

Genome-wide identification and characterization of DCL, AGO and RDR gene families in *Saccharum spontaneum*

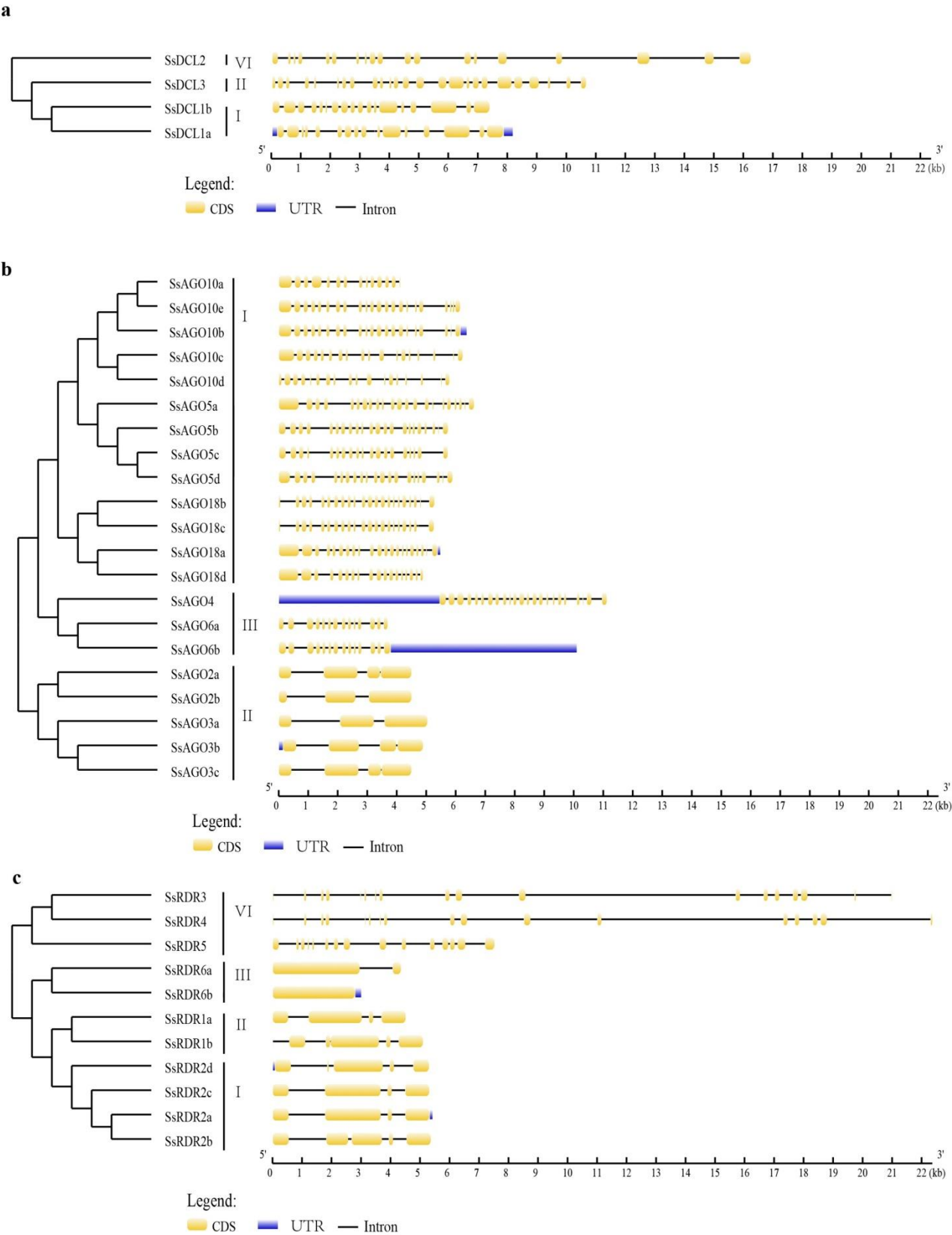
Dong-Li Cui¹, Jian-Yu Meng¹, Xiao-Yan Ren², Jing-Jing Yue³, Hua-Ying Fu¹, Mei-Ting Huang¹, Qing-Qi Zhang² and San-Ji Gao^{1, *}

¹ National Engineering Research Center for Sugarcane, Fujian Agriculture and Forestry University, Fuzhou, Fujian 350002, China

² College of Agriculture, Fujian Agriculture and Forestry University, Fuzhou, Fujian 350002, China

³ FAFU and UIUC-SIB Joint Center for Genomics and Biotechnology, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, Fujian Agriculture and Forestry University, Fuzhou, Fujian 350002, China

* Correspondence should be addressed to Qing-Qi Zhang, 000q010002@fafu.edu.cn; San-Ji Gao, gaosanji@fafu.edu.cn



Supplementary Figure S1. Structure analysis of genes encoding DCL (a), AGO (b), and RDR (c) families in *S. spontaneum*. The full-length cDNA sequences of each gene are analyzed and displayed. Coding sequences (CDS) and untranslated regions (UTR) are indicated by yellow and blue boxes, respectively. Introns are indicated by black lines. The scale bar at the bottom indicates the length of the genes in kilobase (kb).

(a) Dicer-like (DCL)

[illegible]

PAZ

SSAG02a	668	LQLLFCFW.....STKTLKLTDTLTGLITQCLFLSDRANKPMSGDOYMTNALKINKLGGSNVQLFDPLP.....RVGGVPMFIMADNVHPSFGNVE...SPS
SSAG02b	680	LFLFCPMSEQHGP.....STKTLKLTDTLTGLITQCLFLSDRANKPMSGDOYMTNALKINKLGGSNVQLFDPLP.....RVGGVPMFIMADNVHPSFGNVE...SPS
SSAG03a	572	FLFLCPSMHQGL.....STKTLKLTDTLTGLITQCLFLSDRANKPMSGDOYMTNALKINKLGGSNVQLFDPLP.....RVAGAPMFLIADNVHPSFGNVE...SPS
SSAG03b	639	LFLFLCPSMHQGL.....STKTLKLTDTLTGLITQCLFLSDRANKPMSGDOYMTNALKINKLGGSNVQLFDPLP.....RVGGVPMFIMADNVHPSFGNVE...SPS
SSAG03c	670	LQLLFCFW.....STKTLKLTDTLTGLITQCLFLSDRANKPMSGDOYMTNALKINKLGGSNVQLFDPLP.....RVGGVPMFIMADNVHPSFGNVE...SPS
SSAG04	555	FLFLCVLAEKNSDIPGPMKKKCLAEFIVTQCVAPT.....RVNDYQ...LTN.....LGGNSLQIETSPAILRVGVTITILMDGSPHSGHS...DVPS
SSAG05a	700	LTLIVFLPDANASVPSGRKKRLCELTEGIVTCQMPNNVHRRGGYQ...LONALKINKVVGGRVTIDDLNRIHLLDTLTIFGADVHPHAPED...ASPS
SSAG05b	541	LLIILIPDMSG...PSGRKKRLCELTEGLISQCCAPKNRVGGYQ...LENLSKINKVVGGRVTIDDLNRIHLLDTLTCTVFAGADVHPHSGHS...SPS
SSAG05c	518	LLIILIPDMSG...PSGRKKRLCELTEGLISQCCAPKNRVGGYQ...LENLSKINKVVGGRVTIDDLNRIHLLDTLTCTVFAGADVHPHSGHS...SPS
SSAG05d	590	LLIILIPDMSG...PSGRKKRLCELTEGLISQCCAPKNRVGGYQ...LENLSKINKVVGGRVTIDDLNRIHLLDTLTCTVFAGADVHPHSGHS...SPS
SSAG05e	590	LLIILIPDMSG...PSGRKKRLCELTEGLISQCCAPKNRVGGYQ...LENLSKINKVVGGRVTIDDLNRIHLLDTLTCTVFAGADVHPHSGHS...SPS
SSAG06a	535	FLFLCVLPERKNCDDIPGPMKKKRLHEMIVTQCLAPSN...KMDYD...FTNVLKINKALGGNSKALELHRQIMPVVTOITLILMDGSPHSGHRA...DIPS
SSAG06b	528	FLFLCVLPERKNCDDIPGPMKKKRLHEMIVTQCLAPSN...KMDYD...FTNVLKINKALGGNSKALELHRQIMPVVTOITLILMDGSPHSGHRA...DIPS
SSAG07a	629	LFLFLCPSMHQGL.....STKTLKLTDTLTGLITQCLFLSDRANKPMSGDOYMTNALKINKLGGSNVQLFDPLP.....RVGGVPMFIMADNVHPSFGNVE...SPS
SSAG010b	607	LALAILPDNNGP...LSDIKRIKLTDTGLISQCLKTHVFKISKYQ...LVANSKINKVVGGRVTIDDLAISWRILFDLTCTIFGADVHPHETED...SPS
SSAG010c	553	VMP.....LSDIKRIKLTDTGLISQCLKTHVFKISKYQ...LVANSKINKVVGGRVTIDDLAISWRILFDLTCTIFGADVHPHETED...SPS
SSAG010d	392	VMP.....LSDIKRIKLTDTGLISQCLKTHVFKISKYQ...LVANSKINKVVGGRVTIDDLAISWRILFDLTCTIFGADVHPHETED...SPS
SSAG010e	607	LALAILPDNNGP...LSDIKRIKLTDTGLISQCLKTHVFKISKYQ...LVANSKINKVVGGRVTIDDLAISWRILFDLTCTIFGADVHPHETED...SPS
SSAG018a	690	LALAILPDKNNG...LGNFKRIKLTDTGLISQCLKLNKVGAGPYQ...FANVAKINKAAGGGRNLNAPKSLSPVSVSEITPITFGADVHPAALD...TAGS
SSAG018b	445	LALAILPDKNNG...LGNFKRIKLTDTGLISQCLKLNKVGAGPYQ...FANVAKINKAAGGGRNLNAPKSLSPVSVSEITPITFGADVHPKSKGDSAGAS...TAGS
SSAG018c	445	LALAILPDKNNG...LGNFKRIKLTDTGLISQCLKLNKVGAGPYQ...FANVAKINKAAGGGRNLNAPKSLSPVSVSEITPITFGADVHPKSKGDSAGAS...TAGS
SSAG018d	645	LALAILPDKNNG...LGNFKRIKLTDTGLISQCLKLNKVGAGPYQ...FANVAKINKAAGGGRNLNAPKSLSPVSVSEITPITFGADVHPAALD...TAGS

[illegible]

PIWI

[illegible]

SsAGO2a	FARCTKPVSLATPVVYVADLVVRVGRVYVE	965
SsAGO2b	FARCTKPVSLATPVVYVADLVVRVGRVYVE	881
SsAGO3a	FARCTKPVSLATPVVYVADLVVRVGRVYVE	972
SsAGO3b	FARCTKPVSLATPVVYVADLVVRVGRVYVE	923
SsAGO3c	FARCTKPVSLATPVVYVADLVVRVGRVYVE	967
SsAGO4	YQRSTTAISVVFAPICVYHLAAQVGFQIK	854
SsAGO5a	YARCTRSVSVVPPAYVYAHLLAFRFRMYIE	1014
SsAGO5b	YARCTRSVSVVPPAYVYAHLLAFRFRMYIE	874
SsAGO5c	VPPDAYVYAHLLAFRFRMYIE	771
SsAGO5d	YARCTRSVSVVPPAYVYAHLLAFRFRMYIE	888
SsAGO6a	YARCTRSVSVVPPAYVYAHLLAFRFRMYIE	634
SsAGO6b	YARCTRSVSVVPPAYVYAHLLAFRFRMYIE	677
SsAGO10i	YARCTRSVSVVPPAYVYAHLLAFRFRMYIE	757
SsAGO10b	YARCTRSVSVVPPAYVYAHLLAFRFRMYIE	905
SsAGO10c	YARCTRSVSVVPPAYVYAHLLAFRFRMYIE	760
SsAGO10d	YARCTRSVSVVPPAYVYAHLLAFRFRMYIE	599
SsAGO10e	YMLLMQLRPVITLGFCCSTLLITYSLACSPPVQKTLTIPPAYVYAHLLAFRFRMYIE	934
SsAGO18a	YSSCTRSVSVVPPAYVYAHLLAFRFRMYIE	993
SsAGO18b	YSSCTRSVSVVPPAYVYAHLLAFRFRMYIE	753
SsAGO18c	YSSCTRSVSVVPPAYVYAHLLAFRFRMYIE	747
SsAGO18d	YSSCTRSVSVVPPAYVYAHLLAFRFRMYIE	918

(b) Argonaute (AGO)

SsDCL1a	1	MVDVLMTPQILNLRLHSIIKMDAIFLLILDECHRAVKKHPYSIVMSEFYH....TTPKDKRPVFCMTASPVNLK... 73
SsDCL1b	5	SVTYLCLTQVVMTPQILNLRLHSIIKMDAIFLLILDECHRAVKKHPYSIVMSEFYH....TTPKDKRPVFCMTASPVNLK... 83
SsDCL2	6	WNADKWRDMVDGAETVMTQPIQLDNLRLHSFFRLRDIPELLIDECHRAVKKHPYSIVMSEFYH...HPQLNSRSPDPIPTTFGMTASLINKGLAR 96
SsDCL3	34	WNAKRWKEAIGTKIIVMTQPIQLDNLRLHSFFRLRDIPELLIDECHRAVKKHPYSIVMSEFYH...IDS...EWKPAIFGMTASPV... 199
		vmt qill lrh lli dech a py efy p fgmtas

DEAD

SsDCL1a	305	..RVQALIKILLKYQHTEDFRAIIFVERVVTALVLF.....
SsDCL1b	312	PRVQALIKILLKYQHTEDFRAIIFVERVVTALVLFKVFAELPSLGFIRCASLIGHNNNQEMRSCQMDDTIKFRDGRAYSSAFNDVQSAKFHDVQGYIL
SsDCL2	296	...VIMFHALLCMHMQLDRCIVFVDRVITSIVLEALLSTINQMS...GWSVRYMASGRIGGLQHQS.....RNKHAIEIVDSFRSG...
SsDCL3	401	..VETLINIFKSFGEYIKRCFEPANQRTLDLFRAG.....
		v r
SsDCL1a		KVTLLVATSVAEEGDITROCNVIRFDLAKTVLAYIQSRGRARKPG 384
SsDCL1b		ELICVSYTSYKKTSMVGLQVTLVATSVAEEGDITROCNVIRFDLAKTVLAYIQSRGRARKPG 475
SsDCL2		KVHLLIATAILLEGLDVPSCNLIIRFDQAATVCSFTQSRGRARMQ. 415
SsDCL3		KVNLLFTIDVTEEGDVPNCSCVIRFDLPRTVCSYVQSRGRARRSS 480
		v l t eeg d c irfd tv qsrgrar

Helicase_C

SsDCL1a	456	..RVQALIKILLKYQHTEDFRAIIFVERVVTALVLF.....
SsDCL1b	547	PRVQALIKILLKYQHTEDFRAIIFVERVVTALVLFKVFAELPSLGFIRCASLIGHNNNQEMRSCQMDDTIKFRDGRAYSSAFNDVQSAKFHDVQGYIL
SsDCL2	490	...VIMFHALLCMHMQLDRCIVFVDRVITSIVLEALLSTINQMS...GWSVRYMASGRIGGLQHQS.....RNKHAIEIVDSFRSG...
SsDCL3	546	..VETLINIFKSFGEYIKRCFEPANQRTLDLFRAG.....
		v r
SsDCL1a		KVTLLVATSVAEEGDITROCNVIRFDLAKTVLAYIQSRGRARKPG 547
SsDCL1b		ELICVSYTSYKKTSMVGLQVTLVATSVAEEGDITROCNVIRFDLAKTVLAYIQSRGRARKPG 638
SsDCL2		KVHLLIATAILLEGLDVPSCNLIIRFDQAATVCSFTQSRGRARMQ. 573
SsDCL3		KVNLLFTIDVTEEGDVPNCSCVIRFDLPRTVCSYVQSRGRARRSS 633
		v l t eeg d c irfd tv qsrgrar

Dicer_dimer

SsDCL1a	748	DRGRGIFNDYQNGKGLFMADSCWDAKDLAGMVVTAHESCKRFYVDSICYNMNBENSFPKRKEGYLGPLEYSSPADYYKQRYGVGLIYKKOPIIRARG
SsDCL1b	888	DRGRGIFNDYQNGKGLFMADSCWDAKDLAGMVVTAHESCKRFYVDSICYNMNBENSFPKRKEGYLGPLEYSSPADYYKQRYGVGLIYKKOPIIRARG
SsDCL2	586	...SNMMDMBECPCKDGTGIVQTEDGFFCRCMIRNSIVRTFENGMEYAVSGFRDLNANSLMPGSS.....LSYKTHFEKEGDLDTCEPOPLLKACK
SsDCL3	830	NLMYENLCSKRSNKGEDEIHLANKSLHFSAVKDSVLSLSEFRIYSVLDLILDTTADDSFDEMYNG.KVSEFVSFVDYYHCKRYGVIOHFGCOPILLKKQ
		h v h g v a s k g qpl
SsDCL1a		VSYCRNLLSPRFEHSEAGNGEFSENLDKT..YYVYLPPELCLVH...FLPGSLVRGAQRLPSHMRVESMIL 910
SsDCL1b		VSYCRNLLSPRFEHSEAGNGEFSENLDKT..YYVYLPPELCLVH...FLPGSLVRGAQRLPSHMRVESMIL 1050
SsDCL2		LVEVRNLLDKCNYYKEKERSSSKNVVERSSKNVVELPPELCIVVMSEMSAVTLRSFMFIPSHMYRIQCML 746
SsDCL3		SHNARNLLFSKMKYQDGSFGNSLIVEKEQ..IHARVPELILHI..DVTDLKSEYILLPSVMHREVOSLM 995
		n l ppel ps m r l

PAZ

SsDCL1a	925	NKILEALTAAACCEBFCYERAEHLGDAYLKNMWSRFLFLKYPQKHEGGLTRMRQQMVSNMITYOYALNKNLOSYYIQADRBAFSRWAAAPGVLVVFDEETRD
SsDCL1b	1065	NKILEALTAAACCEBFCYERAEHLGDAYLKNMWSRFLFLKYPQKHEGGLTRMRQQMVSNMITYOYALNKNLOSYYIQADRBAFSRWAAAPGVLVVFDEETRD
SsDCL2	793	..ILEALTAKNCEBFSOESLEHLGDSFLKYVTTQHFEGKRHRQHEGGLIKMKKNLISNAAPQLACSNNEV....CLSLFHLHLLCLLLCFQGYIRSE
SsDCL3	1012	SIILEALTLLRCCEBFSIERLEHLGDSVLYKVIIGCDLEFLKEMKHEGGLSDMRSKAVCNATLHRHGIRRSIGYIRDTAGBPRRWVAPGQISLRFPFCNC
		ilea t c e f e e lgd lk v f y heg l m n l l f p
SsDCL1a		SEPSIFDEBFNPSSELQKNSYDDYEDSMQEDGEIEGDSSCYRVLSKTLADVVEALGVIYVAGGKIAANHLMKWIGINAEILDPO 1109
SsDCL1b		SEPSIFDEBFNPSSELQKNSYDDYEDSMQEDGEIEGDSSCYRVLSKTLADVVEALGVIYVAGGKIAANHLMKWIGINAEILDPO 1249
SsDCL2		A.....FNEPFT.....WIVPGVGYDICTSLRK.....LKSRIADSVREAL SAAGEQAAYIFLKSIGIDIEFHKM 950
SsDCL3		G....IKTAFVELS....GVYISDDPSFVVVGKPCDRG...HRWMCSTVSDVEAL VGGGIAAAIWVNRWFGIDVRCDM. 1184
		f p sk d veal g aa gi

RIBOc

SsDCL1a	1291	FVRELQERCCQQAEGLEYKASRTGNVATV..FVFVDGIQIGVAQNPOKKM..QKLAARNALVVLK 1352
SsDCL1b	1430	FVRELQERCCQQAEGLEYKASRTGNVATV..FVFVDGIQIGVAQNPOKKM..QKLAARNALVVLK 1491
SsDCL2	1169	FVRELQERCCDRKAYTMEYVTVIRENGVSSVAVVRTEGTTYKATRTGFSKLDKKLAASSVILDMK 1233
SsDCL3	1383	FVRELQERCSHLGCFIKSKCTSKGEELIITEMVQLRDELLIAQGHDRNRKSKAKAAARILVDL. 1446
		p e l e c v a aa l

DSRM

(c) RNA dependent RNA Polymerase (RDR)

SsRDR1a	379	VTPAKVYFYGPEINVSNRVVRHFYADIDNFLRISFVDECEKLRSDVLSRSTSG..N....DARRTALYNRVLSVLSNGINIGDKHFEFLAFSSSOLR
SsRDR1b	387	VTPAKVYFYGPEINVSNRVVRHFYADIDNFLRISFVDECEKLRSDVLSRSTSG..N....DARRTALYNRVLSVLSNGINIGDKHFEFLAFSSSOLR
SsRDR2a	380	ITPSKIYICLGPEEEVSNYVVKYHSQYASDFARVTFVDEDSKLSNPALSAEQGFFS....TPLKTGLYHRLISILKEGFCIGPKKYEFLAFSASOLR
SsRDR2b	379	ITPSKIYICLGPEEEVSNYVVKYHSQYASDFARVTFVDEDSKLSNPALSAEQ.....T.....ASOLR
SsRDR2c	379	ITPSKIYICLGPEEEVSNYVVKYHSQYASDFARVTFVDEDSKLSNPALSAEQGFFS....TPLKTGLYHRLISILKEGFCIGPKKYEFLAFSASOLR
SsRDR2d	324	ITPSKIYICLGPEEEVSNYVVKYHSQYASDFARVTFVDEDSKLSNPALSAEQGFFS....TPLKTGLYHRLISILKEGFCIGPKKYEFLAFSASOLR
SsRDR3	18	GKKKMVNRGISKALRLIFSLILSKTLT
SsRDR4	25	QGDIKSTSSVR...FSLILSKTLT
SsRDR5	6	KTEKLKEEKKEKNKKVSPVRCYFIRMESGWERDYPYKLAGYTVDEA.....RRFMFHINNAPTAVKLYSLRTILILSKTIK
SsRDR6a	494	ITPTRAYCMPQVRSNRVTRHYHQLSNRFLRVTFVDEGMPLPNSNALSLHVPLVKDSMSCPSQOMTTVYRRVQTILTEGFTMCGRKYSFLAFSTSQLK
SsRDR6b	501	ITPTRAYCMPQVRSNRVTRHYHQLSNRFLRVTFVDEGMPLPNSNALSLHVPLVKDSMSCPSQOMTTVYRRVQTILTEGFTMCGRKYSFLAFSTSQLK
SsRDR1a		DNSAWMFASRQGLTASDIRKMGDFRDIRNVAKYAAARLGQSFSSTETLKVPKYEV.EQIPDITN...GTEYIFSDGVGKISANAEKAVATCKLKL.RFA
SsRDR1b		DNSAWMFASRQGLTASDIRKMGDFRDIRNVAKYAAARLGQSFSSTETLKVPKYEV.EQIPDITN...GTEYIFSDGVGKISANAEKAVATCKLKL.RFA
SsRDR2a		GNSVWMFASNLSLTAESIRRMWGHFEDIRSVSKCAARMGQLFSSSRQTLFVTSYDV.EVIPDIEVTTDGTGYIFSDGIGKISTREARNVAKLIGLDDPAHL
SsRDR2b		GNSVWMFASNLSLTAESIRRMWGHFEDIRSVSKCAARMGQLFSSSRQTLFVTSYDV.EVIPDIEVTTDGTGYIFSDGIGKISTREARNVAKLIGLDDPAHL
SsRDR2c		GNSVWMFASNLSLTAESIRRMWGHFEDIRSVSKCAARMGQLFSSSRQTLFVTSYDV.EVIPDIEVTTDGTGYIFSDGIGKISTREARNVAKLIGLDDPAHL
SsRDR2d		GNSVWMFASNLSLTAESIRRMWGHFEDIRSVSKCAARMGQLFSSSRQTLFVTSYDV.EVIPDIEVTTDGTGYIFSDGIGKISTREARNVAKLIGLDDPAHL
SsRDR3		LVDVSLTIDVILIDDEPCRDDEHGKVV.MSDGKRLIHTDGTGFISENLAKKCPRIK....KKS KDHIHRGETTIVRLFYNGYAVKGTLLVDKRAADGA
SsRDR4		LVDVSLTIDVILIDDEPCRDDEHGKVV.MSDGKRLIHTDGTGFISENLAKKCPRIK....KKS KDHIHRGETTIVRLFYNGYAVKGTLLVDKRAADGA
SsRDR5		LVDVFSKVNVLIDDEPCREESCEIV.VQDGKPLIHTDGTGLISVDLARNCPFSVFKGNLKDADVSKGHQHLTQHFRMFHNGYAVKGTLLADKRLP..
SsRDR6a		QKSAWFFAEDGTTVASIKEMWQGF.P.IRNPAAKHAARMGLCTFSSYATVTMQPCEVNEYLEVDVIH...NGYNFSDGIGMITQDLAEVAERLPLTDNYA
SsRDR6b		QKSAWFFAEDGTTVASIKEMWQGF.P.IRNPAAKHAARMGLCTFSSYATVTMQPCEVNEYLEVDVIH...NGYNFSDGIGMITQDLAEVAERLPLTDNYA
SsRDR1a		PSVFOIRYGGYKGVAVDPSRN...CKLSLRKSMKFSQSENITLDVLAYSQYQPCFLNRQILITLLSTLGVGDSVFEELKQKEVVRQLNRMVTEPQAAREAV
SsRDR1b		PSVFOIRYGGYKGVAVDPSRN...CKLSLRKSMKFSQSENITLDVLAYSQYQPCFLNRQILITLLSTLGVGDSVFEELKQKEVVRQLNRMVTEPQAAREAV
SsRDR2a		PSAFOIRYGGYKGVAVDPTSF...FNLSLRPSMKKFESKSTMLNITNWSKSPQCYVNRETISLSTLGIKDEVEFSMQDDMHESDGLMTNKEAALSVL
SsRDR2b		PSAFOIRYGGYKGVAVDPTSF...FNLSLRPSMKKFESKSTMLNITNWSKSPQCYVNRETISLSTLGIKDEVEFSMQDDMHESDGLMTNKEAALSVL
SsRDR2c		PSAFOIRYGGYKGVAVDPTSF...FNLSLRPSMKKFESKSTMLNITNWSKSPQCYVNRETISLSTLGIKDEVEFSMQDDMHESDGLMTNKEAALSVL
SsRDR2d		PSAFOIRYGGYKGVAVDPTSF...FNLSLRPSMKKFESKSTMLNITNWSKSPQCYVNRETISLSTLGIKDEVEFSMQDDMHESDGLMTNKEAALSVL
SsRDR3		VVQSOLITNIRNNITVIRSMV...KVKADPKLCQMKSLSSLEIVSTSHQPNRTSTSRATIALHYGCVKEEYFMELLHNAIEGVENARYSPRHALKLA
SsRDR4		VVQSOLITNIRNNITVIRSMV...KVKADPKLCQMKSLSSLEIVSTSHQPNRTSTSRATIALHYGCVKEEYFMELLHNAIEGVENARYSPRHALKLA
SsRDR5		PDTIHRSMI...KIYCDENSGFQCSFNSLEVITTSYRPAKRAFTSRITSLHYGRVPAEYFLELLANAEVDVNAKHTLRNSLEVA
SsRDR6a		PSAYOIRYAGFKGVAVSPGQNDGIRMMSLRPSMRKFESHAHTVIEVSWTKFQPAFLNRQILITLLSTLGVDPDAVEWQMEAMLGNLKRILSNNSNVAIKVV
SsRDR6b		PSAYOIRYAGFKGVAVSPGQNDGIRMMSLRPSMRKFESHAHTVIEVSWTKFQPAFLNRQILITLLSTLGVDPDAVEWQMEAMLGNLKRILSNNSNVAIKVV
SsRDR1a		ELMPMGVEVTNVVKELLSCGYQDDEHPYLSMLQOTFRASKLLELKTKSRIFIPEGRAMMCCIDEETRTKYGOVFIQAS.....YCADDDR..K
SsRDR1b		ELMPMGVEVTNVVKELLSCGYQDDEHPYLSMLQOTFRASKLLELKTKSRIFIPEGRAMMCCIDEETRTKYGOVFIQAS.....YCADDDR..K
SsRDR2a		KGIG.GADTKTAAEMLLQGYEBSSEPYLLMIKKAHRANRLTDIRTRCKIHVQKGRVLICIDEETCKLDYGOVYIRITKNRKDQKSEQPFFCNDGDK..T
SsRDR2b		T KGIG.GADTKTAAEMLLQGYEBSSEPYLLMIKKAHRANRLTDIRTRCKIHVQKGRVLICIDEETCKLDYGOVYIRITKNRKDQKSEQPFFCNDGDK..T
SsRDR2c		T KGIG.GADTKTAAEMLLQGYEBSSEPYLLMIKKAHRANRLTDIRTRCKIHVQKGRVLICIDEETCKLDYGOVYIRITKNRKDQKSEQPFFCNDGDK..T
SsRDR2d		T KGIG.GADTKTAAEMLLQGYEBSSEPYLLMIKKAHRANRLTDIRTRCKIHVQKGRVLICIDEETCKLDYGOVYIRITKNRKDQKSEQPFFCNDGDK..T
SsRDR3		SGYA.NMEDSMLEMRMTHSGI..BLEBPYLLSRFSMAKQEMKGFRE.GKLPIDECYVLMCTTDPGTGTLKNEVCVILD.....S
SsRDR4		SGYA.NMEDSMLEMRMTHSGI..BLEBPYLLSRFSMAKQEMKGFRE.GKLPIDECYVLMCTTDPGTGTLKNEVCVILD.....S
SsRDR5		SFNHA.LDDDSMARMLILSGIOPEDBAYLQSQALMAKVERKGLKE.GRIPIDITYYLMCTTDPGTGTLKNDQVCILD.....N
SsRDR6a		TNSC.PEHGS TAGLMLSAGFAPATEPPLRAMLAICASQMKGLLDKTIWIFPKGRWLMGCTDELGLIEQGQCIFRVSTPSLNNRYVNRGSGISPSSEYKNNNA
SsRDR6b		TNSC.PEHGS TAGLMLSAGFAPATEPPLRAMLAICASQMKGLLDKTIWIFPKGRWLMGCTDELGLIEQGQCIFRVSTPSLNNRYVNRGSGISPSSEYKNNNA
SsRDR1a		FVVTCGRVVVAKNPOLHFGDGRVRLQAVDPALYH.....LPHPNCESGSDLDGDIYFVSWDSRLIP.....SRLV
SsRDR1b		FVVTCGRVVVAKNPOLHFGDGRVRLQAVDPALYH.....LPHPNCESGSDLDGDIYFVSWDSRLIP.....SRLV
SsRDR2a		AVIVCGVAITKNPOLHFGDGRVRLQAVDPALDARGLID...CVVFPQGERPHPNCESGGDLGDLFFITWDDKILP.....EKVD
SsRDR2b		AVIVCGVAITKNPOLHFGDGRVRLQAVDPALDARGLID...CVVFPQGERPHPNCESGGDLGDLFFITWDDKILP.....EKVD
SsRDR2c		AVIVCGVAITKNPOLHFGDGRVRLQAVDPALDARGLID...CVVFPQGERPHPNCESGGDLGDLFFITWDDKILP.....EKVD
SsRDR2d		AVIVCGVAITKNPOLHFGDGRVRLQAVDPALDARGLID...CVVFPQGERPHPNCESGGDLGDLFFITWDDKILP.....EKVD
SsRDR3		GQCSGDLVLFKHPCLHFGDIHILTAQINGLEKNFYGSKNAILFFTSGQSLADEMAGSDGDEGFVWSRNHMLKAFAEKQSEFPWVQLLKPKEQTQHKPR
SsRDR4		GQYSGDLVLFKHPCLHFGDIHILTAQINGLEKNFYGSKNAILFFTSGQSLADEMAGSDGDEGFVWSRNHMLKAFAEKQSEFPWVQLLKPKEQTQHKPR
SsRDR5		QQLSGDLVLFKHPCLHFGDIHILTAQINGLEKNFYGSKNAILFFTSGQSLADEMAGSDGDEGFVWSRNHMLKAFAEKQSEFPWVQLLKPKEQTQHKPR
SsRDR6a		EIIVCGVVMKNPOLHFGDGRVRLQAVDPALYH...LVD...CLVFEKKGERPHANBASGSDLDGDIYFVTWDENVLPFG.....RKSC
SsRDR6b		EIIVCGVVMKNPOLHFGDGRVRLQAVDPALYH...LVD...CLVFEKKGERPHANBASGSDLDGDIYFVTWDENVLPFG.....RKSC
SsRDR1a		DEMDDYTPAPAE TLHDHVTIEEIEE..YFTNYIVNESLGI TANAHVVFADKECLKAESPSCIQ LAKLFSIADVFPKGTGPALIPPELHVKEYPDFMEKLDK
SsRDR1b		DEMDDYTPAPAE TLHDHVTIEEIEE..YFTNYIVNESLGI TANAHVVFADKECLKAESPSCIQ LAKLFSIADVFPKGTGPALIPPELHVKEYPDFMEKLDK
SsRDR2a		APMDYTATPRIMDHAVTLEETIQK..HFVSYMINDALGVISTAHLIHADRNPLKARSPECQLQALALHSMADVFAKTGAPAEMLALRPREFPDFMERWER
SsRDR2b		APMDYTATPRIMDHAVTLEETIQK..HFVSYMINDALGVISTAHLIHADRNPLKARSPECQLQALALHSMADVFAKTGAPAEMLALRPREFPDFMERWER
SsRDR2c		APMDYTATPRIMDHAVTLEETIQK..HFVSYMINDALGVISTAHLIHADRNPLKARSPECQLQALALHSMADVFAKTGAPAEMLALRPREFPDFMERWER
SsRDR2d		APMDYTATPRIMDHAVTLEETIQK..HFVSYMINDALGVISTAHLIHADRNPLKARSPECQLQALALHSMADVFAKTGAPAEMLALRPREFPDFMERWER
SsRDR3		GPRDFNESTLERLLFNECITTTTFFIPNVLGLSSDCWLHYMDRFLTEEVDQDEKKSIAAKMIKLVLYYIALDG...HKVNVDRNLRVQAYPHFMEKEGF
SsRDR4		GPRDFNESTLERLLFNECITTTTFFIPNVLGLSSDCWLHYMDRFLTEEVDQDEKKSIAAKMIKLVLYYIALDG...HKVNVDRNLRVQAYPHFMEKEGF
SsRDR5		PKDYSGSELESLLFHEFLRARTFSPVYLGAASNCWLALMDRLTSGVPKSERQAIKEKMLLDVIYIYALDAPKEGNKITVPEELKVKKYPHFMERDEK
SsRDR6a		TPMDYFPAPKTRRLPRDVHQHDTID..FYLESVMVNDLGRICNAHVAHADRSDDGAMDPKCVLELAELAAITAVDSAKTIVRMPSPSPKYEYDFMEKEDA
SsRDR6b		TPMDYFPAPKTRRLPRDVHQHDTID..FYLESVMVNDLGRICNAHVAHADRSDDGAMDPKCVLELAELAAITAVDSAKTIVRMPSPSPKYEYDFMEKEDA
SsRDR1a		VTYVSEGVIGKLYREI 926
SsRDR1b		VTYVSEGVIGKLYREI 934
SsRDR2a		PMYVSNGLVGLKLYRAA 959
SsRDR2b		PMYVSNGLVGLKLYRAA 922
SsRDR2c		PMYVSNGLVGLKLYRAA 958
SsRDR2d		 811
SsRDR3		DFYHSTSLGRYDET 520
SsRDR4		DSYHSTSLGRYDET 518
SsRDR5		KSYTSLVGLKIYDEV 553
SsRDR6a		ISYKSKKILGRLYHQF 1079
SsRDR6b		 997

RdRP

Supplementary Figure S2. Pairwise alignment of amino acid sequence of conserved domains in DCL (a), AGO (b), and RDR (c) families of *S. spontaneum*.

Supplementary Table S1. Primers used for 18 candidate genes encoding DCL, AGO, and RDR families in *S. spontaneum*.

Gene name ^a	Gene ID	Forward primer (5'→3')	Reverse primer (5'→3')
SsDCL3	Sspon.01G0019860-4D	TAAAGGTGGACCTCGGACTG	CAGCAGGTATGTGCAAGGTG
SsDCL2	Sspon.01G0022370-1A	ATGAAACGGCTACTGGAACC	CCGAGTGGCCTTGTATGTAG
SsDCL1a	Sspon.01G0001230-2C	GTCAGTGAAGGAGAAGGCAAG	GCACCAGTTTCAAGGAAAGC
SsAGO5c	Sspon.01G0014460-1A	CCATCAGAGAGGGAGGGTAG	ATTGTCCGATGGAAGGGTTG
SsAGO5a	Sspon.01G0028360-2B	ACTCTGGCATTTCAGGGAACG	AGACTGATCGTGTGCATCGG
SsAGO10e	Sspon.08G0006580-3C	ACTCACCGGACATCAGGAAG	CGACAATCCCATCTGTGTCG
SsAGO18b	Sspon.01G0024300-1A	CACTGCACGTCCAGATCAAG	CAACACTGTGACACGAGACC
SsAGO18d	Sspon.02G0007830-2D	ACGTGACCATCTCACCAGAG	TCTGCCTGTTCTCGGATACC
SsAGO3b	Sspon.05G0030950-1C	TGATCCAGCAGCTTGGTGAG	GATCTTCTGCGGCTTCATGC
SsAGO2b	Sspon.05G0037340-1D	AGCAGTTTACGGCGTCTTTC	AAGACCCTGCAAGACCTCTC
SsAGO6b	Sspon.07G0020900-2D	CTGCTGGAGCTGTCCAAGTG	GGCTTTCCTTCACGACCGTT
SsAGO10c	Sspon.08G0006580-2B	CCACCTCAGGCAGTTCATCG	GCTTCCTGATGTCCGGTGAG
SsRDR1a	Sspon.04G0003980-2D	CCATCTGTGCTGTCAAGCTC	GGAATAATGTCCCGCTCTGC
SsRDR2b	Sspon.05G0009620-2B	AAGAAGCCGCTTTGTCTGTC	TCAGCCTATTAGCCCGATGG
SsRDR6b	Sspon.08G0019290-2D	GGAACAACAACCTGTCGCAAG	TCACTAGGCTGCATTGTCAC
SsRDR2d	Sspon.05G0009620-4D	GGAAAGGCCTCATCCAAATG	TCAACCTTCTCCGGAATCAG
SsRDR5	Sspon.03G0023160-2C	GGCGTATTTGCAATCTCAGC	GAAAGTTGGCCGTTGTCAAG
SsRDR3	Sspon.05G0014430-3C	ACAGGACATCAACGTCAAGG	CCAGATGCAAGCTTTAGTGC
GAPDH		CACGGCCACTGGAAGCA	TCCTCAGGGTTCCTGATGCC

^a GAPDH (reference gene), glyceraldehyde 3-phosphate dehydrogenase.

Supplementary Table S5. The number of genes encoding DCL, AGO, and RDR proteins identified in 21 plant species.

Plant species	DCL	AGO	RDR	Reference
<i>Arabidopsis</i>	4	10	6	Vazquez 2006
rice	8	19	5	Kapoor et al. 2008
maize	5	18	5	Qian et al. 2011
tomato	7	15	6	Bai et al. 2012
grapevine	4	13	5	Zhao et al. 2015
soybean	7	21	7	Liu et al. 2014
sorghum	5	14	7	Liu et al. 2014
foxtail millet	8	19	11	Yadav et al. 2015
poplar	5	15	9	Zhao et al. 2015
<i>B. napus</i>	8	27	16	Cao et al. 2016
<i>B. rapa</i>	4	13	6	Cao et al. 2016
<i>B. oleracea</i>	4	14	7	Cao et al. 2016
cucumber	5	7	8	Gan et al. 2016 & 2017
chickpea	4	13	5	Garg et al. 2017
pigeonpea	4	13	5	Garg et al. 2017
<i>A. duranensis</i>	3	11	5	Garg et al. 2017
<i>A. ipaensis</i>	4	11	5	Garg et al. 2017
coffee	9	11	8	Noronha et al. 2017
pepper	4	12	6	Qin et al. 2018
<i>Citrus sinensis</i>	5	13	7	Sabbione et al. 2019
<i>S. spontaneum</i>	4	21	11	In this study

References

1. Bai, M.; Yang, G.S.; Chen, W.T.; Mao, Z.C.; Kang, H.X.; Chen, G.H.; Yang, Y.H.; Xie, B.Y. Genome-wide identification of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families and their expression analyses in response to viral infection and abiotic stresses in *Solanum lycopersicum*. *Gene* **2012**, *501*, 52–62.
2. Cao, J.Y.; Xu, Y.P.; Li, W.; Li, S.S.; Rahman, H.; Cai, X.Z. Genome-wide identification of Dicer-Like, Argonaute, and RNA-Dependent RNA Polymerase gene families in *Brassica* species and functional analyses of their *Arabidopsis* homologs in resistance to *Sclerotinia sclerotiorum*. *Front. Plant Sci.* **2016**, *7*, 1614.
3. Gan D. F.; Liang D. D.; Wu J.; Zhan M. D.; Yang F.; Xu W. J. Genome-wide identification of the Dicer-Like, Argonaute, and RNA-dependent RNA polymerase gene families in cucumber (*Cucumis sativus* L.). *J. Plant Growth Regul.* **2016**, *35*, 135–150.
4. Gan, D.; Zhan, M.; Yang, F.; Zhang, Q.; Hu, K.; Xu, W.; Lu, Q.; Zhang, L.; Liang, D. Expression analysis of argonaute, Dicer-like, and RNA-dependent RNA polymerase genes in cucumber (*Cucumis sativus* L.) in response to abiotic stress. *J. Genet.* **2017**, *96*, 235–249.

5. Garg, V.; Agarwal, G.; Pazhamala, L.T.; Nayak, S.N.; Kudapa, H.; Khan, A.W.; Doddamani, D.; Sharma, M.; Kavi, K.P.; Varshney, R.K. Genome-wide identification, characterization, and expression analysis of small RNA biogenesis purveyors reveal their role in regulation of biotic stress responses in three legume crops. *Front. Plant Sci.* **2017**, *8*, 488.
6. Kapoor, M.; Arora, R.; Lama, T.; Nijhawan, A.; Khurana, J.P.; Tyagi, A.K.; Kapoor, S. Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA Polymerase gene families and their expression analysis during reproductive development and stress in rice. *BMC Genomics* **2008**, *9*, 451.
7. Liu, X.; Lu, T.; Dou, Y.; Yu, B.; Zhang, C. Identification of RNA silencing components in soybean and sorghum. *BMC Bioinformatics* **2014**, *15*, 4.
8. Noronha, F.C.; Marinho, R.P.; Cherubino, R.T.; Ricon, D.O.R.; Cunha, D.S.C.T.; Rodrigues, D.A.L.; de Souza, G.M.; Chalfun-Junior, A. A genome-wide analysis of the RNA-guided silencing pathway in coffee reveals insights into its regulatory mechanisms. *PLOS ONE* **2017**, *12*, e176333.
9. Qian, Y.; Cheng, Y.; Cheng, X.; Jiang, H.; Zhu, S.; Cheng, B. Identification and characterization of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in maize. *Plant Cell Rep.* **2011**, *30*, 1347–1363.
10. Qin, L.; Mo, N.; Muhammad, T.; Liang, Y. Genome-wide analysis of DCL, AGO, and RDR gene families in Pepper (*Capsicum Annuum L.*). *Int. J. Mol. Sci.* **2018**, *19*, 1038.
11. Sabbione, A.; Daurelio, L.; Vegetti, A.; Talon, M.; Tadeo, F.; Dotto, M. Genome-wide analysis of AGO, DCL and RDR gene families reveals RNA-directed DNA methylation is involved in fruit abscission in *Citrus sinensis*. *BMC Plant Biol.* **2019**, *19*, 401.
12. Vazquez, F. *Arabidopsis* endogenous small RNAs: highways and byways. *Trends Plant Sci.* **2006**, *11*, 460–468.
13. Yadav, C.B.; Muthamilarasan, M.; Pandey, G.; Prasad, M. Identification, characterization and expression profiling of Dicer-Like, Argonaute and RNA-Dependent RNA Polymerase gene families in Foxtail Millet. *Plant Mol. Biol. Report* **2015**, *33*, 43–55.
14. Zhao, H.; Zhao, K.; Wang, J.; Chen, X.; Chen, Z.; Cai, R.; Xiang, Y. Comprehensive analysis of Dicer-Like, Argonaute, and RNA-dependent RNA Polymerase gene families in Grapevine (*Vitis Vinifera*). *J. Plant Growth Regul.* **2014**, *34*, 108–21.
15. Zhao, K.; Zhao, H.; Chen, Z.; Feng, L.; Ren, J.; Cai, R.; Xiang, Y. The Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in *Populus trichocarpa*: gene structure, gene expression, phylogenetic analysis and evolution. *J. Genet.* **2015**, *94*, 317–321.