionised H_2O and sterilize by autoclaving. Allow medium to cool to 60ºC and add glucose to 2% (50mL of a sterile 40% stock solution) and 100mL of the appropriate 10X dropout solution.

2X YPDA
Prepare YPD as above. After the autoclaved medium has cooled to 55ºC add 15mL of a 0.2% adenine hemisulfate solution per liter of medium (final concentration is 0.003%).

Competent Cell Solutions

Solution I (1L)
9.9 g $MnCl_2.4H_2O$
1.5 g $CaCl_2.2H_2O$
150 g glycerol
30 mL KHAc 1M;
adjust pH to 5.8 with HAc, filter through a 0.2µm filter and store at 4ºC

Solution II (1L)
20 mL 0.5M MOPS (pH 6.8)
1.2 g RbCl
11g $CaCl_2.2H_2O$
150 g glycerol;
filter through a 0.2µm filter and store at 4ºC

DNA Solutions

50X TAE Buffer
242 g Tris base
57.1 mL glacial acetic acid
100 mL 0.5M EDTA (pH 8.0)

Loading Buffer (LB)
0.25% bromophenol blue
30% glycerol

STET
8% Sucrose
5% Triton X-100
50 mM Tris-HCl (pH 8,5)
50 mM EDTA

Miniprep Solutions

Solution I
50 mM glucose
25 mM Tris.HCl (pH 8.0)
10 mM EDTA

Solution II
0.2 N NaOH
1% SDS

Solution III
3 M potassium acetate
2 M glacial acetic acid

Midiprep Solutions

Cell Resuspension Solution
50 mM Tris-HCl (pH 7.5)
10 mM EDTA
100 µg/mL RNAase A

Cell Lysis Solution
0.2 M NaOH
1% SDS

Neutralization Solution
4.09 M Guanidine hydrochloride (pH 4.8)
759 mM potassium acetate
2.12 M Glacial acetic acid

Column Wash Solution
60 mM potassium acetate
8.3 mM Tris-HCl (pH 7.5)
0.04 mM EDTA
60 % ethanol

SDS-PAGE and Immunobloting Solutions

Stock Solution 30% Acrylamide/8% Bisacrylamide
Acrylamide 150g
Bisacrylamide 4g
Adjust volume with ddH2O
Keep at 4°C protected from light in aluminium foil

Stock Solution 4xLGB
1.5M Tris-HCl
4% SDS
Adjust volume with ddH2O
Adjust the pH to 8.9 dissolve the SDS and keep at 4°C

Stock Solution 5xUGB
0,6M Tris-HCl
Adjust the pH to 6.8 and adjust volume with ddH2O.
Keep at 4°C

Ammonium Persulfate (APS)
10% APS
Adjust volume with ddH2O
Keep at 4°C

Sodium Dodecyl Sulfate (SDS)
10% SDS
Adjust volume with ddH2O
Keep at RT

10x Running Buffer
0,25M Tris-HCl
1,92M Glycine
1% SDS
You can heat to help dissolving the SDS
Adjust the pH to 8.3 and adjust volume with ddH2O.
Keep at RT

1x Transfer Buffer
25mM Tris-HCl
192M Glycine
Adjust volume with ddH2O
Add 20% Methanol before the transfer
The pH should be between 8.2-8.4
Keep at 4°C some time before doing the transfer

4x Protein Sample Buffer
40% Glycerol
250mM Tris-HCl pH 6.8
8% SDS
2% β-Mercaptoethanol
Adjust volume with ddH2O
Bromophenol Blue
Keep at RT for short period or 4°C for longer periods
You can keep at -20°C but without adding the β-Mercaptoethanol

TBS (1x)
10 mM Tris-HCl (pH 8.0)
150 mM NaCl

TBST (1x)
10 mM Tris-HCl (pH 8.0)
150 mM NaCl
0.05% Tween

Blocking and antibody solution
TBST (1x) in 3% of low-fat milk

ECL Solution Reagent A (USA)
100uM Luminol
2mM 4-iodophenol
50mM Tris-HCl pH 9.35
This solution is mixed with reagent B (hidrogen peroxide 3% or 20V) in a proportion of 100:1 (ex. 2mL A + 20uL B)

Membrane Stripping Solution
2% SDS
62.5 mM Tris-Hcl (pH= 6.7)
100 mM β-Mercaptoethanol

Coomassie blue staining solutions

Fixation solution
40% Methanol
10% SDS
ddH2O
Keep at RT

Staining solution
20% Methanol
0.12% Coomassie Blue G250
ddH2O
Keep at RT

Distaining solution
25% Methanol
Keep at RT

Immunoprecipitation solutions

RIPA Lysis Buffer (10x)
0.5M Tris- HCl (pH 7.4)
1.5M NaCl
10mM EDTA
10% NP-40
2.5% sodium deoxycholate

RIPA Lysis Buffer + Protease inhibitors
Add to RIPA buffer the following concentrations:
5 µM Pepstatin A
2 µM Leupeptin
10mM Benzamidine
1.5 µM Aprotinin
1 mM PMSF

Washing solution
1xPBS in 3%BSA

Solutions for the 2-D gel electrophoresis

Lysis buffer
9M Urea
4% CHAPS

Equilibration buffer
50mM Tris (pH8.8)
6M Urea
30% Glycerol
2% SDS
0.002% Bromophenol blue

Rehydratation solution
8M Urea
2M Thiurea
2% CHAPS
0.002% of bromophenol blue
Supplemented with 2.5µl of IPG buffer (in the 4-7 pH range) and 14mg of DTT

Cell Culture Solutions and Immunocytochemistry

Complete DMEM
For a final volume of 500 mL, add:
50 mL (10% v/v) Fetal Bovine Serum (FBS) (Gibco BRL, Invitrogen)
Antibiotics (5 mL)
100 U/mL penicillin
100 mg/mL streptomycin

PBS (1x)
For a final volume of 500 mL, dissolve one pack of BupH Modified Dulbecco's Phosphate Buffered Saline Pack (Pierce) in deionised H_2O. Final composition:
8 mM Sodium Phosphate
2 mM Potassium Phosphate
40 mM NaCl

10 mM KCl
Sterilize by filtering through a 0.2 μm filter and store at 4 ºC

1 mg/mL Poly-L-ornithine solution (10x)
To a final volume of 100 mL, dissolve in deionised H_2O 100 mg of poly-L-ornithine (Sigma-Aldrich, Portugal).

4% Paraformaldehyde Fixative solution
For a final volume of 100 mL, add 4 g of paraformaldehyde to 25 mL deionised H_2O. Dissolve by heating the mixture at 58 ºC while stirring. Add 1-2 drops of 1 M NaOH to clarify the solution and filter (0.2 μm).
Add 50 mL of 2X PBS and adjust the volume to 100 mL with deionised H_2O.

Immunohistochemistry

Citrate buffer
Buffer A: 1mM $C_6H_8O_7H_2O$
Buffer B: 50 mM $C_6H_5Na_3O_72H_2O$
Citrate buffer: 0.15 mM Buffer A, 8.5 mM Buffer B, pH 6.0

Blocked solution
5% goat serum
1% bovine serum albumin
1x PBS

Permeabilization solution
0.2% Triton X-100
1x PBS

Primer Hybridization

Tris annealing buffer
10mM Tris
Adjust the pH to 8.0 and adjust the volume with ddH_2O.

Genomic DNA extraction

Alkaline lysis buffer
25mM NaOH
2mM EDTA
Adjust the pH to 12.0 and adjust the volume with ddH_2O.

Neutralization buffer
40mM Tris-HCl
Adjust the pH to 5.0 and adjust the volume with ddH_2O.

Sperm proteins extracts

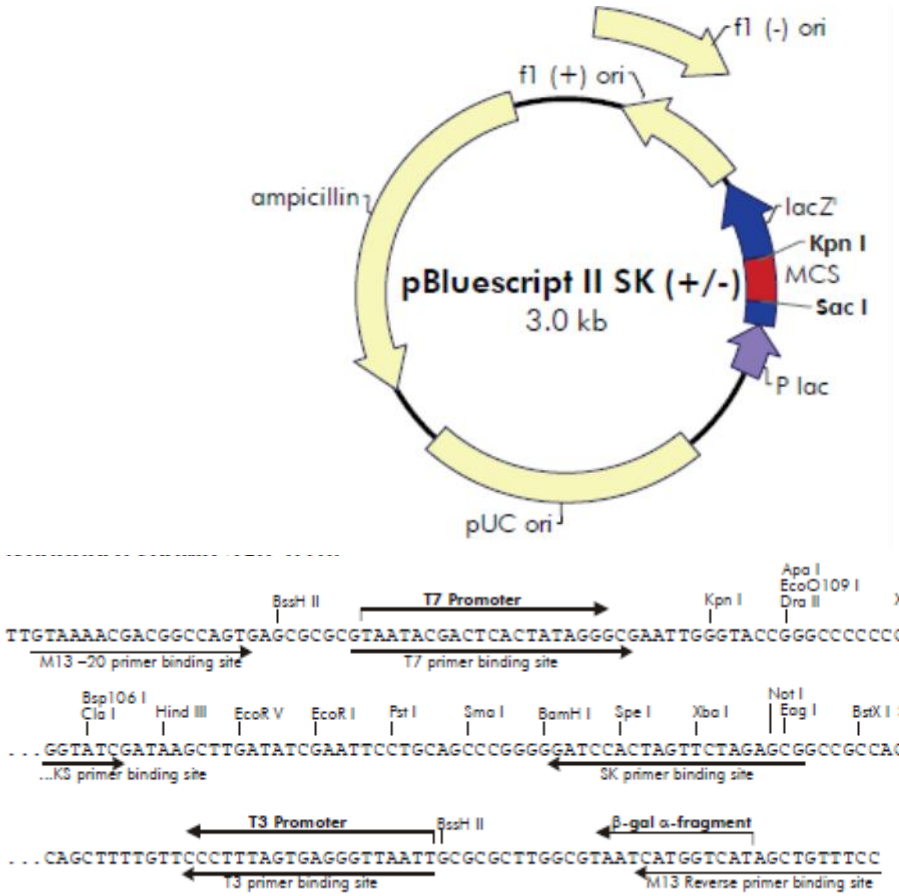
Homogenization buffer plus
10mM Tris-HCl (pH 7.0)
1mM EGTA
0.1mM EDTA

Protease inhibitors
10mM Benzamidine,
1mM PMSF
0.1mM TPCK

BACTERIA AND YEAST STRAINS

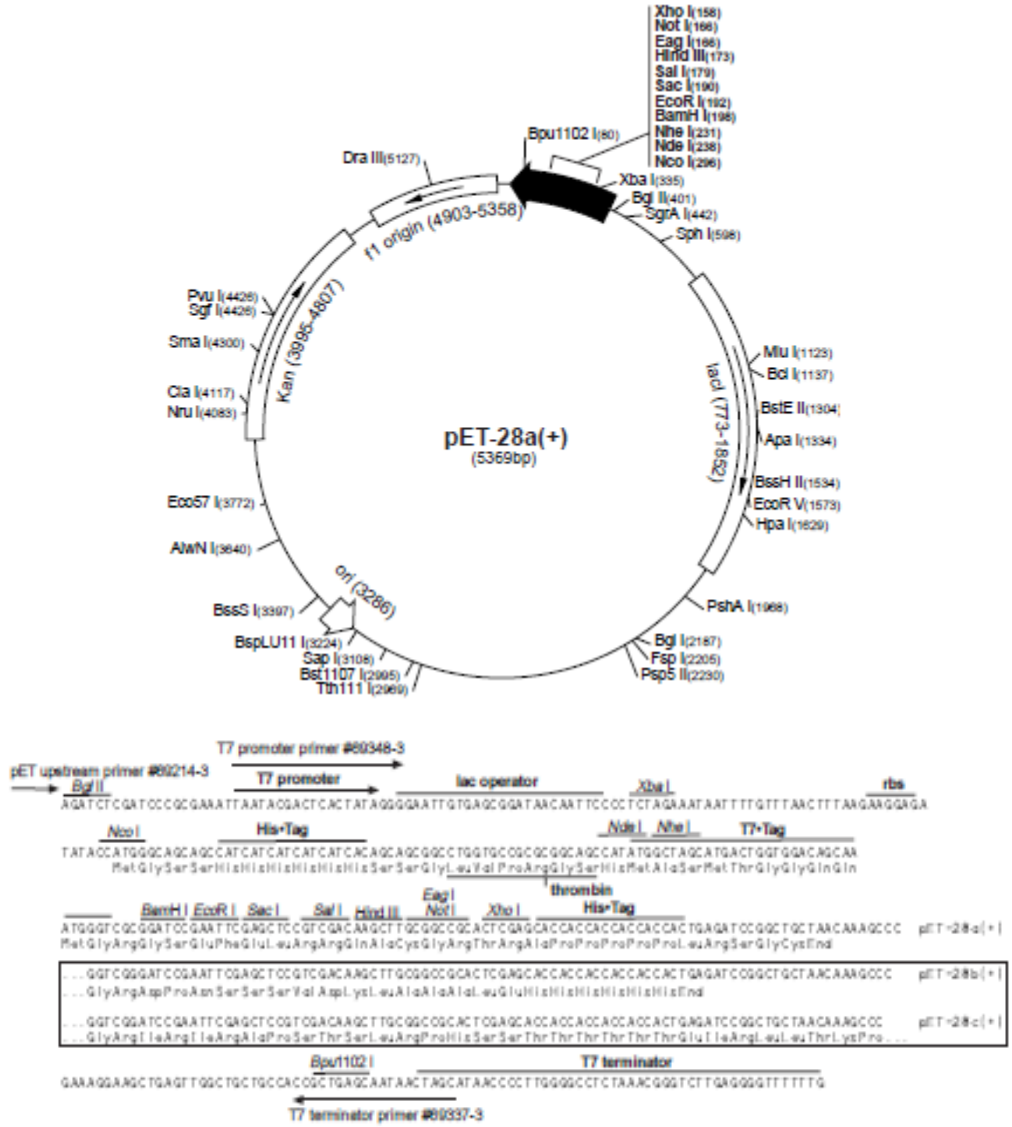
- **E. coli XL1- blue**: *recA endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F' proAB lacZ⊿M15 Tn10(Tet^r)]*
- **Rosetta(DE3)pLysS** : F$^-$ ompT hsdS_B(R_B^- m_B^-) gal dcm λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) pLysSRARE (Cam^R)
- **S. cerevisiae AH109**: MATa, trp1-901, leu2-3, 112□ ura3-52, his3-200, gal4Δ, gal 80Δ, LYS2:: $GAL1_{UAS}$-$GAL1_{TATA}$-HIS3, $GAL2_{UAS}$-$GAL2_{TATA}$-ADE2, URA3::$MEL1_{UAS}$-$MEL1_{TATA}$-lacZ, MEL1
- **S. cerevisiae Y187**: MATα□ ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4Δ, met-, gal 80Δ, URA3:: $GAL1_{UAS}$-$GAL1_{TATA}$-lacZ, MEL1

Plasmids

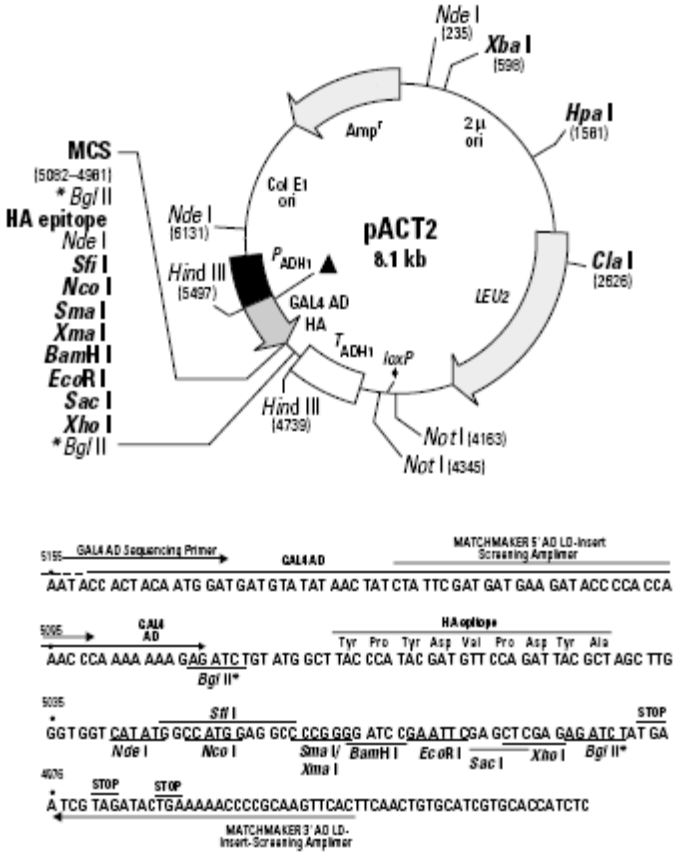


The pBluescript II phagemids (Stratagene) are cloning vectors designed to simplify commonly used cloning and sequencing procedures, including the construction of nested deletions for DNA sequencing, generation of RNA transcripts *in vitro* and site-specific mutagenesis and gene mapping. The pBluescript II phagemids have an extensive polylinker with 21 unique restriction enzyme recognition sites. Flanking the polylinker are T7 and T3 RNA polymerase promoters that can be used to synthesize RNA *in vitro*.1, 2 The choice of promoter used to initiate transcription determines which strand of the insert cloned into the polylinker will be transcribed.

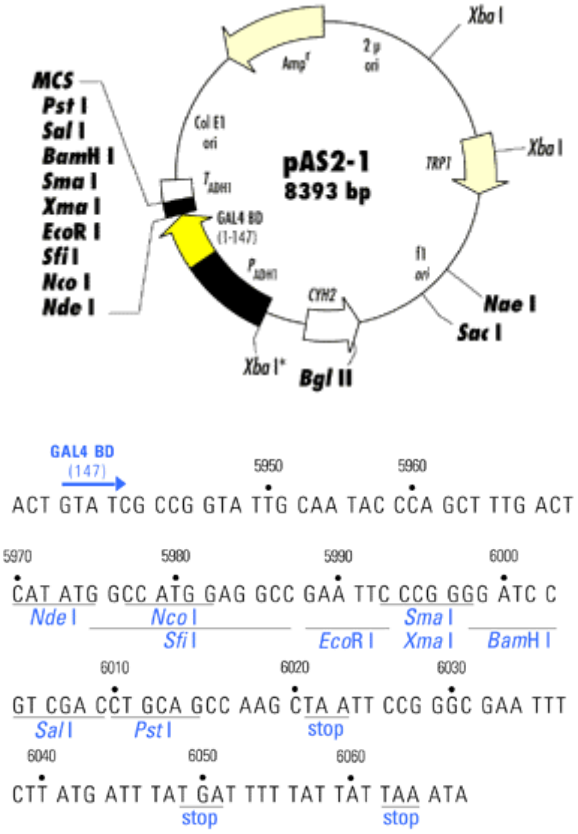
pBluescript II (+) and (–) are available with two polylinker orientations designated as either KS or SK using the following convention: (1) in the KS orientation, the *Kpn* I restriction site is nearest the *lac*Z promoter and the *Sac* I restriction site is farthest from the *lac*Z promoter; and (2) in the SK orientation, the *Sac* I site is the closest restriction site to the *lac*Z promoter and the *Kpn* I site is the farthest.



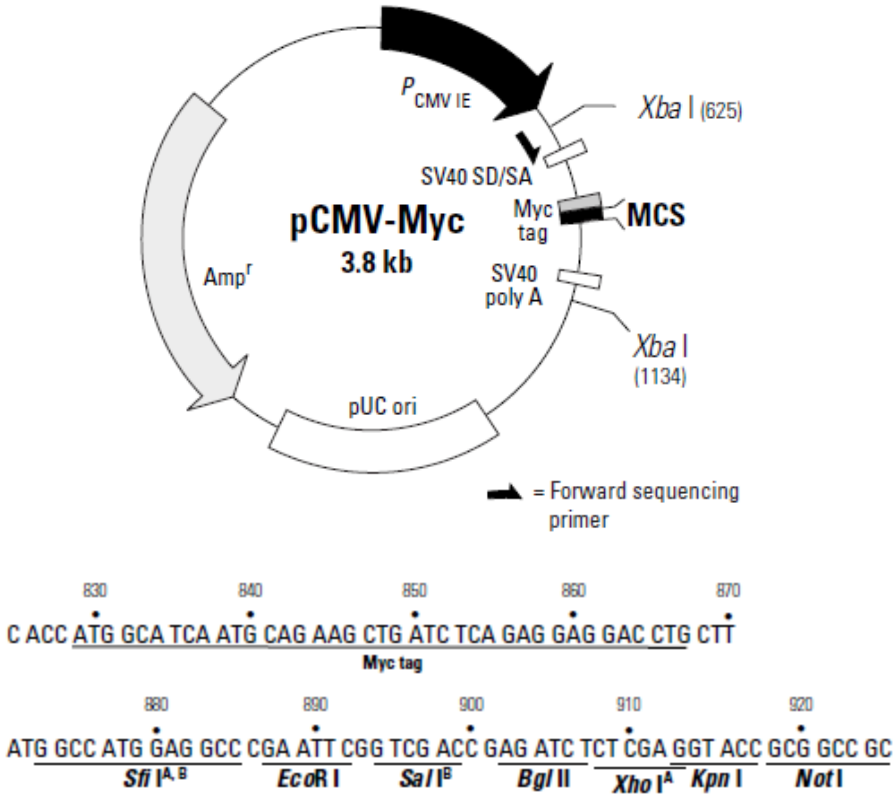
The pET-28a-c(+) vectors (Novagen) are appropriate for bacterial expression and carry an N-terminal His•Tag®/thrombin/T7•Tag® configuration plus an optional C-terminal His•Tag sequence. Unique sites are shown on the circle map. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below. The f1 origin is oriented so that infection with helper phage will produce virions containing single-stranded DNA that corresponds to the coding strand. Therefore, singlestranded sequencing should be performed using the T7 terminator primer.



pACT2 (Clontech) map and MCS. pACT2 is used to generate a hybrid containing the GAL4 AD, an epitope tag and a protein encoded by a cDNA in a fusion library. The hybrid protein is expressed at medium levels in yeast host cells from an enhanced, truncated ADH1 promoter and is target to the nucleus by the SV40 T-antigen nuclear localization sequence. pACT2 contains the LEU2 gene for selection in Leu$^-$ auxotrophic yeast strains.



pAS2-1 (Clontech) map and MCS. Unique sites are coloured blue. pAS2-1 is a cloning vector used to generate fusions of a bait protein with the GAL4 DNA-BD. The hybrid protein is expressed at high levels in yeast host cells from the full-length ADH1 promoter. The hybrid protein is target to the yeast nucleus by nuclear localization sequences. pAS2-1 contains the TRP1 gene for selection in Trp^- auxotrophic yeast strains.



The pCMV-Myc Mammalian Expression Vector (Clontech) expresses proteins containing the N-terminal c-Myc epitope tag. The c-Myc epitope tag is well-characterized and highly immunoreactive. High-level expression in mammalian cells is driven from the human cytomegalovirus immediate early promoter/enhancer (*P*CMV IE). The vector contains an intron (splice donor/splice acceptor); the epitope tag; an MCS; and a polyadenylation signal from SV40. This vector also possesses the ampicillin resistance gene for selection in *E. coli.*

Primers

Primer	Sequence (5'::::: 3')	Nt No.
GAL4 AD	TACCACTACAATGGATG	17
GAL4 BD	TCATCGGAAGAGAGTAG	17
T7 Promotor	AATACGACTCACTATAGG	18
T7 Terminator	GCTAGTTATTGCTCAGCGG	19
3HA Forward	ATTCATGTACCCATACGATGTTCCAGATTACGCTACC GGATACCCATACGATGTTCCAGATTACGCTACCGGA TACCCATACGATGTTCCAGATTACGCTACCGGAG	108
3HA Reverse	AATTCTCCGGTAGCGTAATCTGGAACATCGTATGGG TATCCGGTAGCGTAATCTGGAACATCGTATGGGTAT CCGGTAGCGTAATCTGGAACATCGTATGGGTACATG	108
PGK2 Forward	GCGCACACCTCAGGACTATT	20
SV40 Reverse	CTTGGCGTAATCATGGTGGTACC	23
PPP1R2 Internal Forward	ACTGCAGATGGAGAAAGCATGAATAC	26