

Univerzita Karlova
Přírodovědecká fakulta

Studijní program:

Experimentální biologie rostlin

Studijní obor:

Buněčná a molekulární biologie rostlin



Bc. Tomáš Kašpar

Role AGO-hook domény histonového chaperonu SPT6L v regulaci genové exprese

The role of the AGO-hook domain of histone chaperone SPT6L in regulation of gene expression

Typ závěrečné práce:

Diplomová práce

Školitel:

Mgr. Vojtěch Čermák, Ph.D.

Konzultant:

RNDr. Lukáš Fischer, Ph.D.

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, dne

Bc. Tomáš Kašpar

Poděkování

Především bych chtěl poděkovat svému školiteli Vojtěchu Čermákovi za jeho nekonečnou trpělivost a skvělé vedení během vypracování tohoto projektu. Dále patří dík i vedoucímu naší laboratoře a konzultantovi této práce Lukáši Fischerovi za podněcující konzultace, za možnost být součástí skvělého týmu a za možnost podílet se na tomto projektu. Chtěl bych také poděkovat všem členům naší laboratoře za úžasný kolektiv, jaký tvoří, a za to, že se na ně mohu kdykoliv obrátit se svými všetečnými otázkami.

Dále bych chtěl poděkovat své rodině a přátelům za jejich neustálou oporu. A mé přítelkyni patří dík za obrovskou pomoc a neutuchající podporu při psaní této práce.

Dále bych chtěl poděkovat organizaci GA UK, která poskytla finance nezbytné pro vypracování této práce v rámci projektu č. 382521.

Anotace

AGO-hook domény přítomné u některých eukaryotických proteinů slouží k vazbě proteinů z rodiny ARGONAUTE (AGO). Proteiny AGO fungují v mnoha procesech regulujících expresi genů za pomoci jimi navázané malé RNA (sRNA). sRNA slouží k rozpoznání cílového komplementárního transkriptu. Tato práce si klade za cíl zjistit roli předpokládané AGO-hook domény proteinu SPT6L. SPT6L je transkripční elongační faktor z komplexu RNA polymerázy II (Pol II), kde slouží jako histonový chaperon a účastní se epigenetického značení histonů. SPT6L je u *Arabidopsis thaliana* jedním ze dvou paralogů proteinu SPT6, který je charakterizován právě přítomností AGO-hook domény, což je ryze rostlinné specifikum. Funkce této domény není u rostliny *A. thaliana* známá. Přesto by se dalo předpokládat, že tato doména bude zajišťovat navádění proteinů AGO do komplexu Pol II, čímž poté může umožnit regulaci modifikací chromatinu, či kotranskripčně ovlivňovat transkripty Pol II cílené sRNA.

Tato práce nastiňuje funkci AGO-hook domény proteinu SPT6L u rostliny *A. thaliana*. Získané výsledky ukazují interakci AGO-hook domény proteinu SPT6L s AGO a její vliv na expresi a na zpracování transkriptů.

Klíčová slova

SPT6L, AGO-hook, transkripce, chromatin, *Arabidopsis thaliana*

Annotation

AGO-hook domains present in some eucaryotic proteins are crucial for a binding family of ARGONAUTE proteins (AGO). These AGO proteins are essential in many biological processes regulating gene expression by small RNA (sRNA), which is complementary to the gene that is supposed to be influenced. This thesis claims to find the function of the putative AGO-hook domain of the protein SPT6L. SPT6L is an elongation factor and histone chaperon of a complex of RNA polymerase II (Pol II) where it is acting in the epigenetic marking of histones. SPT6L of *Arabidopsis thaliana* is one of two paralogues of SPT6 proteins, that is characteristic of the presence of the AGO-hook domain. This is a plant specificity of this protein. What's more, the function of this domain remains unknown. Despite this, it could be assumed that this domain is necessary for binding of AGO proteins in the complex of Pol II, and by these interactions, it can enable guidance of regulation of chromatin modification, or it can co-transcriptionally influence nascent transcripts of Pol II by the sRNA.

This thesis casts light on the function of the AGO-hook domain of SPT6L in *A. thaliana* in processes of gene regulation and protein interactions. It claims to confirm the interaction of the AGO-hook domain of SPT6L protein with AGO proteins and its influence on gene expression and transcript processing.

Keywords

SPT6L, AGO-hook, transcription, chromatin, *Arabidopsis thaliana*

Obsah

Prohlášení.....	3
Poděkování.....	4
Anotace.....	5
Annotation.....	6
1 Seznam zkratk.....	10
2 Úvod	13
2.1 Cíle práce.....	14
3 Přehled literatury	15
3.1 RNAi.....	15
3.1.1 Historický úvod RNAi	15
3.1.2 Komponenty RNAi.....	15
3.1.3 PTGS	19
3.1.4 TGS	19
3.1.5 Proteiny AGO	24
3.1.6 AGO-hook doména	26
3.2 Faktory ovlivňující elongaci a transkripty Pol II	29
3.2.1 Chromatin ovlivňující funkci Pol II	29
3.2.2 Alternativní sestřih	30
3.2.3 Transkripční elongační faktory Pol II	31
4 Materiál	34
4.1 Práce s rostlinným materiálem	34
4.1.1 Křížení rostlin	34
4.2 Chemikálie	34
4.3 Protilátky.....	35
4.4 Přístroje.....	35
4.5 Další materiál.....	35

5	Metody	37
5.1	Genotypování rostlin <i>A. thaliana</i>	37
5.1.1	Izolace rostlinné DNA.....	37
5.1.2	PCR	37
5.1.3	Elektroforéza DNA	39
5.2	Fylogenetická analýza	39
5.2.1	Vyhledávání sekvencí	39
5.2.2	Vyhledávání AGO-hook domén	39
5.2.3	Tvorba fylogenetických stromů.....	40
5.3	Příprava RNA pro analýzu transkriptomu	40
5.3.1	Izolace rostlinné RNA	40
5.3.2	Elektroforéza RNA.....	41
5.3.3	Dodatečné kontroly RNA.....	42
5.4	Bioinformatická analýza transkriptomických dat.....	42
5.5	Proteomika	44
5.5.1	Izolace rostlinných proteinů	44
5.5.2	Koimunoprecipitace	45
5.5.3	Detekce proteinů na western blotu	46
5.5.4	Analýza proteinů pomocí MS	49
5.6	Další doplňující materiály a metody.....	50
6	Výsledky	51
6.1	Fylogeneze SPT5L a SPT6L	51
6.2	Proteomika	55
6.2.1	Western-blot	55
6.2.2	Koimunoprecipitace	56
6.3	Transkriptom	58
6.3.1	Diferenciální exprese genů	60

6.3.2	Analýza využití exonů	65
6.3.3	Analýza alternativního sestřihu	66
7	Diskuze	68
8	Závěr	75
9	Reference	76
	Přílohy.....	86

1 Seznam zkratek

A3'/5'SS	Alternativní 3'/5' sestřihové místo
AGO	ARGONAUTE
ATRX	ARABIDOPSIS TRITHORAX
BAH	Bromo adjacent homology
BES1	BRI1 EMS SUPPRESSOR 1
bp	pár bazí
CLSY	SNF2 DOMAIN-CONTAINING PROTEIN CLASSY
CMT	CHROMOMETHYLASE
CTD	C-terminální konec
DCL	DICER-LIKE
DMS3	DEFECTIVE IN MERISTEM SILENCING 1
DRD1	DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1
DRM	DOMAINS REARRANGED METHYLTRANSFERASE
dsRNA	dvouvláknová RNA
ELF6	EARLY FLOWERING 6
ES	Exon Skipping
FACT	FACILITATES CHROMATIN TRANSCRIPTION
FWA	FLOWERING WAGENINGEN
GFP	Green Fluorescent Protein
GO	Gene Ontology
HEN1	HUE ENHACER 1
htsiRNA	siRNA spojené s heterochromatinem
HxKy(me/ac)z	z-metylace/acetylace lysinu y na histonu H x
HYL1	HYPONASTIC LEAVES 1
ChIP	Chromatinová imunoprecipitace
IDN2-IDP	INVOLVED IN DE NOVO 2-IDN2 PARALOGUE
IR	Retence intronu
IWS1	INTERACT WITH SPT6

JMJ	JUMONJI
LLPS	liquid-liquid phase separation
LT	laboratorní teplota
MET1	METHYLTRANSFERASE 1
MID	Middle
miRNA	micro-RNA
mRNA	messenger RNA
MS	hmotnostní spektrometr(ie)
MXE	mutually exclusive exons
natsiRNA	natural antisense transkript siRNA
NERD	NEEDED FOR RDR2-INDEPENDENT DNA METHYLATION
NRPE1	NUCLEAR RNA POLYMERASE E1
nt	nukleotid
NTD	N-terminální konec
PAGE	polyakrylamidová gelová elektroforéza
PAZ	Piwi Argonaut and Zwillie
PCR	polymerázová řetězová reakce
PIWI	P-element Induced Wlmpy
Pol	Polymeráza
pri-miRNA	primární miRNA
PSSM	position-specific score matrix
PTGS	posttranskripční umlčování genů
RACE	Rapid amplification of cDNA ends
RdDM	RNA-directed DNA methylation
RDM1	RNA-DIRECTED DNA METHYLATION 1
RDR	RNA-directed RNA polymerase
REF6	RELATIVE OF EARLY FLOWERING 6
RIN	RNA Integrity Number
RNAi	RNA interference
SDE3	SILENCING DEFECTIVE 3

SE	SERRATE
SET	Su(var)3-9, Enhancer-of-zeste and Trithorax
SGS3	SUPPRESSOR OF GENE SILENCING 3
SH2	Src homology 2
SHH1	SAWADEE HOMEODOMAIN HOMOLOG 1
siRNA	malá interferující RNA
SPT4/5/6(L)	SUPPRESSOR OF TY INSERTION 4/5/6(-LIKE)
SRA	SET and RING associated
sRNA	malá RNA
ssRNA	jednovláknová RNA
SUVH	SU(VAR)3-9 HOMOLOG
tasiRNA	trans-acting siRNA
TE	transpozibilní element
TEF	transkripční elongační faktor
TES	místo konce transkripce
TFIIS	TRANSCRIPT ELONGATION FACTOR IIS
TGS	transkripční umlčování genů
TSS	místo počátku transkripce
VCS	VARICOSE
VIM1	VARIANT IN METHYLATION 1
vsiRNA	virová siRNA
WB	Western Blot

2 Úvod

Funkcí Pol II je především zajišťování transkripce většiny genů u eukaryotických organismů, ale i ovlivňování stavu chromatinu. Jedná se o DNA-dependentní RNA polymerázu a její funkce může být ovlivněna na mnoha úrovních. Mezi ty nejzásadnější faktory ovlivňující funkci této polymerázy je stav chromatinu, transkripční elongační faktory (TEF), ale také i iniciace transkripce, jakožto prvotní krok pro přepis daného lokusu. Stav chromatinu ovlivňuje aktivitu polymerázy a efektivitu transkripce a sestřih poté moduluje, jak vznikající transkript bude ve výsledku vypadat a jaká bude jeho stabilita. TEF ovlivňují funkci a procesivitu samotné Pol II. Jedním z takových TEF u rostliny *Arabidopsis thaliana* je i protein SPT6L. Ten je jedním z mnoha eukaryotických proteinů nesoucích předpokládanou AGO-hook doménu (Gu *et al.*, 2012). Jedná se o doménu obsahující WG/GW motivy, které jsou zásadní pro vazbu proteinů AGO. Proteiny AGO dokáží navázáním sRNA cílit komplementární vlákno RNA a pomocí jeho rozpoznání způsobit jeho štěpení a tím posttranskripční umlčování, nebo dokáží zprostředkovat umlčování pomocí změn v chromatinu, jakou je například metylace DNA.

SPT6L je konzervovaný elongační faktor z komplexu Pol II a je také histonovým chaperonem, který napomáhá v rekonstituci chromatinu za polymerázou po transkripci (Duina, 2011). V rostlině *A. thaliana* jsou přítomny dva jeho paralogy SPT6 a SPT6L, avšak pouze protein SPT6L nese AGO-hook doménu.

Dalšími proteiny s přítomnou AGO-hook doménou u rostliny *A. thaliana* jsou například proteiny NRPE1, největší podjednotka komplexu Pol V, anebo její elongační faktor SPT5L (Karlowski *et al.*, 2010). Oba tyto proteiny jsou zásadními hráči v procesu metylace DNA řízené sRNA (RdDM). K jejich funkci jim napomáhají právě jejich AGO-hook domény, které jim umožňují zachytávat proteiny AGO a skenovat vznikající transkript.

Tato práce si klade za cíl odhalit roli AGO-hook domény proteinu SPT6L. Tato doména by mohla hrát roli při značení chromatinu pomocí sRNA. Měla by mít schopnost přivádět efektorové proteiny pro značení chromatinu, tedy proteiny AGO do komplexu Pol II. Tato doména by zároveň mohla mít specifickou funkci při samotné transkripci, kdy by kotranskripčně mohla ovlivňovat vznikající transkript.

Pro rostliny je specifická přítomnost dvou dodatečných polymeráz příbuzných Pol II, jejichž funkce se samostatně specializovala právě pro dráhu RdDM, těmi jsou Pol IV a V. Proto přítomnost domény AGO-hook proteinu SPT6L, který pracuje právě s Pol II, poukazuje na funkci Pol II v procesech spojených s sRNA. Funkce Pol II jakožto efektorového proteinu v RNAi je poměrně neprobádaným a

poněkud kontroverzním tématem. Výzkum domény AGO-hook v proteinu SPT6L si tak neklade za cíl pouhé popsání funkce této domény, ale mohl by odpovědět na mnohem širší otázku, jak se Pol II zapojuje do procesů RNA interference (RNAi) i na jiné úrovni než na produkci sRNA, kde už tyto funkce popsány byly.

2.1 Cíle práce

Cílem této práce je zjistit funkci AGO-hook domény histonového chaperonu a transkripčního elongačního faktoru proteinu SPT6L u rostliny *A. thaliana*. Cílem je ukázat důležitost AGO-hook domény v proteinu SPT6L v evolučním i funkčním měřítku. Předpokládanou funkcí je vazba proteinů AGO, proto cílem je i potvrzení této interakce. Dalším cílem je ukázat jak se AGO-hook proteinu SPT6L podílí na změnách v transkriptomu a vyvodit, jak ovlivňuje funkci Pol II.

3 Přehled literatury

3.1 RNAi

3.1.1 Historický úvod RNAi

RNA interference je proces umlčování genů zprostředkovaný pomocí malých RNA konzervovaný napříč eukaryoty (Cerutti & Casas-Mollano, 2006). Jako první se v rostlinném organismu s fenoménem RNAi setkali dvě práce snažící se navýšit množství pigmentu v květech petúnií, kde byly použity konstrukty nesoucí geny pro enzymy syntézy flavonoidů pod velice silnými promotory. Předpokládaným výsledkem bylo navýšení celkového množství pigmentů, avšak výsledek získaný byl často opačný (Napoli *et al.*, 1990; van der Krol *et al.*, 1990).

Posléze byla za objev RNAi udělena Nobelova cena v roce 2006, kdy bylo ukázáno, že při dodání dvouvláknové RNA (dsRNA) dochází ke změnám v translaci homologního genu a k odchýlkám ve fenotypu *Caenorhabditis elegans* (Fire *et al.*, 1998). Fire *et al.* v roce 1998 demonstroval, že dsRNA v *C. elegans* spouští specifickou blokadu exprese homologní RNA a tento fenomén je znám jako RNAi. Jen o pár měsíců později došel ke stejnému závěru i Waterhouse na rostlinném modelu *Nicotiana tabaccum*. Ten zjistil, že exprese sense a zároveň antisense virové RNA, která dokáže tvořit duplex, vyvolá u rostliny rezistenci proti virové nákaze (Waterhouse *et al.*, 1998). O rok později byly nalezeny sRNA zodpovědné za RNAi dlouhé asi 25 nt (Hamilton & Baulcombe, 1999). V roce 2000 se ukázalo, že tento fenomén RNAi je přítomen také u *Drosophila melanogaster* (Hammond *et al.*, 2000).

Pro definování funkce RNAi u rostlin, byly zásadní experimenty s rostlinou *A. thaliana*. Ty ukázaly, že zásadními proteiny podílející se na RNAi jsou DCL (DICER-LIKE), RDR (RNA-DEPENDENT RNA POLYMERASE), AGO (ARGONAUTE) ve spolupráci s malými RNA (sRNA).

3.1.2 Komponenty RNAi

3.1.2.1 Proteiny v RNAi

Na vzniku sRNA se podílí celá řada proteinů. Pro jednovláknové sRNA (ssRNA), které nejprve vyžadují dosyntetizování druhého vlákna, je zásadní funkce proteinů RDR. RDR jsou důležité komponenty ve virech, rostlinách a například u *C. elegans*, avšak většina živočichů tyto enzymy v průběhu evoluce ztratila. *A. thaliana* obsahuje celkem šest genů pro tyto enzymy, které jsou ve dvou kladech rozdělené podle sekvence jejich katalytické domény na klady RDR1/2/6 a RDR 3/4/5 (Zong *et al.*, 2009).

RDR1/2/6 vytváří dsRNA z ssRNA prekurzorů vzniklých různými cestami, které poté mohou být procesovány na siRNA. RDR2 specificky zajišťuje přepis transkriptů Pol IV a RDR6 přepis transkriptů Pol II. Specifické procesy budou popsány v kapitolách níže. Proteiny druhého kladu RDR3/4/5 mají dosud blíže nepopsané funkce, ale v rostlinných genomech je vždy alespoň jeden gen tohoto kladu přítomen (Willmann *et al.*, 2011).

Po dosyntetizování druhého vlákna jsou dsRNA naštěpeny na sRNA pomocí proteinů DCL. Proteiny DCL jsou homologní k živočišným proteinům Dicer1/2 a rostlinné genomy obsahují alespoň čtyři odlišné DCL proteiny (DCL1-4; Mukherjee *et al.*, 2013). DCL1 je zodpovědný za akumulaci většiny 21-nt dlouhých miRNA (Reinhart *et al.*, 2002). DCL4 je producentem 21-nt protivirových siRNA a endogenních siRNA. Funkce DCL2 a DCL4 je částečně redundantní. DCL2 dokáže stimulovat tranzitivitu a vytváří tasiRNA (Gascioli *et al.*, 2005; Bouché *et al.*, 2006; Parent *et al.*, 2015). DCL3 produkuje 24-nt dlouhé siRNA asociované s repeticemi a transponovatelnými elementy (TE) a uplatňuje se především v dráze RdDM pro transkripční umlčování (TGS; Pontes *et al.*, 2006).

Po štěpení proteiny DCL vznikají duplexy sRNA s 2 nt dlouhým přesahem na 3' konci, které jsou rozpoznávány proteinem HEN1 (HUE ENHANCER 1). Tento protein metyluje dvouvláknové miRNA a siRNA na 2'-OH skupinách ribózy 3' terminálních nukleotidů. Tato metylace slouží jako ochrana před uridinylací tohoto konce, která je jinak signálem k degradaci dané RNA. Jedná se o univerzální krok v biogenezi sRNA (Li *et al.*, 2005; Yu *et al.*, 2005; Yang *et al.*, 2006).

Poté co vzniknou duplexy sRNA pomocí DCL proteinů, které jsou stabilizovány metylací, je do komplexu s proteinem AGO použito pouze jedno vlákno, takzvaný „guide“. Druhé vlákno takzvaný „passenger“ se dalších procesů neúčastní. Průvodní vlákno není zřejmě zvoleno náhodně, ale na základě termodynamických vlastností, podle stability konců dané sRNA a při rozpoznání sRNA daným AGO je zásadní i povaha 5'-koncového nukleotidu. Proteiny AGO jsou schopné specificky vázat sRNA nejrůznějšího původu a vlastností, a tím specifikovat svou funkci (Khvorova *et al.*, 2003; Mi *et al.*, 2008; Takeda *et al.*, 2008).

3.1.2.2 Druhy sRNA využívané v RNAi

Malé RNA jsou nezbytným činitelem v RNAi, který zajišťuje specifitu působení této dráhy pomocí komplementarity sRNA s daným cílem. Obecně jsou funkce a druhy sRNA odlišovány podle způsobu jejich vzniku, délky a následného efektu. V RNAi klasicky působí sRNA dlouhé 21-24 nt a jsou jimi microRNA (miRNA) a siRNA (small interfering RNA), v rámci kterých lze rozlišovat dále na tasiRNA (trans-acting siRNA), htsiRNA (siRNA spojené s heterochromatinem), natsiRNA (natural antisense transcript siRNA), vsiRNA (virové siRNA) a phasiRNA (phased siRNA; Ghildiyal & Zamore, 2009).

Zpravidla platí, že miRNA jsou odvozeny z dobře definovaných genomových lokusů, jsou evolučně konzervované a jejich přirozenou funkcí je regulace genové exprese. siRNA na druhé straně jsou často produkovány ze sekvencí TE, virů a heterochromatinovými sekvencemi a jejich prekurzory jsou různými cestami vzniklé dvouvláknové úseky RNA a málokdy jsou konzervované. Většinou zprostředkovávají umlčení těch samých lokusů, ze kterých pochází, tedy *in cis*.

Velmi dobře probádanou skupinou sRNA jsou miRNA (obr. 1A). Jedná se o třídu regulačních RNA dlouhých 20-22 nt, které jsou kódovány endogenními *MIR* geny a jsou transkribovány pomocí Pol II do takzvaných pri-miRNA. Takto přepsané tvoří specifickou stem-loop strukturu, kde stem část je tvořena na základě vnitřní komplementarity dané RNA. Po transkripci jsou tyto pri-miRNA štěpeny proteinem DCL1, kdy se vyštěpí duplex miRNA (Reinhart *et al.*, 2002; Dong *et al.*, 2008), z duplexu je poté zainkorporováno guide vlákno do komplexu s proteinem AGO (především AGO1) a může cílit komplementární transkripty. miRNA může negativně regulovat množství cílových proteinů různými způsoby: represí translace (Brodersen *et al.*, 2008), štěpením cílového transkriptu (Baumberger & Baulcombe, 2005) nebo v ojedinělých případech i pomocí inhibice transkripce metylací DNA (Khraiweh *et al.*, 2010). Pro správnou tvorbu a zpracování pri-miRNA jsou zásadní dva proteiny SE (*SERRATE*) a *HYL1* (*HYPONASTIC LEAVES 1*; Dong *et al.*, 2008).

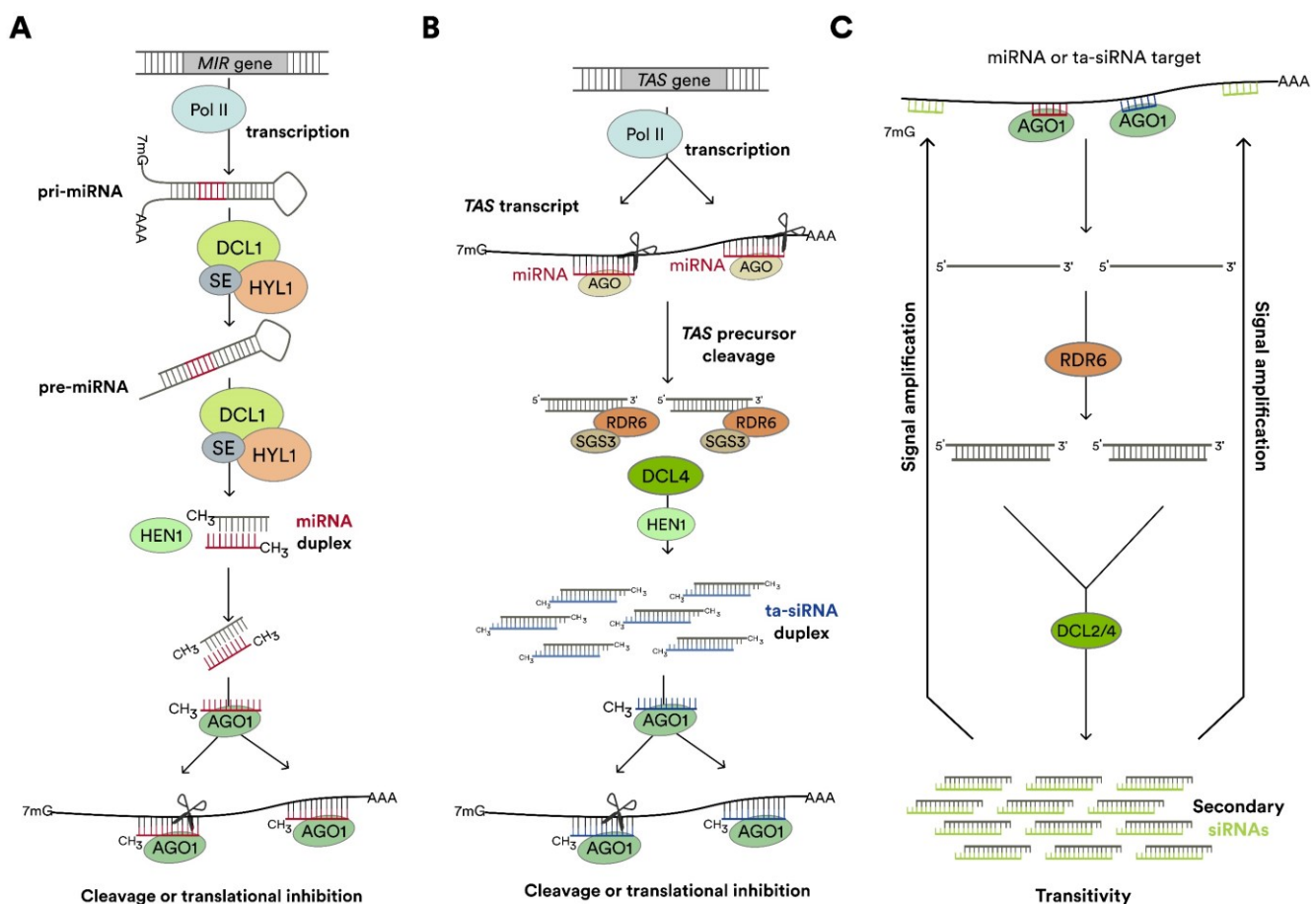
Skupina natsiRNA jsou sRNA, které vznikají z endogenních transkriptů genů *NAT* (Natural Antisense Transkript), a to zejména při stresových situacích. Jejich prekurzory vznikají párováním sense a antisense vláken. Pro jejich vyštěpení z takto vzniklé dsRNA jsou zásadní proteiny DCL1 a DCL3. DCL1 je štěpí na 20-22 nt a DCL3 na 23-28 nt dlouhé sRNA. Převážně působí *in cis* na geny posttranskripčně štěpením cílových sekvencí mRNA. Pro 20-22 nt dlouhé natsiRNA tuto funkci zastává protein AGO2, avšak pro delší natsiRNA o délce 23-26 nt mohou vytvářet komplexy s proteiny AGO1/4, které jsou typické pro štěpení transkriptů a pro metylaci DNA (Zhang *et al.*, 2012).

Další důležitou skupinou jsou sekundární siRNA (obr. 1C), ty vznikají z transkriptů, které jsou primárně cíleny jinou sRNA. Proces, kdy sekundární sRNA vznikají z přilehlé oblasti, která není cílena primární siRNA se nazývá tranzitivita. Proces tvorby sekundárních siRNA vyžaduje proteiny DCL2/4 a syntéza druhého komplementárního vlákna k ssRNA probíhá za pomoci proteinu RDR6. Také je potřeba, aby primární sRNA byla dlouhá 22 nt, buď jako produkt DCL2 nebo vzniklá v procesu biogeneze miRNA (Moissiard *et al.*, 2007; Sanan-Mishra *et al.*, 2021)

Podskupinou sekundárních siRNA jsou tasiRNA (obr. 1B). Ty vznikají z transkriptů nekódujících genů *TAS* (trans-acting siRNA). *TAS* transkripty rozštěpené AGO1 s komplementární miRNA jsou přepsány RDR6 a sekundární siRNA z *TAS* dsRNA jsou poté generovány proteinem DCL4. Jak název napovídá tyto transkripty většinou působí na důležité regulační geny *in trans*. Proteinem důležitým

pro tvorbu tasiRNA je SGS3 (SUPPRESSOR OF GENE SILENCING 3). Tento protein stabilizuje intermediáty RNA při tvorbě tasiRNA (Yoshikawa *et al.*, 2005).

Neméně důležitou skupinou a také velice početnou skupinou sRNA jsou také htsiRNA, ty budou popsány v kapitolách níže. V neposlední řadě mohou sRNA vznikat také pomocí exogenních vlivů jako jsou transgeny či viry. Poté co virus napadne rostlinu, může být jeho RNA přepsána do dsRNA a naštěpena proteiny DCL2/4 na siRNA, které mohou potlačit infekci daným virem. Mechanismům RNAi se viry mohou bránit pomocí supresorů umlčování jako například P38 u viru TCV, který potlačuje funkci DCL4 (Deleris *et al.*, 2006; review o supresorech umlčování k nahlédnutí Burgyán & Havelda, 2011).



Obrázek 1: Schéma s několika příklady vzniku různých druhů sRNA. **A** miRNA vznikají přepisem *MIR* genů a dále se na jejich biogenezi účastní proteiny SE a HYL1 a jejich štěpení na efektorové miRNA zprostředkovává protein DCL1. **B** Druhou možností vzniku sRNA je přepis *TAS* genů, které jsou štěpeny po rozpoznání jinou sRNA. Pro jejich biogenezi jsou zásadní proteiny RDR6, SGS3, DCL4, tasiRNA jsou specifickým typem sekundárních siRNA. **C** Sekundární siRNA mohou vést k amplifikaci signálu, či rozšíření cílené sekvence v procesu tranzitivity. Přepis komplementárního vlákna zajišťuje protein RDR6 a štěpení na siRNA se děje pomocí proteinů DCL.

3.1.3 PTGS

Post-transkripční umlčování (PTGS) je proces, kdy se ovlivňuje cílový transkript v buňce poté, co je přepsán do RNA. K tomu dochází štěpením cílového transkriptu rozpoznáním komplementární sRNA, která je asociovaná s proteinem AGO. Druhou možností PTGS je blokáce translace proteiny AGO.

Štěpit cílové RNA dokáží samy proteiny AGO s asociovanými sRNA, neboť některé z nich mají prokázané endonukleolytické vlastnosti. Nejčastějším efektorovým proteinem štěpení cílové RNA je AGO1 či AGO7 (Baumberger & Baulcombe, 2005; Montgomery *et al.*, 2008). Štěpení cílového transkriptu často může sloužit i k tvorbě dalších biogenních siRNA, sekundárních siRNA, jejichž biogeneze je popsána výše.

Druhou možností, jak může PTGS fungovat, je blokování translace. Tento proces je lépe prostudovaný u živočichů, kde je využíváno proteinů, které nesou AGO-hook doménu, a ty ustanovují interakci s poly-A koncem cílové mRNA a odstraněním 5'-čepičky dochází k postupné degradaci této mRNA (Neumeier & Meister, 2021). U rostlin tento jev byl pozorován také, kdy při rozpoznání komplementárního transkriptu malou RNA na proteinu AGO dochází k represi translace namísto ke štěpení transkriptu. K tomu je využíváno dalších faktorů, jako například VCS (VARICOSE). Ten je důležitý pro odstranění čepičky z RNA (Brodersen *et al.*, 2008). Schopnost reprimovat translaci a ovlivňovat růst rostliny mají i 22 nt dlouhé siRNA vzniklé aktivitou RDR6 a DCL2 nesené na proteinu AGO1 (Wu *et al.*, 2020).

3.1.4 TGS

Proces transkripčního umlčování (TGS) je změna epigenetického značení chromatinu, jež vede ke změněné (snížené) přístupnosti DNA pro transkripční faktory a Pol II, čímž se může ovlivnit míra exprese dané sekvence/genu. Hlavní modifikací chromatinu ovlivňující úroveň exprese daného lokusu je metylace DNA a modifikace histonů.

3.1.4.1 Methylace DNA

Některé rostliny jako například kukuřice (*Zea mays*) mohou ve svém genomu obsahovat i přes 85 % TE, které jsou roztroušené po celém genomu. Ty je potřeba umlčovat, aby byla udržena stabilita genomu (Schnable *et al.*, 2009). Methylace DNA (konkrétně cytosinů) je obecným mechanismem, jak udržovat TE/heterochromatin umlčené. K této metylaci může docházet v různých sekvenčních kontextech DNA, a to konkrétně v CG, CHH i CHG (H = A, T, C). Methylace DNA u *A. thaliana* je udržována v různých drahách (obr. 2A). Proteiny zajišťující metylaci DNA jsou METHYLTRANSFERASE 1 (MET1), CHROMOMETHYLASE 2/3 (CMT2/3) a DOMAINS REARRANGED METHYLTRANSFERASE (DRM1/2). CG kontexty jsou v udržovány metylované

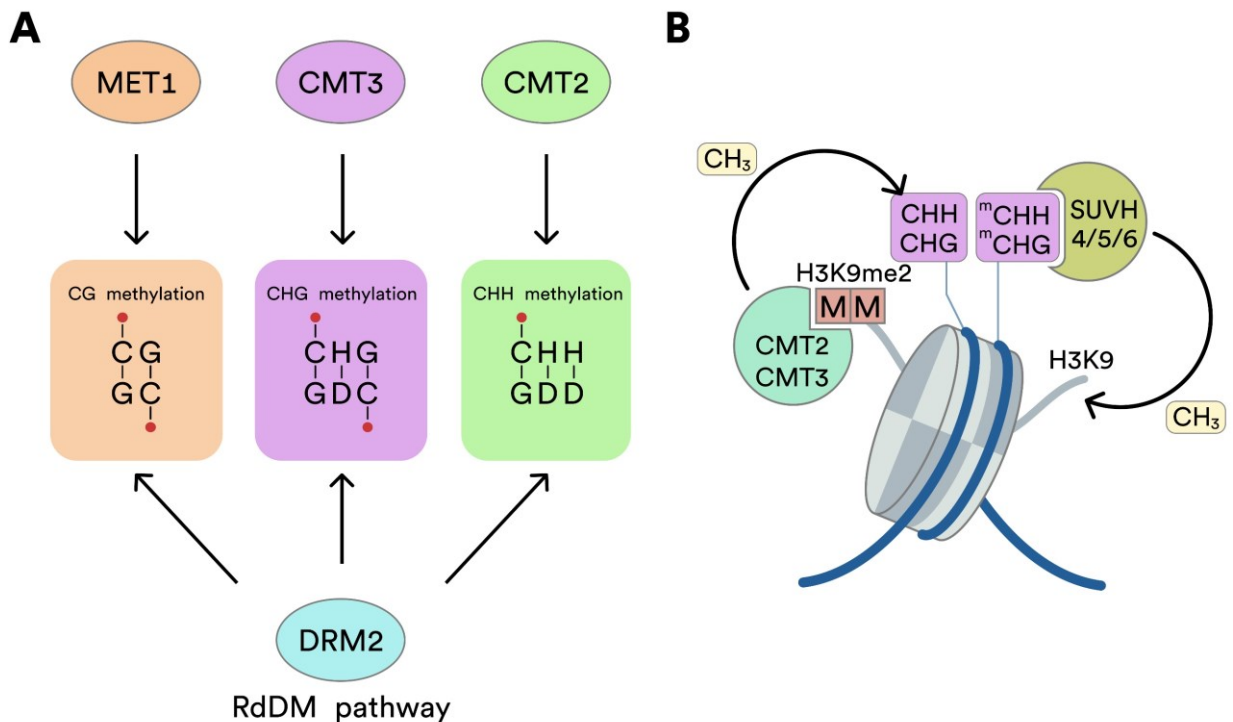
proteinem MET1, zatímco CMT3 primárně cílí metylace na kontextu CHG a kontexty CHH jsou metylovány proteinem CMT2, popřípadě DRM1/2, který je schopen *de novo* metylovat všechny zmíněné kontexty (Kishimoto *et al.*, 2001; Zemach *et al.*, 2013; Stroud *et al.*, 2014).

MET1 udržuje metylaci v symetrickém kontextu CG. V udržování metylace DNA je důležitý partner VIM1 (VARIANT IN METHYLATION 1), který navádí MET1 na hemimetylovanou DNA v symetrickém CG kontextu po replikaci (Kim *et al.*, 2014).

Pro udržování metylace v CHG kontextu existuje zpětnovazebná smyčka propojená s histonovou značkou H3K9me2. CMT3 váže tuto značku na heterochromatinu doménou BAH (Bromo adjacent homology) a chromo-doménou a zavádí zde metylaci DNA v kontextu CHG (Du *et al.*, 2012). Na tuto metylaci DNA se poté váží H3K9 methyltransferázy SUVH4 (SU(VAR)3-9 HOMOLOG 4) a její homology SUVH5/6 pomocí SRA (SET and RING-associated) domény a mohou dimetylovat H3K9. Tím se uzavírá zpětnovazebná smyčka udržovací metylace DNA (CHG) a H3K9me2 (obr. 2B; Johnson *et al.*, 2007; Gouil & Baulcombe, 2016).

Methylace CHH kontextu zavádí na heterochromatinovou DNA u *A. thaliana* CMT2 a to skrze navázání na histonové značky H3K9me2. Homology tohoto proteinu však nebyly nalezeny u *Z. mays*, což naznačuje, že metylace CHH kontextu v těchto rostlinách je plně závislá na RdDM dráze popsané níže (Zemach *et al.*, 2013).

Tyto výše zmíněné methyltransferázy spíše udržují stálou metylaci, na rozdíl od dráhy RdDM, která je schopna tyto značky zavádět i *de novo* proteinem DRM2. RdDM dráha je zásadní pro udržení stability genomu, tím že udržuje metylaci na starých inzercích TE a nové inserce metyluje *de novo*. Tato dráha využívá sRNA k navádění metylačních značek na molekulu DNA. Rozdělujeme ji na dva základní typy, a to kanonickou a nekanonickou, podle způsobu vzniku sRNA (Erdmann & Picard, 2020).



Obrázek 2: Schémata metylace DNA. **A** Proteiny metylace DNA pro různé kontexty. MET1 zajišťuje metylaci CG kontextu DNA, CMT3 na CHG a CMT2 na CHH. Protein DRM2 je zásadní pro metylaci DNA v procesu RdDM. **B** Zpětnovazební smyčka metylace DNA. Pro její fungování jsou za potřebí proteiny SUVH4/5/6, které se váží na metylovanou DNA a navádí dimetylaci na H3K9. H = A, T, C; D = A, T, G

3.1.4.2 Kanonická RdDM

Kanonická dráha je nejlépe charakterizovanou dráhou RdDM. Jedná se o dráhu, která udržuje metylované nejčastěji již umlčené sekvence DNA, tedy ty, s již přítomnou metylací DNA a heterochromatinovými značkami ve své sekvenci. Tato dráha se úzce opírá o prvotní syntézu htsiRNA. Ty jsou produkovány specifickou polymerázou Pol IV. Její transkripty jsou zpracovány na sRNA, které navádějí metylaci DNA podle nascentního transkriptu Pol V (obr. 3). Tyto dvě rostlinně specifické polymerázy (Pol IV a Pol V) mají podjednotky, které jsou paralogní či totožné s Pol II. Bylo ukázáno, že jsou odvozeny od komplexu Pol II a jejich funkce se postupně specializovala na dráhu RdDM. Evoluce těchto dvou polymeráz by mohla být příkladem „escape from adaptive conflict“, kdy po duplikaci došlo k subfunkcionalizaci jednotlivých polymeráz pro funkci v dráze RdDM (Ream *et al.*, 2009; Huang *et al.*, 2015).

Počátek procesu RdDM vyžaduje činnost Pol IV (Herr *et al.*, 2005) ve spolupráci s RDR2 a ty společně produkují dsRNA prekurzory siRNA. Takto vzniklé prekurzory o délce přibližně 30-40 bp jsou poté naštěpeny proteinem DCL3 na 24 nt úseky (Blevins *et al.*, 2015; Li *et al.*, 2015; Zhai *et al.*, 2015; Singh *et al.*, 2019). Pro správné navádění komplexu Pol IV a tvorbu prekurzorů z těchto míst jsou

zásadní SHH1 (SAWADEE HOMEODOMAIN HOMOLOG 1) a CLSY (SNF2 DOMAIN-CONTAINING PROTEIN CLASSY). Protein SHH1 rozpoznává metylaci histonu H3K9 pomocí své SAWADEE domény a nemodifikované H3K4 (Law *et al.*, 2013; Zhou *et al.*, 2018). Podle recentní studie je přepis Pol IV zároveň pevně spjat s přepisem do dsRNA proteinem RDR2, a to nejenom kotranskripčně, ale v rámci jednoho enzymového komplexu. Tyto dvě polymerázy jsou spolu propojeny skrze interakce ve speciálním vazebném místě. Tato interakce vytváří jakýsi RNA kanál, jímž je vznikající RNA protahována a rovnou přepisována do dsRNA. Tato interakce je nezbytná pro správnou biogenezi siRNA (Huang *et al.*, 2021).

Po vytvoření 24-nt dlouhých dsRNA úseků přepsaných Pol IV-RDR2 a štěpených DCL3 následuje jejich stabilizace metylací pomocí proteinu HEN1 (Li *et al.*, 2005; Yang *et al.*, 2006) a navázání jednoho vlákna z těchto sRNA na proteiny AGO4/6/9 (Duan *et al.*, 2015).

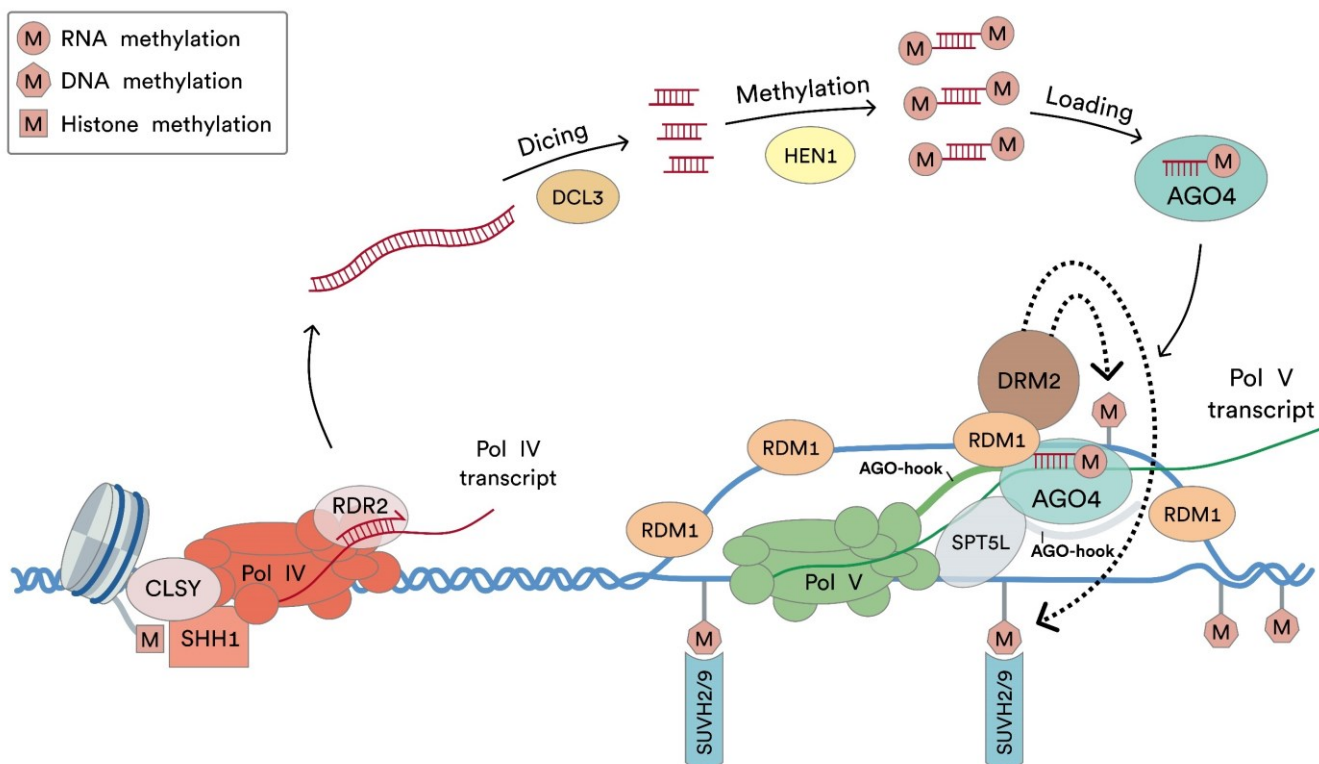
Tyto komplexy AGO s navázanou sRNA jsou poté rekrutovány na nascentní transkript Pol V. S Pol V mimo dalších proteinů spolupracuje v navádění i protein SPT5L (SUPPRESSOR OF TY INSERTION 5-LIKE). Oba proteiny SPT5L i Pol V (na své největší podjednotce NRPE1) obsahují doménu AGO-hook, která napomáhá interakci s proteiny AGO (Rowley *et al.*, 2011; Lahmy *et al.*, 2016). Dále se na této interakci podílí také komplex IDN2-IDP (INVOLVED IN DE NOVO 2-IDN2 PARALOGUE). Do komplexu proteinů spojených s nascentním transkriptem Pol V asociuje i metyltransferáza DRM2, která pomocí interakce s AGO, IDN2 a Pol V metyluje DNA (Böhmdorfer *et al.*, 2014).

Pro vazbu Pol V na cílová místa na chromatinu a správnou tvorbu transkriptu jsou zásadní proteiny z chromatin remodelujícího komplexu DDR, těmi jsou DRD1 (DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1), DMS3 (DEFECTIVE IN MERISTEM SILENCING 1) a RDM1 (RNA-DIRECTED DNA METHYLATION 1; Law *et al.*, 2010; Zhong *et al.*, 2012). Komplex DDR se váže k homologům histonových metyltransferáz SUVH2/9, které však postrádají metyltransferázovou aktivitu. Ty rozpoznávají metylovanou DNA a jsou zásadní pro tvorbu nekódujících transkriptů Pol V na lokusech cílených RdDM (Liu *et al.*, 2014). Protein SUVH9 dále interaguje s histonovou deacetylázou FVE a tato interakce navádí RdDM do specifického subsetu lokusů (Huang *et al.*, 2021).

Pol V na již zametylovaných RdDM místech udržuje spolu se siRNA tyto oblasti dále umlčené. Avšak má schopnost i náhodného skenování euchromatinových oblastí bez metylace, čímž umožňuje vznik transkriptů a tím umlčovat reaktivované či nově zainkorporované TE (Tsuzuki *et al.*, 2020).

Recentní studie od Sigman *et al.*, 2021 ukazuje, že navádění Pol V do cílových míst bez existující předchozí metylace má na starost některý protein z kladu AGO4/6/9. Tato studie ukazuje, že během iniciace prvního kola RdDM funguje Pol II v produkci prvních siRNA z nemetylovaného lokusu, které

jsou zainkorporovány do AGO4/6/9 a následně cíleny *in cis* nebo *in trans* na transkript Pol II, načež je rekrutována Pol V, která již dané místo může udržovat metylované pomocí kanonické RdDM (Sigman *et al.*, 2021).



Obrázek 3: Schéma kanonické dráhy RdDM. Pol IV přepisuje cílový úsek na ssRNA, který je kotranskripčně přepisován na dsRNA pomocí RDR2. Z ní je vyštěpena siRNA proteinem DCL3 a stabilizována metylací prostřednictvím HEN1. Jedno vlákno z této siRNA je zainkorporováno do proteinu AGO. Tento komplex siRNA a proteinu AGO může vyhledávat na nascentním transkriptu Pol V komplementární místo, jehož nalezení spouští DNA metylaci zprostředkovanou pomocí proteinu DRM2.

3.1.4.3 Nekanonická RdDM

V roce 2005 bylo popsáno, že funkce Pol IV navádí umlčování TE, avšak také bylo ukázáno, že tato aktivita není zodpovědná za aktivitu RdDM při virových infekcích, tudíž že dráha kanonická není jedinou, a tudíž že alternativní dráhy musí existovat (Herr *et al.*, 2005). Vedle kanonické dráhy RdDM, využívající 24 nt dlouhé siRNA vzniklé drahou PolIV-RDR2-DCL3, stojí nekanonické dráhy. Ty se liší biogenezí a původem siRNA, které jsou v nich obsaženy. Kanonická dráha využívá především 24 nt dlouhých siRNA, které jsou pro tuto dráhu specifické. Naopak nekanonická dráha RdDM využívá siRNA z různých zdrojů a zahrnuje 24 nt a i 21-22 nt dlouhé siRNA, které umožňují *de novo* metylaci DNA. Tyto siRNA nejsou specifické pro tuto dráhu, ale mají své funkce zároveň i v rámci PTGS a pouze malá část, která se naváže na správné proteiny AGO (především AGO4/6/9) se může podílet na nekanonické RdDM (Erdmann & Picard, 2020).

Stejně jako je tomu u transkriptů Pol IV-RDR2, tak i transkripty Pol II mohou být štěpeny prostřednictvím DCL3 na 24-nt dlouhé sRNA, které se poté mohou podílet na RdDM, a to i nezávisle na aktivitě RDR. Dráha Pol II-DCL3 přispívá k opětovnému umlčování transkribovaných TE (Panda *et al.*, 2016). Molekuly dsRNA, které v této dráze pracují, jsou odvozeny od invertovaných repetitiv (IR), či od *MIR* genů.

Jednou nekanonickou RdDM je dráha, kdy AGO6 váže 21-22 nt dlouhé endogenní siRNA specificky i z aktivních TE, které jsou přepsány do RNA pomocí Pol II a následně mohou navádět RdDM (McCue *et al.*, 2015).

Další nekanonická dráha RdDM, která je schopna cílit lokusy mezi geny a vyžaduje aktivitu Pol IV a Pol V, ale již je nezávislá na RDR2 a AGO4, je takzvaná Pol IV-NERD dráha. Základem pro tuto dráhu je protein NERD (NEEDED FOR RDR2-INDEPENDENT DNA METHYLATION). Tento protein obsahuje AGO-hook doménu, která je schopná vázat protein AGO2. Tato dráha vede k RdDM a transkripčně represivnímu chromatinu na daném lokusu (Pontier *et al.*, 2012).

Další dráha nekanonické RdDM je takzvaná Dicer-independent RdDM. V ní nenaštěpené dsRNA produkované Pol II jsou inkorporovány na AGO4 a později zkráceny na 3' konci na požadovanou velikost činností exonukleáz. Tato dráha produkuje siRNA dlouhé 21-24 nt a cílí komplementární místa pro metylaci DNA (Ye *et al.*, 2016).

Nekanonickým drahám je stále potřeba více porozumět, neboť je vidět, že jsou důležité zejména pro iniciaci umlčování a převedení aktivity na kanonickou dráhu. Jakmile jsou na takovém místě tyto značky ustáleny funkci převezme kanonická dráha RdDM a transkripčně dané místo drží umlčené.

3.1.5 Proteiny AGO

Proteiny AGO byly původně jako první objeveny u rostliny *A. thaliana*. Prvním popsáním proteinem byl AGO1. Proteiny byly pojmenovány podle svého mutanta *ago1*, jehož listy připomínaly chapadla loděnek (Bohmert *et al.*, 1998).

Proteiny AGO jsou zásadní efektorové proteiny pro TGS i PTGS. Všechny eukaryotické proteiny AGO jsou definovány několika doménami. Těmi jsou domény PIWI (P-element Induced WImpy), MID (middle), PAZ (Piwi Argonaut and Zwiller) a variabilní NTD (N-terminální doména).

Funkce N-koncové domény není dodnes příliš jasná, ale mohla by se podílet na rozplétání duplexů sRNA a katalytické aktivitě AGO (Hauptmann *et al.*, 2013). PAZ doména obsahuje takzvaný OB-fold, který je zásadní pro vazbu ssRNA (Song *et al.*, 2003, 2004). MID doména obsahuje rigidní smyčku, označenou jako „nucleotide specificity loop“. Ta je důležitá pro specifitu daného AGO proteinu a jeho

preferenci pro odlišné nukleotidy na 5' konci sRNA (Frank *et al.*, 2012). PIWI doména obsahuje RNase H-like fold a umožňuje některým proteinům AGO štěpení cílového vlákna komplementárního k navázanému (Song *et al.*, 2004).

U rostliny *A. thaliana* bylo zaanotováno 10 proteinů AGO. Tyto proteiny je možno rozdělit na tři klady, a to AGO1/5/10, AGO2/3/7 a AGO4/6/8/9 (Vaucheret, 2008).

3.1.5.1 Větev AGO1/5/10

AGO1 protein je hlavním proteinem AGO pro PTGS, neboť u něj byla prokázána aktivní schopnost štěpit mRNA. Jedná se o hlavní efektorový AGO protein pro miRNA a siRNA, a tak reguluje celou řadu vývojových a fyziologických procesů. AGO1 také pracuje jako efektorový protein v obraně proti virům (Baumberger & Baulcombe, 2005). Protein AGO1 je schopen regulovat sám sebe skrze miR168 (Liu *et al.*, 2020).

AGO5 preferenčně váže sRNA s 5'-C, které jsou odvozeny z intergenových sekvencí. AGO5 je schopen vázat virové siRNA (vsiRNA), což by naznačovalo funkci v antivirové obraně. Na potlačení virové infekce spolupracuje s AGO2 (Mi *et al.*, 2008; Takeda *et al.*, 2008; Brosseau & Moffett, 2015).

AGO10 hraje důležitou roli při udržování nediferencovaných kmenových buněk apikálního meristému prýtu. AGO10 je schopen sekvestrovat miR165/166 a tím blokovat jejich účinky. miR165/166 jsou miRNA, které standardně s pomocí AGO1 cílí transkripty rodiny genů HD-ZIP III, které jsou potřeba pro udržování meristému v nediferencovaném stavu (Zhang & Zhang, 2012; Zhou *et al.*, 2015).

3.1.5.2 Větev AGO2/3/7

Protein AGO2 u *A. thaliana* je zásadní ve virové obraně a bylo ukázáno, že je zásadní pro rezistenci proti některým virovým onemocněním (Brosseau & Moffett, 2015). AGO2 váže virovou vsiRNA s 5'-koncovým A (Mi *et al.*, 2008). K protivirové aktivitě proteinu AGO2 je zapotřebí jeho funkční katalytická doména (Carbonell *et al.*, 2012).

AGO3 je fylogeneticky příbuzný AGO2 a má podobnou funkci ve vazbě vsiRNA (Minoia *et al.*, 2014). AGO3 je však schopen především vázat i 24-nt dlouhé siRNA (stejně jako AGO4), a tak je možné, že by se také mohl podílet na dráze RdDM (Zhang *et al.*, 2016).

AGO7 je partnerem pro vazbu miRNA a je důležitým producentem sekundárních siRNA. U *A. thaliana* AGO7 specificky váže miR390, která cílí transkripty TAS3 a přispívá tak k biogenezi tasiRNA (Montgomery *et al.*, 2008; Carbonell *et al.*, 2012). Ty cílí několik AUXIN RESPONSE FACTORS, které se podílejí na regulaci časování vývoje a vývoje laterálních orgánů (Fahlgren *et al.*, 2006).

3.1.5.3 Větev 4/6/8/9

Proteiny AGO4/6/9 preferenčně váží sRNA dlouhé 24 nt s koncovým 5'-A. Odlišnosti ve specifické vazbě vznikají tím, že jsou jednotlivé proteiny exprimovány různých pletivech. Zatímco AGO4 je exprimován po celé rostlině, AGO6 je exprimován zejména v meristémeh kořenů a prýtu, celkově v pletivech, která jsou bohatá na dělicí se buňky (Havecker *et al.*, 2010; Eun *et al.*, 2011). Všichni členové tohoto kladu z rodiny AGO4/6/8/9 se podílejí na specifikaci prekurzorů samičího gametofytu (Hernández-Lagana *et al.*, 2016).

AGO4 a AGO6 proteiny kolokalizují s největší podjednotkou Pol V v jádře, což ukazuje na to, že hrají společně hlavní roli v RdDM dráze a na jejich funkci je závislá metylace lokusů cílených pomocí RdDM. Jejich funkce v metylaci DNA jsou odlišné, ale vzájemně propojené a neredundantní. Kanonická dráha RdDM zahrnuje 24-nt siRNA, které s těmito proteiny AGO fungují. Dále jsou také schopné pracovat v dráze nekanonické pomocí 21-22 nt dlouhých tasiRNAs. Methylace cílových míst takových tasiRNAs je závislá spíše na funkci AGO6 (Duan *et al.*, 2015; Fang & Qi, 2015).

AGO4 se dále podílí na resistenci proti napadení bakteriemi, avšak ztráta ostatních komponent dráhy RdDM nepostihuje tuto rezistenci, což naznačuje, že AGO4 pravděpodobně funguje nezávisle na RdDM ve zprostředkování bakteriální rezistence (Agorio & Vera, 2007). Dále se AGO4 podílí na obraně rostliny proti DNA virům, kdy vyléčení rostliny vyžaduje AGO4 (Raja *et al.*, 2008).

AGO6 funguje ve specifické dráze RDR6-RdDM a váže 21-22 nt dlouhé endogenní siRNA specificky z aktivních TE, a tak je považován za hráče v iniciaci *de novo* metylace DNA aktivních TE a mohl by propojit transkripční a posttranskripční umlčování TE (McCue *et al.*, 2015).

AGO8 a AGO9 jsou kódovány pomocí dvou sousedních lokusů v genomu a mají vysokou sekvenční identitu. Ačkoliv byl AGO8 obecně považován za pseudogen, ukazuje se, že se také podílí na specifikaci gametofytu (Hernández-Lagana *et al.*, 2016). Protein AGO9 je silně exprimovaný ve vaječných buňkách a prašnicích. AGO9 preferenčně interaguje s 24 nt siRNA odvozenými z aktivních i umlčených TE v samičím gametofytu a napomáhá v jeho správné formaci (Olmedo-Monfil *et al.*, 2010).

3.1.6 AGO-hook doména

AGO-hook je proteinová doména, která je vysoce variabilní, avšak má specifické aminokyselinové složení. Obsahuje typické a jako jediné konzervované WG/GW motivy, které jsou důležité pro vazbu proteinů AGO. Vedle přítomných aminokyselin W a G, jsou zde časté N, S, A, D a naopak jsou znevýhodněny aminokyseliny C, Y, M, I, L. Toto složení ukazuje, že doména AGO-hook preferuje malé hydrofilní a nabitě aminokyseliny, tedy s výjimkou W (Karlowski *et al.*, 2010). Zároveň se jedná o

aminokyseliny podporující neuspořádanost dané domény. To naznačuje, že se jedná o doménu s přirozenou neuspořádaností („intrinsically disordered“; Campen *et al.*, 2008).

AGO-hook doména je složena z tandemových repetice, které jsou hypervariabilní, kdy se během evoluce substitucemi původní konsensus sekvence mění a zachovává se pouze repetitivní charakter. Tyto jednotlivé repetice jsou navíc i taxonomicky specifické a v rámci evoluce se drasticky mění co do sekvence, tak i do počtu. Celá AGO-hook doména (konkrétně proteinu NRPE1) se tak může kompletně přestavět již během 35 miliónů let evoluce (Trujillo *et al.*, 2016).

Mezi proteiny s AGO-hook doménou u *A. thaliana* jsou proteiny NRPE1, SPT5L, SDE3 (SILENCING DEFECTIVE 3) a SPT6L. Doména AGO-hook *per se* není však jen rostlinně specifická a nachází se v rámci všech eukaryot, a dokonce i u některých virů. Nejznámějším příkladem je živočišný protein GW182/TRNC6A a jemu příbuzné proteiny TRNC6B, TRNC6C (Karlowski *et al.*, 2010).

Protein SPT5L (AT5G04290) typický pro dráhu RdDM na rozdíl od svých homologů SPT5 (AT2G34210 a AT4G08350) má na svém C-konci přítomnou AGO-hook doménu s 44 WG/GW motivy. Jedná se tak o nejdelší AGO-hook doménu u *A. thaliana* vůbec. Přítomné motivy jsou zásadní pro vazbu proteinů AGO4 (Bies-Etheve *et al.*, 2009; Lahmy *et al.*, 2016).

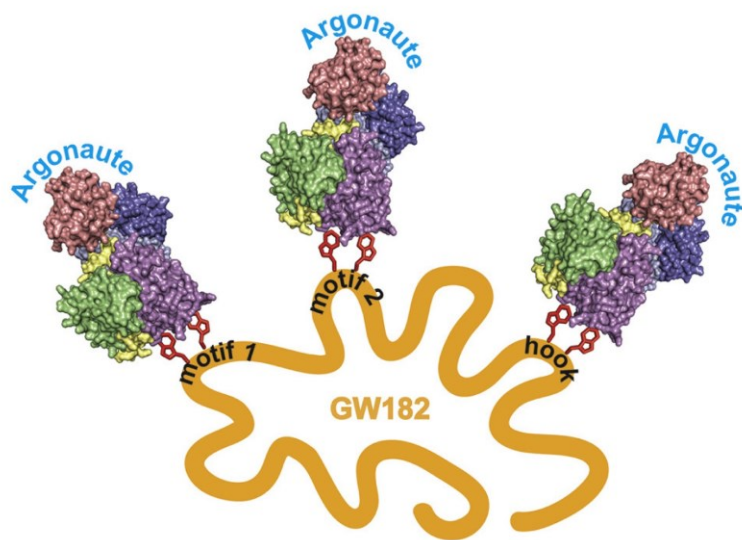
Dalším proteinem z dráhy RdDM je největší podjednotka Pol V, protein NRPE1. Pol V je přítomná u všech suchozemských rostlin a všechny obsahují WG/GW repetice (Huang *et al.*, 2015). Tento motiv je u Pol V potřebný k navázání interakce s AGO4 (El-Shami *et al.*, 2007).

U *A. thaliana* byla u proteinu SDE3 identifikována AGO-hook doména na jeho C-konci obsahující šest motivů WG/GW (Karlowski *et al.*, 2010). SDE3 je obecnou helikázou, která se u krytosemenných rostlin dělí do dvou kladů. A to na homology AT5G35970, které neobsahují AGO-hook a na klad homologů SDE3 s AGO-hook doménou. AGO-hook doména chybí také u nižších rostlin a musela se vyvinout u raných krytosemenných rostlin před rozštěpením na jednoděložné a dvouděložné. AGO-hook proteinu SDE3 je schopen vázat proteiny AGO1/2/4 a tím působit v umlčování virů skrze PTGS (Garcia *et al.*, 2012).

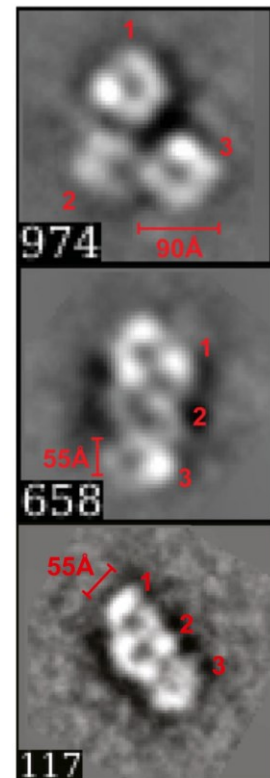
Existují i virové proteiny – supresory umlčování, které tuto doménu využívají k potlačení aktivity AGO proteinů v procesech, kterými se rostlina proti virům brání (prostřednictvím RNAi). Tím je například P1 protein ipomoviru SPMV, který na svém konci nese AGO-hook doménu vázající protein AGO1. Jedná se o tři tyto motivy v podobném aminokyselinovém kontextu, v jakém se nachází i v rostlinných proteinech NRPE1 (Giner *et al.*, 2010). Podobně funguje i protein Nef viru HIV, který je schopen WG/GW motivy vychytávat živočišný protein AGO2 (Aqil *et al.*, 2013).

V lidském GW182 (hsGW182) jsou také WG/GW motivy. Ty dokáží vázat proteiny hsAGO1/2. A navíc váží AGO s navázanou sRNA s větší afinitou (až 8x), jsou tedy selektivní, neboť vazba volného AGO proteinu bez sRNA je neproduktivní. AGO-hook doména hsGW182 je vázána do hydrofobních kapes PIWI domény proteinu AGO. Pro tuto interakci jsou zásadní W. Na hsGW182 jsou přítomna celkem tři AGO vazebná místa (obr. 4), a z nich každé dokáže vázat proteiny AGO samostatně a nezávisle na ostatních (Elkayam *et al.*, 2017).

A



B



Obrázek 4: AGO-hook proteinu hsGW182 **A** Názorné schéma vazby proteinů AGO na AGO-hook. **B** Elektronmikroskopické snímky skutečné vazby tří AGO proteinů na tři AGO-hook motivy u proteinu GW182 (převzato a upraveno z Elkayam *et al.*, 2017).

3.2 Faktory ovlivňující elongaci a transkripty Pol II

U eukaryotických organismů jsou hladiny a charakter mRNA přesně kontrolovány v rámci jednotlivých buněk v závislosti na podmínkách vnějších i vnitřních. Klíčovou roli u eukaryot hraje DNA-dependentní RNA polymeráza II – Pol II. U rostlin se jedná o jednu z pěti DNA-dependentních RNA polymeráz a jedná se i o polymerázu, která dala v evoluci vznik dvěma rostlinně specifickým polymerázám Pol IV a Pol V.

Exprese eukaryotických genů je řízena na mnoha úrovních a zajišťuje přesné množství a charakter přítomných produktů jednotlivých genů. Iniclace je prvním zásadním krokem transkripce pomocí Pol II a dlouho bylo předpokládáno, že po iniciaci Pol II pokračuje ve vytváření transkriptu podle templátu až do doby, než narazí na místo terminace. Avšak i elongace je zásadní fází ve vytváření transkriptu Pol II, která je vysoce řízená. Odlišné fosforylace na C-terminální doméně (CTD) Pol II charakterizují postupné kroky v průběhu elongační fáze, což umožňuje rekrutovat správné složky histon-modifikujících faktorů a faktorů upravujících vznikající transkript do komplexu Pol II (Buratowski, 2009).

3.2.1 Chromatin ovlivňující funkci Pol II

Základní úrovní, kde je možno ovlivnit funkci Pol II je stav chromatinu, ten se obecně dělí na dva základní: euchromatin (přístupný, aktivně transkribovaný) a heterochromatin (represivní, umlčený, netranskribovaný). Ačkoliv toto dělení není ani zdaleka dostatečné pro popsání skutečného stavu a dynamiky a v rámci chromatinu je rozlišována řada dílčích typů chromatinu, bude pro tuto práci použito dále jen uvedené základní členění.

Modifikací histonů, které ovlivňují jeho stav je mnoho a mezi ty nejznámější patří acetylace, metylace, fosforylace a ubiquitylace.

Acetylace histonů je typická pro euchromatin. Acetylace H3K9ac je značkou na aktivně transkribovaných genech a typicky se tedy nachází v rámci euchromatinu (Zhou *et al.*, 2010).

Metylace lysinu je další nejpopsanější značkou. Nejvýznamnější je modifikace na čtyřech lysinových zbytcích u histonu H3. U těch může docházet k monometylaci až trimetylaci. Metylace na H3K4 jsou přítomny pouze v oblasti genů a jejich promotorů. H3K4me3 je typická pro euchromatin a hraje důležitou roli v aktivně se transkribujících genech (Zhang *et al.*, 2009). Tuto značku může odstraňovat specifická Jumonji (JMJ) domain-containing H3K4 demetyláza JMJ14. Tato demetyláza je zásadní pro rostlinu v mnoha ohledech, kdy je například pozitivním regulátorem v obraně proti patogenům a zároveň může být i aktérem v regulaci PTGS (Searle *et al.*, 2010; Li *et al.*, 2020).

Metylace lysinu 36 na H3 je další značkou označující zpravidla aktivní transkripci daného genu. Tuto metylaci zajišťují metyltransferázy, jako je například ARABIDOPSIS THRITHORAX-RELATED PROTEIN 2 (ATRX2). Značka se nachází zpravidla na transkripčním počátku a negativně koreluje s represivní značkou H3K27me3 (Yang *et al.*, 2014; Lee *et al.*, 2017; Mehraj *et al.*, 2021)

Další značky označující lysin na H3 jsou metylace na H3K9 a H3K27, které jsou společně značkami represivními. Pro formaci heterochromatinu napomáhá především přítomnost DNA metylace (například v oblastech TE), kterou rozpoznávají proteiny SUVH4/5/6, načež dimetylují H3K9 a *vice versa* rozpoznáním H3K9me2 je zaváděna metylace DNA pomocí CMT, jak bylo již popsáno výše (viz kapitola Metylace DNA).

Druhá represivní značka je H3K27me1, která je značkou primárně nezávislou na metylaci DNA. Monometylací H3K27 obstarávají proteiny se SET doménou ATRX5/6. H3K27me1 je zásadní pro úspěšné umlčování a represi genů a TE a udržování stability genomu (Jacob *et al.*, 2009; Potok *et al.*, 2022).

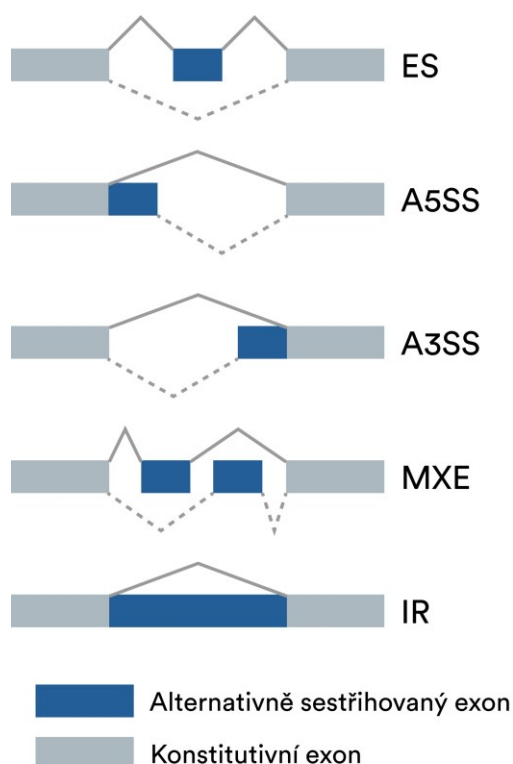
H3K27me3 je represivní značka spojovaná s represí daných genů, a to i v euchromatinových oblastech, a tak hraje důležitou roli ve vývoji, a to nejenom u rostlin. Tato značka je deponována pomocí Polycomb repressive complex 2 (PRC2), který ovlivňuje celou řadu vývojových přechodů v rostlině (Chanvivattana *et al.*, 2004; H. Wang *et al.*, 2016; Yan *et al.*, 2018). Odstraňování této značky mají na starost JMJ histonové demetylázy REF6 (RELATIVE OF EARLY FLOWERING 6), ELF6 (EARLY FLOWERING 6) a JMJ13, které jsou zásadními vývojovými regulátory s pleiotrofními účinky a které kontrolují tkáňově-specifickou aktivaci genů reprimovaných skrze PRC2 (Yan *et al.*, 2018).

Další úrovní, jak ovlivnit vlastnosti chromatinu, je zařazením nekanonických histonových variant, kdy každá varianta je schopna regulovat chromatin jiným způsobem. Například histonová varianta H2A.W podporuje heterochromatinizaci dané sekvence. Zatímco těla aktivně transkribovaných genů mají nabohacení pro kanonickou formu H2A a formu H2A.X a začátky takových genů jsou nabohaceny H2A.Z (Yelagandula *et al.*, 2014).

3.2.2 Alternativní sestřih

Alternativní sestřih je důležitou kotranskripční úpravou transkriptu a regulačním mechanismem, který je schopen navýšit diverzitu proteinů a ovlivnit stabilitu mRNA. Mezi základní typy alternativního sestřihu patří vynechání exonu (ES), alternativní 5'/3' místo sestřihu (A5SS, A3SS), retence intronu (IR) a navzájem se vylučující exony (MXE). Pro přehled působení jednotlivých událostí alternativního sestřihu viz obr. 5. Přibližně asi jedna pětina všech genů je alternativně sestřihována, přičemž nejčastější událostí v rostlinách je IR. Ačkoliv jsou faktory a procesy u eukaryot

značně podobné, mají i rostliny svá specifika. Rostliny mají, nejspíše díky genomovým duplikacím, větší množství proteinů spliceosomu (Wang & Brendel, 2006). Zároveň jsou procesem alternativního sestřihu schopny odpovídat na vnější vlivy a stres, jako je například teplotní stres. Vystavení teplotnímu šoku a vytvoření paměťové stopy je tak ovlivněno i alternativním sestřihem. To by mohlo být částečně zapříčiněno i modulací aktivity Pol II pomocí teploty (Ling *et al.*, 2018; Kindgren *et al.*, 2020). Recentní studie ukázala, že by se v rostlinách na procesech alternativního sestřihu mohly podílet i proteiny AGO cílené do příslušných genů nejspíše vazbou miRNA (Meng *et al.*, 2022).



Obrázek 5: Schémata jednotlivých událostí alternativního sestřihu. Poté co je vytvořena mRNA pomocí Pol II, prochází značná část procesem alternativního sestřihu. Proces ES, kdy se daný intron z transkriptu úplně vystřihuje, je nejčastějším u živočichů, kdežto u rostlin převažuje IR. Jedná se o proces, kdy se v transkriptu ponechává intron mezi dvěma exony, který by byl za normálních okolností vyštěpen. MXE je poměrně vzácnou událostí, kdy se ve finálním transkriptu vybírá ze dvou exonů, které se nemohou vyskytovat společně. Poslední událostí je A3/5SS, kdy se posouvá místo sestřihu na 3' nebo na 5' místě sestřihu.

3.2.3 Transkripční elongační faktory Pol II

Mezi proteiny ovlivňující funkci a procesivitu Pol II patří celá řada transkripčních elongačních faktorů, které jsou potřebné v nejrůznějších situacích. TEF jsou u rostlin obvykle konzervované a je častý jejich cross-talk v rámci odpovědi na nejrůznější stimuly, ať už z prostředí či vývojových. Několik známých TEF asociuje s Pol II v rostlinných buňkách a tvoří zde transkripční elongační komplex, mezi tyto faktory patří například TFIIS, FACT, SPT4/SPT5, SPT6(L) a dále také PAF1C a Elongator (Antosz *et al.*, 2017).

Jedním z nejlépe popsáných TEF je transkripční faktor IIS (TFIIS), který je schopen ovlivňovat vlastnosti komplexu Pol II. U *A. thaliana* bylo ukázáno, že TFIIS není esenciální, neboť jeho mutace nemá žádný zjevný fenotyp za normálních podmínek. Avšak, i přes normální vzezření těchto rostlin je

u mutantů efekt u dormance semen (Grasser *et al.*, 2009). Tento protein je také zásadní při zvládání nejrůznějších stresů, jako je teplotní šok, kde je i navýšena hladina jeho transkriptu. Je poté schopen ovlivňovat regulaci genů pro příslušnou odpověď, a i správnou expresi proteinů teplotního šoku. Tato funkce je konzervovaná v rámci rostlinné říše. TFIIS v tomto procesu hraje roli při ovlivňování alternativního sestřihu (Szádeczky-Kardoss *et al.*, 2022).

Další možností, jak mohou TEF ovlivňovat transkripci Pol II je pomocí zpřístupnění chromatinu pro transkripci či jeho přestavbami. To zajišťuje například komplex FACT (FACILITATES CHROMATIN TRANSCRIPTION). FACT je heterodimerní komplex složený z proteinů SPT16 a SSRP1 (Antosz *et al.*, 2017). FACT je zásadní histonový chaperon, který je schopen interagovat s komponentami nukleozomů i volných dimerů H2A/H2B, ale i H3/H4. Je zásadní pro iniciaci transkripce, ale i pro její elongaci, protože funguje při rozkládání a opětovné skládání histonů v průběhu elongace transkriptu. Je důležitý například při výměně histonových variant (jako H2A.X) za varianty kanonické (Belotserkovskaya *et al.*, 2003; Heo *et al.*, 2008).

Dalšími zmíněnými TEF jsou SPT4/SPT5, které byly původně charakterizovány jako supresory elementu Ty (Spt) v *Saccharomyces cerevisiae* (Winston *et al.*, 1984). Jedná se o heterodimerní TEF. Snížená exprese SPT4 způsobuje rozsáhlé růstové a reprodukční defekty, které jsou odvislé od funkce tohoto proteinu, který je schopen regulovat například geny spojené s auxinem. SPT5 byl zároveň detekován asociovaný s euchromatinem. Rostliny s redukováným množstvím SPT4 mají časté defekty v elongaci transkripce, stejně jako její celkovou sníženou úroveň (Dürr *et al.*, 2014).

3.2.3.1 Protein SPT6L

Mezi TEF komplexu Pol II patří i protein SPT6. Protein Spt6 byl původně anotován jako supresor elementu Ty (Spt) v *S. cerevisiae*. Po původním objevu proteinu se ukázalo, že protein Spt6 hraje důležitou roli v transkripci (Duina, 2011). Jedná se totiž o konzervovaný transkripční faktor, podporující elongaci transkriptu Pol II (Ardehali *et al.*, 2009). Proteiny homologní k Spt6 obsahují tandemovou doménu SH2, která je schopná vázat CTD Pol II na základě interakce s fosforylovanými Ser (Sun *et al.*, 2010; Brázda *et al.*, 2020).

Protein SPT6 je obecným histonovým chaperonem, který je schopen vázat histony H3/H4 a kotranskripčně je znovu ustanovovat na původních pozicích. Během elongační fáze transkripce je SPT6 nezbytný pro zachování struktury chromatinu, a tím prevence začátku transkripce z kryptických promotorů a z intragenových oblastí, a to zejména v genech s vysokou mírou exprese. Pro přestavbu histonů spolupracuje spolu s histonovým chaperonem FACT, který je schopen vázat histony H2A/H2B. Typická struktura proteinu SPT6 obsahuje také záporně nabitý N-konec, pomocí něhož

může interagovat i s proteinem IWS1 (INTERACT WITH SPT6), který je zásadní pro správnou funkci tohoto komplexu (Duina, 2011).

U rostliny *A. thaliana* jsou anotovány dva homology SPT6, a to SPT6 (AT1G63210) a SPT6L (AT1G65440). Protein SPT6 není téměř transkribovaný, avšak i ten byl nalezen v imunoprecipitovných komplexech Pol II. Naproti tomu SPT6L je silněji transkribovaný (Antosz *et al.*, 2017) a obsahuje AGO-hook doménu, ve které se nachází 12 WG/GW motivů. Tento homolog s AGO-hook doménou je konzervovaný v rámci všech cévnatých rostlin. SPT6L je aktivně transkribován především v mladých a vývojově důležitých a aktivně se dělících rostlinných částech. Mutace proteinu SPT6L je letální, neboť je nezbytný pro ustavení apikálně-bazální osy embrya rostliny (Gu *et al.*, 2012).

SPT6L je jaderný protein a je přítomen v aktivně transkribovaných genech, tedy v těch s přítomnými značkami aktivní transkripce, jako je H3K36me2 a acetylace histonů. Zároveň SPT6L na chromatinu koreluje s přítomností Pol II a množství SPT6L je korelováno s aktivitou transkripce v daném místě. To naznačuje, že protein SPT6L u *A. thaliana* má stejné funkce jako SPT6 u ostatních organismů (Chen *et al.*, 2019).

Pro interakci proteinu SPT6L s Pol II je zásadní doména SH2, avšak tato interakce není potřebná pro výskyt SPT6L v místech počátku transkripce (TSS) genů, neboť SPT6L nasedá na DNA nezávisle na Pol II. Pro nabohacení SPT6L na potřebná místa v genomu jsou důležité HtH a YqgF domény tohoto proteinu (Chen *et al.*, 2019).

Známým interakčním partnerem proteinu SPT6L je IWS1. Protein IWS1 je schopen interagovat s proteinem BES1 (BRI1 EMS SUPPRESSOR 1), který je schopen interagovat s REF6 a ELF6, a tím regulovat metylaci H3K27me3 (Li *et al.*, 2010). Lidský protein IWS1 je centrálním proteinem zprostředkovávajícím interakce mnoha různých faktorů včetně SPT6. Tato interakce je umožněna přirozeně neuspořádanými doménami IWS1 a vazebnými motivy, které tyto domény rozpoznávají. Tyto interakce neuspořádaných domén jsou zásadní pro správnou funkci Pol II (Cermakova *et al.*, 2021).

Protein SPT6L obsahuje AGO-hook doménu, která by měla být schopná vázat proteiny AGO. Navíc podle analýz provedených naší skupinou, je protein SPT6L kolokalizován také s Pol V, což ukazuje, že by tento protein mohl mít svou specifickou funkci i v komplexu Pol V (Čermák, nepublikováno). SPT6L je kandidátním proteinem, který by mohl zprostředkovávat regulaci exprese pomocí sRNA jak v komplexu Pol II, tak i Pol V.

4 Materiál

4.1 Práce s rostlinným materiálem

Rostliny *A. thaliana* byly pěstovány na rašelinových peletách (Jiffy) v kultivační místnosti podle potřeby s dlouhodobou periodou 16 h světlo/8 h tma, či s krátkodobou 8 h světlo/16 h tma při teplotě 22 °C ve dne a 19 °C v noci. Ozářenost byla přibližně 100 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$.

Pro sklizení semen byl použit ARA systém (Betatech), který snižuje ztráty semen a zároveň zabraňuje kontaminaci semen z okolních rostlin během pěstování

Jako materiál sloužily rostliny *A. thaliana* ekotyp Col-0, SALK mutant *nrpe1-11* (SALK_022919), *spt6l* (SAIL_569_E05) komplementovaní mutanti SPT6L-GFP a SPT6L Δ Ah-GFP, kde je protein SPT6L zkrácený o AGO-hook doménu, tedy 217 (1430–1647) aminokyselin a CRISPR/Cas9 mutanti *spt6l Δ Ah* o stejné délce delece. Tyto rostliny mi byly poskytnuty od Vojty Čermákem a Magdalény Metličkové.

4.1.1 Křížení rostlin

Pro křížení rostlin byly použity rostliny s mladými květenstvími, z nichž byly nejprve ostříhány všechny šesule a již rozvinuté květy. Květy použité pro vzájemné křížení byly ještě nerozvinuté. Z takových květů byly odstraněny veškeré květní obaly a nedozrálé tyčinky a za pomoci pinzety byly pestíky takových květů opyleny z dospělých květů jiné rostliny mající požadovaný genotyp, který byl ověřen genotypováním.

4.2 Chemikálie

Agar (pro elektroforézu): Gibco

Amonium persulfát: USB 76322

Bromfenolová modř: LACHEMA

EDTA: Sigma

Etanol: LachNer s. r. o.

Formamid: Serva

GelRed: Biotinum

Glycin: Serva

Chloroform: LachNer s. r. o.

Isopropylalkohol: LachNer s. r. o.

Kyselina boritá: LACHEMA

Kyselina octová: PENTA

Metanol: LachNer s. r. o.

Monomery akryl-amid (30%): Sigma

NaCl: LACHEMA

PMSF: Serva

SDS: MP Biomedicals

TEMED: Serva

TRI Reagent: Sigma-Aldrich

Triton-X-100: Merck

TRIS báze: Serva

Xylen cyanol FF: Sigma

TRIS-HCl: Serva

4.3 Protilátky

ADI-SAB 100 anti-Mouse HRP: ENZO

Goat Anti-Rabbit IgG-HRP: ENZO

Monoclonal ANTI-FLAG® M2 F1804: ThermoFischer

Polyclonal ANTI-GFP AS15 2987 Rabbit: Agrisera

4.4 Přístroje

Centrifugy: Hattich centrifuge universal 32 R, Cleaver centrifuge C1301P, Heraeus Biofuge 13

G:BOX (Syngene) program GeneSnap (verze 7.09, Syngene)

Homogenizér: Retsch MM301

Nanodrop 2000: Thermo Sciences

PCR cyklery: Biorad T-100 Thermal cykler, MS Reasearch PTC-200 cykler

pH metr: Microprocessor pH 539

Rotátor: FALC F205

Semidry-blot: Biometra Semidry

Třepačka: Biometra WT 17

Vortex: Gense 2 Scientific Industries

Zdroje: Biometra P30, Pharmacia GPS 200/400, E-C Apparatus Corporation 250-90

4.5 Další materiál

1,5 ml Protein LoBind: Eppendorf®

6x LD Ro611: Thermo Fischer

Arasystém: Aracons 720, Aratubes 360

DreamTaq PCR Master Mix K1071: ThermoFischer

GelRed™: Biotum

GeneRuler 100 bp Plus: Thermo Sciences

GeneRuler 1 kbp: Thermo Sciences

GFP-Trap Magnetic Particles M-270: Chromotek

Kazeta pro vývoj WB: CAWO ABS Casette

Komory pro gelovou elektroforézu nukleových kyselin: Wide Mini-Sub Cell GT Cell: Bio-Rad
Laboratories

Komory pro gelovou elektroforézu: Biometra minigel G41 a minigel twin G42

Kristalon: LACHEMA

Laktino (sušené mléko, odtučněné): Artifex Instant, s. r. o.

Nitrocelulose NC2 71224: SERVA

PageRuler Plus Prestained Ladder 10–250 kDa: Thermo Fisher

Pierce™ ECL WB substrát: Thermo Fisher

Primery: GenериBiotech

Protease inhibitor coctail P9599: Sigma

Rašelinové pelety: Jiffy

Rentgenové filmy: Agfa CP-BU M 100 NIF 18x24cm

Ustalovač: MOFA

Vývojka: MOFA

5 Metody

5.1 Genotypování rostlin *A. thaliana*

5.1.1 Izolace rostlinné DNA

Z třítydenních semenáčků byl odebrán jeden lístek, který byl po přidání 400 μ l extrakčního pufru (tab. 1) a 400 μ l chloroformu homogenizován v 2 ml mikrozkuhavce pomocí dvou ocelových nerezových kuliček v kolovém mlýnku (Retsch MM301) po dobu 1 min a při frekvenci 30 Hz. Homogenizovaný vzorek byl centrifugován 3 min 14000 g. 330 μ l horní vodné fáze bylo přeneseno do čisté 1,5 ml mikrozkuhavky a smícháno s 330 μ l isopropanolu. Po 2 min inkubace byl vzorek centrifugován 5 min 14000 g. Supernatant byl odstraněn a peletka DNA ponechána vyschnout 10 min. DNA byla následně rozpuštěna ve 100 μ l 2mM TRIS pH 8,5.

Takto připravená DNA byla buď skladována při -20 °C anebo následně ihned použita pro PCR.

Tabulka 1: Složení extrakčního pufru pro izolaci DNA

Název	Koncentrace
Tris-HCl pH 7.5	200 mM
NaCl	250 mM
EDTA	25 mM
SDS	0,5 %

5.1.2 PCR

Směs pro PCR byla připravena na ledu. Do 1,5 ml mikrozkuhavky byla namíchána směs pro daný počet vzorků podle tabulky č. 2 a za použití primerů uvedených v tabulce č. 3.

Z takto připravené směsi bylo odebráno 9 μ l do 0,2 ml PCR mikrozkuhovek a do nich byl přidán 1 μ l vyizolované DNA.

Takto připravené reakční směsi se vzorky byly poté vloženy do cykleru a byl spuštěn program PCR popsáný v tabulce č. 4.

Tabulka 2: Složení PCR směsi

Název	Množství (na 10 µl)
DreamTaq PCR Master Mix K1071	5 µl
Primery	0,4 µl (2 µM)
Templát	1 µl
miliQ H ₂ O	3,6 µl

Tabulka 3: Tabulka používaných primerů.

Název	Sekvence 3'-5'
NRPE1-11_LP	TTTTCCTCCAAGTTGTGATCG
NRPE1-11_RP	GTTTTCCCCAAGAACAAGACC
pROK2_R_022062	ACAACACTCAACCCTATCTCG
SPT6L – SI_R	AAATGTTAACCTGAGCACTTGTTG
SPT6L_S6	GAGGACATTGATCGGCTTGT
SPT6L_R2	AAAGGCATTTAGGGAGAAAG

Tabulka 4: Program PCR; T_A – odpovídá teplotě nasedání primerů, která je o 5 °C nižší než jejich teplota tání T_M. Teplota tání byla přibližně odhadnuta dle vzorce: $T_M = 2\Sigma(A+T) + 4\Sigma(G+C)$.

Název	Teplota [°C]	Čas	Opakování kroku
Počáteční denaturace	94	3 min	1x
Denaturace	94	20 s	30x
Nasedání primerů	T _A	30 s	
Syntéza DNA	72	1 min	
Finální syntéza	72	5 min	1x
Chlazení	12	∞	1x

5.1.3 Elektroforéza DNA

Produkty PCR byly následně analyzovány pomocí elektroforézy v TAE pufru (tab. 5) v elektrickém poli o intenzitě 5 V/cm. Používány byly 1,2 % agaróзовé gely s 0,01 % GelRed™. Vzorky byly smíchány s 6x LD Thermo Fischer R0611 a nanoseny do jamek v gelu. Pro odečtení délky fragmentů byl použit 1kbp nebo 100bp marker GeneRuler. Gel byl následně zdokumentován na zařízení G:BOX.

Tabulka 5: Příprava 50x TAE pufru

Název	Množství
EDTA	50 mM
TRIS	2 M
Kyselina octová	1 M

5.2 Fylogenetická analýza

5.2.1 Vyhledávání sekvencí

K získání sekvencí proteinů homologních k SPT6L (NC_003070.9) a SPT5L (NC_003076.8) z *A. thaliana* byl využit program BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Bylo ponecháno základní nastavení programu s PSSM (position-specific scoring matrix) BLOSUM62 a vyhledávání byla postupně provedena pro oba proteiny pro zástupce *Archaeplastida*.

5.2.2 Vyhledávání AGO-hook domén

K identifikaci AGO-hook domén v jednotlivých sekvencích proteinů získaných pomocí programu BLAST byl využit program Wsearch (Zielezinski & Karlowski, 2015). Byla použita PSSM pro rostlinné AGO-hook domény.

```
>> python wsearch.py -q protein_seqs.fasta -p pssm/ago-plants.pssm
```

Výstupem programu jsou predikované AGO-hook domény jednotlivých proteinů spolu s jejich skóre a sekvencemi. Pro následující postup byly jako proteiny obsahující AGO-hook doménu zaanotovány ty, které mají alespoň tři takové motivy se skóre alespoň 6, takové proteiny poté byly označeny W1 (s AGO-hook doménou). Proteiny, které kritéria nesplňovaly byly označeny W0 (neobsahující AGO-hook doménu).

5.2.3 Tvorba fylogenetických stromů

Výše zmíněné získané sekvence poté byly dále zpracovávány v programu Geneious Prime v. 2021.2.2.

Nejprve byly tyto sekvence alignovány pomocí algoritmu MAFFT s nastavením FFT-NS-1 PSSM a 200PAM / k =2. Z těchto alignovaných sekvencí byly vyřazeny zkrácené proteiny a ponechány pouze ty, které splňují podmínky pro správnou funkci daného proteinu. Těmito kritérii je jejich délka a přítomnost esenciálních domén. Mezi esenciální domény patří již v literárním úvodu zmiňované SH2 domény a pak také střední část proteinu SPT6L Tex obsahující domény HtH a YqgF. Očekávána byla variabilita proteinů SPT6(L) na jeho N-konci, kde je přítomna N-terminální doména obsahující kyselé aminokyseliny, ta chybí i u proteinu SPT6 u *A. thaliana*. (Close *et al.*, 2011). Stejně tak variabilita je očekávána na C-konci, kde mohou a nemusí být přítomny AGO-hook domény.

Po tomto filtru byly získané sekvence alignovány znovu nyní již za použití algoritmu MAFFT s nastavením E-INS-i PSSM a 200PAM / k =2.

K sestavení fylogenetických stromů byl poté využit program Geneious Tree Builder s modelem Jukes-Cantor a metodou budování Neighbor-Joining.

5.3 Příprava RNA pro analýzu transkriptomu

5.3.1 Izolace rostlinné RNA

Izolace RNA byla provedena za pomoci TRI Reagentu (Sigma-Aldrich, T9424). Protokol je převzat z oficiálního technického bulletinu a byl optimalizován. Celá práce s RNA probíhala ve sterilních podmínkách se zvláštním důrazem na čistotu práce kvůli RNázám.

100 mg semenáčků bylo zmrazeno v tekutém dusíku a homogenizováno 1 min s frekvencí 30 Hz na mlýnku (Retsch) v 2 ml zkumavkách se zámkem a se dvěma nerezovými kuličkami. K takto homogenizovaným stále zmraženým vzorkům byl přidán 1 ml TRI Reagent a obsah zkumavek byl míchán na vortexu dokud plně nerozmrzl. Následně byly vzorky ponechány ve stojánku při LT po dobu 5 min.

Po rozmrazení bylo ke každému vzorku přidáno 200 μ l chloroformu a vzorky byly ihned po dobu 15 sekund promíchávány na vortexu a následně ponechány 15 minut stát při LT.

Následně byly vzorky centrifugovány při 12 000 g, 15 min při 4 °C. Pro izolaci RNA se odebrala horní vodní fáze do nové 1,5 ml zkumavky. K té bylo přidáno 0,5 ml izopropanolu a opět byly vzorky promíchány na vortexu a ponechány 10 minut při LT.

Poté byly vzorky stočeny při 12000 g a 4 °C po dobu 10 minut. Supernatant byl po centrifugaci odstraněn a peletka na dně následně promyta pomocí 1 ml 75 % ethanolu. Obsah zkumavek byl vortexován a centrifugován při 7500 g a 4 °C po dobu 5 minut (pokud se při promytí peletka uvolnila ode dna probíhá centrifugace při 12000 g). Supernatant byl poté odstraněn a peletka ponechána na vzduchu v horizontální poloze schnout po dobu 10 minut.

Po vysušení peletky byla RNA rozpuštěna ve 25 µl RNase-free vody a inkubována 10 minut při 60 °C. Po inkubaci je vzorek pipetou ještě jednou promíchán. Vzorky poté byly zamraženy a skladovány při -80°C.

5.3.2 Elektroforéza RNA

Pro kontrolu kvality izolované RNA byla provedena RNA elektroforéza na agarovém gelu. Rozdělení vzorků probíhá na 1,2 % agarózovém gelu (0,5x TBE; složení v tabulace č. 6). Do vymyté elektroforetické komory do 0,5x TBE pufru byl vložen gel a jednotlivé vzorky nanášeny do jamek ve množství 2 µl smíchané s 2 µl FDA nanášecí barvy (tab. 7).

Tabulka 6: Složení 5x TBE pufru

Název	Množství (na 1 l)
TRIS	54 g
Kyselina boritá	27,5 g
EDTA 0,5 M pH 8.0	20 ml

Tabulka 7: Složení FDA nanášecí barvy

Název	Množství (na 10 ml)
Deionizovaný formamid	9,8 ml
EDTA	200 µl
Xylen cyanol FF	10 mg
Bromfenolová modř	10 mg

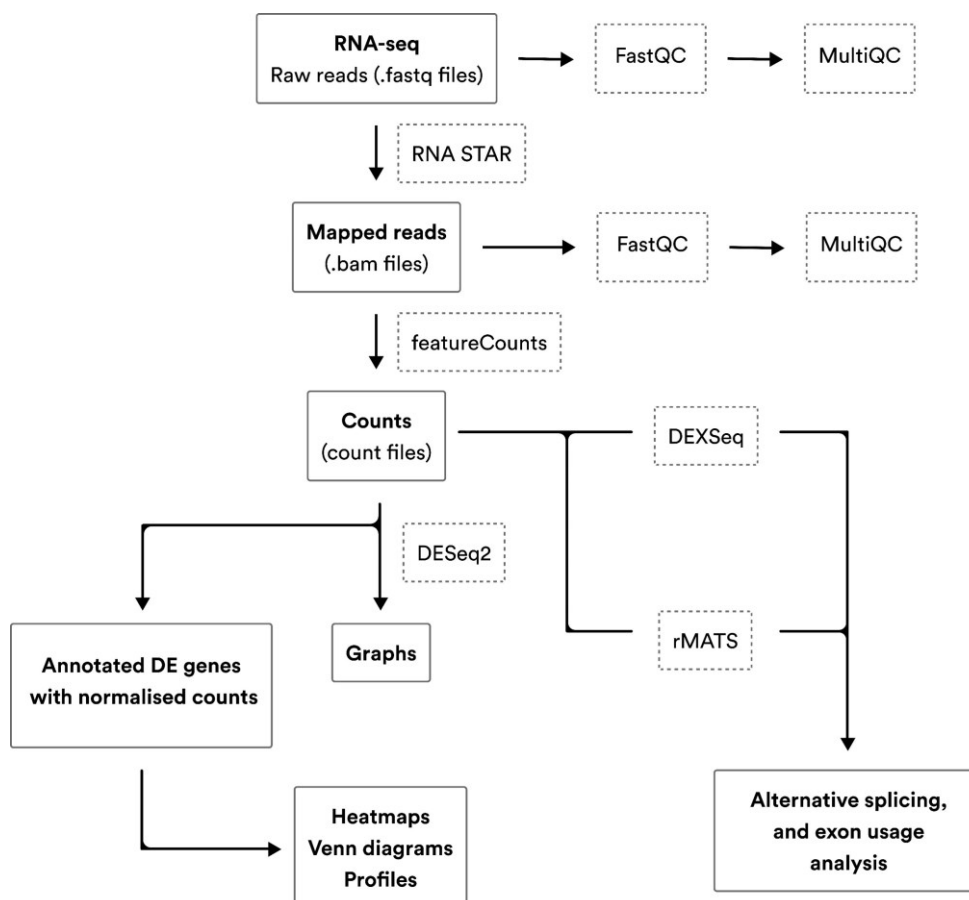
5.3.3 Dodatečné kontroly RNA

Kontrola integrity RNA (RIN) byla provedena pomocí přístroje 2100 Bioanalyzer (Agilent) v Laboratoři sekvenace DNA – Přírodovědecká fakulta UK. Jako vzorky dostatečně čisté byly považovány ty, u kterých byla hodnota vyšší než 6,3.

Dodatečná kontrola kvantity a čistoty byla provedena pomocí Nanodropu (Thermo Sciences), kde u čistých vzorků je poměr absorbance vlnových délek 260/280 blízký hodnotě 2.

5.4 Bioinformatická analýza transkriptomických dat

Příprava knihovny a NGS sekvenování bylo provedeno firmou Novogene. mRNA byla separována poly-T magnetickými kuličkami a po fragmentaci bylo vytvořeno první vlákno cDNA použitím náhodných hexamerů jakožto primerů. Druhé vlákno cDNA bylo nasyntetizováno pomocí dUTP namísto dTTP. Následně byly opraveny konce přidáním 3'-A a na konce byly zaligovány adaptory. Posléze byla knihovna naštěpena enzymem USER s následnou amplifikací a přečištěním. Po získání této směrově specifické knihovny sekvenování probíhalo na Illumina NovaSeq 6000, čtení bylo párové o délce 150 nt (PE150) a o hloubce čtení 40M.



Obrázek 6: Schéma analýzy transkriptomických dat. V plných obdélnících jsou vyznačeny výstupy a v přerušovaných použité metody.

Analýza transkriptomických dat vyžaduje poměrně dost nástrojů a její základní průběh je nastíněn ve schématu (obr. 6). K jejich zpracování byl částečně využit výpočetní server Galaxy (<https://usegalaxy.eu/>) a z části programy volně dostupné z repozitářů github a použity v prostředí Ubuntu 20.04 LTS (GNU/Linux 5.10.16.3). Programy jsou uvedeny v tabulce č. 8.

Tabulka 8: Programy používané při analýze transkriptomických dat.

Název programu	Verze	Reference	Platforma
FastQC	Galaxy Version 0.73	(Andrews <i>et al.</i> , 2010)	Galaxy
MultiQC	Galaxy Version 1.11	(Ewels <i>et al.</i> , 2016)	Galaxy
RNA STAR	Galaxy Version 2.7.8a	(Dobin <i>et al.</i> , 2013)	Galaxy
featureCounts	Galaxy Version 2.0.1	(Liao <i>et al.</i> , 2014)	Galaxy
DESeq2	Galaxy Version 2.11.40.7	(Love <i>et al.</i> , 2014)	Galaxy
DEXSeq	Galaxy Version 1.28.1	(Anders <i>et al.</i> , 2012)	Galaxy
rMATS	rMATS turbo v4.1.2	(Shen <i>et al.</i> , 2014)	Ubuntu

Po obdržení transkriptomických dat je potřeba získané sekvence (reads) zamapovat na referenční genom TAIR10. Jako anotace sloužil Araport11 (Cheng *et al.*, 2017). Mapování bylo provedeno programem STAR s následujícími pozměněnými parametry:

```
>> STAR --sjdbOverhang 149 --twopassMode Basic --outFilterMultimapNmax 100
-- outBAMsortingBinsN 50 --winAnchorMultimapNmax 100
```

Ze získaných readů byly následně získány počty readů pro jednotlivé geny programem featureCounts s následujícími upravenými parametry:

```
>> featureCounts -s 2 -t 'exon' -g 'gene_id' -M
```

Z tabulky počtů readů pro jednotlivé geny (features) poté byla nástrojem DESeq2 získána statistika diferencially exprimovaných genů se základním nastavením.

Pro zjištění rozdílů v alternativním sestřihu a ve využití exonů byly použity nástroje DEXSeq, rMATS.

```
>> python rmats.py --gtf Araport11.gtf -t paired --readLength 150 --
novelSS --allow-clipping
```

```
>> dexseq.R -c 0.05
```

Jednotlivé formáty anotací Araport11 byly získány nástroji gtfToGenePred, gtf2bed.

Pro konverzi formátů souborů BAM byly použity nástroje bamCoverage, bedGraphToBigWig, samtools depth.

Ze sady nástrojů deepTools byly využity nástroje computeMatrix a plotProfile pro vytvoření profilů pokrytí chromatinových značek z volně dostupných ChIPSeq dat. Data byla přepracována Vojtou Čermákem. Datové sety: GSE108673, GSE38286, GSE165877, GSE100010, GSE124546, GSE111813, GSE28398, GSE108960.

5.5 Proteomika

5.5.1 Izolace rostlinných proteinů

Převzato a upraveno podle Chen *et al.*, 2019: 250 mg rostlinného materiálu bylo spolu se dvěma nerezovými kuličkami vloženo do 2 ml ependorfeč se zámkem a okamžitě zmraženo v tekutém dusíku a následně rozdrceno na kulovém mlýnku po dobu 1,5 minuty s frekvencí 30 Hz. Následně do zkumavek bylo přidáno 750 μ l extrakčního pufru podle tabulky č. 9 a obsah byl ponechán na ledu, dokud se vzorky kompletně nerozpustily. Poté byly odstraněny nerezové kuličky pomocí magnetu a vzorky byly poté centrifugovány 14 000 g, 10 min při 4 °C.

Oddělený supernatant byl poté přenesen do nových 1,5 ml mikrozkušavek. Pokud proteiny sloužily pro následnou proteinovou elektroforézu, byly ihned denaturovány přidáním 4x SDS nanášecího pufru (tab. 9) a zahřátím při 95 °C po dobu 5 minut a následně naneseny na SDS-PAGE. Pokud byly použity později byly skladovány v -80 °C.

Tabulka 9: Tabulka složení extrakčního pufru pro izolaci proteinů; tučně vyznačené složky pufru se přidávají až těsně před použitím.

Název	Koncentrace
Tris-Cl pH 8	20 mM
NaCl	100 mM
EDTA	2,5 mM
Triton X-100	0,1 % (V/V)
DTT	1 mM
PMSF	0,1 mM
Protease inhibitor cocktail (Sigma, P9599)	1 % (V/V)

Tabulka 10: Tabulka složení 4x SDS vzorkového pufru.

Název	Koncentrace
Tris-HCl	200 mM
DTT	0.4 M
SDS	8 %
Bromfenolová modř	6 mM
Glycerol	4.3 M

5.5.2 Koimunoprecipitace

Koimunoprecipitace byla provedena pomocí GFP-Trap® Magnetic Particles M-270 a byl použit protokol od firmy Chromotek.

Kuličky byly promíchány v zásobní nádobce jejím otočením a následně bylo 10 µl magnetických kuliček přeneseno do 1,5 ml Eppendorf® Protein LoBind zkumavky. Ke kuličkám bylo přidáno 500 µl

vychlazeného (4 °C) promývacího pufru (tab. 11). Po promíchání byly kuličky separovány pomocí magnetu, dokud nebyl supernatant čistý, ten byl následně odebrán.

K takto připraveným kuličkám bylo přidáno 750 µl roztoku izolovaných proteinů. Inkubace s kuličkami probíhala při 4°C na otočném rotátoru po dobu 1 hodiny. Po inkubaci s proteiny byl izolát proteinů odebrán a kuličky 3x promyty promývacím roztokem (tab. 11). Pokud byly vzorky určeny pro analýzu na hmotnostním spektrometru (MS), tak poslední promytí probíhá bez přidání detergentu (Triton-X). Takto připravené kuličky s navázanými proteiny jsou připraveny pro analýzu MS/MS.

Pokud byly vzorky dále analyzovány na western blotu, byly nejprve imunoprecipitované proteiny z kuliček eluovány. Kuličky byly z roztoku vychytány magnetem a resuspendovány v 50 µl 1x SDS-extrakčním pufru (tab. 10). Poté byly vzorky zahřáty 5 min na 95 ° a supernatant s imunoprecipitovanými proteiny byl přenesen do nových 0,5 ml zkumavek. Tento eluát poté mohl být nanášen na SDS-PAGE a následně analyzován na western blotu.

Tabulka 11: Složení promývacího pufru.

Název	Množství (pro 1 l promývacího pufru)
Tris-Cl pH 7.5 (pH upraveno při 4°C)	10 mM
NaCl	150 mM
Triton-X	0.05 %
EDTA	0.5 mM

5.5.3 Detekce proteinů na western blotu

5.5.3.1 SDS-PAGE

Směs pro dělicí gel se připraví podle návodu v tabulce 12. AP (amonium persulfát) a TEMED (tetra-metyl-etylen-diamin) se přidávají až těsně před nalíváním gelu. Směs je třeba promíchat a poté bez bublin nalít gel zhruba 2 cm od horního okraje a opatrně převrstvit vodou. Gel polymerizuje asi 45 minut, kdy je poté možno horní vrstvu vody slít a gel osušit.

Poté lze nalévat i zaostřovací gel, který se smíchá opět podle tabulky 12. Směs se následně nalije do formy po okraj a vloží se hřebínek a tento gel poté tuhne 30 minut. Po zatuhnutí gelu lze opatrně odstranit svorky a těsnění a pokračovat s gelem na vlastní elektroforézu.

Tabulka 12: Příprava SDS-PAGE gelu. TEMED a AP přidat až těsně před naléváním gelu.

Název	10% dělicí gel	8% dělicí gel	4% zaostřovací gel
Monomery akryl-amid a bis-akryl-amid (30%)	2 ml	1,6 ml	350 μ l
Tris-HCl	1,5 M pH 8,8 1,5 ml	1,5 M pH 8,8 1,5 ml	0.5M pH 6,8 625 μ l
dH ₂ O	2,4 ml	2,8 ml	1,5 ml
10% SDS	60 μ l	60 μ l	25 μ l
40% AP	7,5 μl	7,5 μl	5 μl
TEMED	2 μl	2 μl	1,25 μl

Takto připravené gely byly vloženy do elektroforetické komory, zality elektroforetickým pufrem (tab. 13) a vzorky nanесeny do jamek. Jako marker byl použit PageRuler Plus Prestained Ladder 10-250 kDa, 26619 o objemu 5 μ l. Proteinová elektroforéza trvala zhruba 3 h při 100 V a 25 mA/gel.

Tabulka 13: Návod pro přípravu 10x elektroforetického pufru.

Název	Množství (pro 1 l pufru)
Glycin	144,4 g
Tris-báze	30,3 g
SDS	10 g

5.5.3.2 Western blot

Pokud byl gel určen pro detekci na western blotu (WB), byl nejprve vložen pro ekvilibraci do transferového pufru (tab. 14). Do tohoto pufru byly namočeny i filtrační papíry (Whatman) a membrány (pokud byla použita PVDF membrána byla nejprve na 5 minut položena na hladinu metanolu pro její aktivaci (smočení) a až poté do transferového pufru).

Tabulka 14: Návod pro přípravu transferového pufru (1 L)

Název	Množství (pro 1 l pufru)
Glycin	14,4 g
Tris-báze	12,1 g
Metanol	200 ml

Pro přípravu blotovacího sendviče pro „semi-dry“ blotování se nejprve vložili 3 namočené filtrační papíry na katodu, na ní poté připravená membrána, gel a opět 3 filtrační papíry. To celé se zalilo malým množstvím transferového pufru. Skleněnou tyčinkou byly válením odstraněny případné bubliny, a nakonec se celý sendvič přiklopil druhou elektrodou a spustil se přenos proteinů z gelu na membránu s hodnotami 300 V, 300 mA, 10 W po dobu 3 h. Takto připravené bloty byly buď ihned použity pro detekci, či byly skladovány po zaschnutí mezi filtračními papíry.

5.5.3.3 Imunodetekce

Pro imunodetekci WB je prvním krokem je blokování membrány. Blokování probíhalo ve zhruba 30 ml TBS pH 7.4 bez Tritonu (tab. 15) s 5% sušeným odtučněným mlékem po dobu 1 h při LT na kývačce.

Po zablokování membrány se navázala primární protilátka, kdy se membrána nechá po dobu 1 h na třepačce v roztoku TBS (s pH dle dané protilátky) s protilátkou. Pro detekci FLAG-tagu se použila protilátka Sigma F1804 anti-FLAG v TBS pH 8,0 s 0,5% sušeným mlékem, ředění protilátky je 1:1000. Pro detekci GFP se použila Agrisera AS15 2987 Rabbit anti-GFP v TBS pH 7,4 s 0,5% sušeným mlékem, ředění protilátky 1:4 000.

Membrána se poté 3x promyla po dobu 10 min v TBS-T na kývačce při LT.

Poté se navázala sekundární protilátka. Membrána byla 1 h při LT na kývačce inkubována v TBS pH 7.4 se sekundární protilátkou. Pro detekci primární protilátky anti-FLAG se použila Sigma anti-Mouse HRP v TBS pH 7,4 s 0,5% sušeným mlékem s ředěním 1:1000. Pro detekci primární protilátky anti-GFP se použila Goat Anti-Rabbit IgG-HRP pH 7,4 s 0,5% sušeným mlékem s ředěním 1:5 000.

Poté byla membrána promyta 3-4x na 10 min v TBS-T na kývačce. Na takto připravené membráně se navázané protilátky ihned detekovaly pomocí ECL.

Tabulka 14: Složení TBS pufru, potažmo TBS-T přidáním detergentu – zvýrazněno tučně.

Název	Množství
Tris	20 mM
NaCl	150 mM
Triton-X	0,1 %

5.5.3.4 ECL systém

Na rovné sklo byl napipetováno 1 ml reakční směsi 1 a 1 ml reakční směsi 2 (Pierce™ ECL Western Blotting Substrate 32106) a směsi byly promíchány pipetováním na skle. Membrána byla poté ponořena stranou s proteiny do této směsi a ponechána 1 minutu inkubovat. Poté byla vložena do eurofólie a do kazety.

Poté se luminiscence detekovala exponováním rentgenového filmu. Do kazety s WB v temné komoře byl přiložen film. Pro potřeby této práce se vytvářely dva snímky o různé dlouhé expozici, první kontrolní s expozicí 15 minut a druhý s expozicí 1 hodiny.

Po expozici se film v temné komoře ponořil nejprve do vývojky, dokud nebyly vidět obrysy membrány, poté byl opláchnut v destilované vodě a následně v ustalovači zhruba jednu minutu a opět opláchnut v destilované vodě a následně usušen.

5.5.4 Analýza proteinů pomocí MS

Analýza MS/MS byla provedena Karlem Harantem v Laboratoři hmotnostní spektrometrie BIOCEV.

Data byla zpracována v programu Perseus v2.0.3.0 (MaxQuant).

Detekované proteiny byly vždy pro jednotlivá porovnání vyfiltrovány na signifikantně změněné a dále na proteiny, které jsou anotovány jako jaderné. K tomu byla použita anotace Gene Ontology (GO) Term Enrichment dostupná z portálu TAIR (<https://www.arabidopsis.org/>).

GO enrichment analýzy byly prováděny za pomoci webové aplikace ShinyGO v.0.75 (Ge *et al.*, 2020).

5.6 Další doplňující materiály a metody

Pro běžnou práci s tabulkami byl využit kancelářský balíček Microsoft Office 365. Pro úpravy a práci s tabulkami byly používány následující jazyky: Python 3.9.5, R a pro finální práci s grafy byl používán program GraphPad v. 8.0.1, tabulka č. 15 ukazuje hodnoty statistických signifikancí. Ke zpracování NGS dat byla využita webová aplikace Galaxy (<https://usegalaxy.eu/>). Statistika byla prováděna v programu GraphPad a případně v jazyce R. Pro prohlížení genomických dat byl použit program IGV v. 2.11.3.

Tabulka 15: Tabulka statistických signifikancí.

Označení	Hodnota P	Výsledek
*	0,01 - 0.05	Signifikantní
**	0,001 - 0,01	Velmi signifikantní
***	<0,001	Extrémně signifikantní
ns	>0,05	Nesignifikantní

6 Výsledky

Experimentální práce byla navržena tak, aby byla schopna postihnout funkci AGO-hook domény proteinu SPT6L porovnáním jeho nativní formy a formy s deletovanou AGO-hook doménou *spt6lΔAh*. Tato porovnání by měla ukázat, na jakých procesech se doména podílí.

Experimenty byly rozděleny do tří samostatných pracovních bloků. Nejprve analýza fylogeneze AGO-hook domény proteinu SPT6L, která je schopna ukázat a podchytit evoluční vývoj této domény u proteinů rodiny SPT6/SPT6L v rostlinách. Druhým pracovním blokem byla proteomická analýza interakčních partnerů SPT6L. Ta by měla ukázat nativní interaktory proteinu SPT6L a porovnáním s mutantním proteinem s deletovanou AGO-hook doménou ukázat i partnery, jejichž vazba je na této doméně závislá. Těmi by měly být zejména předpokládané proteiny AGO. Třetím balíčkem je analýza transkriptomu. Porovnáním transkriptomů u mutantů s deletovanou AGO-hook doménou ukáže, jak je tato doména zapojená do procesů transkripce a zpracování RNA. Do tohoto bloku byly ještě přidání mutanti *nrpe1-11* a *nrpe1-11 x spt6lΔAh*. Ty by měly ukázat případnou funkci této domény proteinu SPT6L v komplexu Pol V. Při porovnání imunoprecipitovaných proteinů byli použiti komplementovaní mutanti SPT6L-GFP a SPT6LΔAh-GFP, pro analýzu transkriptomu pak byl základem crispr mutant *spt6lΔAh*.

U proteinu SPT6L zároveň předpokládáme interakci s Pol V na základě analýz provedených naší skupinou (Čermák, nepublikováno). Proto byly experimenty upraveny tak, aby nám poskytly více informací o funkci SPT6L v komplexu Pol II, ale i v komplexu Pol V, kdy transkriptomické experimenty byly obohaceny o mutanty *nrpe1-11* a *nrpe1-11 x spt6lΔAh*.

6.1 Fylogeneze SPT5L a SPT6L

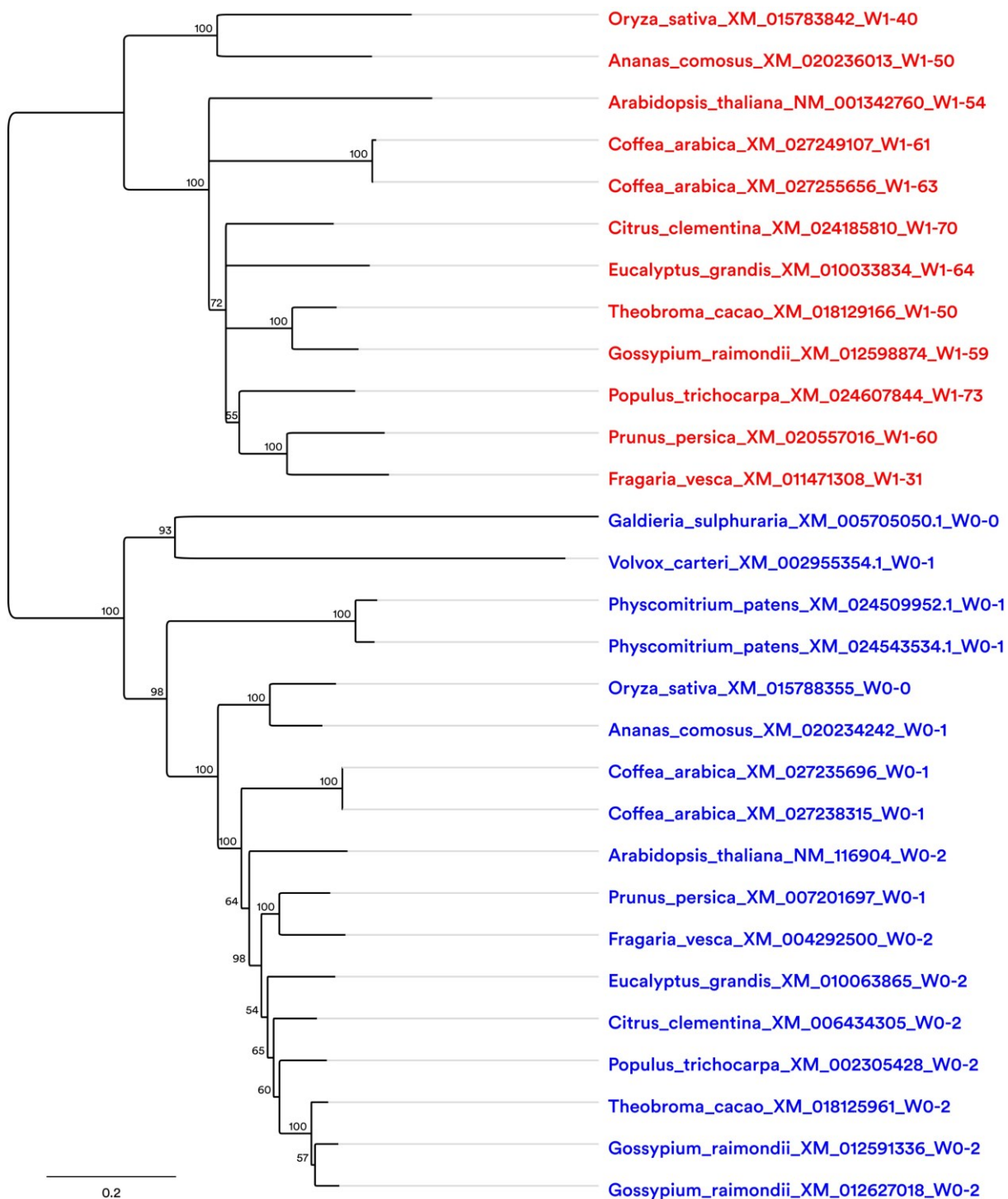
AGO-hook doména je důležitá pro rekrutování proteinů AGO. U živočichů proteiny SPT6 interagují s dalšími proteiny jako jsou RPB1 a SPT5. Tyto proteiny jsou však u rostlin diverzifikované a v komplexu Pol V mají své vlastní unikátní homology. Například SPT5L je známým homologem SPT5, který funguje jako TEF Pol II. Avšak SPT5L, který se odlišuje přítomností AGO-hook domény, pracuje v komplexu Pol V. Stejně tak jsou u rostliny *A. thaliana* přítomny dva homology SPT6, který mají také rozdílné doménové složení, kdy SPT6L má navíc AGO-hook doménu. Z toho důvodu nás zajímala fylogeneze těchto proteinů u rostlin, zdali je zde podobná diverzifikace jako je tomu u SPT5L a SPT5.

Pro porovnání fylogeneze SPT6(L) v závislosti na přítomnosti domény AGO-hook domény byly vybudovány fylogenetické stromy proteinů SPT5(L) a SPT6(L). Z výsledného stromu pro SPT5(L) je vidět, že protein diverzifikoval v průběhu evoluce na dvě samostatné větve (obr. 7). Jedna větev SPT5, je tvořena proteiny bez přítomné AGO-hook domény, které se zřejmě účastní elongace

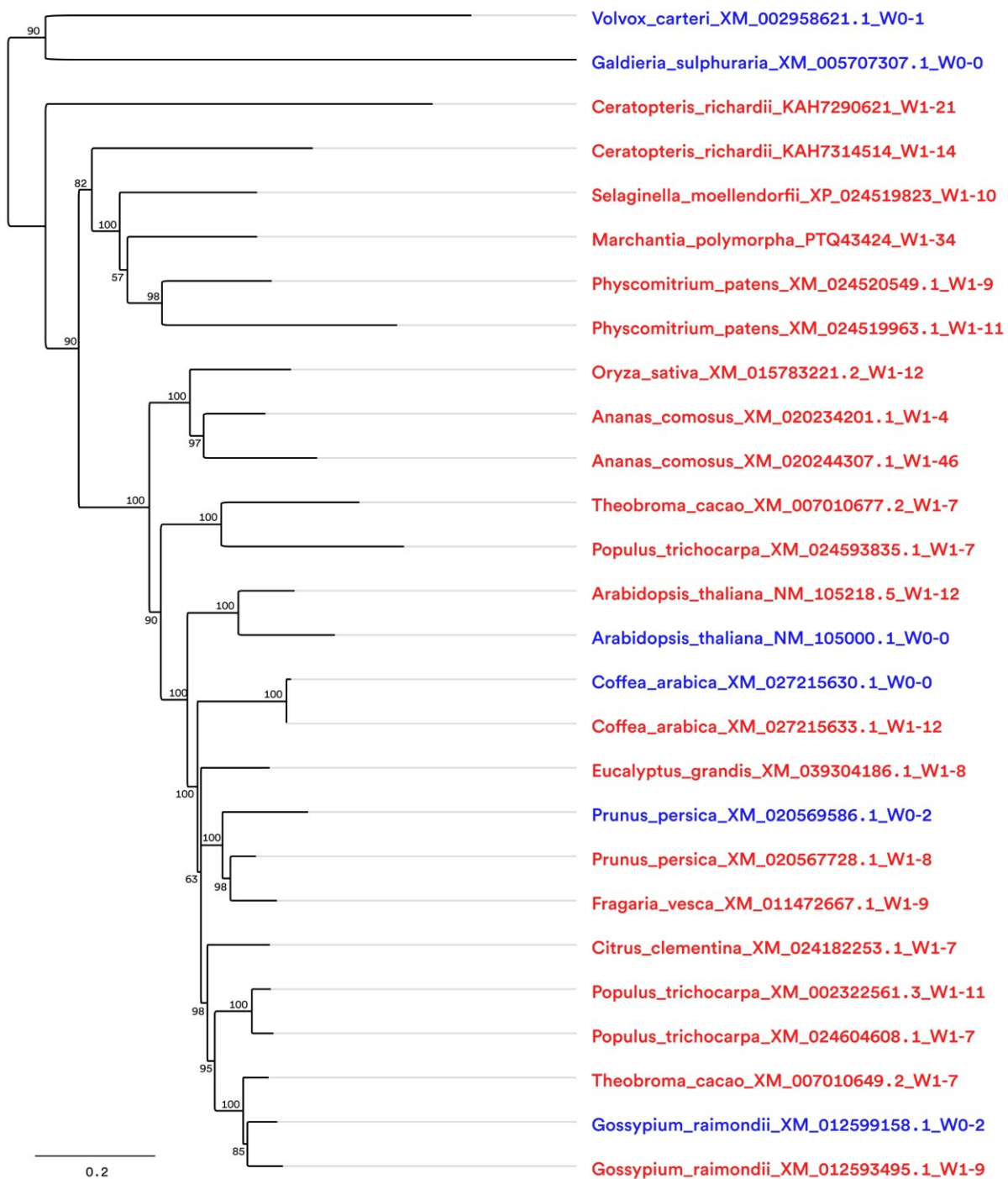
transkripce v komplexu Pol II. Druhá větev proteinů SPT5L s AGO-hook doménou obsahuje proteiny, které se účastní procesů RdDM v komplexu Pol V (Lahmy *et al.*, 2016). Tato diverzifikace proteinu SPT5L proběhla již v rané fázi vývoje cévnatých rostlin, ale mechorosty ještě proteiny SPT5L nemají.

Tak tomu však není u rodiny proteinů SPT6, kde podle provedené analýzy nejsou u většiny druhů přítomné obě verze proteinu, které by divergovaly na SPT6L s AGO-hook doménou a SPT6 bez této domény, které by mohly zastávat specifické role v komplexech Pol II a Pol V jako SPT5/SPT5L. Z fylogenetického stromu je vidět, že existence SPT6 bez AGO-hook domény je spíše výjimkou, zatímco všechny vyšší rostliny z našeho výběru (s výjimkou *Carica papaya*) mají ve svém genomu alespoň jeden gen pro protein SPT6L, který ve své sekvenci obsahuje AGO-hook doménu (obr. 8). Protein SPT6L, který má AGO-hook doménu je přítomen již u mechorostů (*Physcomytrium patens*). Je vidět, že domény AGO-hook jsou kratší, než je tomu u proteinu SPT5L a obsahují i méně WG/GW motivů, avšak jsou i výjimky, které mají počty podobné, jako je tomu třeba u *Ananas comosus*, který má 46 takových motivů.

Celkově bylo pro tvorbu stromů využito 176 sekvencí proteinů SPT5L a 124 SPT6L a tyto stromy dále potvrzují rozložení fylogeneze těchto proteinů (přílohy 1 a 2).



Obrázek 7: Zjednodušený strom fylogeneze rodiny proteinů SPT5, kdy byl náhodně vybrán jeden zástupce na čeleď. Červeně vyznačené jsou proteiny s přítomnou AGO-hook doménou, modře označené bez AGO-hook domény. Názvy proteinů jsou složeny následovně: nejprve latinské rodové a druhové jméno, poté kód použité sekvence a nakonec W0/1-X, kde W0 označuje proteiny bez AGO-hook domény a W1 s AGO-hook doménou a číslo X značí počet WG/GW motivů v daném proteinu.



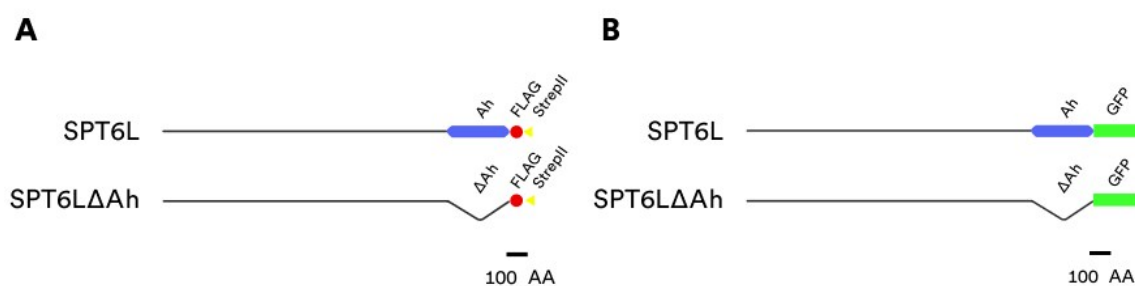
Obrázek 8: Zjednodušený strom fylogeneze SPT6L, kdy byl náhodně vybrán jeden zástupce na čeleď. Červe-ně vyznačené jsou proteiny s přítomnou AGO-hook doménou, modře označené bez AGO-hook domény. Názvy proteinů jsou složeny následovně: nejprve latinské rodové a druhové jméno, poté kód použité sekvence a nakonec W0/1-X, kde W0 označuje proteiny bez AGO-hook domény a W1 s AGO-hook doménou a číslo X značí počet WG/GW motivů v daném proteinu.

6.2 Proteomika

Předpokládanou funkcí AGO-hook domény proteinu SPT6L je vazba některého z proteinů rodiny AGO. Stejně tak jako již bylo prokázáno u proteinů NRPE1 a SPT5L, či u hsGW182 a mnoha dalších proteinů s potvrzenou AGO-hook doménou. Pro potvrzení této interakce byly vytvořeny v rámci dřívější práce V. Čermáka a M. Metličkové konstrukty pro koimunoprecipitaci SPT6L s jeho interakčními proteiny. U těch proteinů, které jsou schopné se vázat na protein SPT6L nezávisle na AGO-hook doméně, by se mohlo jednat o potencionální interakční partnery SPT6L jakožto elongačního faktoru (jako například proteiny komplexu Pol II). U proteinů, kde se interakce ztratí při delecí AGO-hook domény, by se mohlo jednat o předpokládané interakční partnery, jejichž vazba na protein SPT6L je odvislá právě od této domény (například některý protein AGO).

6.2.1 Western-blot

Pro ověření přítomnosti proteinu SPT6L v izolátu proteinů bylo nejprve žádoucí jej detekovat na WB. Pro detekci proteinu SPT6L byly dostupné dvě linie (obr. 9). První linie značená duálním StrepII--FLAG tagem a druhá linie značená fúzí s GFP, v obou případech na C-konci.

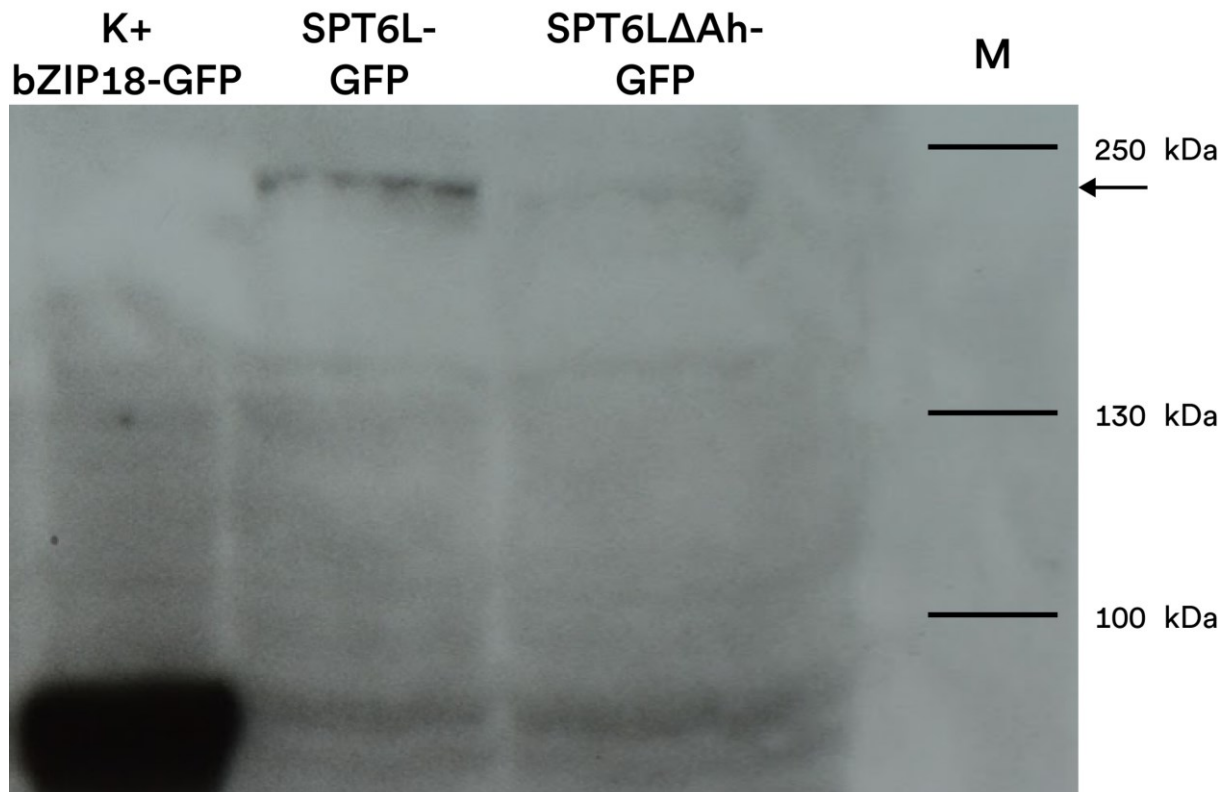


Obrázek 9: Schéma proteinů SPT6L v používaných liniích v proteomických experimentech. **A** první linie nese duální tag StrepII a FLAG tag **B** druhá používaná linie má na svém konci GFP protein.

Jako první byla detekce prováděna na mutantech s duálním tagem, kde i přes četné iterace a úpravy protokolu nebyl protein úspěšně detekován. Při optimalizacích bylo vyzkoušeno široké spektrum variací protokolu, kdy byla vyzkoušena specifita protilátek za různých podmínek (ředění, teplota, složení pufru). Protokol byl měněn i pomocí změny intenzity (stringence) jednotlivých promytí, výměnou membrán (PVDF/nitrocelulóza) i způsoby blokování membrány. Dále byly změny v izolaci proteinů, ve složení transferového pufru a délce expozice. Bylo vyzkoušeno i barvení stříbrem a dot-bloty, avšak bez úspěchu.

Z důvodu přetrvávajících neúspěchů s proteinem značeným duálním tagem, byla následně použita linie s komplementujícím proteinem SPT6L a SPT6LΔAh značeným pomocí GFP. Zde byl protein SPT6L-GFP úspěšně identifikován na WB, a to jak ve své nativní formě, tak i ve zkrácené formě

SPT6L Δ Ah-GFP (obr. 10). Optimalizovaný protokol, který byl úspěšně schopen detekovat protein jako je SPT6L na WB, využil jako vstupní materiál šešule a květy a stringentní promývání při imunodetekci. Dále, transfer na membránu trval po dobu 3 h a doba expozice výsledného WB na film byla 1 h.



Obrázek 10: WB s úspěšně detekovaným proteinem SPT6L s dobou expozice 1 hodina. K+ značí pozitivní kontrolu, kterou je zde bZIP18-GFP (Wiese *et al.*, 2021), M značí marker, kterým je PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa. Šipka označuje přibližnou polohu detekovaného proteinu SPT6L nativního i zkráceného o AGO-hook doménu ve fúzi s GFP. Na blotu je znatelný i drobný posun ve velikosti mezi jednotlivými verzemi proteinu SPT6L způsobený delecí AGO-hook domény.

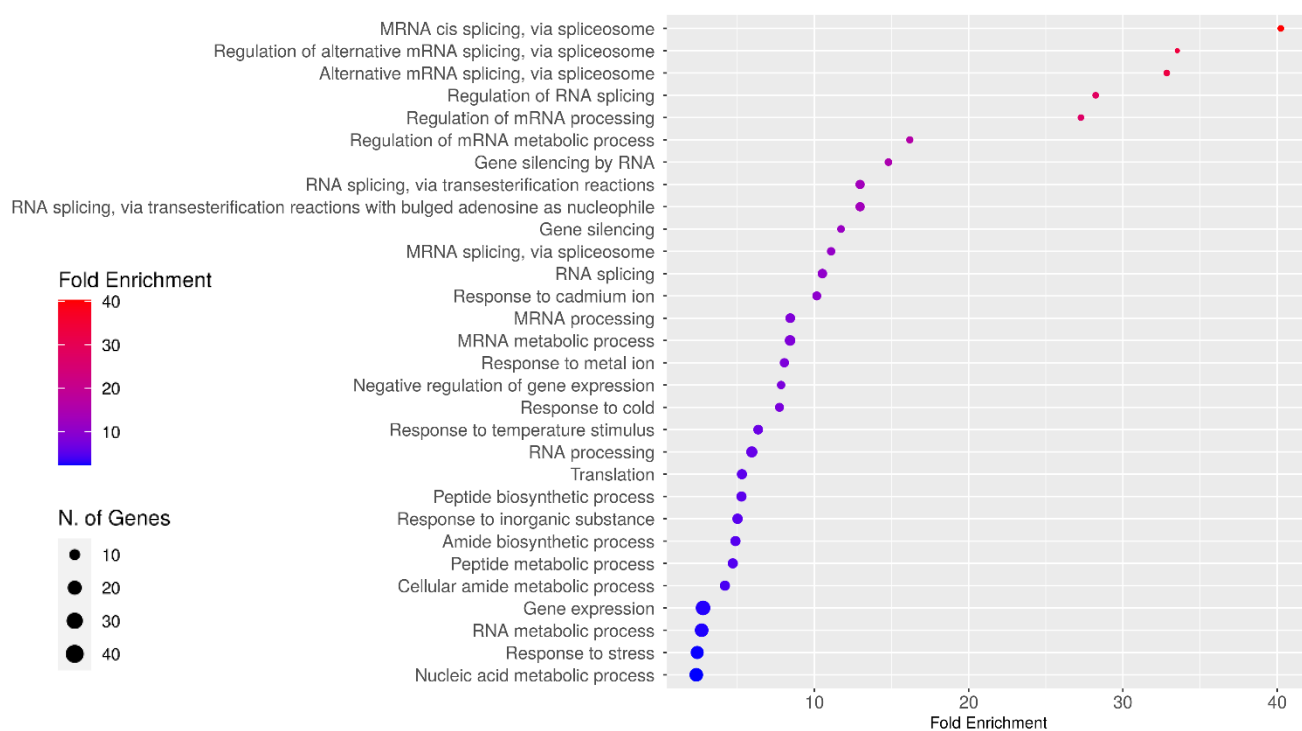
Na blotech je také patrný rozdíl ve velikostech obou proteinů. Nativní protein SPT6L je velkým proteinem o molekulové hmotnosti 185 kDa, na WB je protein SPT6L ve fúzi s GFP o přibližné velikosti 27 kDa. To dohromady dává protein o velikosti zhruba 212 kDa a delecí AGO-hook se hmotnost sníží zhruba o 20 kDa, tedy na 192 kDa.

6.2.2 Koimunoprecipitace

Po úspěšné detekci proteinu SPT6L na WB a úspěšném ověření funkčnosti tagu i protilátek byla provedena koimunoprecipitace pomocí GFP-Trap magnetických kuliček. Jako negativní kontrola sloužily rostliny s exprimovaným volným GFP, jejichž proteinový izolát byl 20x naředěn izolátem proteinů z rostlin WT, kvůli vysoké expresi volného GFP. Tato kontrola by měla být vhodná pro nalezení (a následné odfiltrování) proteinů nespecificky navázaných na GFP.

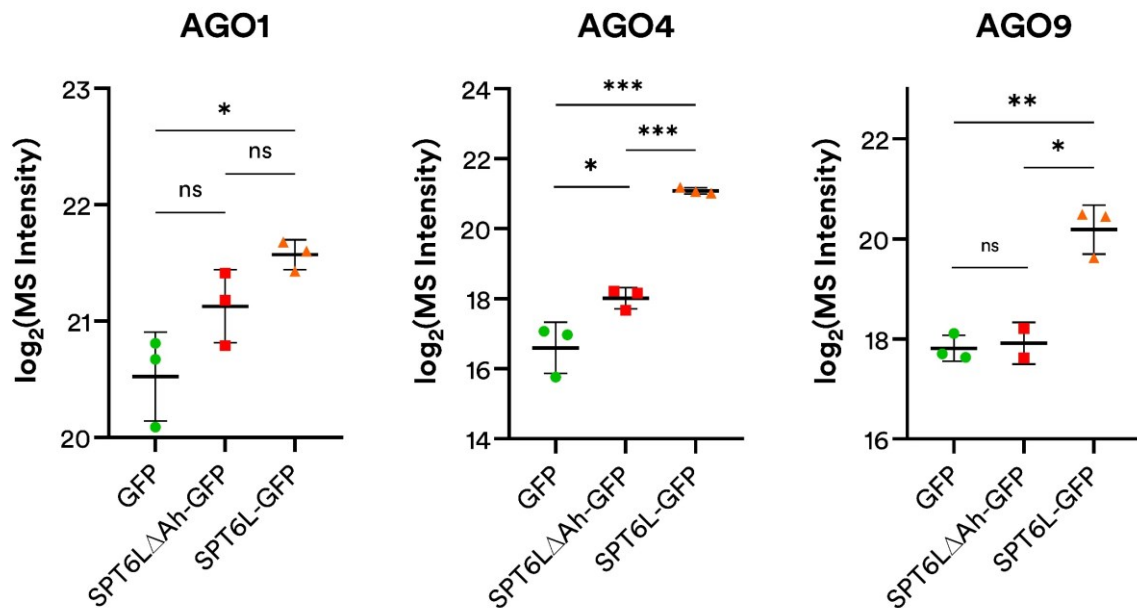
Ve vzorcích byl úspěšně detekován protein SPT6L-GFP i SPT6L Δ Ah-GFP a s nimi i jejich interakční partneri. Interakčních partnerů, kteří byli signifikantně nabohaceni ve vzorku SPT6L-GFP oproti negativní kontrole, bylo 216. Proteiny, u kterých je interakce nabohacena u kontrolních vzorků volného GFP, jsou odfiltrovány jako kontaminace, jedná se o proteiny, které interagují buď s proteinem GFP nebo s magnetickými kuličkami a protilátkami na nich. Poněvadž je SPT6L proteinem jaderným, byly identifikované proteiny dále filtrovány na proteiny jaderné, takových zde zůstalo 103. Mezi těmi nejzásadnějšími a již známými je 11 podjednotek komplexu Pol II (Antosz *et al.*, 2017), ve kterém je předpokládána funkce proteinu SPT6L.

Mezi dalšími významnými koimunoprecipitovanými interakčními proteiny bylo i SPT5L, protein z komplexu Pol V obsahující největší AGO-hook doménu v rostlině *A. thaliana*. Proteiny, které mají interakci s proteinem SPT6L, podle GO (obr. 11) patří do skupiny proteinů spolupracujících na zpracování RNA, v transkripci, v alternativním sestřihu a v regulaci těchto procesů, dále pracují na umlčování genů pomocí sRNA. Dále byly spolu s SPT6L koimunoprecipitovány histony H4, H1.2 a H2A.W.7.



Obrázek 11: GO enrichment proteinů interagujících s nativní SPT6L i s SPT6L Δ Ah. Největší nabohacení proteinů je v procesech sestřihu mRNA, ale i v rámci umlčování genů pomocí sRNA.

Pro ověření funkce AGO-hook domény byly mezi sebou porovnány vzorky SPT6L-GFP a SPT6L Δ Ah-GFP. Proteinů, u kterých vymizí signál (tedy i předpokládána interakce) v závislosti na přítomné AGO-hook doméně bylo signifikantně nabohaceno u SPT6L-GFP 92 a z toho 33 jaderných. Mezi těmito proteiny byly i proteiny AGO4 a AGO9 (obr. 12).



Obrázek 12: Grafy intenzity MS signálu pro proteiny AGO u jednotlivých koimunoprecipitovaných vzorků. Jako negativní kontrola slouží volné GFP. U vzorků s SPT6L proteinem tagovaným pomocí GFP (SPT6LΔAh-GFP a SPT6L-GFP) ukazují AGO4/9 signifikantní interakci s SPT6L závislou na přítomnosti AGO-hook domény proteinu SPT6L. AGO1 má nabohacení pouze při porovnání negativní kontroly a SPT6L-GFP. Signifikance byly vypočítány z pomoci nepárového t-testu.

Z těchto výsledků je tedy patrné, že protein SPT6L interaguje s proteiny AGO4/9 a tato interakce je závislá na přítomnosti AGO-hook domény. Dalším AGO proteinem identifikovaným na MS je AGO1. Protein SPT6L v nativní formě by s tímto proteinem mohl interagovat, avšak závislost této interakce na doméně AGO-hook, nebyla prokázána jako signifikantní.

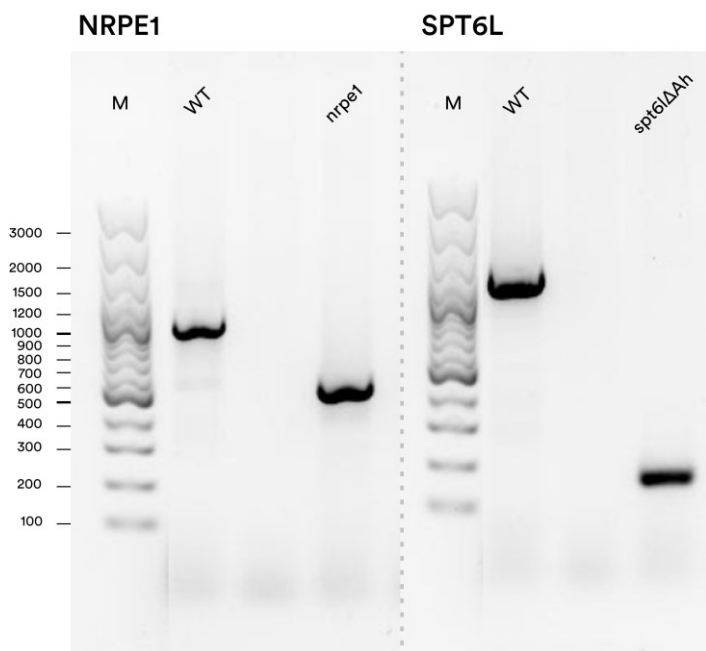
GO enrichment jednotlivých proteinů, kde je interakce závislá na přítomnosti AGO-hooku je uveden v příloze 3.

6.3 Transkriptom

Analýza transkriptomu umožní podívat se, jak ovlivňuje podobu a hladinu vznikajících transkriptů AGO-hook doména proteinu SPT6L. Zároveň díky zjištění, že by protein SPT6L mohl spolupracovat v komplexu Pol V, byly analýzy rozšířeny o mutanta podjednotky komplexu Pol V, který by mohl ukázat, jakým způsobem se na funkci SPT6L podílí Pol V a jak tato mutace interaguje s mutací AGO-hook SPT6L.

Pro experiment byly vybrány rostliny WT, *nrpe1-11*, *spt6lΔAh* a dvojitý mutant *nrpe1-11 x spt6lΔAh*. Jako WT byl použit ekotyp *A. thaliana* Col-0, jednoduché mutanty, *spt6lΔAh* a *nrpe1-11* jsem získal od

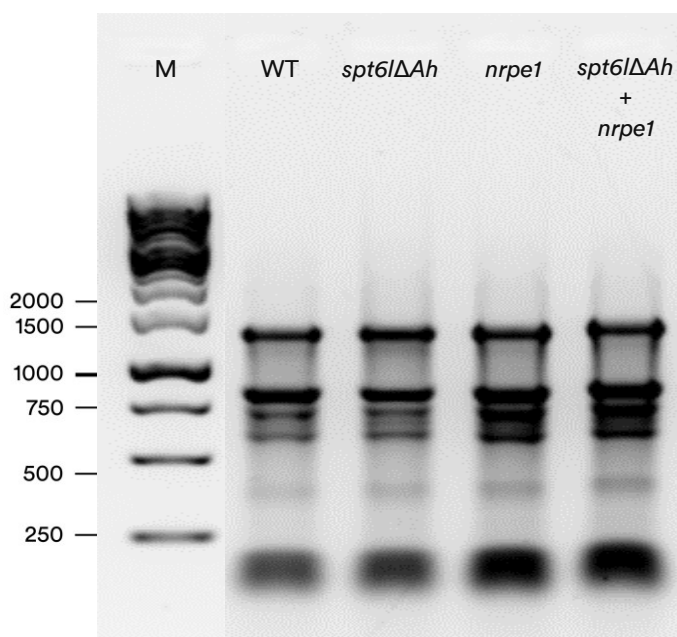
Vojtěcha Čermáka. Dvojitého mutantu *nrpe1-11 + spt6lΔAh* jsem úspěšně získal křížením předešlých zmíněných mutantů. Mutace v obou genech jsem ověřil genotypováním (obr. 13).



Obrázek 13: Genotypování dvojitého mutantu *nrpe1 + spt6lΔAh*. Vzorky na gelech si pořadím odpovídají. Použitý marker je 100 bp GeneRuler (DNA). (Gel byl rozdělen kvůli přítomnosti nerelevantních vzorků)

Z těchto mutantů byla izolována RNA a její kvalita byla překontrolována na elektroforéze. Na gelu nebyly patrné žádné větší kontaminace či degradace (obr. 14). Koncentrace byly též naměřené na Nanodropu, naměřené hodnoty 260/280 na byly v normě okolo hodnoty 2 (nejnižší 1,89). Vzorky byly ještě přeměřeny pomocí Agilent 2100 Bioanalyzer, kde hodnoty RIN odpovídaly požadovaným standardům pro RNA-seq, tedy alespoň 6.3 (nejnižší hodnota byla 7,6). Hodnoty jsou k nahlédnutí v doplňujících datech 2.

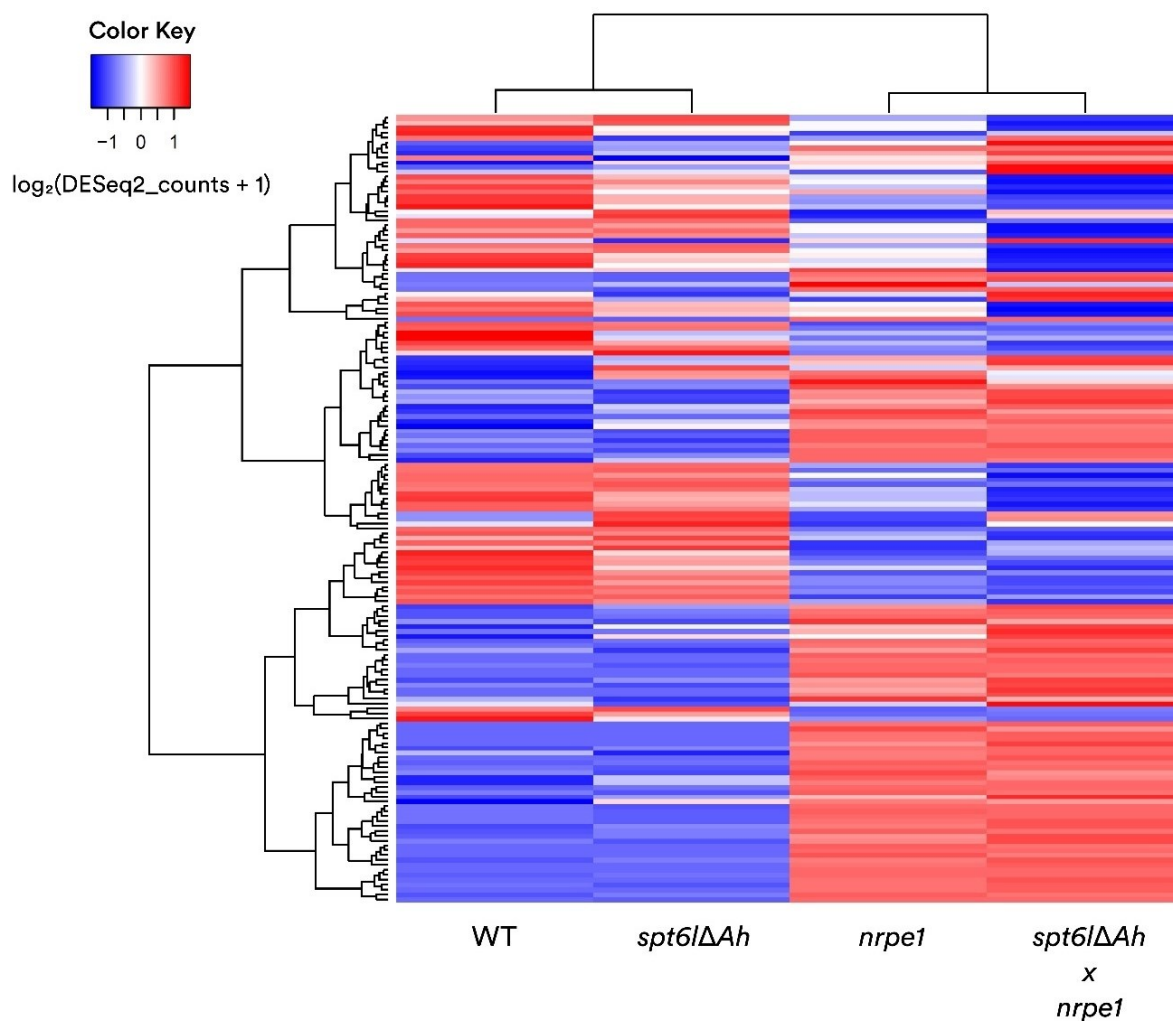
Obrázek 14: Kontrola integrity RNA na elektroforézovém gelu. Použitý marker je 1 kbp GeneRuler (DNA). Na gelu jsou jasné proužky 25S rRNA, 18S rRNA a menší proužky pro rRNA rostlinných organel či 5.8S a 5S rRNA. (Gel byl rozdělen kvůli přítomnosti nerelevantních vzorků)



6.3.1 Diferenciální exprese genů

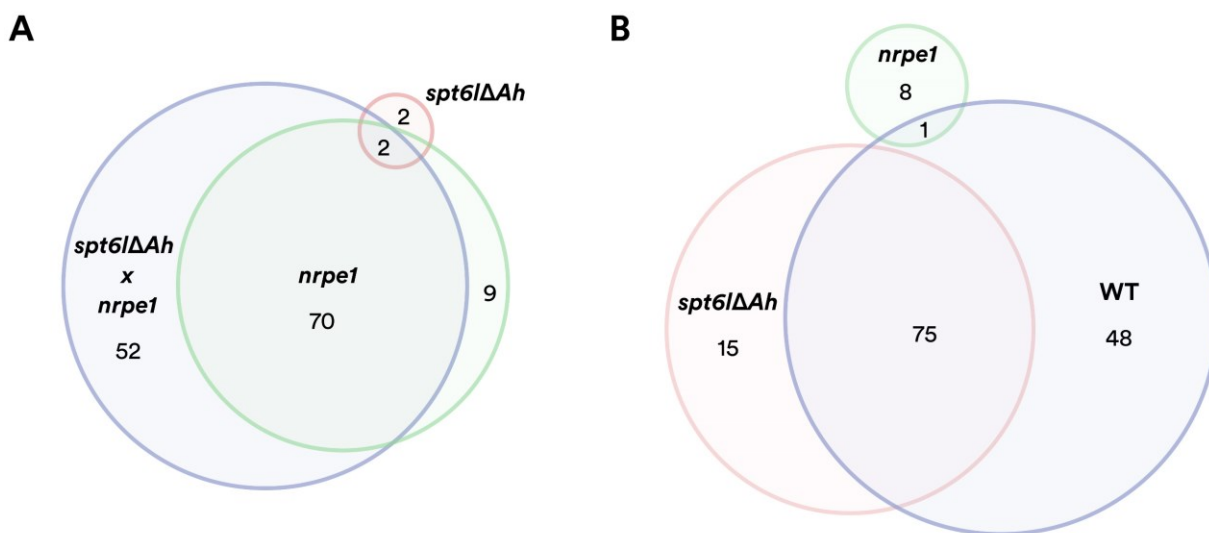
Pro zjištění role AGO-hook domény SPT6L v transkripci byla takto vyizolovaná RNA poslána na sekvenování NGS (Illumina NovaSeq 6000). Kvalita RNA byla znovu překontrolována a opět vzorky splňovaly doporučené standardy. Souhrnná statistika výstupů sekvenování a mapování je k dispozici k nahlédnutí v příloze 4. Výsledný transkriptom byl podroben analýze diferenciální exprese nástrojem DESeq2.

Z heatmapy diferenciální exprese genů lze vidět, že spolu klastrují WT s *spt6lΔAh* a dvojitý mutant s jednoduchým mutantem *nrpe1* (obr. 15).



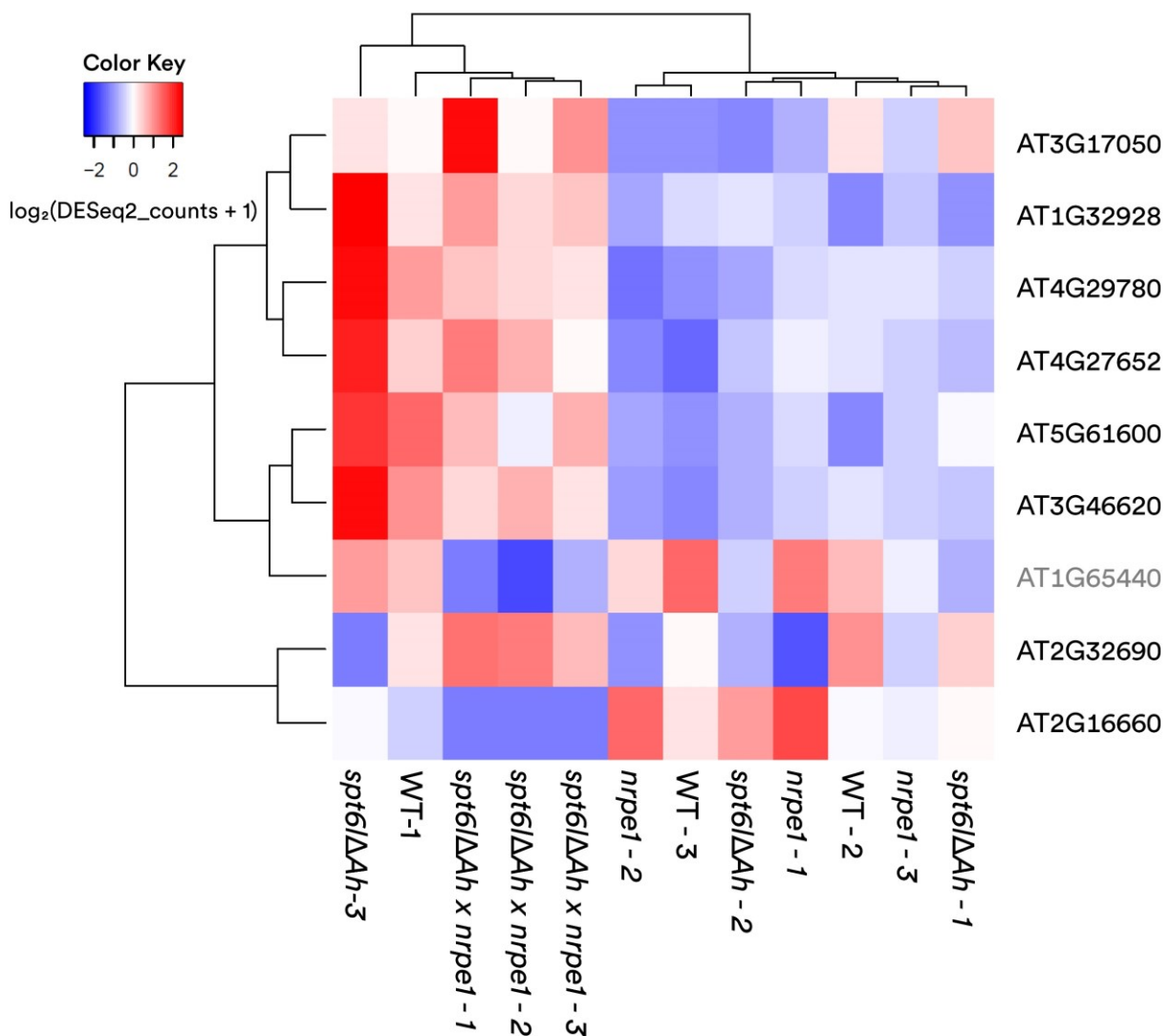
Obrázek 15: Heatmapa všech signifikantně DE genů od všech genotypů. Triplikáty jednotlivých počtů readů normalizovaných nástrojem DESeq2 jsou spojeny do společného průměru. Genotypy byly klastrovány pomocí metody hierarchického klastrování.

Tato podobnost se dá předpokládat i na základě počtů diferenciólně exprimovaných genů, neboť u mutantu *spt6lΔAh* jsou diferenciólně exprimované jen 4 geny a dva z toho specifické pro delecí AGO-hook domény, těmito geny jsou AT2G34600, AT5G11330, AT4G04223, AT2G06995. Jedná se o dvojici genů pro nekódující RNA a dva protein-kódující (obr. 16A).



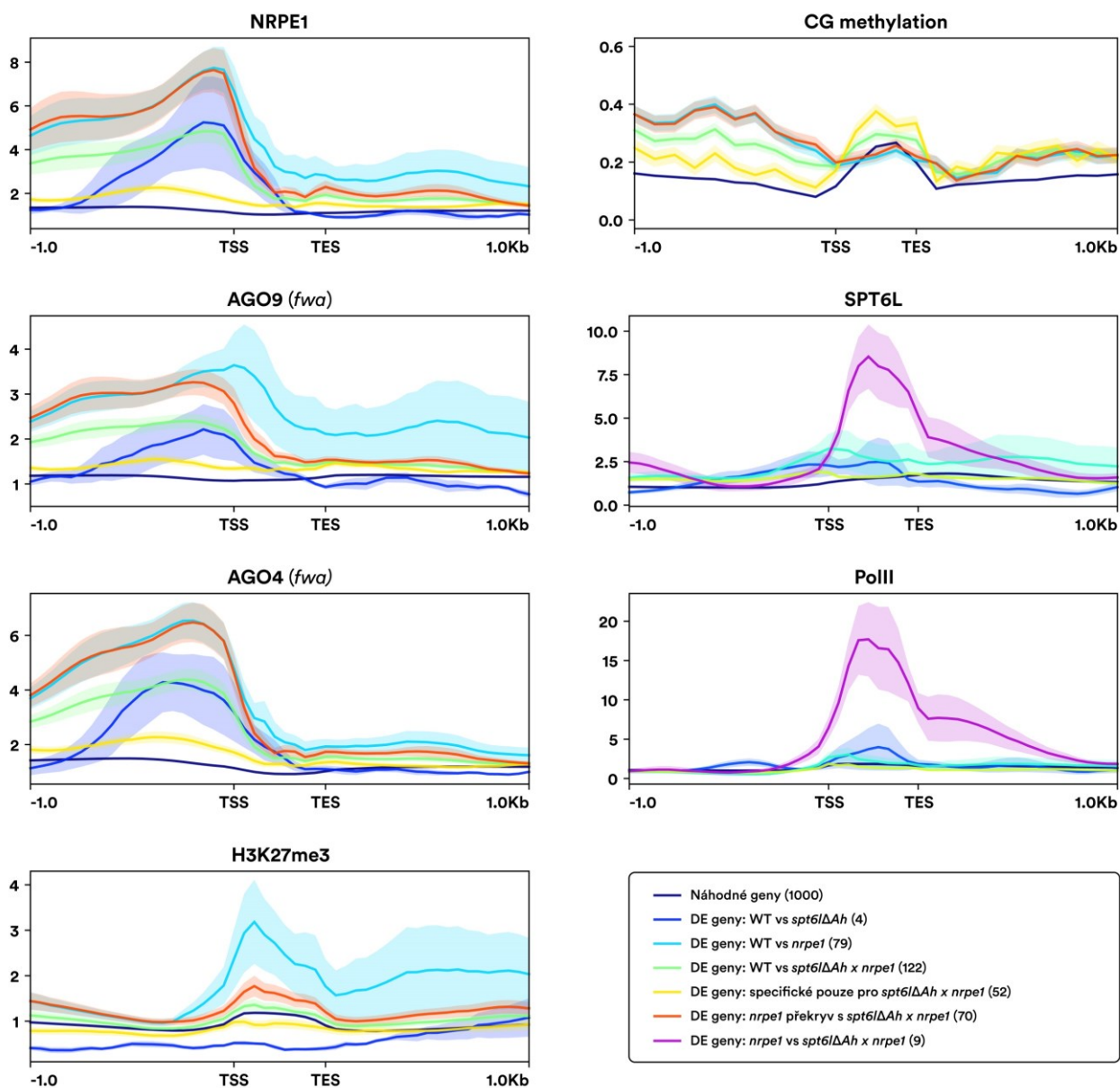
Obrázek 16: Vennovy diagramy s počty diferenciólně exprimovaných genů u různých kombinací porovnání genotypů. **A** Vennův diagram DE genů při porovnání oproti WT. **B** Vennův diagram DE genů při porovnání oproti dvojitému mutantovi *spt6lΔAh* x *nrpe1*.

Dvojitý mutant má diferenciólně exprimováno 124 genů. Z toho 52 genů není diferenciólně exprimováno v žádném jiném mutantovi, tyto geny jsou specifické pro delecí AGO-hook domény na pozadí mutantu *nrpe1* (obr. 16A). Při porovnání dvojitého mutantu oproti mutantovi *nrpe1* je diferenciólně exprimováno 9 genů. Jedním z těchto genů je protein SPT6L, kvůli porovnání WT proteinu a proteinu s deletovanou doménou AGO-hook. Zbylých osm genů je specifických pro toto porovnání. Ty jsou až na jeden více exprimovány u dvojitého mutantu. Při porovnání míry exprese těchto genů u dalších genotypů je vidět, že tato specifická byla dána tím, že se jedna replika u WT a *spt6lΔAh* silně odlišuje expresí těchto genů od zbylých dvou replik a nelze je tedy považovat za relevantní (obr. 17).

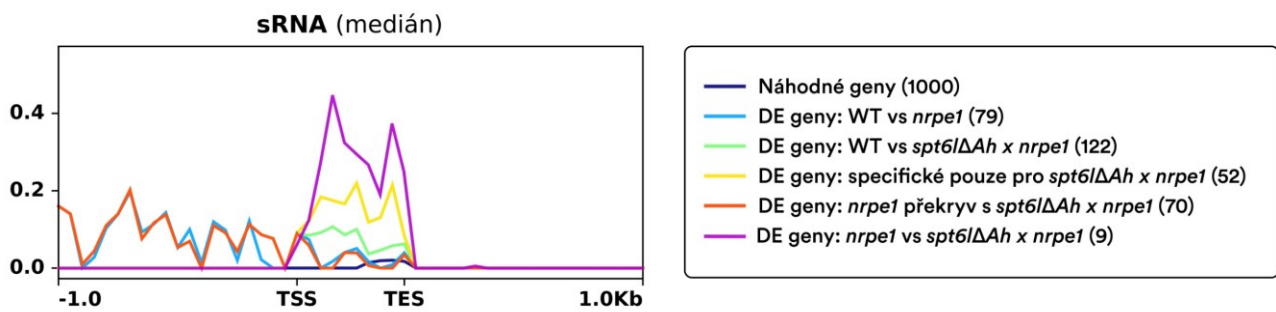


Obrázek 17: Heatmapa míry exprese pro DE geny z porovnání mutanta *nrpe1* a dvojitého mutanta *spt6lΔAh x nrpe1* (stejně geny jako na obr. 16B v zeleném kruhu). Šedě označený je protein SPT6L.

Diferenciálně exprimované geny specifické pro mutaci *nrpe1* a geny v dvojitém mutantovi překrývající se *nrpe1* mutantem (viz obr. 16A zelený kruh) mají nabohacení na chromatinu pro Pol V, AGO4 a AGO9 (obr. 18). Podobné nabohacení je i u 2 genů u mutanta *spt6lΔAh*, které tvoří překryv s mutantem *nrpe1*, avšak již se ztrácí u zbylých dvou genů a u genů specifických pouze pro dvojitého mutanta. Geny specificky diferenciálně exprimované pouze v dvojitém mutantovi mají zvýšenou úroveň CG metylace v těle genu a mapují do nich sRNA (obr. 18 a 19 žlutá linka).



Obrázek 18: Profily pro proteiny či chromatinové značky z dostupných ChIPseq dat vytvořené pro sady diferenciólně exprimovaných genů získaných z transkriptomických dat. Profily pro proteiny AGO jsou vytvořeny z datasetu ChIPseq dat z epimutanta *fwa* (*flowering wageningen*) ostatní ChIPseq data jsou na WT pozadí. Linie v grafu ukazují průměr dané značky/proteinu pro sady diferenciólně exprimovaných (DE) genů vypsáných v legendě. Vystínované oblasti poté doplňují linie a udávají standardní chybu průměru. Profily jsou vytvářeny z oblastí genů, které jsou vztaženy a proloženy do oblasti grafu odpovídající velikosti 500 nt a jsou doplněny oblastmi \pm 1kbp upstream a downstream od místa transkripčního počátku (TSS) a místa ukončení transkripce (TES). Počty genů pro dané linky nabohacení jsou uvedeny v legendě za popisem genotypu.

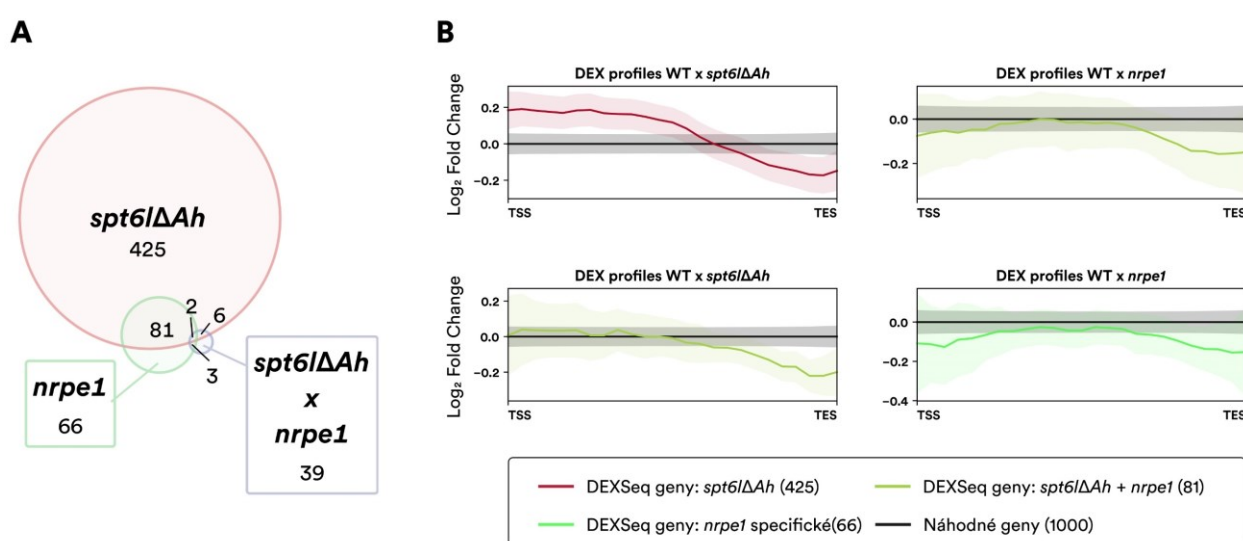


Obrázek 19: Profily sRNA (20-25 nt dlouhých) mapovaných do diferenciólně exprimovaných (DE) genů uvedených v legendě. Linie značí medián pro sRNA. Profily jsou vytvářeny z oblastí genů, které jsou vztaženy a proloženy do oblastí grafu odpovídající velikosti 500 nt a jsou doplněny oblastmi ± 1kbp upstream a downstream od místa transkripčního počátku (TSS) a místa ukončení transkripce (TES). Počty genů pro dané linky nabohacení jsou uvedeny v legendě za popisem genotypu.

6.3.2 Analýza využití exonů

Poněvadž je protein SPT6L důležitým elongačním faktorem komplexu Pol II a z imunoprecipitačních dat vyplývá, že interaguje s proteiny spliceosomu, další analýza se soustředila na změny v jednotlivých transkriptech na úrovni exonů. Pro zjištění diferenciálního využití exonů u jednotlivých genotypů byl použit nástroj DEXSeq.

U mutantu *spt6lΔAh* je vidět obrovský nárůst genů, u kterých dochází k různému využití exonů, naopak u dvojitého mutantu tento jev vymizí, a to i pod hladinu jednoduchého mutantu *nrpe1* (obr. 20A). Navíc má tato změna u mutantu *spt6lΔAh* specifický průběh v rámci jednotlivých genů, kdy se výrazně méně využívají exony, které jsou blíže u TES oproti exonům přiléhajícím k TSS (obr. 20B).



Obrázek 20: Obrázky ukazující změnu ve využití exonů u jednotlivých mutantů při porovnání s WT
A Vennův diagram s geny se změnou mírou využití exonů u jednotlivých genotypů při porovnání s WT.
B Profily vyobrazující změny ve využívání exonů podél genů. Linie ukazují míru, s jakou exony daný set genů využívá ve srovnání s WT. Hodnoty pro dané geny jsou převzaty z výstupu programu DEXSeq pro porovnání genotypů, které jsou napsány v nadpisu pro daný graf. Jsou zde ukázány tři sady genů ve dvou relevantních porovnáních genotypů. První sadou genů, označených červenou linkou jsou geny se signifikantní změnou ve využití exonů u mutantu *spt6lΔAh*, ten je vyobrazen v prvním grafu pouze. Poté jsou tmavě zelenou linkou vyneseny dva grafy pro geny se signifikantní změnou ve využívání exonů u genů, které mají společný překryv u mutantů *spt6lΔAh* a *nrpe1* (na Vennově diagramu překryv tvořen 81 geny) a hodnoty využití exonů jsou vyneseny pro porovnání obou relevantních genotypů, tedy WT vs. *spt6lΔAh* i *nrpe1*. Poslední sadou genů jsou geny se signifikantní změnou ve využití exonů u mutantu *nrpe1* (pouze) na pozadí porovnání WT x *nrpe1*. Profily pro dvojitého mutantu neukázány pro malý počet genů. Geny jsou normalizovány na délku odpovídající 500 bp od počátku transkripce (TSS) až po konec (TES). Počty genů pro dané linky nabohacení jsou uvedeny v legendě za popisem genotypu. Vystínované oblasti poté doplňují linie a udávají standartní chybu průměru.

6.3.3 Analýza alternativního sestřihu

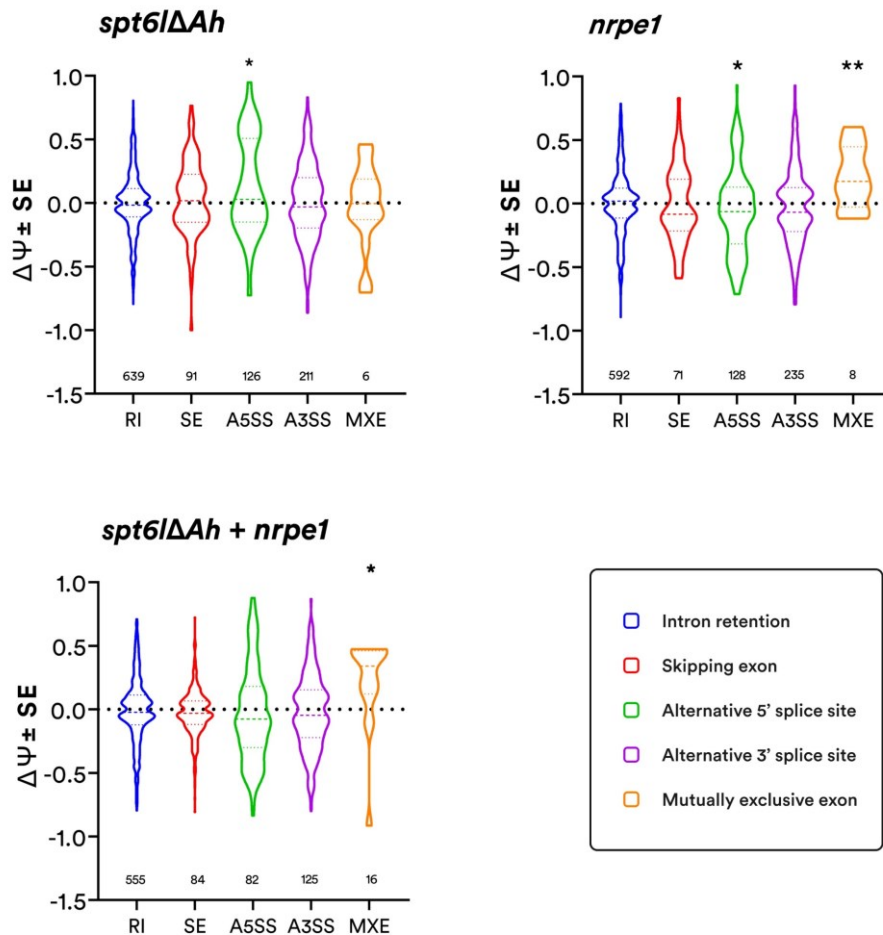
Poněvadž DEXSeq je schopen ukázat pouze využití jednotlivých exonů, a již ne dynamiku s jakou tato změna nastává, byl využit program rMATS, který je schopen odhalit změny v alternativním sestřihu u těchto genů.

Tento nástroj dokáže odlišit i jednotlivé typy událostí: exon skipping, alternativní 5' nebo 3' místo sestřihu, retence intronu nebo navzájem se vylučující exony (grafický přehled v kapitole alternativní sestřih v literárním přehledu), a to přímo z NGS readů i včetně nově zjištěných míst sestřihu.

Z obrázku 21 je vidět, že signifikantně změněné události byly zaznamenány při porovnání daných mutantů oproti WT byly zaznamenány u všech mutantů. Jak se dalo očekávat největší zastoupení mají události RI (retence intronu), které jsou u rostlin nejčastější, naopak nejméně bylo zaznamenáno událostí MXE (navzájem se vylučující exony). Celkem 126 událostí bylo zaznamenáno pro A5SS (alternativní 5' místo sestřihu) u *spt6lΔAh*, u kterého tato událost vychází signifikantně rozdílná. Jak je vidět z grafu je zde přítomná populace genů, u kterých se při porovnání s WT ztrácí události A5SS. Příklad takové signifikantní události je možné vidět na obrázku 22, kdy se u mutantu *spt6lΔAh* vytrácí ready specifické pro tento alternativní sestřih.

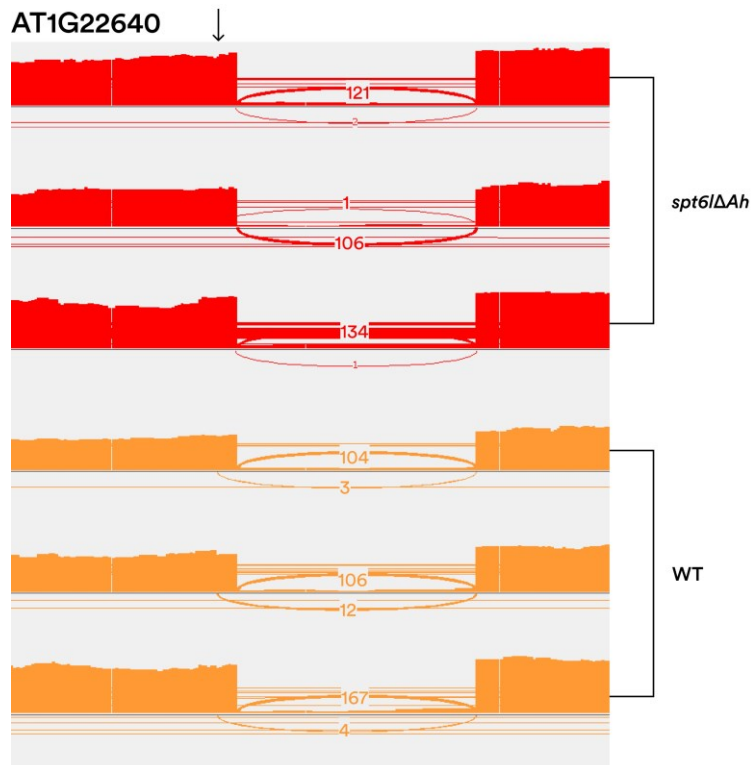
Podobný, avšak opačný trend pro dané události A5SS je signifikantní i u mutantu *nrpe1*. Kde je ještě navíc signifikantní navýšení pro MXE události, které jsou však vzácné a jejich četnosti jsou tak nízké.

U dvojitého mutantu se kompletně vytrácí signifikantní změny pro A5SS události a zůstávají zde signifikantní pouze události MXE.



Obrázek 21: Grafy s výsledky analýzy alternativního sestřihu z programu rMATS pro porovnání mutanta vs WT. Jednotlivé události jsou filtrovány $p < 0.05$; $FDR < 0.05$. Na svislé ose je vyobrazeno $\Delta\Psi$ spolu se standardní chybou průměru. $\Delta\Psi$ (PSI = percent spliced in index) vyjadřuje rozdíl v četnosti dané události alternativního sestřihu při porovnání dvou daných genotypů, kdy kladné číslo tedy říká, že je událost častější u WT a *vice versa*. Signifikance byly vypočteny za pomoci t-testu. Počty zaznamenaných události při dané hladině významnosti jsou zaznamenány pod grafem.

Obrázek 22: Sashimi plot se zamapovanými ready vygenerovaný pomocí programu IGV pro gen AT1G22640 (*MYB3*), který vyšel se signifikantní změnou A5SS při porovnání WT vs. *spt6lΔAh*. Šipka ukazuje na alternativní 5' sestřihové místo pro tuto konkrétní událost. Pod sebou jsou ukázány jednotlivé repliky pro daný genotyp.



7 Diskuze

Protein SPT6L je esenciálním proteinem, který rostlina potřebuje ke svému životu, neboť mutace v tomto genu je letální (alespoň u *A. thaliana*). Naproti tomu homolog tohoto proteinu SPT6 u *A. thaliana* není svou mutací pro rostlinu letální, ačkoliv byl také nalezen (koimunoprecipitován) v komplexu Pol II (Antosz *et al.*, 2017). To poukazuje na důležitost SPT6L proteinu nesoucího AGO-hook doménu v komplexu Pol II a naznačuje to, že protein SPT6 by mohl být potencionálním pseudogenem, zároveň je i možné, že má regulační úlohu specifickou pro *A. thaliana*, i když je mutace bez fenotypového projevu. To podporuje i fylogenetická analýza, kde takové proteiny SPT6 bez AGO-hook domény jsou v rámci rostlinné říše spíše výjimečné a každá rostlina, již včetně bezcévných kóduje protein SPT6L s přítomnou AGO-hook doménou. U *Brassicaceae* je patrná diverzifikace proteinů SPT6(L), kdy každý tvoří samostatnou větev (příloha 1). Je tedy možné, že protein SPT6 vznikl recentní duplikací v rámci tohoto taxonu.

Jedinou výjimkou z fylogenetické analýzy vyšel protein SPT6 u *Carica papaya*, který nemá AGO-hook doménu a jsou zde přítomné pouze dva WG/GW motivy. I ty by však mohly být dostatečné pro vazbu proteinu AGO, poněvadž u proteinu GW182 se každé takové vazebné místo schopné vázat protein AGO skládá ze dvou WG/GW motivů (Elkayam *et al.*, 2017). Není tedy zatím zcela jasné, zdali doména tohoto SPT6 je funkční či funkci AGO-hook domény proteinu SPT6 u této rostliny přejímá jiný protein, nebo tuto doménu v proteinu SPT6 vůbec nepotřebuje. Také je možné, že protein homologní k SPT6L u *C. papaya* nebyl ještě anotován.

Evoluce AGO-hook domény SPT6L u rostlin se objevuje paralelně se vznikem dráhy RdDM a vznikem dvou specifických polymeráz Pol IV a Pol V, kde Pol V má na své největší podjednotce také přítomen AGO-hook. Protein SPT6L s AGO-hook doménou je přítomen již u kapradin a vraneček a nachází se už i v rámci taxonu *Bryophyta*. Vznik AGO-hook domény proteinu SPT6L lze tedy datovat podobně jako vznik řady proteinů RdDM: RDR2, AGO4, DCL3, NRPE1 a je tak dokonce i vývojově starší než AGO-hook doména proteinu SPT5L, která je přítomna až u *Angiosperm* (Pei *et al.*, 2019).

Fylogenetická analýza tedy ukazuje, že AGO-hook doména proteinu SPT6L je jeho zásadní součástí. Tato doména sice zřejmě souvisí s RdDM dráhou, což ukazuje i analýza ChIP dat (Antosz *et al.*, 2017, Čermák nepublikováno), ale rostliny mají potřebu mít stejný protein jak v komplexu Pol II, tak i v komplexu Pol V. Podobně je tomu i u proteinu SPT4, který je také nediverzifikovaný a objevuje se v komplexech obou polymeráz (Köllén *et al.*, 2015). Kdyby byla AGO-hook doména potřeba jen v komplexu Pol V, pak by proteiny SPT6 diverzifikovaly na protein SPT6L účastníci se procesů v rámci komplexu Pol V a SPT6 v komplexu Pol II podobně jako je tomu v případě proteinu SPT5L a SPT5.

Že je protein SPT6L součástí komplexu Pol II bylo známo již dříve ze studie Antosz *et al.*, 2017 a bylo to potvrzeno i naším koimunoprecipitačním experimentem, kde se podařilo spolu s proteinem SPT6L precipitovat i proteiny z komplexu Pol II, celkem jejích 11 podjednotek. S proteinem SPT6L se také podařilo koimunoprecipitovat tři histonové jednotky H4, H1.2 a H2A.W.7. To jenom potvrzuje interakci SPT6L s chromatinem. Zajímavé je specifické nabohacení histonové varianty H2A.W.7, která je typická pro heterochromatin. Jedná se o histonovou variantu, která je deponována nezávisle na metylaci DNA či H3K9me2. Tento histon je schopen spolupracovat s histonem H1 a spolu regulují přístupnost heterochromatinu a metylaci DNA (Yelagandula *et al.*, 2014; Bourguet *et al.*, 2021). Přítomnost těchto dvou specifických histonů by mohla poukazovat na výskyt SPT6L v heterochromatinových sekvencích.

Dále se podařilo při imunoprecipitačním experimentu zaznamenat i signifikantní navýšení proteinu SPT5L, který je specifickým proteinem z komplexu Pol V. Tento interakční partner je tak jedinou nepřímou indicií, k již pozorovanému faktu, že se protein SPT6L nachází i v komplexu Pol V (Čermák, nepublikováno), ale může i znamenat, že protein SPT5L se nachází i v komplexu Pol II. Je to právě Pol V, která je zásadní pro udržování metylace a *de novo* metylace řady lokusů (především TEs), a to za pomoci navádění proteiny AGO. Ty byly také při koimunoprecipitaci zaznamenány. Protein SPT6L interaguje s proteiny AGO4 a 9, které se podílejí na umlčování TE a AGO4 je zásadní pro kanonickou RdDM dráhu.

Sigman *et al.*, 2021 ukazuje, že Pol V do míst bez předchozí existující metylace je naváděna některým proteinem z kladu AGO4/6/9. SPT6L je zajímavým hráčem, který je schopen okupovat místa na DNA nezávisle na Pol II (Chen *et al.*, 2019) a zároveň vázat výše zmíněné potřebné proteiny AGO.

Trochu jiné ne však navzájem se vylučující vysvětlení je, že Pol V je schopna náhodně skenovat euchromatinové oblasti bez předchozí metylace DNA a dokáže tak umlčovat nově reaktivované či zainkorporované TE (Tsuzuki *et al.*, 2020). Opět je zde SPT6L vhodným proteinem, který by mohl na taková místa Pol V přivádět a zprostředkovávat navigaci komplexu Pol V pro toto náhodné skenování euchromatinových lokusů. A to ať už výše zmíněným atrahováním proteiny AGO, či díky obojené funkci v obou polymerázách by mohl protein SPT6L sám sobě mohl taková místa okupovat a napomáhat interakci s Pol V. Protein SPT6L by již pozorovanou nespecifickou vazbou transkripčních počátků (Chen *et al.*, 2019) mohl navádět nejenom Pol II ale i Pol V. Kdy po transkripci tohoto úseku by mohl protein SPT6L zůstat navázaný na daném lokusu a pomocí vazby proteinu AGO4, do tohoto lokusu náselzně atrahovat i Pol V.

Zároveň bylo ukázáno, že protein AGO4 zprostředkovává bakteriální rezistenci u rostliny *A. thaliana*, a to nezávisle na dráze RdDM. AGO4 je v této rezistenci zásadní v aktivaci efektivní obrany rostliny za pomoci změny v metylaci na specifických lokusech (Agorio & Vera, 2007). Je tedy možné, že by protein SPT6L s navázaným AGO4 cílil do specifického subsetu lokusů v návaznosti například na biotický stres, jakým je i bakteriální infekce a zde by mohl ovlivnit úroveň metylace.

Jedna z nekanonických drah, takzvaná dicer-independent dráha, tvrdí že k proteinům AGO se přidružují dlouhé transkripty Pol II (autoři článku samy polemizují, že transkript Pol II by mohl být přepsán do dsRNA pomocí proteinu RDR6), a ty jsou poté zkracovány od konců funkcí specifických endonukleáz exosomu (Ye *et al.*, 2016). Recentní studie, která studovala velké množství komplexů u mnoha různých rostlinných druhů, ukázala, že protein SPT6L byl přítomen v komplexu proteinů exosomu (McWhite *et al.*, 2020). Tím by se protein SPT6L mohl stát velice dobrým kandidátním proteinem, který by mohl přivádět transkripty Pol II přímo na proteiny AGO, neboť představuje protein z komplexu této polymerázy a zároveň interaguje s proteinem AGO4, který je v této dráze popsán, a navíc je známo, že protein AGO4 kolokalizuje s Pol II (Duan *et al.*, 2015). Avšak koimunoprecipitace nezaznamenala žádné proteiny z komplexu exosomu.

To však neznamená, že by se nemohl v tomto komplexu nacházet, vzhledem k tomu, že nebyl koimunoprecipitován ani známý interakční partner proteinu SPT6L IWS1. Právě ten jak bylo ukázáno, je centrálním proteinem, který udržuje transkripční faktory v interakci, a to včetně SPT6 v lidských buňkách (Cermakova *et al.*, 2021). Tato interakce je zprostředkovaná doménami s vnitřní neuspořádanou strukturou. Takové domény, které postrádají stabilní uspořádanou strukturu jsou běžné pro eukaryotické TF.

Mnoho takových neuspořádaných struktur (včetně domény AGO-hook) napomáhá k fázovému oddělení, takzvaným „liquid-liquid phase separation“ (LLPS). Takové LLPS kompartmenty jsou často definovány pomocí RNA, speciálně hlavně dlouhých nekódujících RNA (Guo *et al.*, 2021). Ty zde mají důležitou funkci, kdy rekrutují proteiny jakožto molekulární lešení, které pak dokáží zprostředkovat LLPS. U hsAGO2 a proteinu hsGW182 dochází k LLPS pomocí přítomné AGO-hook domény. Toto fázové přidružení spolupracujících proteinů poté podporuje tvorbu komplexu proteinu hsAGO2 a miRNA (Sheu-Gruttadauria & MacRae, 2018). Jedním z rostlinných proteinů, který vytváří kompartmenty procesem LLPS je i protein SE, který vytváří dicing-bodies, do kterých je schopen kondenzovat proteiny DCL1, HYL1 a pri-/pre-miRNA, které jsou zde zpracovány a transportovány dále (Xie *et al.*, 2021).

Je tedy velmi pravděpodobné, že se i SPT6L svou AGO-hook doménou může podílet na LLPS, čímž může napomáhat cílené lokalizaci AGO a dalších proteinů na chromatin a k nascentnímu transkriptu, tam kde je jich potřeba, a to v komplexu Pol V nebo Pol II. Tak by mohl napomáhat sdružování potřebných faktorů, do již známých tělísek jako jsou například P-bodies pro zpracování mRNA, či do úplně odlišných nových kompartmentů ovlivňujících procesivitu a transkripci Pol II, či shromažďující faktor pro specifické činitele v komplexu Pol V.

Ačkoliv byly potvrzeny interakce AGO4/9 s AGO-hook doménou proteinu SPT6L, tak interakce závislá na této doméně nebyla signifikantní u AGO1, ačkoliv je u něj signifikantní navýšení u nativního proteinu SPT6L-GFP oproti negativní kontrole. Je ale možné, že AGO1, přítomný v relativně vysokých hladinách, mohl být na AGO-hook SPT6L navázán až v průběhu izolace.

Interakci s proteiny AGO je proto v plánu dále ověřit *in vivo* značením bioID (Roux *et al.*, 2018) a následné koimunoprecipitace (ke stejnému účelu by případně šlo využít crosslinkování). Tato metoda by měla být schopna dále potvrdit pozorovanou interakci proteinů AGO s AGO-hook doménou proteinu SPT6L. Koimunoprecipitační experimenty je dále plánováno rozšířit o dvojité mutanty *nrpe1* x SPT6L-GFP a *nrpe1* x SPT6L Δ Ah-GFP. Ty by mohly říci více o interakcích proteinu SPT6L v komplexu Pol V a funkci jeho AGO-hook domény v komplexu této polymerázy.

Protože připojené GFP je poměrně velký protein použitý jako tag k proteinu SPT6L, mohlo by stericky zabraňovat některým interakčním partnerům navázat se na protein SPT6L. Proto by do budoucna bylo případně i vhodné ještě vylepšit protokol izolace proteinu SPT6L značeného duálním tagem StrepII-FLAG, který je mnohonásobně menší nežli použité GFP, a opět provést koimunoprecipitaci s tímto proteinem.

Z výsledků analýz transkriptomu je vidět, že AGO-hook doména nemá zásadní dopad na genovou expresi, a samostatná delece této domény vede k diferencíální expresi pouhých 4 genů. Jedním z těchto genů, specifickým a nesdíleným s jinými zde studovanými mutacemi je ale gen *JAZ7*. Ten je u mutantu *spt6l Δ Ah* upregulovaný. Jedná se o gen z jasmonátové dráhy, který je schopen ovlivňovat alternativní sestřih regulovaný pomocí miRNA (Feng *et al.*, 2020). Specifita změny exprese tohoto diferencíálně exprimovaného genu by odpovídala i specificitě změněného počtu událostí AS, či využití odlišných exonů u *spt6l Δ Ah* mutantu ukázané nástrojem DEXSeq. Mezi imunoprecipitovanými proteiny byl u nativní formy SPT6L signifikantně navýšen i protein AGO1, který je hlavním regulátorem v procesech spojených s miRNA. Avšak jeho interakce nebyla signifikantně závislá na přítomnosti AGO-hook domény.

Geny, které jsou diferenciálně exprimované u mutantů *nrpe1 x spt6lΔAh* a *nrpe1*, mají ve svém chromatinu přítomné Pol IV a Pol V s proteiny AGO, což naznačuje, že jsou tyto geny regulovány procesy spojenými s RdDM. Tato data odpovídají již dříve publikovaným výsledkům pro jednoduchého mutantu *nrpe1* (Zhong *et al.*, 2012).

Genů, které byly diferenciálně exprimovány při porovnání *nrpe1* mutantu s *nrpe1 x spt6lΔAh* je celkem 8, které jsou specifické pro toto porovnání. Tyto geny nebyly nalezeny diferenciálně exprimované u žádných jiných porovnáních dvou genotypů. To znamená, že se jedná o geny specificky ovlivněné nepřítomností AGO-hook domény proteinu SPT6L na pozadí *nrpe1* mutantu. Tyto geny mají specifické nabohacení pro SPT6L a Pol II, což ukazuje na jejich aktivní transkripci. Zároveň však je do sekvencí těchto genů zamapováno vysoké množství sRNA. Je možné, že protein SPT6L do takových genů přivádí protein AGO, který je poté schopen podpořit míru transkripce takových genů. Tento jev je znám u proteinu AGO1, který interakcí s chromatinem dokáže pozitivně ovlivnit transkripci. Ve vazbě na chromatin mu navíc napomáhá chromatin remodelující komplex SWI/SNF. Takto AGO1 dokáže regulovat i geny, které jsou zodpovědné za odpověď na některé biotické stresy v odpovědi zprostředkované pomocí fytohormonů, například i jasmonáty (Liu *et al.*, 2018).

U dvojitého mutantu *nrpe1 x spt6lΔAh* bylo diferenciálně exprimováno vícero genů než u samostatného mutantu *nrpe1*. Genů specifických pro tohoto mutantu, u kterých není překryv s žádnými jinými sadami diferenciálně exprimovaných genů, je celkem 52. Tyto geny nemají specifické nabohacení na chromatinu jako je vidět u genů, které má tento mutant společné s mutantem *nrpe1*. Chybí zde nabohacení pro Pol V, Pol IV a proteiny AGO4/9. To znamená, že se jedná o geny, jejichž exprese není ovlivněna drahou RdDM, alespoň ne napřímo. U těchto genů je však typické nabohacení CG metylace v rámci těla genů, takzvaná „gene body methylation“. Funkce této metylace po délce genů je doposud poměrně neznámá, avšak existuje hypotéza, že funguje při stabilizaci transkripce na daném místě, napomáhá správnému sestřihu transkriptů a zabraňuje aberantní iniciaci transkripce. (X. Wang *et al.*, 2016; Horvath *et al.*, 2019). To také může souviset s typickou výlučností metylace DNA s histonovou variantou H2A.Z. Tato histonová varianta je obecně spojena s méně stabilními nukleozomy a tím i aktivní transkripcí (Zilberman *et al.*, 2008).

Z transkriptomických dat je dále vidět, že protein SPT6L je důležitý v některých z mnoha kroků při zpracování mRNA. To podporují koimunoprecipitované proteiny pro nativní SPT6L-GFP, které ukazují nabohacení pro proteiny, které se účastní zpracování a sestřihu mRNA. Na tomto procesu by se mohla podílet i samotná AGO-hook doména, skrze proteiny AGO, u kterých bylo ukázáno, že hrají svou roli v sestřihu intronů (Meng *et al.*, 2022). Ale je i možné, že se jedná o sekundární efekt diferenciální exprese proteinu JAZ7.

Analýza sekvenčních dat nástrojem DEXSeq ukázala na důležitost tohoto proteinu v procesech ovlivňujících finální podobu transkriptu, neboť je zde výrazné nabohacení genů, u kterých je diferenciální využití exonů pozorováno. Tento trend však znatelně vymizí u dvojitého mutantu *nrpe1 x spt6lΔAh*, a to dokonce pod úroveň samotného mutantu *nrpe1*. Pro tyto geny byly také vytvořeny profily sumarizující změnu využívání daných exonů podél těchto genů. Ta má u genů změněných u mutantu *spt6lΔAh* specifický průběh, kdy na počátku genů je jejich exprese navýšena a tento trend postupně klesá k jejich konci.

Tato změna by mohla znamenat snížené využívání těchto exonů, zároveň však díky povaze programu DEXSeq by se mohlo jednat o změnu způsobenou předčasnou terminací. Mohlo by to tedy znamenat, že Pol II vyžaduje funkci AGO-hook domény k úspěšnému dokončení přepisu některých genů. To, že tento trend vymizí u dvojitého mutantu v Pol V je prozatím nejasné, ale mohlo by souviset např. s cílením uvedených genů pomocí RdDM nebo tím, že RdDM je zdrojem sRNA, které transkripci těchto genů ovlivňují.

Výsledky získané z programu rMATS ukazují, že delece AGO-hook domény má značné účinky na procesy alternativního sestřihu. Ovlivněn je zejména alternativní sestřih 5' místa. Podobný, avšak opačný trend je vidět i mutantu *nrpe1*. U dvojitého mutantu přestává tento efekt mutací být signifikantně změněný a nejspíše se nejedná o pouhé spojení efektů obou mutací, kvůli změněnému počtu genů a vymizení populace patrné na grafu pro mutantu *spt6lΔAh*.

Podobně se v procesech alternativního sestřihu chovají i homology TRNC6, které jsou schopny vázat proteiny hsAGO2 a skrze miRNA ovlivňovat procesy alternativního sestřihu (Chu *et al.*, 2021). Tak se protein SPT6L nabízí jako faktor, který by mohl kontrolovat sestřih transkriptů Pol II a ovlivňovat finální výstup transkriptu svou AGO-hook domény a na ní navázaného proteinu AGO. Mezi geny, které jsou ovlivněny v alternativním sestřihu u těchto mutantů jsou i cíle miRNA (z rodiny miRNA169). Tomu by napovídalo nabohacení AGO1 proteinu, který je hlavním aktérem pro miRNA, avšak je také častým polutantem v koimunoprecipitačních experimentech.

Výsledky z nástroje DEXSeq a rMATS ukazují, že delece AGO-hook má značný vliv na podobu finálního transkriptu. Pozorovaná změna ve využívání jednotlivých exonů by mohla odpovídat změnám v rámci alternativního sestřihu, ačkoliv jsou překryvy mezi těmito datasey genů malé, kdy pouze 10% genů má překryv při porovnání signifikantně změněných genů při použití nástrojů rMATS a DEXSeq, avšak takto malé překryvy pro tyto nástroje jsou typické (Mehmood *et al.*, 2020). Tento malý překryv může být způsobený tím, že nástroj rMATS dokáže efektivněji podchytit změny v AS, které už nemusí být dostatečné pro zaznamenání nástrojem DEXSeq. Zároveň také ne všechny změněné geny při použití nástroje DEXSeq vznikají podstatou z alternativního sestřihu, ale například

vznikajícími aberantními transkripty, kdy AGO-hook doména může být kritická pro správnou procesivitu Pol II zejména ke konci transkriptů.

Neboť je známo, že Pol IV a V dokáží ovlivňovat 3' konce/hranice pro transkripci pomocí Pol II, kdy zpracování těchto konců u mutantů je méně efektivní a zároveň také ovlivňují zpracování a efektivitu alternativního sestřihu pre-mRNA (McKinlay *et al.*, 2018). Podobné trendy jsou pozorovatelné i u rostlin s deletovanou AGO-hook doménou, avšak společný efekt této delece s mutací *nrpe1* je opačný. Je tedy možné, že AGO-hook doména interaguje s Pol V a společně definují 3' hranice pro transkripci Pol II, což by vysvětlilo pokles procesivity u 3' konců, pozorovaný nástrojem DEXSeq.

Co se týče transkriptomických dat, jedná se o data veliká a náročná na zpracování, zejména z časového hlediska. V návaznosti na předchozí analýzy je plánováno dodatečné zjištění alternativních míst polyadenylace, předčasné terminace transkriptů a zjištění využívání kryptických promotorů. Tyto procesy je následně plánováno ověřit pomocí 5' a 3' RACE či RT-PCR.

8 Závěr

Cílem této práce bylo získat nové poznatky o funkci AGO-hook domény histonového chaperonu a transkripčního elongačního faktoru SPT6L u rostliny *A. thaliana*, a to za pomoci tří metodických balíčků: koimunoprecipitace, fylogenetická analýza a analýzy transkriptomických dat z mutantů.

Podařilo se sestavit fylogenetické stromy rodiny proteinů SPT5(L) a SPT6(L), které umožnily hlubší náhled do diverzifikace a evoluce těchto proteinů u rostlin. AGO-hook doména je u rostlin nezbytnou součástí proteinu SPT6L.

Koimunoprecipitace proteinu SPT6L potvrdila jeho přítomnost v komplexu Pol II. Zároveň se podařilo určit, že předpokládanými interagujícími AGO partnery této domény jsou AGO4/9.

Analýza transkriptomických dat (spolu s imunoprecipitačními) naznačuje, že protein SPT6L je důležitým proteinem při zpracování mRNA a že se podílí na finální podobě vzniklých transkriptů. Zároveň bylo ukázáno, že na pozadí mutantů v Pol V je dopad mutace AGO-hook domény pozměněn. Kdy mutace AGO-hook domény na pozadí mutantů *nrpe1* utlumuje dopad alternativního sestřihu a diferenciálního využití exonů. Naopak zvyšuje počet diferenciálně exprimovaných genů specifických pro kombinaci těchto dvou mutací.

9 Reference

- Agorio, A. and Vera, P. (2007) 'ARGONAUTE4 is required for resistance to *Pseudomonas syringae* in *Arabidopsis*', *Plant Cell*, 19(11), pp. 3778–3790. doi: 10.1105/tpc.107.054494.
- Anders, S., Reyes, A. and Huber, W. (2012) 'Detecting differential usage of exons from RNA-seq data', *Genome Research*, 22(10), pp. 2008–2017. doi: 10.1101/gr.133744.111.
- Andrews, S. and others (2010) *FastQC: a quality control tool for high throughput sequence data*. 2010, *Babraham bioinformatics*. Available at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (Accessed: 14 April 2022).
- Antosz, W. *et al.* (2017) 'The composition of the arabidopsis RNA polymerase II transcript elongation complex reveals the interplay between elongation and mRNA processing factors', *Plant Cell*, 29(4), pp. 854–870. doi: 10.1105/tpc.16.00735.
- Aqil, M. *et al.* (2013) 'The HIV-1 Nef Protein Binds Argonaute-2 and Functions as a Viral Suppressor of RNA Interference', *PLoS ONE*, 8(9), p. e74472. doi: 10.1371/journal.pone.0074472.
- Ardehali, M. B. *et al.* (2009) 'Spt6 enhances the elongation rate of RNA polymerase II in vivo', *EMBO Journal*, 28(8), pp. 1067–1077. doi: 10.1038/emboj.2009.56.
- Baumberger, N. and Baulcombe, D. C. (2005) 'Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs', *Proceedings of the National Academy of Sciences of the United States of America*, 102(33), pp. 11928–11933. doi: 10.1073/pnas.0505461102.
- Belotserkovskaya, R. *et al.* (2003) 'FACT facilitates transcription-dependent nucleosome alteration', *Science (New York, N.Y.)*, 301(5636), pp. 1090–1093. doi: 10.1126/SCIENCE.1085703.
- Bies-Etheve, N. *et al.* (2009) 'RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family', *EMBO Reports*, 10(6), pp. 649–654. doi: 10.1038/emboj.2009.31.
- Blevins, T. *et al.* (2015) 'Identification of pol IV and RDR2-dependent precursors of 24 nt siRNAs guiding de novo DNA methylation in arabidopsis', *eLife*, 4(e09591), pp. 1–22. doi: 10.7554/eLife.09591.
- Böhmdorfer, G. *et al.* (2014) 'RNA-directed DNA methylation requires stepwise binding of silencing factors to long non-coding RNA', *Plant Journal*, 79(2), pp. 181–191. doi: 10.1111/tpj.12563.
- Bohmert, K. *et al.* (1998) 'AGO1 defines a novel locus of Arabidopsis controlling leaf development', *EMBO Journal*, 17(1), pp. 170–180. doi: 10.1093/emboj/17.1.170.
- Bouché, N. *et al.* (2006) 'An antagonistic function for Arabidopsis DCL2 in development and a new function for DCL4 in generating viral siRNAs', *EMBO Journal*, 25(14), pp. 3347–3356. doi: 10.1038/sj.emboj.7601217.
- Bourguet, P. *et al.* (2021) 'The histone variant H2A.W and linker histone H1 co-regulate heterochromatin accessibility and DNA methylation', *Nature Communications* 2021 12:1, 12(1), pp. 1–12. doi: 10.1038/s41467-021-22993-5.
- Brázda, P. *et al.* (2020) 'Yeast Spt6 Reads Multiple Phosphorylation Patterns of RNA Polymerase II C-Terminal Domain In Vitro', *Journal of Molecular Biology*, 432(14), pp. 4092–4107. doi: 10.1016/j.jmb.2020.05.007.
- Brodersen, P. *et al.* (2008) 'Widespread translational inhibition by plant miRNAs and siRNAs', *Science*, 320(5880), pp. 1185–1190. doi: 10.1126/science.1159151.
- Brosseau, C. and Moffett, P. (2015) 'Functional and genetic analysis identify a role for arabidopsis

- ARGONAUTE5 in antiviral RNA silencing', *Plant Cell*, 27(6), pp. 1742–1754. doi: 10.1105/tpc.15.00264.
- * Buratowski, S. (2009) 'Progression through the RNA Polymerase II CTD Cycle', *Molecular Cell*. NIH Public Access, pp. 541–546. doi: 10.1016/j.molcel.2009.10.019.
- * Burguán, J. and Havelda, Z. (2011) 'Viral suppressors of RNA silencing', *Trends in Plant Science*, pp. 265–272. doi: 10.1016/j.tplants.2011.02.010.
- Campen, A. *et al.* (2008) 'TOP-IDP-Scale: A New Amino Acid Scale Measuring Propensity for Intrinsic Disorder', *Protein & Peptide Letters*, 15(9), pp. 956–963. doi: 10.2174/092986608785849164.
- Carbonell, A. *et al.* (2012) 'Functional analysis of three Arabidopsis argonautes using slicer-defective mutants', *Plant Cell*, 24(9), pp. 3613–3629. doi: 10.1105/tpc.112.099945.
- Cermakova, K. *et al.* (2021) 'A ubiquitous disordered protein interaction module orchestrates transcription elongation', *Science*, 374(6571), pp. 1113–1121. doi: 10.1126/science.abe2913.
- * Cerutti, H. and Casas-Mollano, J. A. (2006) 'On the origin and functions of RNA-mediated silencing: From protists to man', *Current Genetics*. NIH Public Access, pp. 81–99. doi: 10.1007/s00294-006-0078-x.
- Chanvivattana, Y. *et al.* (2004) 'Interaction of Polycomb-group proteins controlling flowering in Arabidopsis', *Development*, 131(21), pp. 5263–5276. doi: 10.1242/dev.01400.
- Chen, C. *et al.* (2019) 'RNA polymerase II-independent recruitment of SPT6L at transcription start sites in Arabidopsis', *Nucleic Acids Research*, 47(13), pp. 6714–6725. doi: 10.1093/nar/gkz465.
- Cheng, C. Y. *et al.* (2017) 'Araport11: a complete reannotation of the Arabidopsis thaliana reference genome', *The Plant journal : for cell and molecular biology*, 89(4), pp. 789–804. doi: 10.1111/TPJ.13415.
- Chu, Y. *et al.* (2021) 'Argonaute binding within human nuclear RNA and its impact on alternative splicing', *RNA*, 27(9), pp. 991–1003. doi: 10.1261/rna.078707.121.
- Close, D. *et al.* (2011) 'Crystal structures of the *S. cerevisiae* Spt6 core and C-terminal tandem SH2 domain', *Journal of Molecular Biology*, 408(4), pp. 697–713. doi: 10.1016/j.jmb.2011.03.002.
- Deleris, A. *et al.* (2006) 'Hierarchical Action and Inhibition of Plant Dicer-Like Proteins in Antiviral Defense', *Science*, 313(5783), pp. 68–71. doi: 10.1126/science.aaf2336.
- Dobin, A. *et al.* (2013) 'STAR: Ultrafast universal RNA-seq aligner', *Bioinformatics*, 29(1), pp. 15–21. doi: 10.1093/bioinformatics/bts635.
- Dong, Z., Han, M. H. and Fedoroff, N. (2008) 'The RNA-binding proteins HYL1 and SE promote accurate in vitro processing of pri-miRNA by DCL1', *Proceedings of the National Academy of Sciences of the United States of America*, 105(29), pp. 9970–9975. doi: 10.1073/pnas.0803356105.
- Du, J. *et al.* (2012) 'Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants', *Cell*, 151(1), pp. 167–180. doi: 10.1016/j.cell.2012.07.034.
- Duan, C. *et al.* (2015) 'Specific but interdependent functions for Arabidopsis AGO 4 and AGO 6 in RNA-directed DNA methylation', *The EMBO Journal*, 34(5), pp. 581–592. doi: 10.15252/embj.201489453.
- * Duina, A. A. (2011) 'Histone Chaperones Spt6 and FACT: Similarities and Differences in Modes of Action at Transcribed Genes', *Genetics Research International*, 2011, pp. 1–12. doi: 10.4061/2011/625210.
- Dürr, J. *et al.* (2014) 'The transcript elongation factor SPT4/SPT5 is involved in auxin-related gene expression in Arabidopsis', *Nucleic Acids Research*, 42(7), pp. 4332–4347. doi: 10.1093/nar/gku096.

- El-Shami, M. *et al.* (2007) 'Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components', *Genes and Development*, 21(20), pp. 2539–2544. doi: 10.1101/gad.451207.
- Elkayam, E. *et al.* (2017) 'Multivalent Recruitment of Human Argonaute by GW182', *Molecular Cell*, 67(4), pp. 646–658. doi: 10.1016/j.molcel.2017.07.007.
- *Erdmann, R. M. and Picard, C. L. (2020) 'RNA-directed DNA Methylation', *PLoS Genetics*, 16(10), p. e1009034. doi: 10.1371/journal.pgen.1009034.
- Eun, C. *et al.* (2011) 'AGO6 functions in RNA-mediated transcriptional gene silencing in shoot and root meristems in *Arabidopsis thaliana*', *PLoS ONE*, 6(10), p. 25730. doi: 10.1371/journal.pone.0025730.
- Ewels, P. *et al.* (2016) 'MultiQC: Summarize analysis results for multiple tools and samples in a single report', *Bioinformatics*, 32(19), pp. 3047–3048. doi: 10.1093/bioinformatics/btw354.
- Fahlgren, N. *et al.* (2006) 'Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA Affects Developmental Timing and Patterning in *Arabidopsis*', *Current Biology*, 16(9), pp. 939–944. doi: 10.1016/j.cub.2006.03.065.
- *Fang, X. and Qi, Y. (2015) 'Rnai in plants: An argonaute-centered view', *Plant Cell*, pp. 272–285. doi: 10.1105/tpc.15.00920.
- Feng, G. *et al.* (2020) 'Jasmonate induced alternative splicing responses in *Arabidopsis*', *Plant Direct*, 4(8), pp. 1–18. doi: 10.1002/pld3.245.
- Fire, A. *et al.* (1998) 'Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*', *Nature*, 391(6669), pp. 806–811. doi: 10.1038/35888.
- Frank, F. *et al.* (2012) 'Arabidopsis Argonaute MID domains use their nucleotide specificity loop to sort small RNAs', *EMBO Journal*, 31(17), pp. 3588–3595. doi: 10.1038/EMBOJ.2012.204.
- Garcia, D. *et al.* (2012) 'Ago Hook and RNA Helicase Motifs Underpin Dual Roles for SDE3 in Antiviral Defense and Silencing of Nonconserved Intergenic Regions', *Molecular Cell*, 48(1), pp. 109–120. doi: 10.1016/j.molcel.2012.07.028.
- Gascioli, V. *et al.* (2005) 'Partially redundant functions of *Arabidopsis* DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs', *Current Biology*, 15(16), pp. 1494–1500. doi: 10.1016/j.cub.2005.07.024.
- Ge, S. X. *et al.* (2020) 'ShinyGO: A graphical gene-set enrichment tool for animals and plants', *Bioinformatics*, 36(8), pp. 2628–2629. doi: 10.1093/bioinformatics/btz931.
- *Ghildiyal, M. and Zamore, P. D. (2009) 'Small silencing RNAs: An expanding universe', *Nature Reviews Genetics*. Nature Publishing Group, pp. 94–108. doi: 10.1038/nrg2504.
- Giner, A. *et al.* (2010) 'Viral protein inhibits RISC activity by argonaute binding through conserved WG/GW motifs', *PLoS Pathogens*, 6(7), pp. 1–13. doi: 10.1371/journal.ppat.1000996.
- Gouil, Q. and Baulcombe, D. C. (2016) 'DNA Methylation Signatures of the Plant Chromomethyltransferases', *PLoS Genetics*, 12(12), p. e1006526. doi: 10.1371/journal.pgen.1006526.
- Grasser, M. *et al.* (2009) 'Transcript Elongation Factor TFIIS Is Involved in *Arabidopsis* Seed Dormancy', *Journal of Molecular Biology*, 386(3), pp. 598–611. doi: 10.1016/j.jmb.2008.12.066.
- Gu, X. L. *et al.* (2012) 'SPT6L encoding a putative WG/GW-repeat protein regulates apical-basal polarity of embryo in *Arabidopsis*', *Molecular Plant*, 5(1), pp. 249–259. doi: 10.1093/mp/ssr073.
- *Guo, Q., Shi, X. and Wang, X. (2021) 'RNA and liquid-liquid phase separation', *Non-coding RNA*

Research, 6(2), pp. 92–99. doi: 10.1016/j.ncrna.2021.04.003.

Hamilton, A. J. and Baulcombe, D. C. (1999) 'A species of small antisense RNA in posttranscriptional gene silencing in plants', *Science*, 286(5441), pp. 950–952. doi: 10.1126/science.286.5441.950.

Hammond, S. M. *et al.* (2000) 'An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells', *Nature*, 404(6775), pp. 293–296. doi: 10.1038/35005107.

Hauptmann, J. *et al.* (2013) 'Turning catalytically inactive human Argonaute proteins into active slicer enzymes', *Nature Structural & Molecular Biology* 2013 20:7, 20(7), pp. 814–817. doi: 10.1038/nsmb.2577.

Havecker, E. R. *et al.* (2010) 'The arabidopsis RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci', *Plant Cell*, 22(2), pp. 321–334. doi: 10.1105/tpc.109.072199.

Heo, K. *et al.* (2008) 'FACT-Mediated Exchange of Histone Variant H2AX Regulated by Phosphorylation of H2AX and ADP-Ribosylation of Spt16', *Molecular Cell*, 30(1), pp. 86–97. doi: 10.1016/j.molcel.2008.02.029.

Hernández-Lagana, E. *et al.* (2016) 'A multigenic network of ARGONAUTE4 clade members controls early megaspore formation in arabidopsis', *Genetics*, 204(3), pp. 1045–1056. doi: 10.1534/genetics.116.188151.

Herr, A. J. *et al.* (2005) 'RNA Polymerase IV Directs Silencing of Endogenous DNA', *Curr. Cancer Drug Targets*, 308, pp. 118–120. doi: 10.1126/science.1068999.

Horvath, R. *et al.* (2019) 'Single-cell expression noise and gene-body methylation in *Arabidopsis thaliana*', *Heredity*, 123(2), pp. 81–91. doi: 10.1038/s41437-018-0181-z.

Huang, K. *et al.* (2021) 'Pol IV and RDR2: A two-RNA-polymerase machine that produces double-stranded RNA', *Science*, 374(6575), pp. 1579–1586. doi: 10.1126/science.abj9184.

Huang, Y. *et al.* (2015) 'Ancient origin and recent innovations of RNA polymerase IV and V', *Molecular Biology and Evolution*, 32(7), pp. 1788–1799. doi: 10.1093/molbev/msv060.

Jacob, Y. *et al.* (2009) 'ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing', *Nature Structural and Molecular Biology*, 16(7), pp. 763–768. doi: 10.1038/nsmb.1611.

Johnson, L. M. *et al.* (2007) 'The SRA Methyl-Cytosine-Binding Domain Links DNA and Histone Methylation', *Current Biology*, 17(4), pp. 379–384. doi: 10.1016/j.cub.2007.01.009.

Karlowski, W. M. *et al.* (2010) 'Genome-wide computational identification of WG/GW Argonaute-binding proteins in *Arabidopsis*', *Nucleic Acids Research*, 38(13), pp. 4231–4245. doi: 10.1093/nar/gkq162.

Khraiwesh, B. *et al.* (2010) 'Transcriptional Control of Gene Expression by MicroRNAs', *Cell*, 140(1), pp. 111–122. doi: 10.1016/j.cell.2009.12.023.

Khvorova, A., Reynolds, A. and Jayasena, S. D. (2003) 'Functional siRNAs and miRNAs exhibit strand bias', *Cell*, 115(2), pp. 209–216. doi: 10.1016/S0092-8674(03)00801-8.

Kim, J. *et al.* (2014) 'Arabidopsis VIM proteins regulate epigenetic silencing by modulating DNA methylation and histone modification in cooperation with MET1', *Molecular Plant*, 7(9), pp. 1470–1485. doi: 10.1093/mp/ssu079.

Kindgren, P., Ivanov, M. and Marquardt, S. (2020) 'Native elongation transcript sequencing reveals temperature dependent dynamics of nascent RNAPII transcription in *Arabidopsis*', *Nucleic Acids*

Research, 48(5), pp. 2332–2347. doi: 10.1093/nar/gkz1189.

Kishimoto, N. *et al.* (2001) 'Site specificity of the Arabidopsis MET1 DNA methyltransferase demonstrated through hypermethylation of the superman locus', *Plant Molecular Biology*, 46(2), pp. 171–183. doi: 10.1023/A:1010636222327.

Köllén, K. *et al.* (2015) 'The zinc-finger protein SPT4 interacts with SPT5L/KTF1 and modulates transcriptional silencing in Arabidopsis', *FEBS Letters*, 589(21), pp. 3254–3257. doi: 10.1016/j.febslet.2015.09.017.

van der Krol, A. R. *et al.* (1990) 'Inhibition of flower pigmentation by antisense CHS genes: promoter and minimal sequence requirements for the antisense effect', *Plant Molecular Biology*, 14(4), pp. 457–466. doi: 10.1007/BF00027492.

Lahmy, S. *et al.* (2016) 'Evidence for ARGONAUTE4–DNA interactions in RNA-directed DNA methylation in plants', *Genes and Development*, 30(23), pp. 2565–2570. doi: 10.1101/gad.289553.116.

Law, J. A. *et al.* (2010) 'A Protein Complex Required for Polymerase V Transcripts and RNA-Directed DNA Methylation in Arabidopsis', *Current Biology*, 20(10), pp. 951–956. doi: 10.1016/j.cub.2010.03.062.

Law, J. A. *et al.* (2013) 'Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1', *Nature*, 498(7454), pp. 385–389. doi: 10.1038/nature12178.

Lee, K., Park, O. S. and Seo, P. J. (2017) 'Arabidopsis ATXR2 deposits H3K36me3 at the promoters of LBD genes to facilitate cellular dedifferentiation', *Science Signaling*, 10(507), p. eaan0316. doi: 10.1126/scisignal.aan0316.

Li, D. *et al.* (2020) 'JMJD14 encoded H3K4 demethylase modulates immune responses by regulating defence gene expression and piperolic acid levels', *New Phytologist*, 225(5), pp. 2108–2121. doi: 10.1111/nph.16270.

Li, J. *et al.* (2005) 'Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis', *Current Biology*, 15(16), pp. 1501–1507. doi: 10.1016/j.cub.2005.07.029.

Li, L. *et al.* (2010) 'Arabidopsis IWS1 interacts with transcription factor BES1 and is involved in plant steroid hormone brassinosteroid regulated gene expression', *Proceedings of the National Academy of Sciences of the United States of America*, 107(8), pp. 3918–3923. doi: 10.1073/pnas.0909198107.

Li, Shaofang *et al.* (2015) 'Detection of pol IV/RDR2-dependent transcripts at the genomic scale in Arabidopsis reveals features and regulation of siRNA biogenesis', *Genome Research*, 25(2), pp. 235–245. doi: 10.1101/gr.182238.114.

Liao, Y., Smyth, G. K. and Shi, W. (2014) 'FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features', *Bioinformatics*, 30(7), pp. 923–930. doi: 10.1093/bioinformatics/btt656.

Ling, Y. *et al.* (2018) 'Thermoprimering triggers splicing memory in Arabidopsis', *Journal of Experimental Botany*, 69(10), pp. 2659–2675. doi: 10.1093/jxb/ery062.

Liu, C. *et al.* (2018) 'Arabidopsis ARGONAUTE 1 Binds Chromatin to Promote Gene Transcription in Response to Hormones and Stresses', *Developmental Cell*, 44(3), pp. 348–361.e7. doi: 10.1016/J.DEVCEL.2017.12.002.

Liu, X. *et al.* (2020) 'miR168 targets Argonaute1A mediated miRNAs regulation pathways in response to potassium deficiency stress in tomato', *BMC Plant Biology*, 20(477). doi: 10.1186/s12870-020-02660-5.

- Liu, Z. W. *et al.* (2014) 'The SET Domain Proteins SUVH2 and SUVH9 Are Required for Pol V Occupancy at RNA-Directed DNA Methylation Loci', *PLoS Genetics*, 10(1). doi: 10.1371/journal.pgen.1003948.
- Love, M. I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology*, 15(12). doi: 10.1186/s13059-014-0550-8.
- McCue, A. D. *et al.* (2015) 'ARGONAUTE 6 bridges transposable element m RNA -derived si RNA s to the establishment of DNA methylation', *The EMBO Journal*, 34(1), pp. 20–35. doi: 10.15252/embj.201489499.
- McKinlay, A. *et al.* (2018) 'RNA polymerases IV and V influence the 3' boundaries of Polymerase II transcription units in Arabidopsis', *RNA Biology*, 15(2), pp. 269–279. doi: 10.1080/15476286.2017.1409930.
- McWhite, C. D. *et al.* (2020) 'A Pan-plant Protein Complex Map Reveals Deep Conservation and Novel Assemblies', *Cell*, 181(2), pp. 460-474.e14. doi: 10.1016/j.cell.2020.02.049.
- Mehmood, A. *et al.* (2020) 'Systematic evaluation of differential splicing tools for RNA-seq studies', *Briefings in Bioinformatics*, 21(6), pp. 2052–2065. doi: 10.1093/bib/bbz126.
- Mehraj, H. *et al.* (2021) 'Characterization of Histone H3 Lysine 4 and 36 Tri-methylation in Brassica rapa L.', *Frontiers in Plant Science*, 12, p. 785. doi: 10.3389/fpls.2021.659634.
- Meng, Y. *et al.* (2022) 'The novel activity of Argonautes in intron splicing: A transcriptome-wide survey in plants', *Journal of Plant Physiology*, 270, p. 153632. doi: 10.1016/j.jplph.2022.153632.
- Mi, S. *et al.* (2008) 'Sorting of Small RNAs into Arabidopsis Argonaute Complexes Is Directed by the 5' Terminal Nucleotide', *Cell*, 133(1), pp. 116–127. doi: 10.1016/j.cell.2008.02.034.
- Minoia, S. *et al.* (2014) 'Specific Argonautes Selectively Bind Small RNAs Derived from Potato Spindle Tuber Viroid and Attenuate Viroid Accumulation In Vivo', *Journal of Virology*, 88(20), pp. 11933–11945. doi: 10.1128/jvi.01404-14.
- Moissiard, G. *et al.* (2007) 'Transitivity in Arabidopsis can be primed, requires the redundant action of the antiviral Dicer-like 4 and Dicer-like 2, and is compromised by viral-encoded suppressor proteins', *RNA*, 13(8), pp. 1268–1278. doi: 10.1261/rna.541307.
- Montgomery, T. A. *et al.* (2008) 'Specificity of ARGONAUTE7-miR390 Interaction and Dual Functionality in TAS3 Trans-Acting siRNA Formation', *Cell*, 133(1), pp. 128–141. doi: 10.1016/j.cell.2008.02.033.
- Mukherjee, K., Campos, H. and Kolaczowski, B. (2013) 'Evolution of Animal and Plant Dicers: Early Parallel Duplications and Recurrent Adaptation of Antiviral RNA Binding in Plants', *Molecular Biology and Evolution*, 30(3), pp. 627–641. doi: 10.1093/MOLBEV/MSS263.
- Napoli, C., Lemieux, C. and Jorgensen, R. (1990) 'Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans.', *The Plant Cell*, 2(4), pp. 279–289. doi: 10.1105/tpc.2.4.279.
- * Neumeier, J. and Meister, G. (2021) 'siRNA Specificity: RNAi Mechanisms and Strategies to Reduce Off-Target Effects', *Frontiers in Plant Science*. Frontiers Media S.A., p. 2196. doi: 10.3389/fpls.2020.526455.
- Olmedo-Monfil, V. *et al.* (2010) 'Control of female gamete formation by a small RNA pathway in Arabidopsis', *Nature*, 464(7288), pp. 628–632. doi: 10.1038/nature08828.
- Panda, K. *et al.* (2016) 'Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation', *Genome Biology*, 17(1). doi:

- 10.1186/s13059-016-1032-y.
- Parent, J. S. *et al.* (2015) 'Respective contributions of Arabidopsis DCL2 and DCL4 to RNA silencing', *Plant Journal*, 81(2), pp. 223–232. doi: 10.1111/tpj.12720.
- Pei, L. *et al.* (2019) 'Tracing the origin and evolution history of methylation-related genes in plants', *BMC Plant Biology*, 19(1). doi: 10.1186/s12870-019-1923-7.
- Pontes, O. *et al.* (2006) 'The Arabidopsis Chromatin-Modifying Nuclear siRNA Pathway Involves a Nucleolar RNA Processing Center', *Cell*, 126(1), pp. 79–92. doi: 10.1016/j.cell.2006.05.031.
- Pontier, D. *et al.* (2012) 'NERD, a Plant-Specific GW Protein, Defines an Additional RNAi-Dependent Chromatin-Based Pathway in Arabidopsis', *Molecular Cell*, 48(1), pp. 121–132. doi: 10.1016/j.molcel.2012.07.027.
- Potok, M. E. *et al.* (2022) 'The role of ATXR6 expression in modulating genome stability and transposable element repression in Arabidopsis', *Proceedings of the National Academy of Sciences of the United States of America*, 119(3). doi: 10.1073/pnas.2115570119.
- Raja, P. *et al.* (2008) 'Viral Genome Methylation as an Epigenetic Defense against Geminiviruses', *Journal of Virology*, 82(18), pp. 8997–9007. doi: 10.1128/jvi.00719-08.
- Ream, T. S. *et al.* (2009) 'Subunit Compositions of the RNA-Silencing Enzymes Pol IV and Pol V Reveal Their Origins as Specialized Forms of RNA Polymerase II', *Molecular Cell*, 33(2), pp. 192–203. doi: 10.1016/j.molcel.2008.12.015.
- *Reinhart, B. J. *et al.* (2002) 'MicroRNAs in plants', *Genes and Development*, 16(13), pp. 1616–1626. doi: 10.1101/gad.1004402.
- Roux, K. J. *et al.* (2018) 'BioID: A Screen for Protein-Protein Interactions', *Current Protocols in Protein Science*, 91(1), pp. 19.23.1-19.23.15. doi: 10.1002/cpps.51.
- Rowley, M. J. *et al.* (2011) 'Independent chromatin binding of ARGONAUTE4 and SPT5L/KTF1 mediates transcriptional gene silencing', *PLoS Genetics*, 7(6), p. e1002120. doi: 10.1371/journal.pgen.1002120.
- Sanan-Mishra, N. *et al.* (2021) 'Secondary siRNAs in Plants: Biosynthesis, Various Functions, and Applications in Virology', *Frontiers in Plant Science*, 12, p. 110. doi: 10.3389/fpls.2021.610283/BIBTEX.
- Schnable, P. S. *et al.* (2009) 'The B73 maize genome: Complexity, diversity, and dynamics', *Science*, 326(5956), pp. 1112–1115. doi: 10.1126/science.1178534.
- Searle, I. R. *et al.* (2010) 'JMJ14, a JmjC domain protein, is required for RNA silencing and cell-to-cell movement of an RNA silencing signal in Arabidopsis', *Genes and Development*, 24(10), pp. 986–991. doi: 10.1101/gad.579910.
- Shen, S. *et al.* (2014) 'rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data', *Proceedings of the National Academy of Sciences of the United States of America*, 111(51), pp. 5593–5601. doi: 10.1073/pnas.1419161111.
- Sheu-Gruttadauria, J. and MacRae, I. J. (2018) 'Phase Transitions in the Assembly and Function of Human miRISC', *Cell*, 173(4), pp. 946-957.e16. doi: 10.1016/j.cell.2018.02.051.
- Sigman, M. J. *et al.* (2021) 'An siRNA-guided ARGONAUTE protein directs RNA polymerase V to initiate DNA methylation', *Nature Plants*, 7(11), pp. 1461–1474. doi: 10.1038/s41477-021-01008-7.
- Singh, J. *et al.* (2019) 'Reaction Mechanisms of Pol IV, RDR2, and DCL3 Drive RNA Channeling in the siRNA-Directed DNA Methylation Pathway', *Molecular Cell*, 75(3), pp. 576-589.e5. doi: 10.1016/j.molcel.2019.07.008.

- Song, J. J. *et al.* (2003) 'The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes', *Nature Structural Biology*, 10(12), pp. 1026–1032. doi: 10.1038/nsb1016.
- Song, J. J. *et al.* (2004) 'Crystal structure of argonaute and its implications for RISC slicer activity', *Science*, 305(5689), pp. 1434–1437. doi: 10.1126/science.1102514.
- Stroud, H. *et al.* (2014) 'Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis', *Nature Structural and Molecular Biology*, 21(1), pp. 64–72. doi: 10.1038/nsmb.2735.
- Sun, M. *et al.* (2010) 'A tandem SH2 domain in transcription elongation factor Spt6 binds the phosphorylated RNA polymerase II C-terminal repeat domain (CTD)', *Journal of Biological Chemistry*, 285(53), pp. 41597–41603. doi: 10.1074/jbc.M110.144568.
- Szádeczky-Kardoss, I. *et al.* (2022) 'Elongation factor TFIIIS is essential for heat stress adaptation in plants', *Nucleic Acids Research*, 50(4), pp. 1927–1950. doi: 10.1093/nar/gkac020.
- Takeda, A. *et al.* (2008) 'The mechanism selecting the guide strand from small RNA duplexes is different among Argonaute proteins', *Plant and Cell Physiology*, 49(4), pp. 493–500. doi: 10.1093/pcp/pcn043.
- Trujillo, J. T., Beilstein, M. A. and Mosher, R. A. (2016) 'The Argonaute-binding platform of NRPE1 evolves through modulation of intrinsically disordered repeats', *New Phytologist*, 212(4), pp. 1094–1105. doi: 10.1111/nph.14089.
- Tsuzuki, M. *et al.* (2020) 'Broad noncoding transcription suggests genome surveillance by RNA polymerase V', *Proceedings of the National Academy of Sciences of the United States of America*, 117(48), pp. 30799–30804. doi: 10.1073/pnas.2014419117.
- * Vaucheret, H. (2008) 'Plant ARGONAUTES', *Trends in Plant Science*, pp. 350–358. doi: 10.1016/j.tplants.2008.04.007.
- Wang, B. B. and Brendel, V. (2006) 'Genomewide comparative analysis of alternative splicing in plants', *Proceedings of the National Academy of Sciences of the United States of America*, 103(18), pp. 7175–7180. doi: 10.1073/pnas.0602039103.
- Wang, H. *et al.* (2016) 'Arabidopsis Flower and Embryo Developmental Genes are Repressed in Seedlings by Different Combinations of Polycomb Group Proteins in Association with Distinct Sets of Cis-regulatory Elements', *PLoS Genetics*, 12(1), p. e1005771. doi: 10.1371/journal.pgen.1005771.
- Wang, X. *et al.* (2016) 'DNA Methylation Affects Gene Alternative Splicing in Plants: An Example from Rice', *Molecular Plant*, pp. 305–307. doi: 10.1016/j.molp.2015.09.016.
- Waterhouse, P. M., Graham, M. W. and Wang, M.-B. (1998) *Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA*, *Plant Biology*. doi: 10.1073/pnas.95.23.13959.
- Wiese, A. J. *et al.* (2021) 'Arabidopsis bZIP18 and bZIP52 accumulate in nuclei following heat stress where they regulate the expression of a similar set of genes', *International Journal of Molecular Sciences*, 22(2), pp. 1–24. doi: 10.3390/ijms22020530.
- Willmann, M. R. *et al.* (2011) 'The Functions of RNA-Dependent RNA Polymerases in Arabidopsis', *The Arabidopsis book*, 9(e0146). doi: 10.1199/tab.0146.
- Winston, F. *et al.* (1984) 'Mutations affecting Ty-mediated expression of the HIS4 gene of *Saccharomyces cerevisiae*', 107, pp. 179–197. doi: 10.1093/genetics/107.2.179.
- Wu, H. *et al.* (2020) 'Plant 22-nt siRNAs mediate translational repression and stress adaptation',

- Nature*, 581(7806), pp. 89–93. doi: 10.1038/s41586-020-2231-y.
- Xie, D. *et al.* (2021) ‘Phase separation of SERRATE drives dicing body assembly and promotes miRNA processing in Arabidopsis’, *Nature Cell Biology*, 23(1), pp. 32–39. doi: 10.1038/s41556-020-00606-5.
- Yan, W. *et al.* (2018) ‘Dynamic and spatial restriction of Polycomb activity by plant histone demethylases’, *Nature Plants*. Palgrave Macmillan Ltd., pp. 681–689. doi: 10.1038/s41477-018-0219-5.
- Yang, H., Howard, M. and Dean, C. (2014) ‘Antagonistic roles for H3K36me₃ and H3K27me₃ in the cold-induced epigenetic switch at Arabidopsis FLC’, *Current Biology*, 24(15), pp. 1793–1797. doi: 10.1016/j.cub.2014.06.047.
- Yang, Z. *et al.* (2006) ‘HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide’, *Nucleic Acids Research*, 34(2), pp. 667–675. doi: 10.1093/nar/gkj474.
- Ye, R. *et al.* (2016) ‘A Dicer-Independent Route for Biogenesis of siRNAs that Direct DNA Methylation in Arabidopsis’, *Molecular Cell*, 61(2), pp. 222–235. doi: 10.1016/j.molcel.2015.11.015.
- Yelagandula, R. *et al.* (2014) ‘The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in arabidopsis’, *Cell*, 158(1), pp. 98–109. doi: 10.1016/J.CELL.2014.06.006.
- Yoshikawa, M. *et al.* (2005) ‘A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis’, *Genes and Development*, 19(18), pp. 2164–2175. doi: 10.1101/gad.1352605.
- Yu, B. *et al.* (2005) ‘Methylation as a crucial step in plant microRNA biogenesis’, *Science*, 307(5711), pp. 932–935. doi: 10.1126/science.1107130.
- Zemach, A. *et al.* (2013) ‘The arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin’, *Cell*, 153(1), pp. 193–205. doi: 10.1016/j.cell.2013.02.033.
- Zhai, J. *et al.* (2015) ‘A One Precursor One siRNA Model for Pol IV-Dependent siRNA Biogenesis’, *Cell*, 163(2), pp. 445–455. doi: 10.1016/J.CELL.2015.09.032.
- Zhang, X. *et al.* (2009) ‘Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in Arabidopsis thaliana’, *Genome Biology*, 10(6), p. 62. doi: 10.1186/gb-2009-10-6-r62.
- Zhang, X. *et al.* (2012) ‘Genome-wide analysis of plant nat-siRNAs reveals insights into their distribution, biogenesis and function’, *Genome Biology*, 13(3), p. R20. doi: 10.1186/gb-2012-13-3-r20.
- Zhang, Z. *et al.* (2016) ‘Arabidopsis AGO3 predominantly recruits 24-nt small RNAs to regulate epigenetic silencing’, *Nature Plants* 2016 2:5, 2(5), pp. 1–7. doi: 10.1038/nplants.2016.49.
- Zhang, Z. and Zhang, X. (2012) ‘Argonautes compete for miR165/166 to regulate shoot apical meristem development’, *Curr Opin Plant Biol.*, 15(6), pp. 652–658. doi: 10.1016/j.pbi.2012.05.007.
- Zhong, X. *et al.* (2012) ‘DDR complex facilitates global association of RNA polymerase v to promoters and evolutionarily young transposons’, *Nature Structural and Molecular Biology*, 19(9), pp. 870–875. doi: 10.1038/nsmb.2354.
- Zhou, J. *et al.* (2010) ‘Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in arabidopsis reveals correlation between multiple histone marks and gene expression’, *Plant Molecular Biology*, 72(6), pp. 585–595. doi: 10.1007/s11103-009-9594-7.
- Zhou, M., Palanca, A. M. S. and Law, J. A. (2018) ‘Locus-specific control of the de novo DNA methylation pathway in Arabidopsis by the CLASSY family’, *Nature Genetics*, 50(6), pp. 865–873. doi: 10.1038/s41588-018-0115-y.
- Zhou, Y. *et al.* (2015) ‘Spatiotemporal sequestration of miR165/166 by arabidopsis argonaute10

promotes shoot apical meristem maintenance', *Cell Reports*, 10(11), pp. 1819–1827. doi: 10.1016/J.CELREP.2015.02.047.

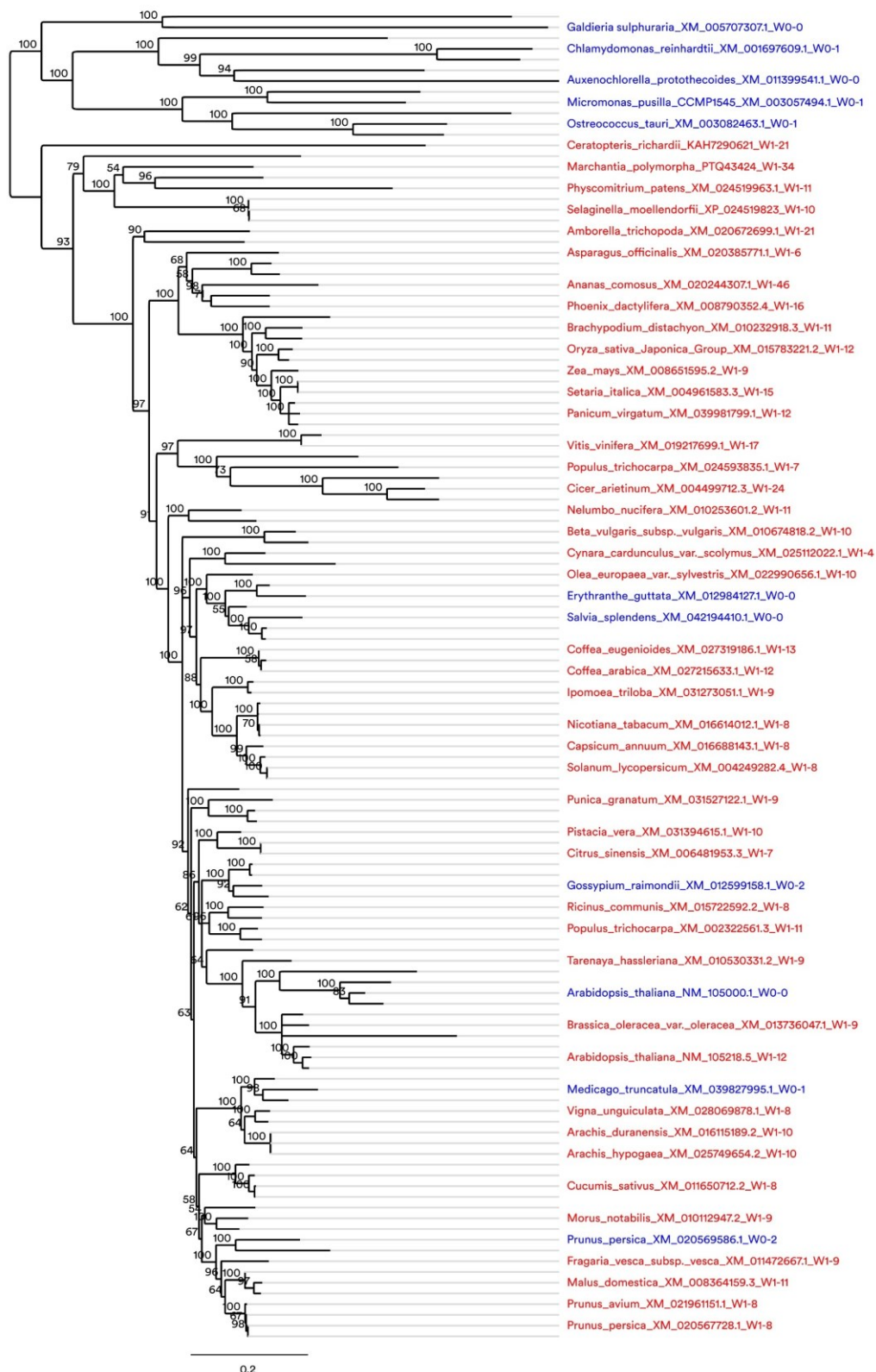
Zielezinski, A. and Karlowski, W. M. (2015) 'Integrative data analysis indicates an intrinsic disordered domain character of Argonaute-binding motifs', *Bioinformatics*, 31(3), pp. 332–339. doi: 10.1093/bioinformatics/btu666.

Zilberman, D. *et al.* (2008) 'Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks', *Nature*, 456(7218), pp. 125–129. doi: 10.1038/nature07324.

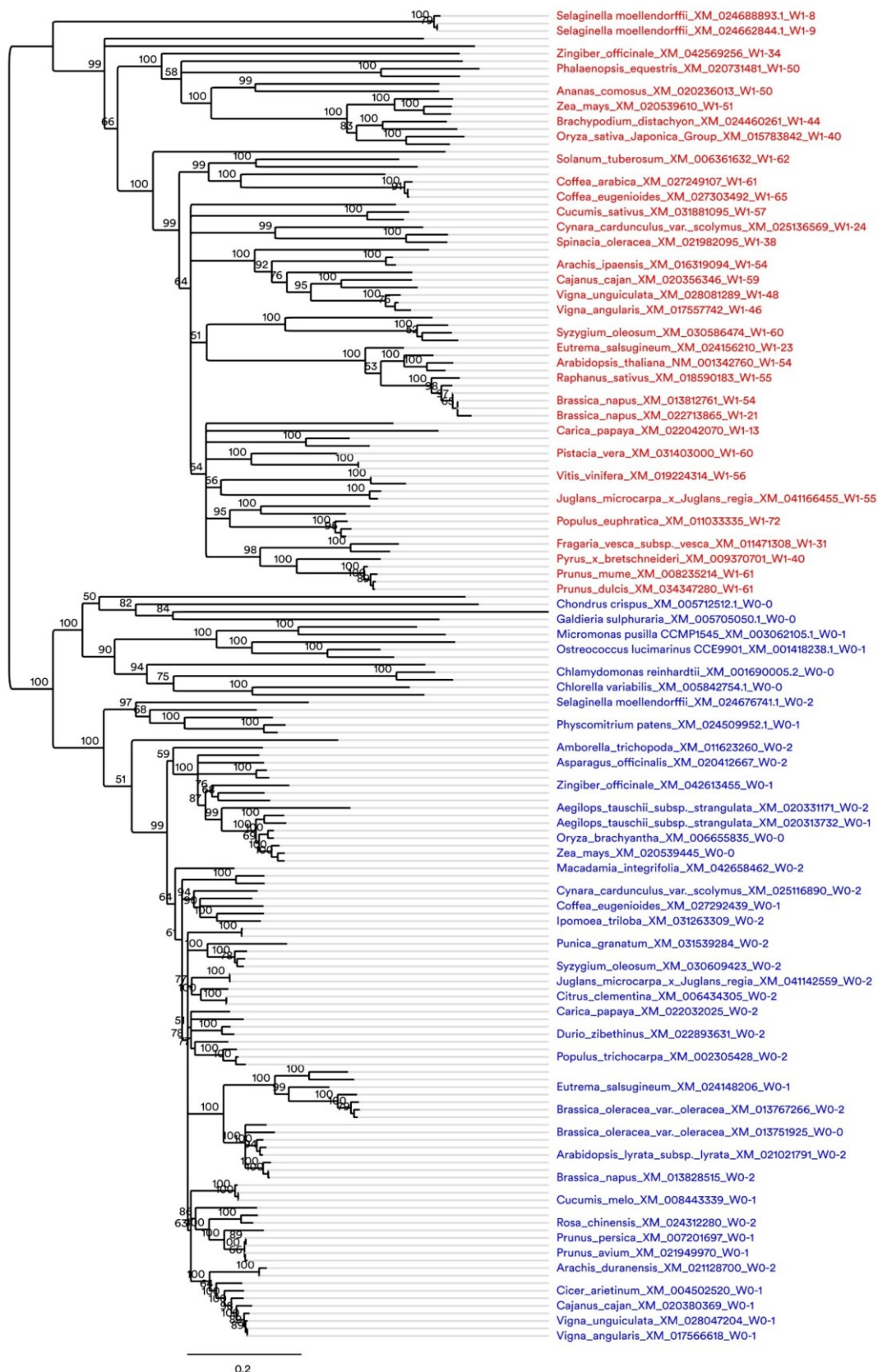
Zong, J. *et al.* (2009) 'Evolution of the RNA-dependent RNA polymerase (RdRP) genes: Duplications and possible losses before and after the divergence of major eukaryotic groups', *Gene*, 447(1), pp. 29–39. doi: 10.1016/j.gene.2009.07.004.

* označuje přehledové články

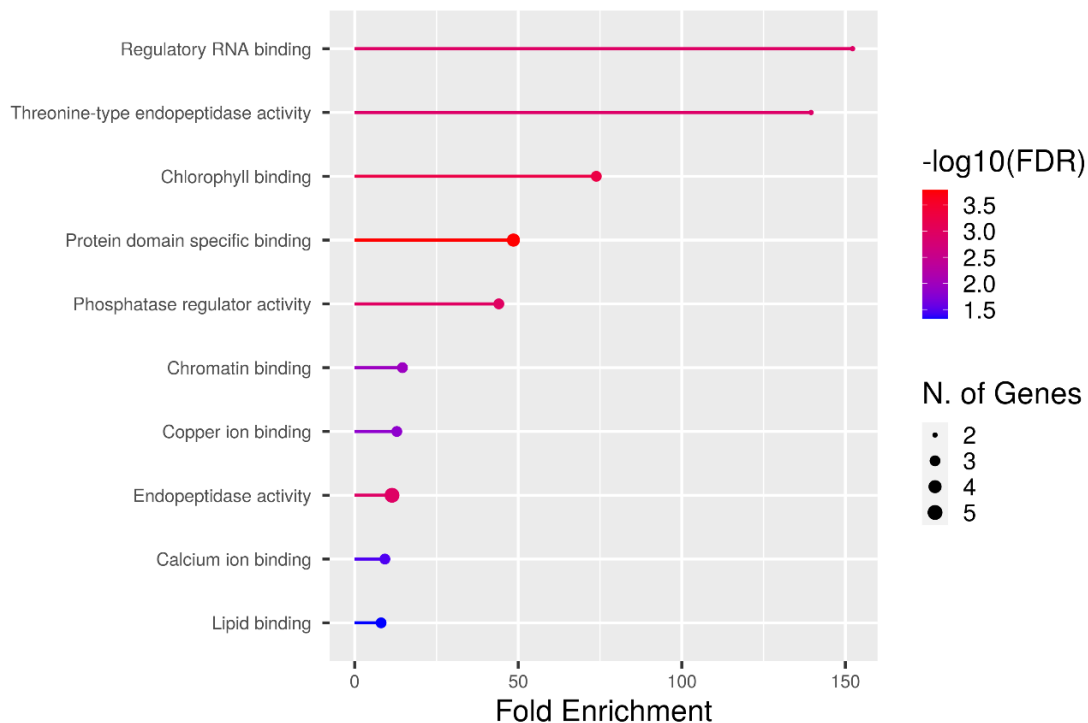
Přílohy



Příloha 1: Strom fylogeneze SPT6L obsahující 124 proteinů. Z prostorových důvodů nejsou ukázány všechny. Červeně vyznačené jsou proteiny s přítomnou AGO-hook doménou, modře označené bez AGO-hook domény. Názvy proteinů jsou složeny následovně: nejprve latinské r rodové a druhové jméno, poté kód použité sekvence a nakonec W0/1-X, kde W0 označuje proteiny bez AGO-hook domény a W1 s AGO-hook doménou a číslo X značí počet WG/GW motivů v daném proteinu.



Příloha 2: Strom fylogeneze SPT5L obsahující 176 proteinů. Z prostorových důvodů nejsou ukázány všechny. Červeně vyznačené jsou proteiny s přítomnou AGO-hook doménou, modře označené bez AGO-hook domény. Názvy proteinů jsou složeny následovně: nejprve latinské r rodové a druhové jméno, poté kód použité sekvence a nakonec W0/1-X, kde W0 označuje proteiny bez AGO-hook domény a W1 s AGO-hook doménou a číslo X značí počet WG/GW motivů v daném proteinu.

A**B**

Příloha 3: GO enrichment molekulárních funkcí koimunoprecipitovaných proteinů. **A** Graf pro geny, které interagují s proteinem SPT6LΔAh-GFP (n=58). **B** Graf pro geny, u nichž se specificky ztrácí interakce s SPT6L proteinem při delecii AGO-hook domény (n=33). Mezi těmito proteiny jsou i AGO proteiny jakožto interakční partneři AGO-hook domény.

Příloha 4: Kontrola kvality a výsledky sekvenování RNA pro jednotlivé vzorky vyizolované RNA. Raw reads ukazuje počet všech readů získaných ze sekvenování. Sloupeček Effective ukazuje (clean reads/raw reads)*100%. Error ukazuje základní chybovost sekvenování. Q20 a Q30 ukazují hodnoty přesnosti mapování a GC ukazují počet GC nukleotidů v readech. Poslední tři sloupečky tabulky říkají úspěšnost mapování a počet unikátně zamapovaných a multimapujících readů.

Genotype	Raw reads	Effective [%]	Error [%]	Q20 [%]	Q30 [%]	GC [%]	260/280	260/230	RIN	Uniquely mapped reads	Unmapped reads	Multimapping reads
WT	43,5M	99,2	0,03	97,42	92,44	46,85	1,91	2,3	7,9	93,87%	1,21%	4,91%
WT	41M	99,17	0,03	97,76	93,27	46,58	1,91	2,09	8	96,73%	1,16%	2,11%
WT	39,8M	98,47	0,03	97,73	93,17	46,41	1,94	1,89	7,7	96,25%	1,62%	2,13%
<i>stp6lΔAh x nrpe1</i>	47,7M	98,87	0,03	97,76	93,25	47,19	1,94	2,02	8,1	93,22%	1,38%	5,4%
<i>stp6lΔAh x nrpe2</i>	39M	99,2	0,03	97,69	93,06	46,48	1,94	2,2	7,6	96,70%	1,18%	2,12%
<i>stp6lΔAh x nrpe3</i>	40,2M	99,13	0,03	97,11	91,88	46,44	1,91	2,11	8	96,56%	1,47%	1,97%
<i>spt6lΔAh</i>	39,9M	99,17	0,03	97,38	92,47	46,99	1,95	1,96	8,1	93,38%	1,37%	5,25%
<i>spt6lΔAh</i>	39,6M	99,25	0,03	97,62	92,91	46,72	1,92	2,23	8	96,74%	1,23%	2,02%
<i>spt6lΔAh</i>	45,7M	99,18	0,03	97,61	92,81	46,77	1,92	2,13	8,1	93,68%	1,27%	4,87%
<i>nrpe1</i>	41,4M	99,06	0,03	97,68	93,03	46,41	1,91	2,27	8,1	96,86%	1,26%	1,88%
<i>nrpe2</i>	46,8M	99,09	0,03	97,65	92,92	46,79	1,92	2,13	8,3	96,80%	1,30%	1,90%
<i>nrpe3</i>	44,6M	99,16	0,03	97,74	93,09	46,61	1,89	1,98	8,3	96,29%	1,52%	2,18%