Spliceosome nanomachinery is coded by more than **100 genes** 

Every human cell contains ~ 100.000 spliceosomes

It takes **20-30 s to transcript and excise** a short intron (1.3 - 1.4 Kb)

**35 %** of human genetic disorders are caused by splicing alterations

Mutations in spliceosome have been linked to retinitis pigmentosa and cancers

Discovered in 1977 by Philip Sharp and Richard J. Roberts (Nobel 1993)

# The Spliceosome

U2 protein complex and a closer look into Splicing Factor 3a





Ona Baguer, Carlota Cerrillo, Florian Conrad, Clara Escala

# Contents

### 1. Splicing: what are we talking about?

- a. Reaction
- b. Origin and evolution
- c. Cryo-EM

### 2. Structure of the human major spliceosome

- a. The Major Spliceosome
- b. Spliceosome cycle
- c. Yeast versus human major splicesosomes
- 3. SF3a: structure, function and evolution
- 4. Conclusions

# Splicing: what are we talking about?



Protein-coding genes in humans contain typically 9-10 introns, but some > 100



Step 1. 2'OH (branch-point adenosine)  $\rightarrow$  PO<sub>4</sub> (between 5' exon<sub>1</sub> and intron)

• Formation of lariat harboring a 2'-5' phosphodiester bond

Step 2. 3'OH (5'  $exon_1$ )  $\rightarrow$  PO<sub>4</sub> (between intron and 3'  $exon_2$ )

Extracted from: Weaver R.F. Molecular Biology. Boston: McGraw-Hill, 1999.

#### **Origin & evolution**



Extracted from: Hang J, Wan R, Yan C, Shi Y. Structural basis of pre-mRNA splicing. Science. 2015;349(6253):1191-8.





#### Group II introns

#### Spliceosomal introns



**Eukaryote-specific** RNA-protein complex

**5 snRNPs\*** with over **100 proteins** 

Major and minor spliceosome

Major spliceosome: ~ 99.5% of human intron removal

\*snRNP: small nuclear ribonucleoprotein particles



**Origin & evolution** 

Cryo-EM

# Cryo- Electron Microscopy



Cryo-EM micrograph of the yeast spliceosomal complex. Scale bar, 30 nm



2D class averages of the yeast spliceosome



EM density for the yeast spliceosome. The resolution goes to 2.9-3.2 Å in the center of the spliceosome

Extracted from: Yan, C., Hang, J., Wan, R., Huang, M., Wong, C. and Shi, Y. Structure of a yeast spliceosome at 3.6-angstrom resolution. Science. 2015; 349: 1182-1191.





and the second s			
Adenovirus 26	70S Ribosome	IDH1	Hemoglobin
at 3.7Å res.	at 3.6Å res.	at 3.8Å res.	at 3.6Å res.
150MDa	2.3MDa	93kDa	64kDa
EMD-8471	EMD-3618	EMD-3192	EMD-3650

Extracted from: Song C, Murata K. Cryo-electron Microscopy for Structural Analysis of Dynamic Biological Macromolecules. Journal of Computer Chemistry, Japan. 2018; 17(1): 38-45.



Extracted from: Yan C, Wan R, Shi Y. Molecular Mechanisms of pre-mRNA Splicing through Structural Biology of the Spliceosome. Cold Spring Harb Perspect Biol. 2019:11(1).

# The Human Major Spliceosome





NTC: Nine Teen Complex

NTR: NTC-related





Extracted from: Wahl M, Will C, Lührmann R. The Spliceosome: Design Principles of a Dynamic RNP Machine. Cell. 2009; 136(4): 701-718.



Adapted from: Galej WP. Structural studies of the spliceosome: past, present and future perspectives. Biochem Soc Trans. 2018;46(6):1407-1422.



Adapted from: Galej WP. Structural studies of the spliceosome: past, present and future perspectives. Biochem Soc Trans. 2018;46(6):1407-1422.



Adapted from: Galej WP. Structural studies of the spliceosome: past, present and future perspectives. Biochem Soc Trans. 2018;46(6):1407-1422.



Adapted from: Absmeier E, Santos KF, Whal MC. Functions and regulation of the Brr2 RNA helicase during splicing. Cell Cycle. 201616;15(24):3362-3377.



Adapted from: Absmeier E, Santos KF, Whal MC. Functions and regulation of the Brr2 RNA helicase during splicing. Cell Cycle. 201616;15(24):3362-3377.

# A complex



Adapted from: Kuwasako K et al. Solution Structures of the SURP Domains and the Subunit-Assembly Mechanism within the Splicing Factor SF3a Complex in 17S U2 snRNP. Structure. 2006;14(11):1677-89.



Yeast vs Human





Spliceosome Cycle

Yeast vs Human

## Sm and LSm complexes







B complex



PDB ID: 6AHD

U5 snRNF

**B** complex U2 snRNP U6 snRNP (pre-mRNA) U2/BPS U2/U6 U2 snRNA U4/U6 pre-mRNA 5'-exon/Loop I 5' (pre-mRNA) pre-mRNA 5'SS/U6 U4 snRNA U5 snRNA U6 snRNA

Zhan X, Yan C, Zhang X, Lei J, Shi Y. Structures of the human pre-catalytic spliceosome and its precursor spliceosome. *Cell Research*. 2018; 28(12): 1129-1140.

Resolution: 3,8 Å PDB ID: 6AHD Spliceosome Cycle

#### Yeast vs Human

# B complex



Resolution: 3,8 Å PDB ID: 6AHD











**Spliceosome Cycle** 

Yeast vs Human

#### S. cerevisiae pre-B complex



### H. sapiens pre-B complex


**Spliceosome Cycle** 

Yeast vs Human



**Spliceosome Cycle** 

Yeast vs Human

## S. cerevisiae pre-B complex

## H. sapiens pre-B complex





- $\rightarrow$  Crucial component of functionally active 17S U2 complex
- $\rightarrow$  Stabilization of U2
- $\rightarrow$  Its displacement initiates the first step of the splicing reaction



Adapted from: Kuwasako K et al. Solution Structures of the SURP Domains and the Subunit-Assembly Mechanism within the Splicing Factor SF3a Complex in 17S U2 snRNP. Structure. 2006;14(11):1677-89.



















Extracted from: Protein Structures [Internet]. Available from: Peptideweb.com











	Pi	rp21			Prp11
	<i>a</i> 1	a2		<i>a</i> 3	
E-NLLETRRSLL	EE-EIIENAIAERIORN	VPELYYHYIQESS	SKVFPDTKLI	PRSSLIAENKI	
METILEQORRYH	~ EEKERLMDVMAKEMLTH	~ KKSTLRDQINSDF	IRTRAMODRY	YMEVSGNLRDL	
	α4		α5		
YKFKKVKR <mark>KRKQ</mark>	IILQQHEINIFLRDYQE	EKQQTFNKINF <mark>E</mark> F	RKLQQLEKEI	LKNEDENFELD	
YDDKDGLRKEEL	NAISGPNEFAEFYNRLE	KQIKEFHRKH <mark>FE</mark> E	LLKARENPS	SEEAQNLVEFT	
α6				α7	
IN <mark>SKKDKYALF</mark> S	SSSDPSRRTNILSDRAF	RDLDLNEIFTRDE	EQYGEY-ELH	EQFHSLWLNVI	
DEEGYGRYLDL-	H-Y-INLKASEKLDYIT	FYLSIFDQLFDIE	PKERKNAEYI	KRYLEMLLEYL	
αξ	3		<b>α9</b>		
K <mark>RGDC</mark> SLLQFLD	ILELFLDDEKYLLTPP-	-DRKNDRY-AFLI	LKLSKYVETI	FFFKSYALLDA	
QDYTDR <mark>VKPLQD</mark>	QNEL-F-E-KKWENGTH	FPGWPKETSSAL1	THAGAHLDLS	SAFSSWEELAS	
<i>α</i> 10	β <mark>1</mark> β2		α11	α12a	
AAVENLIKSDFE	HSYCR-GSL <mark>RSE</mark> AKGIN	CP-FCSRWFKTS	SVFESHLV	GKIHKKNESKR	
LGLDRLKSALLA	LGLKCGGTLEERAQRLE	FSTKGKSLESLDI	SLFAKNPKS	SKGTKRDTERN	
α12b	<i>a</i> 12c	α	:13		
RNFVYSEYKLHR	YLKYLNDEFSRTRSFVE	GRKLAFTANER-A	AE-DIL'I'QK	YEAPAYDSTEK	
KDIAF LEAQIYE	YVEILGEQRHLTHENVÇ	QRKQAR'I'GEEREE	SEEEEQISE:	SESEDEENIPY	
ECIEOUD					
EGAEQVD					
MTIVTUGTNINI	NCEICGNIIIKGEVAL	ZUUL AFMKUAUQN	ALCTOTEN I V	AULANAIÕTED	
		_			
AVSLWAKLKEDS	SGNVVNKKTYEDLKROO			Church attacks and a list	
	E-NLLETRRSLL METILEQQRRYH YKFKKVKRKRKQ YDDKDGLRKEEL <b>a6</b> INSKKDKYALFS DEEGYGRYLDL- <b>a10</b> AAVENLIKSDFE LGLDRLKSALLA <b>a12b</b> RNFVYSEYKLHR KDIAFLEAQIYE EGAEQVD WLYKLHGLNINY	a1   E-NLLETRRSLLEE-EIIENAIAERIQRM   METILEQQRRYHEEKERLMDVMAKEMLTH   a4   YKFKKVKRKRKQIILQQHEINIFLRDYQH   YDDKDGLRKEELNAISGPNEFAEFYNRLH   a6   INSKKDKYALFSSSSDPSRRTNILSDRAH   DEEGYGRYLDL-H-Y-INLKASEKLDYITH   a8   KRGDCSLLQFLDILELFLDDEKYLLTPP-   QDYTDRVKPLQDQNEL-F-E-KKWENGTH   a10 β1   β2   AAVENLIKSDFEHSYCR-GSLRSEAKGIN   LGLDRLKSALLALGLKCGGTLEERAQRLH   a12b a12c   RNFVYSEYKLHRYLKYLNDEFSRTRSFVE   KDIAFLEAQIYEYVEILGEQRHLTHENVQ   EGAEQVD	α1 α2   E-NLLETRRSLLEE-EIIENAIAERIQRNPELYYHYIQESS   METILEQQRRYHEEKERLMDVMAKEMLTKKSTLRDQINSDH   α4   YKFKKVKRKRKQIILQQHEINIFLRDYQEKQQTFNKINFER   YDDKDGLRKEELNAISGPNEFAEFYNRLKQIKEFHRKHFER   α6   INSKKDKYALFSSSSDPSRRTNILSDRARDLDLNEIFTRDE   DEEGYGRYLDL-H-Y-INLKASEKLDYITYLSIFDQLFDIF   α8   KRGDCSLLQFLDILELFLDDEKYLLTPP-DRKNDRY-AFLI   QDYTDRVKPLQDQNEL-F-E-KKWENGFFPGWPKETSSALT   α10 β1   β2   AAVENLIKSDFEHSYCR-GSLRSEAKGIYCP-FCSRWFKT   LGLDRLKSALLALGLKCGGTLEERAQRLFSTKGKSLESLDT   α12b α12c   α12b α12c   WIYKLHRYLKYLNDEFSRTRSFVERKLAFTANER-Z   KDIAFLEAQIYEYVEILGEQRHLTHENVQRKQARTGEEREE   EGAEQVD-   WLYKLHGLNINYNCEICGNYTYRGPKAFQRHFAEWRHAHGN	a1 a2   E-NLLETRRSLEE-EIIENAIAERIQRNPELYHYIQESSKVFPDTKLIMETILEQQRRYHEEKERLMDVMAKEMLTKKSTLRDQINSDHRTRAMQDRS   a4 a5   YKFKKVKRKRQIILQQHEINIFLRDYQEKQQTFNKINFERKLQQLEKEIYDDKDGLRKEELNAISGPNEFAEFYNRLKQIKEFHRKHFEELLKARENPS   a6   INSKKDKYALFSSSSDPSRTNILSDRARDLDLNEIFTRDEQYGEY-ELDEEGYGRYLDL-H-Y-INLKASEKLDYITYLSIFDQLFDIPKERKNAEYI   a8 a9   KRGDCSLLQFLDILELFLDDEKYLLTPP-DRKNDRY-AFLLKLSKYVETIQDYTDRVKPLQDQNEL-F-E-KKWENGTFPGWPKETSSALTHAGAHLDLS   a10 β1 β2 a11   AVVENLIKSDFEHSYCR-GSLRSEAKGIYCP-FCSRWFKTSSVFESHLVQLGLDRLKSALLALGLKCGGTLEERAQRLFSTKGKSLESLDTSLFAKNPKS a12b a12c a13   RNFVYSEYKLHRYLKYLNDEFSRTRSFVERKLAFTANER-AE-DILTQKXKDIAFLEAQIYEYVEILGEQRHLTHENVQRKQARTGEEREEEEEQISES EGAEQVD	a1 a2 a3   F-NLLETRRSLEES I EINALAERIQRNPELYJI QESSKVFPDTKLPRSSLIAENKI MET LEQQRRYHEEKERIMDVMAKEMITKKSTLRDQ INSDHRTRAMQDRYMEVSGNLRDL   a4 a5   VKFKKVKRKRKQ I I LQQHE INI FLRDYQEKQQTFNK INFERKLQQLEKELKNEDENFELD VDCKDGLRKEELNAI SGPNEFAEFYNRLKQI KEFHRKHFEELLKARENPSEEAQNLVEFT   a6 a7   INSKKDKYALFSSSDPSRTNI LSDRARDLDLNEI FTRDEQYGEY - ELEQFHSLWLNVI DEEGYGRYLDL-H-Y-INLKASEKLDY ITYLSI FDQLFDI PKERKNAEYKRYLEMLLEYL   a8 a9   KRGDC SLLQFLDI LELFLDDEKYLLTPP DRKNDRY - AFLLKNSKYVETFFKSYALLDA QUYDRVKPLQDQNEL - F-E-KKWENGTFPGWPRETSSALTHAGAHLDLSAFSSWEELAS   a10 β1 β2 a10 a126   AVENLIKSDEFHSYCR-GSLKSEAKGTYCP - FCSRWFKTSSVEFSHLVGKIHKKNESKR LGLDRLKSALLALGLKCGGTLEERAQRLFSTKGKSLESLDTSLFAKNPKSKGTKRDTERN a126 a126   MFVYSEYKLHRYLKYLNDEFSRTRSFVERKLAFTANER - AE-DILTQKYEAPAYDSTEK KDIAFLEAQI YEYVEI LGEQRHLTHENVQRKQAR TGEEREEEEEEQI SESESEDEENIPY EGAEQVD

Structural alignment (STAMP)

	Prp9	Prp21		Prp11
yeast_prp9 human_SF3a60 mouse_SF3a60 droso_SF3a60	1 .MNLLETRRS LLEEMEIIEN AIAERIQRNP E METILEQQRR YHEEKERLMD VMAKEMLTKK S METILEQQRR YHEEKERLMD VMAKEMLTKK S METLLEQQRR LHEERERLVK LMVDEHATKK P	60 LYYHYIQES SKVFPDTKLP RSSLIAENKI TLRDQINSD HRTRAMQDRY MEVSGNLRDL TLRDQINSD HRTRAMQDRY MEVSGNLRDL GEKERIHSE HRLKYLMELH HNSTSQLRDL	301 FKTSSVFESH LVGK SLDTSLFAKN PKSKC SLDTSLFAKN PKSKC TLDPALMAKK PSAKTASA	360 HK KNESKRRNFV YSEYKLHRYL KYLNDEFSRT RSFVERKLAF RTK RDTERNKDIA FLEAQIYEYV EILGEQRHLT HENVQRKQAR RTK RDTERNKDIA FLEAQIYEYV EILGEQRQLT HENVQRKQAR QS REHERHKEIA QLEALLYKYA DLLSEQRAAT KENVQRKQAR
yeast_prp9 human_SF3a60 mouse_SF3a60 droso_SF3a60	61 YKFKKVKRKR KQIILQQHEI NIFLRDYQEK Q YDDKDGLRKE ELNAISGPNEFAEFYNR L YDDKDGLRKE ELNAISGPNEFAEFYNR L YEDKDNERKA EIAALSGPNEFNEFYAR L	120 QTFNKINRP EETQEDDKDL PNFERKLQQL KQIKEFHRK HPNEICVPMS VEFEELLKAR KQIKEFHRK HPNEICVPMS VEFEELLKAR KQIKQFYKS HPAEVSVPLS VEFDEMIRVY	361 TANERMAEMD ILTQKYEA TGEE TGEE TGGE	420 PA YDSTEKEGAE QVDGEQRDGQ LQEEHLSGKS FDMPLGPDGL REE.EE EEQISESEE DEENEIIYNP KNLPLGWDGK REE.EE EEQISESES DEENEIIYNP KNLPLGWDGK RDDSDV EASESDNEDD PDADDVPYNP KNLPLGWDGK
yeast_prp9 human_SF3a60 mouse_SF3a60 droso_SF3a60	121 EKELKNEDEN FELDINSK KDKYALFSSS S ENPSEEAQNL VEFTDEEG YGRYLDLHDC Y ENPSEEAQNL VEFTDEEG YGRYLDLHDC Y NNP.DDMSAL VEFTDEEG GGRYLDLNEC Y	180 DPSRRTN ILSDRARDLD LNEIFTRDEQ LKYINLKAS EKLDYITYLS IFDQLFDIPK LKYINLKAS EKLDYITYLS IFDQLFDIPK ELYLNLRSV EKLDYITYLM SFDHVFDIPR	421 PMPYWLYKLH GLDREYRC PIPYWLYKLH GLNINYNC PIPYWLYKLH GLNINYNC PIPYWLYKLH GLNISYNC	480 EEI CSNKVYNGRR TFERHFNEER HIYHLRCLGI EPSSVFKGIT EEI CGNYTYRGPK AFQRHFAEWR HAHGMRCLGI PNTAHFANVT EEI CGNYTYRGPK AFQRHFAEWR HAHGMRCLGI PNTAHFANVT EEI CGNFTYKGPK AFQRHFAEWR HAHGMRCLGI PNTAHFANVT
yeast_prp9 human_SF3a60 mouse_SF3a60 droso_SF3a60	181 YGEYMELEQF HSLWLNVIKR GDCSLLQFLD I ERKNAEYKRY LEMLLEYLQD YTDRVKPLQD Q ERKNAEYKRY LEMLLEYLQD YTDRVKPLQD Q ERKNREYRIY IETLNDYLHH FILRIQPLLD L	240 LELFLDDEKYLLTP PMDRKNDRYM NELFGKIQA EFEKKWENGT FPGWP.KETS NELFGKIQT DFEKKWDNGT FPGWP.KETS EGELLKVEL DFQRQWLMGT FPGFSIKETE	481 KIKEAQELWK NMQGQSQL QIEDAVSLWA KLKLQKAS QIEDAVSLWA KLKLQKAS QIEDAITLWE KLKSQKQS	540 TS IAAVPPKPNP SQLKVPTELE LEEEDEEGNV MSKKVYDELK E
yeast_prp9 human_SF3a60 mouse_SF3a60 droso_SF3a60	241 AFLLKLSKYV ETFFFKSYAL LDAAAVENLI K SALTHAGAHL DLSAFSSWEE LASLGLDRLK S SALTHAGAHL DLSAFSSWEE LASLGLDRLK S SALANTGAHL DLSAFSSWEE LASLGLDRLK S	300 SDFEHSY.C RGSLRSEAKG I\ <mark>CPFCSR.W</mark> ALLALGLKC GGTLEERAQR LFSTKGKSLE ALLALGLKC GGTLEERAQR LFSTKGKSLE ALVALGLKC GGTLEERAQR LFSTKGKS	541 545 KQGLV RQGLL RQGLL RQGLL	

Multiple Sequence Alignment (ClustalW)

Prp9			Prp21		Prp11		
yeast_prp9 human_SF3a60 mouse_SF3a60 droso_SF3a60	241 AFLLKLSKYV SALTHAGAHL SALTHAGAHL SALANTGAHL	ETFFFKSYAL DLSAFSSWEE DLSAFSSWEE DLSAFSSWEE	LDAAAVENLI LASLGLDRLK LASLGLDRLK LASLGLDRLK	KSDFEHSY.C SALLALGLKC SALLALGLKC SALVALGLKC	RGSLRSEAKG GGTLEERAQR GGTLEERAQR GGTLEERAQR	300 IYCPFCSR.W LFSTKGKSLE LFSTKGKSLE LFSTKGKS	
yeast_prp9 human_SF3a60 mouse_SF3a60 droso_SF3a60	301 FKTSSVFESH SLDTSLFAKN SLDTSLFAKN TLDPALMAKK	LVGKIHK PKSKGTK PKSKGTK PSAKTASAQS	KNESKRRNFV RDTERNKDIA RDTERNKDIA REHERHKEIA	YSEYKLHRYL FLEAQIYEYV FLEAQIYEYV QLEALLYKYA	KYLNDEFSRT EILGEQRHLT EILGEQRQLT DLLSEQRAAT	360 RSFVERKLAF HENVQRKQAR HENVQRKQAR KENVQRKQAR	

Prp9		Prp21		Prp11		
241					300	
yeast_prp9 AFLLKL	SKYV ETFFFKSYAL	LDAAAVENLI K	SDFEHSY.C	RGSLRSEAKG	IYCPFCSR.W	
human_SF3a60 SALTHA	GAHL DLSAFSSWEE	LASLGLDRLK S	ALLALGLKC	GGTLEERAQR	LFST <mark>K</mark> G <mark>K</mark> SLE	
mouse_SF3a60 SALTHA	GAHL DLSAFSSWEE	LASLGLDRLK S	SALLALGLKC	GGTLEERAQR	LFST <mark>K</mark> G <mark>K</mark> SLE	
droso_SF3a60 SALANT	GAHL DLSAFSSWEE	LASLGLDRLK S	SALVALGLKC	GGTLEERAQR	LFST <mark>K</mark> G <mark>K</mark> S	
301					360	
yeast_prp9 FKTSS\	/FES <mark>H</mark> LVG <mark>KIHK</mark>	KNESKRRNFV Y	SEYKLHRYL	KYLNDEFSRT	RSFVERKLAF	
human_SF3a60 SLDTSL	.FA <mark>K</mark> N PKS <mark>KGTK</mark>	RDTERNKDIA F	LEAQIYEYV	EILGEQRHLT	HENVQRKQAR	
mouse_SF3a60 SLDTSL	.FA <mark>K</mark> N PKS <mark>KGT</mark> K	RDTERNKDIA F	FLEAQIYEYV	EILGEQRQLT	HENVQRKQAR	
droso_SF3a60 TLDPAL	.MA <mark>KK</mark> PSA <mark>K</mark> TASAQS	REHERHKEIA Q	QLEALLYKYA	DLLSEQRAAT	KENVQRKQAR	

























Prp9	Prp21	Prp11		
<b>αA</b> prp21 SRRD-EVIKLTARYY SF3a120 SAFDLDVVKLTAQFV prp21 SFTGQEIKK <mark>SKRTIL</mark> SF3a120 IPPKGLFS-KL-1	α <b>Β</b> α AKDK-SIVEQ-ISKDGEARL-NF-NSSHPLHKTFTDH ARNGRQFLTQLMQKEQRNYQFDFLRPQHSLFNYFTKI α <b>D</b> DNCFERTQYWEFEKDKDREHDKLVELCKIQFAAIPWI DQVCYRVEWAKFQERERKKEEEEKEKERVAYAQIDWH	SURP domain SURP domain SURP domain		
prp21 DLEQ-RLRR SF3a120 NFPPPTTPE				

## Structural alignment (STAMP)

Prp9 Prp21 Prp11
------------------

E-value	score	bias	E-value	score	bias	exp	Ν	Model	Description
1.5e-52	178.3	9.1	1.5e-52	178.3	9.1	2.3	3	PRP21_like_P	Pre-mRNA splicing factor PRP21 like protein
2e-47	158.4	5.4	6.3e-25	86.4	0.1	2.7	2	Surp	Surp module
4.5e-20	70.5	0.0	1.1e-19	69.3	0.0	1.6	1	ubiquitin	Ubiquitin family

Prp9	Prp21	Prp11						
E-value score bias E-value sc	ore bias exp N Model I	Description						
1.5e-52 178.3 9.1 1.5e-52 17	8.3 9.1 2.3 3 PRP21 like P	Pre-mRNA splicing factor PRP21 like protein						
2e-47 158.4 5.4 6.3e-25 8 4.5e-20 70.5 0.0 1.1e-19 6	6.4 0.1 2.7 2 Surp 9.3 0.0 1.6 1 ubiquitin	Surp module Ubiquitin family						
	Prp9			Prpzi			Prp11	
----------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------	------------------------------------------------------------------	-----------------------------------------------------------------	-----------------------------------------------------------------------------------------	------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------	-------------------------------------------------------------
human_SF3a120 bovin_SF3a120 mouse_SF3a120 yeast_prp21	1 MPAGPVQAVP PPPPVPTEPK QPT MPAGPVQAVP PPPPAATEPK QPT MQAGPVQAVP PPPPVATESK QP1 MEP	TEEEASSK EDSAPSKF TEEEASSK EDSTPSKF IEEEASSK EDPTPSKF EDTQLK ED	VV GIIYPPPEVR VV GIIYPPPEVR VV GIIYPPPEVR	60 NIVDKTASFV NIVDKTASFV NIVDKTASFV IKTTVNYI	301 PEELGARILI QERYEKF PEELGARILI QERYEKF PEELGARILI QERYEKF	GES EEVEMEVESD GES EEVEMEVESD GES EEVEMEVESD	EEDDKQEKAE EPPSQLDQDT EEDEKQEKAE EPPSQLDQDT EEDQEKAE ETPSQLDQDT	360 QVQDMDEGSD QVQDMDEGSD QVQDMDEGSD
human_SF3a120 bovin_SF3a120 mouse_SF3a120 yeast_prp21	61 ARNGPEFEAR IRQNEINNPK FNF ARNGPEFEAR IRQNEINNPK FNF ARNGPEFEAR IRQNEINNPK FNF KQHGVEFENK LLEDER FSF	FLNPNDPY HAYYRHKY FLNPNDPY HAYYRHKY FLNPNDPY HAYYRHKY FIKKDDPL HEYYTKLM	SE FKEGKAQEPS SE FKEGKAQEPS SE FKEGKAQEPS NE PTDTVSGE	120 AAIPKVMQQQ AAIPKVMQQQ AAIPKVMQQQ DN	361 DEEEGQKVPP PPETPMP DEEEGQKVPP PPETPMP DEEEGQKVPP PPETPMP	PPL PPTPDQVIVR   PPL PPTPDQVIVR   PPL PPTPDQVIVR	KDYDPKASKP LPPAPAPDEY KDYDPKASKP LPPAPAPDEY KDYDPKASKP LPPAPAPDEY	420 LVSPITGEKI LVSPITGEKI LVSPITGEKI
human_SF3a120 bovin_SF3a120 mouse_SF3a120 yeast_prp21	121 QQTTQQQLPQ KVQAQVIQET IVF QQASQQQLPQ KVQAQVIQET IVF QQATQQQLPQ KVQAQVIQET IVF DRKSEREIAR	PKEPPPEF EFIADPPS PKEPPPEF EFIADPPS PKEPPPEF EFIADPPS PPDF LFSQYDTG	IS AFDLDVVKLT IS AFDLDVVKLT IS AFDLDVVKLT IS RRDMEVIKLT	180 AQFVARNGRQ AQFVARNGRQ AQFVARNGRQ ARYYAKD.KS	421 PASKMQEHMR IGLLDPR PASKMQEHMR IGLLDPR PASKMQEHMR IGLLDPR	WLE QRDRSIREKQ WLE QRDRSIREKQ WLE QRDRSIREKQ	SDDEVYAPGL DIESSLKQLA SDDEVYAPGL DIESSLKQLA SDDEVYAPGL DIESSLKQLA EDTEIFEGSL DLEQMRLRRV	480 ERRTDIFGVE ERRTDIFGVE ERRTDIFGVE QTGIKLFDSI
human_SF3a120 bovin_SF3a120 mouse_SF3a120 yeast_prp21	181 FLTQLMQKEQ RNYQFDFLRP QHS FLTQLMQKEQ RNYQFDFLRP QHS FLTQLMQKEQ RNYQFDFLRP QHS IVEQMISKDG .EARLNFMNS SHF	SLENYETK LVEQYTKI SLENYETK LVEQYTKI SLENYETK LVEQYTKI FLHKTETD EVAQYKRV	LI PPKGLFSKLK LI PPKGLFTKLK LI PPKGLFSKLK YS FTGQ	240 KEAE NPREVL KEAE NPREVL KEAE NPREVL EIKK SKRTIL	481 ETAIGKKIGE EEIQKPE ETAIGKKIGE EEIQKPE ETAIGKKIGE EEIQKPE KPTNEE EKIVSDQ	EKV TWDGHSGSMA I EKV TWDGHSGSMA I EKV TWDGHSGSMA I GKQ KGGDSKGKKR I	RTQQAAQANI TLQEQIEAIH RTQQAAQANI TLQEQIEAIH RTQQAAQANI TLQEQIEAIH KIRAVGETRL KKSKK	540 KAKGLVPEDD KAKGLVPEDD KAKGLVPEDD
human_SF3a120 bovin_SF3a120 mouse_SF3a120 yeast_prp21	241 DQVCYRVEWA KFQERERKKE EEE DQVCYRVEWA KFQERERKKE EEE DQVCYRVEWA KFQERERKKE EEE DNCFERTQYW EFEKDKDREH DKL	EKEKERVA YAQIDWH EKEKERVA YAQIDWH EKEKERVA YAQIDWH LVELCKIQ FAAIPWDH	FV VVETVDFQPN FV VVETVDFQPN FV VVETVDFQPN FT QVAKFSIP	300 EQGNFPPPTT EQGNFPPPTT EQGNFPPPTT	541 TKEKIGPSKP NEIPQQP TKEKIGPSKP NEIPQQP TKEKIGPSKP NEIPQQP	PPP SSATNIPSSA   PPP SSATNIPSSA   PPP SSATNIPSSA	PPITSVPRPP TMPPPVRTTV PPITSVPRPP AMPPPVRTTV PPITSVPRPP AMPPPVRTTV	600 VSAVPVMPRP VSAVPVMPRP VSAVPVMPRP

Sequence alignment (ClustalW)

D..... 1 1

Prp9		Prp21		Prp11
human_SF3a120 bovin_SF3a120 mouse_SF3a120 yeast_prp21	1 MPAGPVQAVP PPPPVPTEPK MPAGPVQAVP PPPPAATEPK MQAGPVQAVP PPPPVATESK MEP.	QPTEEEASSK EDSAPSKPVV ( QPTEEEASSK EDSTPSKPVV ( QPIEEEASSK EDPTPSKPVV ( EDTQLK ED	GIIYPPPEVR NIVDK GIIYPPPEVR NIVDK GIIYPPPEVR NIVDK 	60 TASFV TASFV TASFV TVNYI
human_SF3a120 bovin_SF3a120 mouse_SF3a120 yeast_prp21	61 ARNGPEFEAR IRQNEINNPK ARNGPEFEAR IRQNEINNPK ARNGPEFEAR IRQNEINNPK KQHGVEFENK LLEDER	FNFLNPNDPY HAYYRHKVSE I FNFLNPNDPY HAYYRHKVSE I FNFLNPNDPY HAYYRHKVSE I FSFIKKDDPL HEYYTKLMNE I	FKEGKAQEPS AAIPK FKEGKAQEPS AAIPK FKEGKAQEPS AAIPK PTDTVSGE	120 VMQQQ VMQQQ VMQQQ DN
human_SF3a120 bovin_SF3a120 mouse_SF3a120 yeast_prp21	121 QQTTQQQLPQ KVQAQVIQET QQASQQQLPQ KVQAQVIQET QQATQQQLPQ KVQAQVIQET DRKSEREIAR	IVPKEPPPEF EFIADPPSIS IVPKEPPPEF EFIADPPSIS IVPKEPPPEF EFIADPPSIS PPDF LFSQYDTGIS	AFDLDVVKLT AQFVA AFDLDVVKLT AQFVA AFDLDVVKLT AQFVA RDMEVIKLT ARYYA	180 RNGRQ RNGRQ RNGRQ KD.KS
human_SF3a120 bovin_SF3a120 mouse_SF3a120 yeast_prp21	181 FLTQLMQKEQ RNYQFDFLRP FLTQLMQKEQ RNYQFDFLRP FLTQLMQKEQ RNYQFDFLRP IVEQMISKDG .EARLNFMNS	QHSLFNYFTK LVEQYTKILI QHSLFNYFTK LVEQYTKILI QHSLFNYFTK LVEQYTKILI SHFLHKTFTD FVAQYKRVYS	PPKGLFSKLK KEAEN PPKGLFTKLK KEAEN PPKGLFSKLK KEAEN TGQ EIKKS	240 PREVL PREVL PREVL KRTIL
human_SF3a120 bovin_SF3a120 mouse_SF3a120 yeast_prp21	241 DQVCYRVEWA KFQERERKKE DQVCYRVEWA KFQERERKKE DQVCYRVEWA KFQERERKKE DNCFERTQYW EFEKDKDREH	EEEKEKERVA YAQIDWHDFV EEEKEKERVA YAQIDWHDFV EEEKEKERVA YAQIDWHDFV DKLVELCKIQ FAAIPWDKFT (	/VETVDFQPN EQGNF /VETVDFQPN EQGNF /VETVDFQPN EQGNF QVAKFSIP	300 PPPTT PPPTT PPPTT

Prp9	Prp21		Prp1	1
1 human_SF3a120 MPA bovin_SF3a120 MPA mouse_SF3a120 MQA yeast_prp21	GPVQAVP PPPPVPTEPK QPTEEEASS GPVQAVP PPPPAATEPK QPTEEEASS GPVQAVP PPPPVATESK QPIEEEASS BDTQL	K EDSAPSKPVV GIIYPPPE K EDSTPSKPVV GIIYPPPE K EDPTPSKPVV GIIYPPPE K ED	60 EVR NIVDKTASFV EVR NIVDKTASFV EVR NIVDKTASFV IKTTVNYI	Conserved Prp21 residues involved in
61 human_SF3a120 ARN bovin_SF3a120 ARN mouse_SF3a120 ARN yeast_prp21 KQH	GPEFEAR IRQNEINNPK FNFLNPNDF GPEFEAR IRQNEINNPK FNFLNPNDF GPEFEAR IRQNEINNPK FNFLNPNDF GVEFENK LLEDER FSFIKKDDF	Y HAYYRHKVSE FKEGKAQE Y HAYYRHKVSE FKEGKAQE Y HAYYRHKVSE FKEGKAQE L HEYYTKLMNE PTDTVSGE	120 EPS AAIPKVMQQQ EPS AAIPKVMQQQ EPS AAIPKVMQQQ EDN	nteractions oetween Prp21 - Prp9 * hydrophobic
121 human_SF3a120 QQT bovin_SF3a120 QQA mouse_SF3a120 QQA yeast_prp21 DRK	TQQQLPQ KVQAQVIQET IVPKEPPPE SQQQLPQ KVQAQVIQET IVPKEPPPE TQQQLPQ KVQAQVIQET IVPKEPPPE SEREIARPPD	* * F EFIADPPSIS AFDLDVVK F EFIADPPSIS AFDLDVVK F EFIADPPSIS AFDLDVVK F LFSQYDTGIS RRDMEVIK	180 KLT AQFVARNGRQ KLT AQFVARNGRQ KLT AQFVARNGRQ KLT ARYYAKD.KS	⁺ ionic * Hbond
181 human_SF3a120 FLT bovin_SF3a120 FLT mouse_SF3a120 FLT yeast_prp21 IVE	* QLMQKEQ RNYQFDFLRP QHSLFNYFT QLMQKEQ RNYQFDFLRP QHSLFNYFT QLMQKEQ RNYQFDFLRP QHSLFNYFT QMISKDG .EARLNFMNS SHPLHKTFT	K LVEQYTKILI PPKGLFSK K LVEQYTKILI PPKGLFSK K LVEQYTKILI PPKGLFSK D FVAQYKRVYS FTG	240 KLK KEAENPREVL KLK KEAENPREVL KLK KEAENPREVL Q EIKKSKRTIL	
241 human_SF3a120 DQV bovin_SF3a120 DQV mouse_SF3a120 DQV yeast_prp21 DNC	CYRVEWA KFQERERKKE EEEKEKERV CYRVEWA KFQERERKKE EEEKEKERV CYRVEWA KFQERERKKE EEEKEKERV FERTQYW EFEKDKDREH DKLVELCKI	YA YAQIDWHDFV VVETVDFQ YA YAQIDWHDFV VVETVDFQ YA YAQIDWHDFV VVETVDFQ YQ FAAIPWDKFT QVAKFSIP	300 QPN EQGNFPPPTT QPN EQGNFPPPTT QPN EQGNFPPPTT P	

Prp9			Prp21				Prp11
yeast_prp21 human_SF3a120 bovin_SF3a120 mouse_SF3a120	61 arngpefea varngpefe arngpefea	mepedtq r irqneinnpk a rirqneinnp r irqneinnpk	lkedikttvn fnflnpndpy kfnflnpndp fnflnpndpy	yikqhgvefe hayyrhkvse yhayyrhkvs hayyrhkvse	nkllederfs fkegkaqeps efkegkaqep fkegkaqeps	120 fikkddplhe aaipkvmqqq saaipkvmqq aaipkvmqqq	Ī
yeast_prp21 human_SF3a120 bovin_SF3a120 mouse_SF3a120	121 yytklmnep qqttqqqlp qqqasqqql qqatqqqlp	t dtvsgedndr q kvqaqviqet p qkvqaqviqe q kvqaqviqet	ksereiarpp ivpkepppef tivpkepppe ivpkepppef	dflfsqydtg efiadppsis fefiadppsi efiadppsis	isrrdmevik afdldvvklt safdldvvkl afdldvvklt	180 ltaryyakdk aqfvarngrq taqfvarngr aqfvarngrq	
yeast_prp21 human_SF3a120 bovin_SF3a120 mouse_SF3a120	181 siveqmisk fltqlmqke qfltqlmqk fltqlmqke	d gearlnfmns q rnyqfdflrp e qrnyqfdflr q rnyqfdflrp	shplhktftd qhslfnyftk pqhslfnyft qhslfnyftk	fvaqykrvys lveqytkili klveqytkil lveqytkili	ftgqe ppkglF ippkgLF ppkglF	240 IKKS SKLKKE.AEN TKLKKE.AEN SKLKKE.AEN	
yeast_prp21 human_SF3a120 bovin_SF3a120 mouse_SF3a120	241 KRTILDNCF PREVLDQVC PREVLDQVC PREVLDQVC	E RTQYWEFEKD Y RVEWAKFQER Y RVEWAKFQER Y RVEWAKFQER	KDREHDKLVE ERKKEEEEKE ERKKEEEEKE ERKKEEEEKE	LCKIQFAAIP KERVAYAQID KERVAYAQID KERVAYAQID	WDKFTQVAKF WHDFVVVETV WHDFVVVETV WHDFVVVETV	300 SIPEDT.EIF DFQPNEQGNF DFQPNEQGNF DFQPNEQGNF	Sequence
yeast_prp21 human_SF3a120 bovin_SF3a120 mouse_SF3a120	301 EGSLDLEQM PPPTTPEEL PPPTTPEEL PPPTTPEEL	R LRRVQTGI G ARILIQERye G ARILIQER G ARILIQER	KL.FD kfgesEE.VE YEkFG YEkFG	SIKPTNEEE. MEVESDEEDD ESEEVE.MEV ESEEVEME.V	KQeka ESDEEDEkqe ESDEEDQeka	360 eeppsqld kaeeppsqld eetpsqld	alignment (HMM fetched from PFAM)

Prp9	Prp21		Prp11
61 yeast_prp21 human_SF3a120 arngp bovin_SF3a120 varngp mouse_SF3a120 arngp	mepedtq lkedikttvn y efear irqneinnpk fnflnpndpy h befea rirqneinnp kfnflnpndp y efear irqneinnpk fnflnpndpy h	120 yikqhgvefe nkllederfs fikkddplhe nayyrhkvse fkegkaqeps aaipkvmqqq yhayyrhkvs efkegkaqep saaipkvmqqq hayyrhkvse fkegkaqeps aaipkvmqqq	Conserved Prp21 residues involved in interactions
121 yeast_prp21 yytklu human_SF3a120 qqttq bovin_SF3a120 qqqasq mouse_SF3a120 qqatq	nnept dtvsgedndr ksereiarpp d qqlpq kvqaqviqet ivpkepppef e qqqlp qkvqaqviqe tivpkepppe f qqlpq kvqaqviqet ivpkepppef e	180 Jflfsqydtg isrrdmevik ltaryyakdk efiadppsis afdldvvklt aqfvarngrq fefiadppsi safdldvvkl taqfvarngr efiadppsis afdldvvklt aqfvarngrq	between Prp21 - Prp9 / 11 * hydrophobic * ionic
181 yeast_prp21 sivequ human_SF3a120 fltqlu bovin_SF3a120 qfltq mouse_SF3a120 fltqlu	miskd gearlnfmns shplhktftd f mqkeq rnyqfdflrp qhslfnyftk l lmqke qrnyqfdflr pqhslfnyft k mqkeq rnyqfdflrp qhslfnyftk l	240 Fvaqykrvys ftgqeIKKS lveqytkili ppkglF SKLKKE.AEN «lveqytkil ippkgLF TKLKKE.AEN lveqytkili ppkglF SKLKKE.AEN	* Hbond *solvent exposed
241 yeast_prp21 KRTIL human_SF3a120 PREVL bovin_SF3a120 PREVL mouse_SF3a120 PREVL	DNCFE RTQYWEFEKD KDREHDKLVE L DQVCY RVEWAKFQER ERKKEEEEKE K DQVCY RVEWAKFQER ERKKEEEEKE K DQVCY RVEWAKFQER ERKKEEEEKE K	***** 300   CKIQFAAIP WDKFTQVAKF SIPEDT.EIF   KERVAYAQID WHDFVVVETV DFQPNEQGNF   KERVAYAQID WHDFVVVETV DFQPNEQGNF   KERVAYAQID WHDFVVVETV DFQPNEQGNF	
301 yeast_prp21 EGSLD human_SF3a120 PPPTT bovin_SF3a120 PPPTT mouse_SF3a120 PPPTT	LEQMR LRRVQTGIKL.FD S PEELG ARILIQERye kfgesEE.VE M PEELG ARILIQERYEkFG E PEELG ARILIQERYEkFG E	360 SIKPTNEEEQeka eeppsqld ESEEVE.MEV ESDEEDEkqe kaeeppsqld ESEEVEME.V ESDEEDQeka eetpsqld	

















β2β3β4prp11GSVGAIQVNYSSEVKENSVDSDDKAK-VPPLIRIVSGLELSDT-KQKGKKFLVIAYEPFESF3a66VKIGKVTKQRDSEMY-PEIAEGIMPRHFFMSAYEQRIEPPDRRWQYLLMAAEPYEβ5a2β6β5a2β6prp11NIAIELPPNEILFSENNDDNNNDGVDELMKKCTFWDAISKLYYVQSF3a66TIAFKVPSREIDKAEGKFTHWNRETKQFFLQFHFKMEKPPAPPSL	Prp9		F	Prp21			Prp11	
	prp11 SF3a66 prp11 SF3a66	β2 GSVG <mark>AIQVN</mark> VKIGK <mark>VTKQ</mark> β5 NIAIELPPNE TIAFKVPSRE	YSSEVKENSVDSDDKAH RDSEMY-PEIAH EILFSENNDDNNNDGVI EIDKAEGKFTHWNRETH	β <b>3</b> K-VPPLIRIVSGLELSI CGIMPRHRFMSAYEQR α <b>2 β6</b> DELNKKCTFWDAISKL XQFFLQFHFKMEKPPAI	DT-KQF IEPPDF <b>β7</b> YYVQ PPSL	<b>β4</b> KGKK <mark>FLVIAY</mark> EPFE RRWQ <mark>YLLMAA</mark> EPYE		

Structural alignment (STAMP)

		0	
100	5	0	
		2	
	~	-	

	1					60	
human_SF3a66	MDFQHRPGGK	TGSGGVASSS	ESNRDRRERL	RQLALETIDI	N	KDPYFMK	
mouse_SF3a66	MDFQHRPGGK	TGSGGVASSS	ESNRDRRERL	RQLALETIDI	N	KDPYFMK	
dicdi_SF3a66	MSEYGK	AGSGGLQSSQ	YDNIDRRERQ	KQLVLEHVDV	S	KDPYIIS	
yeast_prp11	MNYLEGVGSK	KGGGGIASES	QFNLQRRKEV	ESLLSKGENV	PYTFQDEKDD	QVRSNPYIYK	
	61					120	
human_SF3a66	NHLGSYECKL	CLTLHNNEGS	YLAHTQGKKH	QTNLARRAAK	EAKEAPA	.QPAPEK	
mouse_SF3a66	NHLGSYECKL	CLTLHNNEGS	YLAHTQGKKH	QTNLARRAAK	EAKEAPA	.QPAPEQ	
dicdi_SF3a66	NHIGSFECRL	CLTVHNNVGN	YLAHTQGKKH	QTHLARRAAK	EQRENPS	.VSKNNYIQT	
yeast_prp11	NHSGKLVCKL	CNTMHMSWSS	VERHLGGKKH	GLNVLRRGIS	IEKSSLGREG	QTTHDFRQQQ	
	121					180	
human_SF3a66	VKVEVKKFVK	IGRPGYKV	TKQRDSEMGQ	QSLLFQIDYP	.EIAEG	IMPRH	
mouse_SF3a66	VKVEVKKFVK	IGRPGYKV	TKQRDTEMGQ	QSLLFQIDYP	.EIAEG	IMPRH	
dicdi_SF3a66	TRVIHKKTIK	IGRPGYKI	IKQRDSKTGQ	LSLLFQIDYP	.EIESG	LQPRH	
yeast_prp11	KIIEAKQSLK	NNGTIPVCKI	ATVKNPKNGS	VGLAIQVNYS	SEVKENSVDS	DDKAKVPPLI	
	<b>1</b> 81		* *			240	
human_SF3a66	RFMSAYEQRI	EPPDRRWQYL	LMAAEPYETI	AFKVPSREID	KAEGKFWTHW	NRETKQFFLQ	
mouse_SF3a66	RFMSAYEQRI	EPPDRRWQYL	LMAAEPYETI	AFKVPSREID	KAEGKF	FFLQ	
dicdi_SF3a66	RIMSAFEQRV	EQPNKDYQYL	LFAAEPYETI	AFKIPNKEID	RTTGPD	GK	
yeast_prp11	RIVSGLELSD	TK.QKGKKFL	VIAYEPFENI	AIELPPNEIL	FSENND	••••••	
	241					300	
human_SF3a66	FHFKMEKPPA	PPSLPAGPPG	VKRPPPPLMN	GLPPRPPLPE	SLPPPPPGGL	PLPPMPPTGP	
mouse_SF3a66	FHFKMEKPPA	PPSLPAGPPG	VKRPPPPLMN	GLPPRPPLPD	ALPPPPPGGL	PLPPMPPTGP	
dicdi_SF3a66	FFTHWDRN		K	TFTLQLYFKE			
yeast prp11	MDNNND	G	VD	ELNKKCTFWD	AISKLYYVOF	<b>FFKOAEOEOA</b>	

Conserved Prp11 residues involved in interactions between Prp21 - Prp11

\* hydrophobic\* Hbond

Sequence alignment (ClustalW)



Spliceosome has evolved from the Group II introns found in bacteria

Cryo-EM has contributed to understand spliceosome 3D structure and its dynamics

This cell machinery is a dynamic multi-subunit RNA-protein complex Major Spliceosome is composed by 5 snRNPs: U1, U2, U4, U5 and U6 SF3a is a crucial component of U2 and its displacement enables the first catalytic reaction

A deep and clear global understanding of the spliceosome is still lacking



Carroni M, Saibil H. Cryo electron microscopy to determine the structure of macromolecular complexes. Methods. 2016;95:78-85.

Chen W, Moore MJ. Spliceosomes. Cell. 2015;25(5).

Galej WP. Structural studies of the spliceosome: past, present and future perspectives. Biochem Soc Trans. 2018;46(6):1407-1422.

Kuwasako K, He F, Inoue M, Tanaka A, Sugano S, Güntert P et al. Solution Structures of the SURP Domains and the Subunit-Assembly Mechanism within the Splicing Factor SF3a Complex in 17S U2 snRNP. Structure. 2006;14(11):1677-1689.

Lin PC, Xu RM. Structure and assembly of the SF3a splicing factor complex of U2 snRNP. EMBO J. 2012;31(6):1579-90.

Papasaikas P, Valcárcel J. The Spliceosome: The Ultimate RNA Chaperone and Sculptor. Trends in Biochemical Sciences. 2016;41(1):33-45.

Song C, Murata K. Cryo-electron Microscopy for Structural Analysis of Dynamic Biological Macromolecules. Journal of Computer Chemistry, Japan. 2018; 17(1): 38-45.

Vosseberg J, Snel B. Domestication of self-splicing introns during eukaryogenesis: the rise of the complex spliceosomal machinery. Biology Direct. 2017;12(1).

Wahl M, Will C, Lührmann R. The Spliceosome: Design Principles of a Dynamic RNP Machine. Cell. 2009; 136(4): 701-718.

Yan C, Wan R, Shi Y. Molecular Mechanisms of pre-mRNA Splicing through Structural Biology of the Spliceosome. Cold Spring Harb Perspect Biol. 2019:11(1).

Zhan X, Yan C, Zhang X, Lei J, Shi Y. Structures of the human pre-catalytic spliceosome and its precursor spliceosome. Cell Research. 2018; 28(12): 1129-1140.

Zhan X, Yan C, Zhang X, Li L, Lei J, Shi Y. Structure of the human activated spliceosome in three conformational states. Cell. 2018 Mar; 28(3): 307-322.

## 1. How many snRNP does the spliceosome have?

- a) < 5
- b) > 5
- c) 5
- d) Major spliceosome has 5 and minor 3
- e) None of the above
- 2. Which is the best technique to obtain a representation of the whole spliceosome?
  - a) X-ray diffraction
  - b) NMR
  - c) a and b
  - d) Cryo Electron Microscopy
  - e) All of the above
- 3. Regarding Major Spliceosome,
  - a) U2 snRNA interacts with pre-mRNA and U6 snRNA
  - b) U6 snRNA interacts with pre-mRNA and U4 snRNA
  - c) U1 interacts with pre-mRNA
  - d) All of the above
  - e) None of the above
- 4. snRNP has:
  - a) snRNA, 1 Sm/LSm complex and other splicing factors
  - b) snRNA, 2 Sm/LSm complexes and other splicing factors
  - c) snDNA, 1 Sm/LSm complex and other splicing factors
  - d) snDNA, 2 Sm/LSm complexes and other splicing factors
  - e) snRNA, snDNA, 2 Sm/LSm complexes and other splicing factors
- 5. Regarding the Sm complex,
  - a) It is present only in U2, U5 and U6
  - a) It has eight subunits
  - b) It is not well conserved amongst species
  - c) Mutations in this proteins can be lethal
  - d) Doesn't play any crucial role in splicing

- 6. Which protein-complex does SF3a belong to?
  - a) U1

Questions

- b) U2
- c) a and b
- d) U6
- e) none of the above
- 7. Regarding SF3a, select the correct answer:
  - a) Its displacement allows the first splicing reaction to happen
  - b) It is a protein that stabilizes the U2 complex
  - c) a and b
  - d) It is located in the catalytic core of the spliceosome
  - e) All of the above
- 8. Regarding SF3a, select the correct answer:
  - a) Not a lot of crystal structures are available, due to the complexity of the spliceosome
  - b) It is very conserved, due to its essential role in splicing
  - c) a and b
  - d) It is formed by three subunits
  - e) All of the above
- 9. The zinc finger domain of Prp9:
- a) Is also found in SF3a60
- b) Typically binds DNA
- c) Is a beta/beta/alpha structure
- d) Is highly conserved amongst other species
- 10. The SURP domain of Prp21:
  - a) Forms the lariat

b)

C)

d)

e)

- Interacts with the alpha 5 helix of Prp9
- Interacts with the beta 5 helix of Prp9
- Contains only one alpha helix
- Doesn't have a crucial role in U2

## Thank you for your attention!



Ona Baguer, Carlota Cerrillo, Florian Conrad, Clara Escala

