

Figure S1 **Peptide coverage maps of each Pol IV subunit and other co-purifying components.** Listed below are the gene names and AGI numbers for each protein listed in Table 1 followed by their full length protein sequence. The peptides recovered from the MS analysis are listed and mapped onto the protein sequence. Regions that are crossed out correspond to peptides that also maps to another gene while regions that are highlighted in yellow correspond to a unique peptide ( i.e. a peptide that maps only to the listed protein sequence) and regions in green correspond to two unique peptides.

**NRPD1 (At1g63020)**

MEDDCEELQV **PVGTLSIGFSISNNNDRDK** **MSVLEVEAPNQVTDSRLGLPNPDSVCRTCGSKD**  
 R **KVCEGHFGVINFAYSIIINPYFLK** **EVAALLNKICPGCKYIRKK** **QFQITEDQPERCR** **YCTLNTGYPL**  
**MK**FRVTTKEVFRR **SGIVVEVNEESLMK**LKKR **GVLTLP**PPDYWSFLPQDSNIDESCLKPTRRIITHA  
 QVYALLLGIDQRLIKKDIPMFNSLGLTSFPVTPNGYR **VTEIVHQFN**GARLIFDERTRIIYK **KL**VGFE  
**GNTLELSSR** **VMECMQYSR**LFSETVSSSKDSANPYQKKS **DTPKLCGLRFMKDVLLGKRSDHTF**  
 R **TVVVGDP**SLKLNEIGIPESIAK **R** **LQVSEHLNQC**NK **ERL**VTSFVPTLLDNK **EMH**VRRGDRL **LVAIQ**  
**VNDLQ**TGDKIFRSLMDGDTVLMNRP **PSIHQHS**LIAMTVRILPTTSV **VSLNPICCLPFR** **GDFD**GDC  
**LHG**YVPQSIQAK **VELDEL**VALDKQLINRQNGR **NLLSLGQDSLTAAYLVNVEK** **NCYL**NRAQM **QQ**  
 LQMYCPFQLPPP **AIKASPS**STEPQWTGMQLFGMLFPPGFDYTYPL **NNVVV**SNGELLSFSEGS  
 AWLR **DGEG**NFIERLLKHDKGKVL **DIYSAQ**EMLSQWLLMRGLSVSLADLYLSSDLQSR **KNL**TEE  
**ISYGLR**EAEQVCNK **QQLM**VESWRDFLAVNGEDK **EEDSV**SDLARFCYERQK **SATL**SELAVSAFK  
**DAYR**DVQALAYRYGDQSN **SFLIM**SKAGSKGNIGKLVQHSMCIGLQNSAVSL **SFGFP**RELTCAA  
**WNP**NSPLR **GAKG**KDSTTTESYVPYGV **IENSFLTGLN**PLESFVHSVTSR **DSS**FSGNADLP **GTLSR**  
 R **LMFF**MR **DIYA**AYDGTVRNSFGNQLVQFTYETDGPVEDITGEALGSL **SACAL**SEAAYSALD  
 QPISLLETSP **LLNLK** **NVLE**CGSKKGQREQTMSLYLSEYLSKKKHGFEYGSLEIK **NHLE**KLSFSEI  
 VSTSMIIFSPSSNTKVPLSPWVCHFHISEKVLKRK **QLSAES**VVSSLNEQYKSRNRELKLDIVDLDI  
 QNTNH **CSSDDQAMKDD**NVCITVTVVEASK **HSVLE**DAIRLVLIPFLLDSPVKGDQGIKKVNILWT  
 DRPKAPKR **NGNH**LAGELYLKVTMYGDRGKRNCWTALLETCLPIMDMIDWGRSHPDNIR **QCCS**  
**VY**GIDAGRSIFVANLES **AVSD**TGKEILREHLLL **VADSL**SVTGEFVALNAKGWSKQR **QV**ESTPAP  
**FTQAC**FSSPSQCFLKAAKEGVR **DDLQGS**IDALAWGKVPGFGTGDQFEIIISP **KVHG**FTTPVDVY  
 DLLSSTKTMRR **TNSAPK**SDKATVQPFGLLHSAFLDKIKVLDGKGIPMSLLRTIFTWK **NI**ELLSQSL  
**KR**ILHSYEINELLNERDEGLVK **MVLQ**LHPNSVEKIGPGVKGIRVAKSK **HGD**SCCFEVVRIDGTFE  
**DFS**YHKCVLGATKIIAPK **KMNFYK**SKYLKNGTLESGGFSEN

Length 1453aa  
 Unique 552/1453=38%  
 Total=38%

- |                                |                                  |                                  |  |
|--------------------------------|----------------------------------|----------------------------------|--|
| V. <b>PVGTLSIGFSISNNNDR</b> .D | R. <b>VMECMQYSR</b> .L           | R. <b>GDFDGDCLHG</b> YVPQSIQAK.V | R. <b>DFLAVNGEDK</b> .E                |
| K. <b>MSVLEVEAPNQVTDSR</b> .L  | R. <b>LFSETVSSSK</b> .D          | K. <b>VELDEL</b> VALDK.Q         | R. <b>DFLAVNGEDK</b> EEDSVSDLAR.F      |
| R. <b>LGLPNPDSVCR</b> .T       | R. <b>LFSETVSSSK</b> .D          | K. <b>VELDEL</b> VALDK.Q         | R. <b>DFLAVNGEDK</b> EEDSVSDLAR.F      |
| K. <b>EVAALLNK</b> .I          | R. <b>TVVVGDP</b> SLK.L          | K. <b>NCYL</b> NR.A              | K. <b>SATL</b> SELAVSAFK.D             |
| K. <b>EVAALLNK</b> .I          | R. <b>TVVVGDP</b> SLK.L          | K. <b>NCYL</b> NR.A              | K. <b>SATL</b> SELAVSAFK.D             |
| K. <b>QFQITEDQPER</b> .C       | K. <b>LNEIGIP</b> ESIAK.R        | R. <b>DGEG</b> NFIER.L           | K. <b>DAYR</b> DVQALAYR.Y              |
| K. <b>QFQITEDQPER</b> .C       | K. <b>LNEIGIP</b> ESIAK.R        | R. <b>DGEG</b> NFIER.L           | R. <b>DVQALAYR</b> .Y                  |
| R. <b>YCTLNTGYPLMK</b> .F      | R. <b>LQVSEHLNQC</b> NK.E        | R. <b>KNL</b> TEEISYGLR.E        | R. <b>DVQALAYR</b> .Y                  |
| R. <b>SGIVVEVNEESLMK</b> .L    | R. <b>LVT</b> SFVPTLLDNK.E       | K. <b>NL</b> TEEISYGLR.E         | R. <b>YGDQSN</b> SFLIMSK.A             |
| R. <b>VTEIVHQFN</b> GAR.L      | R. <b>LVAIQVNDLQ</b> TGDK.I      | K. <b>NL</b> TEEISYGLR.E         | R. <b>YGDQSN</b> SFLIMSK.A             |
| K. <b>LVGFEGNTLELSSR</b> .V    | R. <b>LVAIQVNDLQ</b> TGDK.I      | K. <b>QQLM</b> VESWR.D           | R. <b>ELTCAAWNP</b> NSPLR.G            |
| R. <b>VMECMQYSR</b> .L         | R. <b>GDFDGDCLHG</b> YVPQSIQAK.V | K. <b>QQLM</b> VESWR.D           | R. <b>DSS</b> FSGNADLP <b>GTLSR</b> .R |

R.DIYAAYDGTVR.N  
R.DIYAAYDGTVR.N  
K.NVLECGSK.K  
K.NVLECGSK.G  
R.EQTMSLYLSEYLSK.K  
K.HGFEYGSLEIK.N  
K.HGFEYGSLEIK.N

K.QLSAESVSSSLNEQYK.S  
K.QLSAESVSSSLNEQYK.S  
K.HSVLELDAIR.L  
K.HSVLELDAIR.L  
R.NGNHLAGELYLK.V  
K.VTMYGDR.G  
K.VTMYGDR.G

R.QCCSVYGDAGR.S  
R.SIFVANLESASVSDTGK.E  
R.QVESTPAPFTQACFSSPSQCFLK.A  
R.QVESTPAPFTQACFSSPSQCFLK.A  
R.DDLQGSIDALAWGK.V  
K.NIELLSQSLK.R  
K.NIELLSQSLK.R

K.MVLQLHPNSVEK.I  
K.HGDSCCFEVVR.I  
R.IDGTFEDFSYHK.C

### NRPD2/E2 (AT3G23780)

MPDMDIDVKDLEEF EATTGEINLSELGEGFLQSFCKKKAATSFFDKYGLISHQLNSYNYFIEHGLQ  
NVFQSFGEMLVEPSFDVVKKKDNDWRYATVKFGEVTVEKPTFFSDDKELEFLPWHARLQNMT  
YSARIKVNVQVEVFKNTVVKSDKFKTGQDNYVEKKILDVKKQDILIGSIPVMVKSILCKTSEKGE  
NCKKGDCAFDQGGYFVIKGAEKVFIAQEQMCTKRLWISNSPWTV/SFRSENKRNR/FIVRLSENE  
KAEDYKRREKVLTVYFLSTEIPVWLLFFALGVSSDKEAMDIAFDGDDASITNSLIASIHVADAVC  
EAFRCGNNALTYVEQQIKSTKFPPAESVDECLHLYLFPGLQSLKKKARFLGYMVKCLLNSYAGK  
RKCENRDSFRNKRIELAGELLEREIRVHLAHARRKMTRAMQKHLSGDGLKPIEHYLDASVITN  
GLSRAFSTGAWSHPFKMERVSGVVANLGRANPLQTLIDLRRTRQQVLYTGKVGDARYPHPS  
HWGRVCFLSTPDGENCGLVKNMSLLGLVSTQSLESVVEKLFACGMEELMDDTCTPLFGKHKV  
LLNGDWVGLCADSESFVAELKSRRRQSELPREMEIKRDKDDNEVRIFTDAGRLLRPLLVENL  
QKLKQEKPSQYFPDHLLDHGILELIGIEEEEDCNTAWGIKQLLKEPKIYTHCELDLSFLLGVSCAV  
VPFANHHDHGRRVLYQSQKHCQQAIGFSSTNPNI RCDTLSQQLFYQKPLFKTLASECLKKEVLF  
NGQNAIVAVNVHLGYNQEDSIVMNKASLERGMFRSEQIRSYKAEVDAKDSEKRKKMDELVQF  
GKTHSKIGKVDSEDDGFPFIGANMSTGDIVIGRCTESGADHSIKLKHTEGIVQKVVLSSNDEG  
KNFAAVSLRQVRSPCLGDKFSSMHGQKGVLYGYLEEQNFPTIQGIVPDIVINPHAFPSRQTPG  
QLLEAALSKGIACPIQKEGSSAAYTKLTRHATPFSTPGVTEITEQLHRAGFSRWGNERNVYNGRS  
GEMMRSMIFMGPTFYQRLVHMSEDKVKFRNTGPVHPLTRQPVADRKRFGGKFGEMERDCLI  
AHGASANLHERLF<sup>T</sup>LS<sup>D</sup>SSQM<sup>H</sup>ICR<sup>K</sup>CKTYANVIERTPSSGRKIRGPYCRVCVSSDHVVRVYV  
PYGAKLLCQELFSMGITLNFDTKLC

Length 1172aa

Unique 287/1172=24.5%

Total=24.5%

M.PDMDIDVK.D

K.VFIAQEQMCTK.R

R.PLLVENLQK.L

K.NFAAVSLR.Q

M.PDMDIDVK.D

R.CGNNALTYVEQQIK.S

R.VLYQSQK.H

R.QTPGQLLEAALS.G

K.KAATSFFDK.Y

R.FLGYMVK.C

R.VLYQSQK.H

K.GIACPIQK.E

K.KAATSFFDK.Y

K.CLLNSYAGK.R

K.TLASECLK.K

K.GIACPIQK.E

K.KAATSFFDK.Y

K.CLLNSYAGK.R

K.TLASECLK.K

R.LVHMSEDK.V

K.FGEVTVEKPTFFSDDK.E

K.RIELAGELLER.E

K.TLASECLKK.E

R.LVHMSEDK.V

R.LQNMTYSAR.I

R.IELAGELLER.E

K.KMDELVQFGK.T

R.DCLIAHGASANLHER.L

R.LQNMTYSAR.I

R.IELAGELLER.E

K.MDELVQFGK.T

R.DCLIAHGASANLHER.L

K.VNVQVEVFK.N

R.VSGVVANLGR.A

K.MDELVQFGK.T

K.TYANVIER.T

K.VNVQVEVFK.N

R.VSGVVANLGR.A

K.VVLSSNDEGK.N

K.TYANVIER.T

K.GDCAFDQGGYFVIK.G

R.ANPLQTLIDLR.R

K.VVLSSNDEGK.N

R.VCVSSDHVVR.V

K.GDCAFDQGGYFVIK.G

R.QQVLYTGK.V

K.VVLSSNDEGKNFAAVSLR.Q

R.VYVPYGAK.L

K.VFIAQEQMCTK.R

R.VCFLSTPDGENCGLVK.N

K.NFAAVSLR.Q

Underlined peptides also map to the predicted protein product of the NRPD2b pseudogene (At3g18090), but to no other proteins and are thus included in the unique % coverage calculation as was done in [1].

### NRPB3/D3/E3A (At2g15430)

MDGATYQRFPKIKIRELKDDYAKFELRETDVSMANALRRVMISEVPTVAIDLVEIEVNSSVLNDE  
FIAHRLGLIPLTSERAMSMRFSRDCDACDGDGQCEFCSEVFR LSSKCVTDQTLDVTSRDLYS  
DPTVTPVDFTIDSSVSDSSEHKGIIIVKLRRGQELKLRAIARKGIGKDHAKWSPAATVTFMYEPDII  
INEDMMDTLDSEEKIDLISSPTKVFGMDPVTRQVVVVDPEAYTYDEEVIKKAEMGKPGLEIS  
PKDDSFITVESTGAVKASQLVLNAIDLLKQKLDVRLSDDTVEADDQFGELGAHMRGG

Length 319aa

Unique 141/319= 44.2% ; The remaining 13.5% also map to NRPD3B/E3B (AT2G15400).

Total 141+43=184/319=57.7%

K.IRELKDDYAK.F	R.DCDACDGDGQCEFCSEVFR.L	K.VFGMDPVTR.Q	K.AEAMGKPGLEISPK.D
R.ELKDDYAK.F	R.DCDACDGDGQCEFCSEVFR.L	K.VFGMDPVTR.Q	K.PGLEISPK.D
R.ETDVSMANALR.R	K.CVTDQTLDVTSR.D	R.QVVVVDPEAYTYDEEVIK.K	K.DDSFITVESTGAVK.A
R.ETDVSMANALR.R	R.DLYSADPTVTPVDFTIDSSVSDSSEHK.G	R.QVVVVDPEAYTYDEEVIK.K	K.QKLDVRL.S
R.LGLIPLTSER.A	K.IDLISSPTK.V	R.QVVVVDPEAYTYDEEVIK.A	R.LSDDTVEADDQFGELGAHMR.G
R.LGLIPLTSER.A	K.IDLISSPTK.V	R.QVVVVDPEAYTYDEEVIK.A	

### NRPD3B/E3B (At2g15400)

MDGVTYQRFPTVKIRELKDDYAKFELRETDVSMANALRRVMISEVPTMAIHLVKIEVNSSVLNDE  
EFIAQRLSLIPLTSERAMSMRFCQDCEDCNGDEHCEFCSEVFP LSAKCVTDQTLDVTSRDLYS  
ADPTVTPVDFTSNSSTSDSSEHKGIIIAKLRRGQELKLKALARKGIGKDHAKWSPAATVTYMYEP  
DIIINEEMMNTLTDEEKIDLISSPTKVFGIDPVTGQVVVVDPEAYTYDEEVIKKAEMGKPGLEI  
HPKHDSFVFTVESTGALKASQLVLNAIDILKQKLDAIRLSNTVEADDQFGELGAHMREG

Length 319aa

Unique 79/319=24.76% ; The remaining 13.5% also map to NRPB3/D3/E3A (At2g15430).

Total 79+43=122/319=38.24%

K.IRELKDDYAK.F	K.IEVNSSLNDEFIAQR.L	K.CVTDQTLDVTSR.D	K.IDLISSPTK.V
R.ELKDDYAK.F	K.IEVNSSLNDEFIAQR.L	* R.DLYSADPTVTPVDFTSNSSTSDSSEHK.G	
R.ETDVSMANALR.R	R.LSLIPLTSER.A	* R.DLYSADPTVTPVDFTSNSSTSDSSEHK.G	
R.ETDVSMANALR.R	R.FCQDCEDCNGDEHCEFCSEVFP LSAK.C	K.IDLISSPTK.V	

### NRPD4/E4 (At4g15950)

MSEKGGKGLKSSLKSKDGGKDGSSSTKLKGRKIHFDQRTPPANYKILNVSSDQQPFQSSAAK  
GKSDKPTKSSKNSLHSFELKDLPENAECEMMDCEAFQILDGIGKQLVGLSEDPSIKIPVSYDRALA  
YVESCVMHTNPQSVRKVLEPLKTYGISDGEMCVIANASSESVDEVLAFLIPSLKTKKEVINQPLQD  
ALEELSKLKKSE

Length 205 aa

Unique 26/205=12.7%

Total=12.7%

K.ILVSSDQQPFQSSAAK.C	K.NSLHSFELK.D	K.NSLHSFELK.D
----------------------	---------------	---------------

### NRPB5/D5 (At3g22320)

MLTEEEKRLYRIQKTLMQMLRDRGYFIADSELTMTKQQFIRKHGDNMKREDLVTLKAKRNDN  
SDQLYIFFPDEAKVGVKTMKMYTNRMKSENVFRAILVVQQNLTPFARTCISEISSKFHLEVFQEA  
EMLVNIKEHVLVPEHQVLTTEEKKTLLERYTVKETQLPRIQVTDPIARYFGLKRGQVVKIIRPSET  
AGRYVTYRYVV

Length 205aa

Unique 92/205=44.9%

Total=44.9%

-MLTEEEK.R	R.GYFIADSELTMTK.Q	R.TCISEISSK.F	R.IQVTDPIAR.Y
-MLTEEEK.R	K.EDLVTLK.A	R.TCISEISSK.F	K.IIRPSETAGR.Y
-MLTEEEKR.L	R.EDLVTLK.A	K.EHVLVPEHQVLTTEEK.K	
R.DRGYFIADSELTMTK.Q	R.NDNSDQLYIFFPDEAK.V	R.IQVTDPIAR.Y	

### NRPD5B/E5B (At2g41340)

MEGKGKEIVVGHHSISKSSVECHKYYLARRTTMEMLRDRGYDVSDEDINLSLQQFRALYGEHPD  
VDLLRISAKHRFDSSKKISVFCGTGIVKVNAMRVIAADVLSRENITGLILVLQSHITNQALKAVEL  
FSFKVELFEITDLLVNVSKHVLRPKHQVLNDKEKESLLKKSIEEKQLPRLSSKDPIVRYYGLETG  
QVMKVITYKDELSESHVTYRCVS

Length 218aa

Unique 78/218=35.8%

Total=35.8%

R.DRGYDVSDEDINLSLQQFR.A	R.ALYGEHPDVDLLR.I	K.AVELFSFK.V	R.LSSKDPIVR.Y
R.GYDVSDEDINLSLQQFR.A	R.VIAADVLSR.E	K.HQVLNDKEK.E	R.YYGLETGQVMK.V

### NRPB6A/D6A/E6A (At5g51940)

MADEDYNDVDDLGYEDEPAEPEIEEGVEEDVEMKENDDVNGEPIEAEDKVETEPVQRPRKTS  
KFMTKYERARILGTRALQISMNAPVMVELEGETDPLEIAMKELRQRKIPFTIRRYLPDGSFEEWG  
VDELIVEDSWKRQVGGD

Length 144aa

Unique 25/144=17.4%

Total=17.4%

K.ENDDVNGEPIEAEDKVETEPVQRPR.K	K.VETEPVQRPR.K
-------------------------------	----------------

The VETEPVQRPR peptide also maps to NRPB6B/E6B (AT2G04630), but no peptides unique to this protein were identified.

### **NRPD7A (At3g22900)**

MFIKVKLPWDVTIPAEDMDTGLMLQRAIVIR **LLEAFSK**EKATK**DLGYLITPTILENIGEGK**IKEQTG  
EIQFPVVFNGICFKMFKGEIVHGVVHKVHKTGVFLKSGPYEIIYLSHMKMPGYEFIPGENPFFMN  
QYMSRIQIGARVR **FVVLDTIEW**REA EKDFMALASIDGDNLGP

Length 174aa  
Unique 34/174=19.5%  
Total=19.5%

R. **LLEAFSK**.E

K. **DLGYLITPTILENIGEGK**.I

R. **FVVLDTIEW**.E

### **NRPD7B/E7B (At4g14660)**

MFLKVQLPWNVMIPAENMDAKGLMLKRAILVELLEAFASKKATKELGYVAVTTLDKIGEGKIRE  
HTGEVLFPMFSGMTFKIFKGEIHHGVVHKVHLKHGVFMR **CGPIENVYLSYTK**MPDYK **YIPGENPI**  
**FMNEK**TSRIQVETTVRVVIGIKWMEVEREFQALASLEGDYLGPLSEE

Length 178aa  
Unique 26/178=14.6%  
Total=14.6%

R. **CGPIENVYLSYTK**.M

K. **YIPGENPIFMNEK**.T

### **NRPB9A/D9A/E9A (At3g16980)**

MSTMKFCRE **ECNNILYPKEDK**EQKILLYACR **NCDHQEVADNSCVYRNEVHHSVSERTQILTDVA**  
**SDPTLPR**TKAVRCSKCQHR **EAVFFQATARGE**EGMTLFFVCCNPNCGHRWRE

Length 114aa  
Unique 50/114=43.9% ; The remaining 16.7% also map to NRPB9B/D9B/E9B (AT4G16265)  
Total 50+19=69/114=60.5%

R. **ECNNILYPK**.E

R. **ECNNILYPKEDK**.E

R. **NCDHQEVADNSCVYR**.N

R. **TQILTDVASDPTLPR**.T

R. **ECNNILYPK**.E

K. **ILLYACR**.N

R. **NEVHHSVSER**.T

R. **EAVFFQATAR**.G

### **NRPB9B/D9B/E9B (At4g16265)**

MSTMKFCRE **ECNNILYPKEDK**EQSILLYACR **NCDHQEAADNNCVYRNEVHHSVSEQTQILSDVA**  
**SDPTLPR**TKAVRCAKQCQHG **EAVFFQATARGE**EGMTLFFVCCNPNC **SHRWRE**

Length 114aa  
Unique 53/114=46.5%; The remaining 7.9% also map to NRPB9A/D9A/E9A (At3g16980).  
Total 53+9=62/114=54.4%

R. **ECNNILYPK**.E

R. **ECNNILYPKEDK**.E

S. **ILLYACR**.N

R. **NEVHHSVSEQTQILSDVASDPTLPR**.T

R. **ECNNILYPK**.E

K. **EDKEQSILLYACR**.N

R. **NCDHQEAADNNCVYR**.N

### **NRPB10/D10/E10 (At1g11475)**

MIIPVRCFTCGKVIGNKWDQYLDLLQLDYTEGDALDALQLVRYCCRRMLMTHVDLIEK **LLNYNT**  
**LEK**SDNS

Length 71aa  
Unique 9/71=12.7%  
Total=12.7%

K. **LLNYNTLEK**.S

K. **LLNYNTLEK**.S

### NRPB11/D11/E11 (At3g52090)

MNAPERYERFVVPEGTKKVSYDRDTKIINAASFTVEREDHTIGNIVRMLQHRDENVLFAGYQLP  
HPLKYKIIVRDPKDWFRNLKVVYGCLETVNISWFKLLVLTHTTSQSSPMQAYNQAINDLDKELD  
YLNQFEAEVAKFSNQF

Length 146aa  
Unique 38/146=26%  
Total=26%

R.FVVPEGTK.K	K.IINAASFTVER.E	K.NQFEAEVAK.F
K.IINAASFTVER.E	R.EDHTIGNIVR.M	K.NQFEAEVAK.F

### NRPB12/D12/E12 (At5g41010)

MDPAPEPVTYVCGDCGQENTLKSGDVIQCRECGYRILYKKRTRRVVQYEAR

Length 51aa  
Unique 38/51=74.5%  
Total=74.5%

-MDPAPEPVTYVCGDCGQENTLK.S	K.SGDVIQCR.E	R.RVVQYEAR.-	R.VVQYEAR.-
D.PAPEPVTYVCGDCGQENTLK.S	K.SGDVIQCR.E	R.VVQYEAR.-	

### RDR2 (At4g11130)

MVSETTTNRSTVKISNVPQTIVADELLRFLELHLGEDTVFALEIPTTRDNWKPRDFARVQFTTLE  
VKSRACLSSQSKLLFKTHNLRLEAYDDIIPRPVDPRKRLDDIVLTVGFPESDEKRFCALEKWD  
GVRWCWILTEKRRVEFWVWESGDCYKIEVRFEDIETLSCCVNGDASEIDAFLLKLYGPKVFKR  
VTVHIATKFKSDRYRFCKEDFDFMWIRTTDFSGSKSIGTSTCFCLEVHNGSTMLDIFSGLPYYR  
EDTSLSLTYVDGKTFASAAQIVPLLNAAILGLEFPYEILFQLNALVHAQKISLFAASDMELIKILRGM  
SLETALVILKHLHQSSICYDPVFFVKTQMQSVVKKMKHSPASAYKRLTEQNIMSCQRAYVTPS  
KIYLLGPELETANYVVKNF AEHVSDFMRVTFVEEDWSKLPANALSVNSKEGYFVKPSRTNIYNR  
VLSILGEGITVGPKRFEFLAFSASQLRGNSVWMFASNEKVKAEDIREWMGCFRKIRSISKCAAR  
MGQLFSASRQTLIVRAQDVEQIPDIEVTTDGDADYCFSDGIGKISLAFKQVAQKCGLSHVPSAF  
QIRYGGYKGVIAVDRSSFRKLSLRDSMLKFDSNNRMLNVTRWTE SMP CFLNREIICLLSTLGIED  
AMFEAMQAVHLSMLGNMLEDRDAALNVLQKLSGENSKNLLVKMLLQGYAPSSPEPYLSMMLRV  
HHESQLSELKSR CRILVPKGRILIGCMDGILEYGVVYRVTLTKAELKSRDQSYFRKIDEETS  
VVIGKVVVTKNPCLHPGDIRVLDIAYEVHFEKGYLDCCIIFPQKGERPHPNECSGGDLGDQFF  
VSWDEKIIPSEMDPPMDYAGSRPRLMDHDVTL EEIHKFFVDY MISDTLGVISTAHLVHADRDPE  
KARSQKCLELANLHSRAVDFAKTGAPAEMPYALKPREFPDFLERFEKPTYISESVFGKLYRAVK  
SSLAQRKPEAESEDTVAYDVTLEEAGFESFIETAKAHRDMYGEKLTSLMIYYGAANEEIILT GIL  
KTKEMY LARDNRRY GDMKDRITLSVKDLHK EAMGWFEK SCEDEQQKKKLASAWYYVTYNPN  
HRDEKLTFLSFPWIVGDVLLDIKAENAQRQSV E EKT SGLVSI

Length 1133aa  
Unique 337/1133=29.7%  
Total=29.7%

K.ISNVPQTIVADELLR.F	R.LEAYDDIIPR.P	R.EDTSLTYVDGK.T	K.LPANALSVNSK.E
R.VQFTTLEVK.S	R.LEAYDDIIPRPVDPR.K	R.LTEQNIMSCQR.A	K.LPANALSVNSK.E
R.VQFTTLEVK.S	R.LDDIVLTVGFPESDEK.R	K.IYLLGPELETANYVVK.N	K.EGYFVKPSR.T
R.AQLSSQSK.L	R.CWILTEK.R	R.VTFVEEDWSK.L	R.VLSILGEGITVGPK.R
R.AQLSSQSK.L	K.EDFDFMWIR.T	R.VTFVEEDWSK.L	R.GNSVWMFASNEK.V

R.EWMGCFR.K	R.DAALNVLQK.L	K.GERPHPNCESSGGDLGDQFFVSWDEK.I	R.FEKPTYISESVFGK.L
R.MGQLFSASR.Q	R.DAALNVLQK.L	K.IIPSEMDPPMDYAGSRPR.L	K.EAMGWFEK.S
R.MGQLFSASR.Q	R.KDEETSVVIGK.V	R.LMDHDVTLLEEIHK.F	K.EAMGWFEK.S
R.AQDVVEQIPDIEVTTDGADYCFSDGIGK.I	K.DEETSVVIGK.V	R.EFPDFLER.F	
R.AQDVVEQIPDIEVTTDGADYCFSDGIGK.I	K.NPCLHPGDIR.V	R.EFPDFLER.F	

### RDM4 (At2g30280)

MDGVGESSTQNEVEEKPVIVRVKRKVGQSLLDFAFWLEINERPLKRPTLDFSKLSISDSGERGPS  
 VAEDVKPKKVLVRHLETVTDSETTADIIHSFFESDHNEKSCSKGKFEERKIAFKKDNRKEQRLTK  
 SVQKQQIASENARFEQIWRSRKGNKEGIHEKCHFFDVIRVDTEERRNAQEFTSLEDQKMLAS  
 FLPLLRECIPTAAEEIEADIQSSHTEEYVYDYAVNEEMDISEDSSKNQFPLVIVEDEEEFCGDS  
 DESDYDSEDSNAEDHPKTDYPEEEEEEEEEEDDDDDDDDESEEEKSEASDESDEETSKRHVR  
 SVLGDDEFDDYAEDVYGYSESDEEFES

Length 346 aa  
 Unique 40/346=11.6%  
 Total=11.6%

K.LSISDSGER.G	R.GPSVAEDVKPK.K	R.RNAQEFTSLEDQK.M	R.DNAQEFTSLEDQK.M
K.LSISDSGER.G	R.FEQIWR.S	R.DNAQEFTSLEDQK.M	

### CLASSY1 (At3g42670)

MKRKHFEFNFHPNCPFEVFCWGTWKAVEYLRIENGTMTMRLLENGQVLDDIKPFQRLRIRS  
 RKATLIDCTSFLRPGIDVCLYQRDEETPEPVVVDARVLSIERKPHSECLCTFHVSVIDQGC  
 GLEKHRMNKVPVLVGLNEIAILQKFCKEQSLDRYRWRVYSEDCSSLVKTRLNLGKFLPDLTWLL  
 VTSVLKNIVFQIRTVHEKMOVYQIVTDEDCEGSSSSLSAMNITVEDGVVMSKVLFNPAEDTCQD  
 SDVKEEIEEEMELRRSKRRSRPRERYGDSEIQPDSKDGWVRMMPYRYNIWNVSSDDDDDEE  
 EDCEDDKDTHLPLSHLLRKKGSKKGFSSKDKQR EIVLVDKTERKKRKKTEGFSRCELSVI  
 PFTPVFEPIPLEQFGLNANSLCGVSGNLMDEIDKYRSKAAKYGKKKKKKIEMEEMESDLGWN  
 GPIGNVVHKRNGPHSRIRSVSR ETGVSEEPQIYK KRTLSAGAYNK LIDSYMSRIDSTIAAKDKAT  
 NVVEQWQGLK NPASFSIEAEERLSEEEEDDGETSENEILWREMECLASSYILDDHEVRVDNE  
 AFHKATCDCEHDYELNEEIGMCCRLCGHVGT EIKHVSAPFARHKKWTTETKQINEDDINTTIVN  
 QDGVESHFTIPVASSDMPSAEE SDNVWSLIPQLKRKLHLHQKKAFFLWKNLAGSVVPAMMD  
 PSSDKIGGCVVSHTPGAGKTFLIIAFLASYLKIFPGKRPLVLAPKTTLYTWYKEFIKWEIPVPVHLL  
 HGRRTYCMSKEKTIQFEGIPKPSQDVMHVLDCLDKIQKWAQPSVLMGYTSFLTLMREDSKF  
 AHRKYMAKVLRES PGLLVLDEGHNPRSTKSRLRKALMKVDTDLRILLSGTLFQNNFCEYFNTLC  
 LARPKFVHEVLVELDKKFQTNQAEQKAPHLENRARKFFLDIIAKKIDTKVGDERLQGLNMLRN  
 MTSGFIDNYEGSGSGSGDVL PGLQIYTLLMNSTDVQHKSLTKLQNMSTYHGYPLELELLITLAAI  
 HPWLVKTTTCCAFFNPQELLEIEKLDKDAKKGSKVMFVLNLVFRVVKREKILIFCHNIPIRFL  
 ELFENVFRWKRGRELLTLDLELFRGR VIDKFEPPGGQSR VLLASITACAEGISLTAASRVIM  
 LDSEWNPSKTKQAIARAFRPGQKVVVYVYQLLSRGTLEEDKYRRTTWKEWVSSMIFSEEFVE  
 DPSQWQAEKIEDDVLREIVEEDKVK SFHMIMKNEKASTGG

Length 1256aa  
 Unique 101/1256=8%; The remaining 2.9% also map to CLASSY2 (AT5G20420).  
 Total 101+36=137/1256=10.9%

R.DEETPEPVVVDAR.V	R.ETGVSEEPQIYK.K	K.ATNVVEQWQGLK.N	R.VIDKFEPPGGQSR.V
R.YSEDCSSLVK.T	K.LIDSYMSR.I	R.VDNEAFHK.A	R.VIMLDSEWNPSK.T
R.EIVLVDK.T	K.DKATNVVEQWQGLK.N	R.ESPGLLVLDEGHNPR.S	R.GTLEEDKYR.R

K.IEDDVLR.E

R.EIVEEDKVK.S

**CLASSY2 (AT5G20420)**

MKKRGFYNLKHPFDPCPFEEFFCSGTWKPVEYMRIEDGMMTIRLLENGYVLEDIRPFQRLRLRS  
RKAALSDCICFLRPDIDVCVLYR**IHEDDLEPVVVDAR**IVSIERKPHESECSCCKINVRIYIDQGCIGS  
**EK**QRINRDSVVIGLNQISILQKFYKEQSTDQFYRWRFSEDCTSLMKTRLSLGKFLPDLWSLTVT  
STLKSIVFQIRTVQTKM VYQIVTDEEGSSSTLSSMNITLEDGVSLSKVVK**FNPADILDDSQDLEIK**  
**QETDYYQEEDEVVELR**RSKRNRVDPDIYTGCDEYPTIDGWVRMMPYQFGKCAVNVSEDEDE  
DDNNEGDGTNDDLYIPLSRLFIKKTNSREAKPKSRKGEIVVIDKRRVHGFGRKERKSELSVIP  
FTPVFEPIPLEQFGLNANSFGGGGFSRSQYFDETEKYRSKGMKYGKKMTEMEEMMEADLC  
WKGPNQVKSQKRTSRSSRSVAPKTEDSDEPRVYKK**VTL**SAGAYNK**LIDTYMNNIESTIAAKDE**  
**PTSVVDQWEELK**TNFAFKLHGDMEK**NLSE**GEGETSENEMLWR**EMEL**CLASSYILDDNEVR  
VDNEAFEKARSGCEHDYRLEEEIGMCCRLCGHVGSEIKDVSAPFAEHKKWTIETKHIEEDDIKT  
KLSHKEAQTDFSMISDSSEMLAAEESDNVWALIPKLRKRLHVHQRRAFEFLWRNVAGSVEPS  
LMDPTSGNIGGCVISHSPGAGKTFLIAFLTSYLKLFPGKRPLVLAPKTTLYTWYKEFIKWEIPVP  
VHLIHGRRTYCTFKQNKTVQFNGVPKPSRDVMHVLDLCLEKIQKWHAPSVLVMGYTSFTTLMR  
EDSKFAHRKYMALVRES**PGLLVLDEGH**NPRSTKSRLRKALMKVGTDLRILLSGTLFQNNFCE  
YFNTLCLARPKFIHEVLMELDQKFKTNHGVNKAPHLEENRARKLFLDIIAKKIDASVGDERLQGL  
NMLKNMTNGFIDNYEGSGSGDALPGLQIYTLVMNSTDIQHKILTKLQDVIKTYFGYPLEVELQ  
ITLAAIHPWLVTSSNCCTK**FFNPQELSEIGK**LKHDACKGSKVMFVNLNIFRVVKREKILIFCHNIAP  
RMFTELFENIFRWQRGREILTGTGDELFERGR**VIDKFE**PGNPSR**VLLASITACAEGISLTAASR**  
**VIMLDSEWNPSK**TKQAIARAFRPGQQKV VYVYQLLSRG**TLEEDKYR**RTTWKEWVSCMIFSEEF  
VADPSLWQAEK**I**EDDILR**EIVGEDK**VKS FHMIMKNEKASTG

Length 1261aa

Unique 149/1261=11.8% ; The remaining 2.85% also map to CLASSY1 (At3g42670).

Total 149+36=185/1261=14.7%

- |                                 |                                  |                              |                        |
|---------------------------------|----------------------------------|------------------------------|------------------------|
| * R. <b>IHEDDLEPVVVDAR</b> .I   | * K. <b>VTL</b> SAGAYNK.L        | R. <b>ESPGLLVLDEGH</b> NPR.S | R. <b>GTLEEDKYR</b> .R |
| * R. <b>IYDQGCIGSEK</b> .Q      | * K. <b>LIDTYMNNIESTIAAK</b> .D  | * K. <b>FFNPQELSEIGK</b> .L  | * K. <b>I</b> EDDILR.E |
| * K. <b>FNPADILDDSQDLEIK</b> .Q | * K. <b>DEPTSVVDQWEELK</b> .T    | * R. <b>VIDKFE</b> PGNPSR.V  |                        |
| * K. <b>QETDYYQEEDEVVELR</b> .R | * K. <b>NLSE</b> GEGETSENEMLWR.E | R. <b>VIMLDSEWNPSK</b> .T    |                        |

**CHR31 (At1g05490)**

MECIGKRVKRSWQRLQAVNKRKKMETVAPVTSPPKRRQKPKKNYDSIEDITPTCNDVSPV  
PQVSNMYSVPNNSVKESFSRIMRDLNVEKSGPSSSRLTDGSEQNPCLKERSFRVSDLGVEK  
**KCSPEITDLDVGIPVPR**FSKLKDVSEQKNTCLMQSSPEIADLDLISVPSSSVLKDVSEEIRFLK  
DKCSPEIRGLVLEKSVPGIEILSDSESETEARRRASAKKLFEESSRIVESISDGEDSSSETDEE  
EEENQDSEDNNTKDNVTVESLSSDPSSSSSSSSSSSSSSSSSSSSSSDDES YVKEVVDNRDDD  
DLRKASSPIKRVSLVERKALVRYKRSGSSLTKPRERDNKIQKLNHREEEKERQREVVRVVK  
QPSNVVYTCAHCGKENTGNPESHSSFIRPHSIRDEIEDVNNFASTNVSKYEDSVSINSGKTTGA  
PSRPEVENPETGKELNTPKPSISRPEIFTTEKAIDVQVPEEPSRPEIYSSEKAK**EVQAP**EMPSR  
**PEVFSSEK**AKEIQV**PEMPSIPEIQNSEK**AKEVQANNRMGLTTPAVAEGLNKSVVTNEHIEDSD  
SSISSGDGYESDPTLKDKEVKINNHSWRLNGNKEVDLFRLLVNSVWEK**GQLGEE**DEADEL  
**VSSAEDQS**EQAR**EDHRKYDDAGLLI**RPPPLIEKFGVEEPQSPVSEIDSEEDRLWEELAFF  
TK**SNDIGG**NELFSNVEKNISANET**PAAQCK**KGKHDLCIDLEVGLKCMHCGFVEREIR**SMDVSE**  
**WGEK**TTRERRKFDREEEEEGSSFIGKLGFDAPNNSL**NEGCVSSEGT**VW**DK**IPGVKSQMYPHQ  
QEGFEFIWKNLAGT**IMLNEL**KDFENSDETGGCIMSHPGTGKTRLTIIFLQAYLQCFDCKPVIIA  
PASLLLTWAEFKWNISIPFHNLSLDFTGKENSAA**LGLLMQKNATARS**NEIRMVKIYSWIKS  
KSILGISYNLYEKLAGVKDEDKTKMVREVKPDKELDDIREILMGRPGLLVLDEAHTPRNQRS



WKTLSKVETQKRILLSGTPFQNNFLELCNVLGLARPKYLERLTSTLKKSGMTVTKRGKK**NLGNE**  
**INNR**GIEELKAVMLPFVHVHKGSLQSSLPGLR**ECVVVLNPPPELQR**RVLESIEVTHNRKTKNVFE  
TEHKLSLVSVHPSLVSRCIKSEKERLSIDEALLAQLKKVRLDPNQSVKTRFLMEFVELCEVIKEKV  
LVFSQYIDPLKLIKHLVSRFKWNPGEVLYMHGKLEQKQR**QTLINEFNDPK**SKAKVFLASTKA  
**CSEGISLVGASR**VILLDVVWNPAPERQAISRAYRIGQKRIVYTYHLVAKGTPEGPKYCKQAQKD  
**RISELVFACSSRHDKGKEKIAEAVTEDKVLDTMVEHSKLGDMFDNLIVQPK**EADLVEGFSILMP

Length 1410aa

Unique 239/1410=16.9; The remaining 1% also map to CHR40 (At3g24340)

Total 239+13=252/1410=17.9%

K. <b>CSPEITDL</b> DVGIPVPR.F	K. <b>SNDIGGNELFSNVEK</b> .N	K. <b>LGFDAPNNSLNEGCVSSEGTVWDK</b> .I	K. <b>ACSEGISLVGASR</b> .V
K. <b>EVQAPEMP</b> SRPEVFSSEK.A	K. <b>NISANETPA</b> AQCK.K	K. <b>NLAGTIML</b> NELK.D	R. <b>ISELVFACSSR</b> .H
K. <b>EIQVPEMP</b> SIPEIQNSEK.A	R. <b>SMDVSEWGEK</b> .T	K. <b>NLGNEINNR</b> .G	K. <b>IAEAVTE</b> DK.V
K. <b>GQLGEEDE</b> ADLVSSAEDQSSEQAR.E	R. <b>SMDVSEWGEK</b> .T	R. <b>ECVVVLNPPPELQR</b> .R	K. <b>VLDTMVEH</b> SK.L
K. <b>GQLGEEDE</b> ADLVSSAEDQSSEQAR.E	R. <b>FEEEGSS</b> FIGK.L	R. <b>QTLINEFNDPK</b> .S	K. <b>LGDMFDNLIVQPK</b> .E

### CHR40 (At3g24340)

MDMTSCVARRRSRTESYLNLSILNKSK**GISGEEEDQSLGCVNSR**TEKRRVNMRDACSPSPRK  
KKRRRRK**DDDDDDVV**FVRTEYPEGKRDDENVGSTSGNLQSK**SDFDGD**RVCDFDADDRNLGCE  
EK**ASNFPID**DDDDVV**FVGT**VQRENDHVEDDDNVGSASVISPRVCFDEDDAKVSGKENPLS  
PDDDDVVFLGTIAGENQHVEDVNAGSEVCDILLDDANLRGEEKTYVSDEVVSLSSSSDDEED  
PLEELGTDSREEVSGEDRDSGESDMDDEDANDSDSSDYVGESSDSSDVESDSDSDFVCSDEE  
GGTRDDATCEKNPSEKVVYHHKKSRTFRRKHNFDVINLLAKSMLESKDVFKEIDFSWDKIAEVDS  
REDPVVRESSSEKVNHEGKPRRRSFHRVREKNHNLNGESFYGGEEKLCDGEETINYSTEDSPP  
LNL**RFGCEEP**V**LI**EKTTEEKELDSLWEDMNVALTLEGMHSSTPDKNGDMLCSK**GTHDFVLD**DE  
**IGL**KCVHCAYYAVEIK**DISPAM**DKYRPSVNDNKKCSDRKGDPNRLPFDASDPSSFVAPLDNI  
EGTVWQYVPGIKDTLYPHQQEGFEFIWKNLAGTTK**INELNSV**GVK**KG**SGGCIISHKAGTGKTRLT  
VVFLQSYLKRFPNSHPMVIAPATLMRTWEDEVKWNVNIPFYNMNSLQLSGYEDAEAVSRLEG  
NRHHNSIRMVKLVSWWKQKSILGISYPLYEKLAANK**NTEGMQV**FR**RML**VELPGLLVLDEGHTP  
RNQSSLIWKVLTEVRTEKRIFLSGTLFQNNFKELSNVLCLARPADKDTISSRIHELKCSQEGEH  
GRVNEENRIVDLKAMIAHFVHVHEGTILQESLPGLRDCVVVLNPPFQQKKILDRIDTSQNTFEFE  
HKLSAVSVHPSLYLCCNPTKK**EDLVIG**PAT**LGL**TKRLRLKYEEGVKTKFLIDFIRISGTVKEK**VLV**  
**YSQYID**TLK**LIME**QLIAECDWTEGEQILLMHGKVEQRDRQH MIDNFNKPDSGSKVLLASTKACS  
**EGISLVGASR**VVILDVVWNPVSVESQAISRAFRIGQKRAVFIYHLMVKDTSEWNKYCKQSEKHRI  
**SELVFS**STNEKDKPINNE**VVSK**DRILDEMVRHEKCLKHIFEKILYHPKSDMNTSFF

Length 1132 aa

Unique 156/1132=13.8%; The remaining 1.1% also map to CHR31 (At1g05490).

Total 156+13=169/1132=14.9%

K. <b>GISGEEEDQSLGCVNSR</b> .T	K. <b>ASNFPID</b> DDDDVV <b>FVGT</b> VQR.E	K. <b>INELNSV</b> GVK.G	K. <b>ACSEGISLVGASR</b> .V
K. <b>DDDDDDVV</b> FVR.T	R. <b>FGCEEP</b> V <b>LI</b> EK.T	K. <b>NTEGMQV</b> FR.R	R. <b>ISELVFS</b> STNEK.D
K. <b>SDFDGD</b> R.V	K. <b>GTHDFVLD</b> DEIGLK.C	K. <b>EDLVIG</b> PAT <b>LGL</b> TK.R	K. <b>DKPINNE</b> V <b>VSK</b> .D
K. <b>ASNFPID</b> DDDDVV <b>FVGT</b> VQR.E	K. <b>DISPAM</b> DK.Y	K. <b>VLVYSQYID</b> TLK.L	

## SHH1 (At1g15215)

MAASDDSSHYFTEFTLSEIVDMENLYKELGDQSLHKDFCQTVASTFSCSVNRNGKSSITWKQV  
QIWFQEKLKHQSQPKSKTLPSPPLQIHDLNPNSSYASNASNATFVGNSTFVQTRKGGKASDLAD  
LAFEAKSARDYAWYDVSSFLTYRVLRTGELEVRVRFSGFDNRHDEWVNVKTSVRERSIPVEPS  
ECGRVNVGDLLLCFQEREDQALYCDGHVLNIRGIHDHARCNCVFLVRYELDNTTEESLGLERIC  
RRPEE

Length 258aa

Unique 65/258=25.2%

Total=25.2%

K.ELGDQSLHK.D

R.FSGFDNR.H

R.EDQALYCDGHVLNIR

K.ASDLADLAFEAK.S

R.HDEWVNVK.T

R.YELDNTTEESLGLER.I

1. Ream TS, Haag JR, Wierzbicki AT, Nicora CD, Norbeck AD, 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. Mol Cell 33: 192-203.