

Characterization of two members of the AN subfamily, NtANs, from *Nicotiana tabacum*

Jing Wang

College of Tobacco Science, Henan Agricultural University, Zhengzhou, 450002, P. R. China



Introduction





Arabidopsis



Ipomoea nil



Zea mays

Oryza sativa



Marchantia



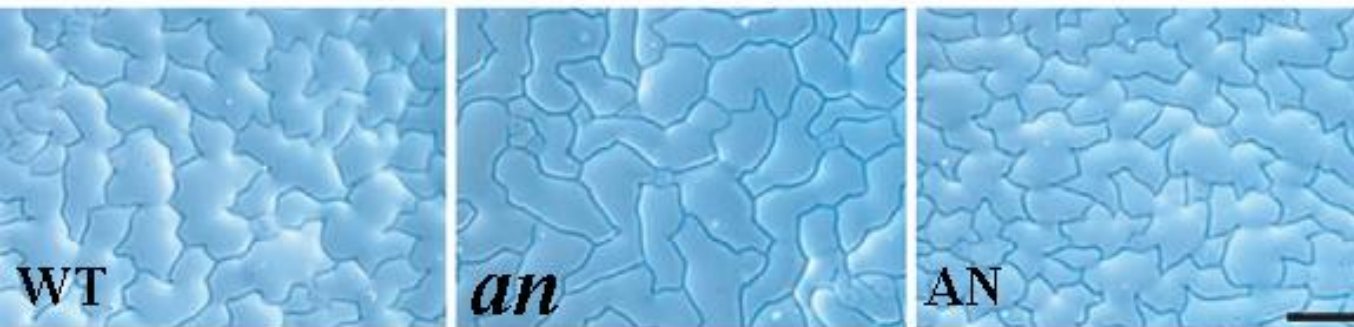
Bryophyta

Larix spp





WT : wild type
an: *an* mutation
AN: AN transgenic



Shapes of adaxial epidermal cells



Trichome branching patterns



Materials

Tetraploid specie:

K326

Diploid specis:

N. sylvestris

N. tomentosiformis



Methods

EST amplification :

Degenerate PCR

5'/3' cDNA ends amplification :

RACE PCR

Promoter amplification :

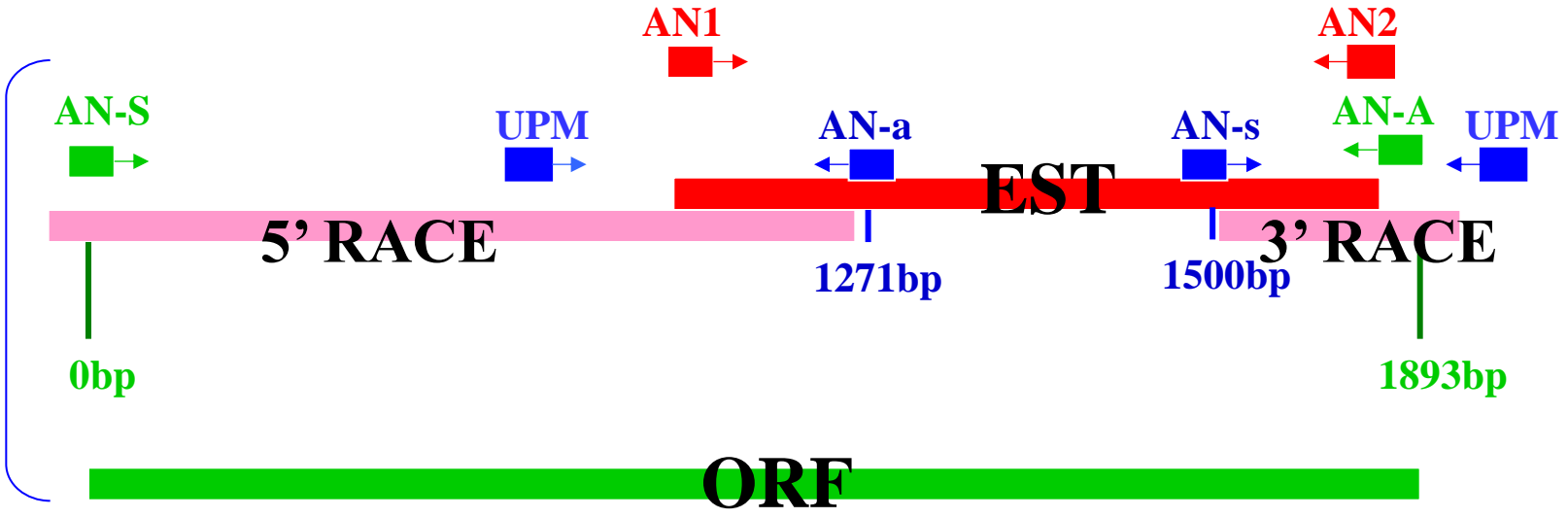
Reverse PCR

Gene structure amplification :

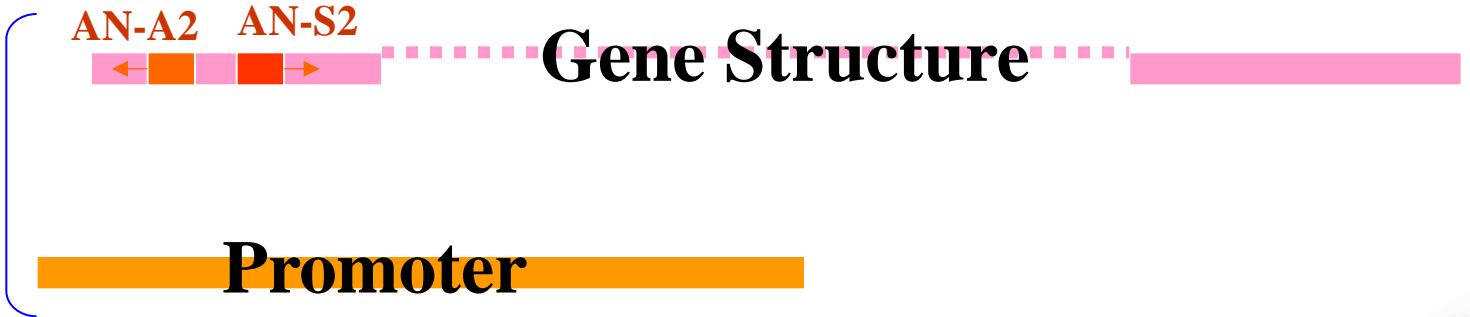
LD PCR



cDNA

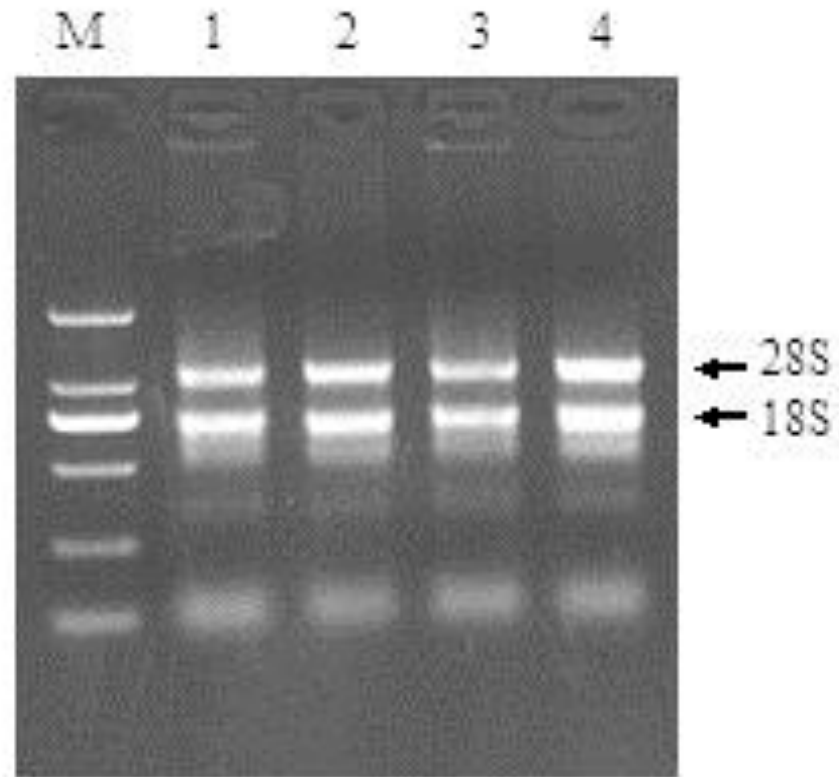


DNA



Results

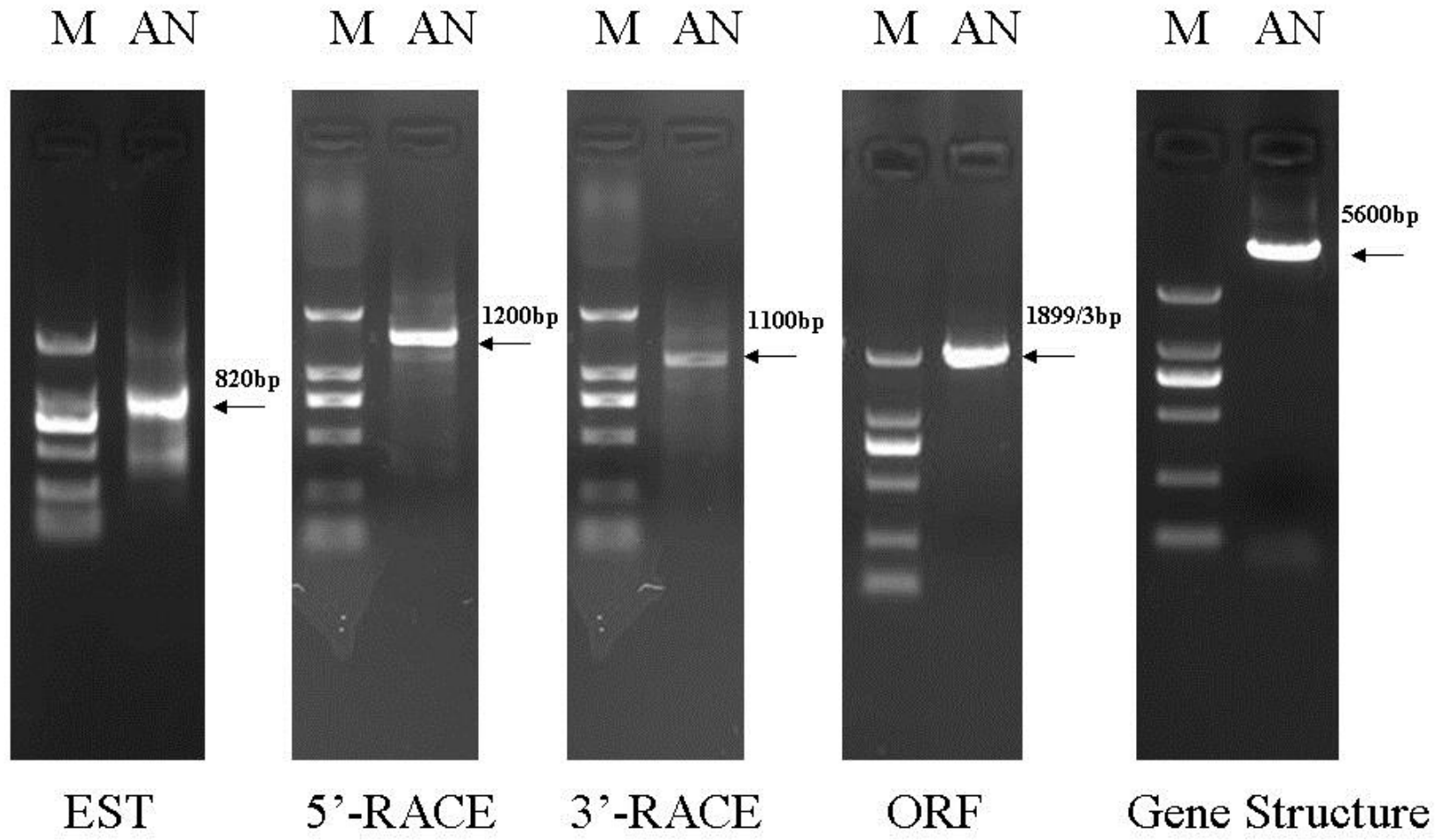
1. RNA Extraction



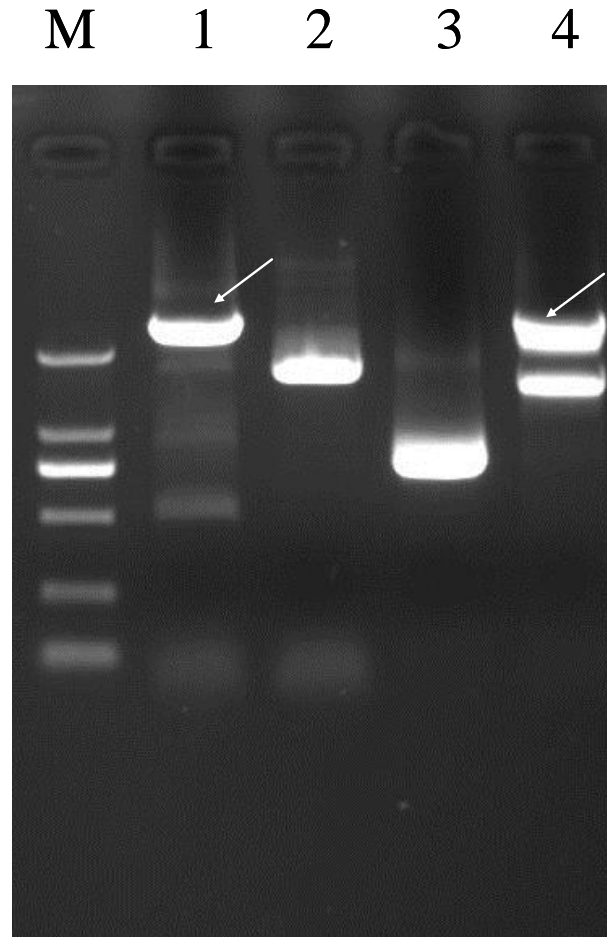
RNA Extraction of Tobacco



2. AN Gene Clone



3. AN gene promoter clone



Promoter



4. Protein Sequence Analysis

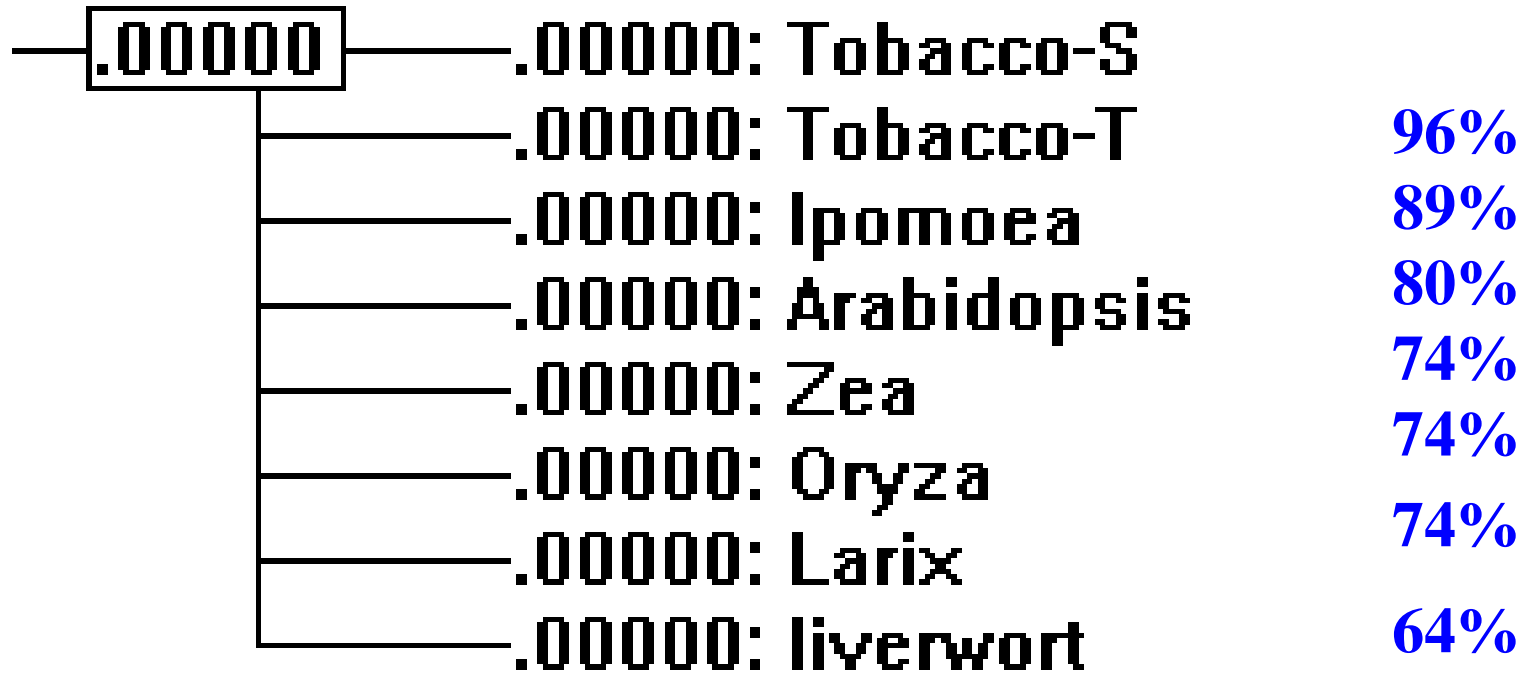
```
Tobacco-S : ---MAHHNKTTSSSVT-----QQVPLVITLNCIEDTALAEQECLSGIAVIEHVPLSRLAEARIESATAVLLHS : 64
Tobacco-T : ---MAHHNKTTSSSVT-----QQVPLVITLNCIEDTALAEQECLSGIAVIEHVPLSRLAEARIESATAVLLHS : 64
Ipomoea : MSWKKRENQSSSSSRSPAPME---LPNNPTLEKQKMPVIVTTLNCIEDTAFEQDCLAGIVLVHVPVLSRLAEARIESASAVLLHS : 81
Arabidopsi : ---MSKIRSSATMPHRDQPS-----PASPHVVTLNCIEDCALEQDSLAVGAVGVYVPLSRIADGKIESATAVLLHS : 68
Zea : ---MAHSPAPSG-----GGGGGPIPLVSLNCLDDLSLEQEGLAGVAAVEHVPLSVAACGRIEAASAVLLPS : 64
Oryza : ---MLHGPAHSPPAAAAYAV-----AGGGGPEPLVITLNCIEDPSMEQEVLAGAAAVEHAEPLSALSSGRVAAAAYVLLTS : 73
Larix : ---MTKMEAVNSNGRDGDRDP-----AGGPIVITLNCIEDDCRFEEETLRGIALVQHVPLSRIADGRIESAMAVLLHS : 70
liverwort : ---MGKGVTMPPEASSGRVDGGRVAERNPGSRQIGKMRSDGTVRKSKGYEPLSMGLPIVVALNCLMDDCRAEAEALEGVAVVEHVGLAQVGGKIEAAVAVLVQS : 101
m P166 LNC6 D E e L G a 6 hvpLs 6a 46E A AVL6 S
```

```
Tobacco-S : LAFLPRAAQRRLRSWQLILCLGSSDRAVDSALASDLGLSRLVHVDVSRABEAVDTVMALLLGLLRRTHLLSRHTLS----ASGWLGSVQPLCRGMRRRCRGLVL : 163
Tobacco-T : LAFLPRAAQRRLRSWQLILCLGSSDRAVDSALASDLGLSRLVHVDVSRABEAVDTVMALLLGLLRRTHLLSRHTLS----ASGWLGSVQPLCRGMRRRCRGLVL : 163
Ipomoea : LAFLPRAAQRRLRPWQLILCLGSSDRAVDSALAADLGLTRLVHVDCSRABEAVDTVMALLLGLLRRTHLLSRHALS----ASGWLGSVQPLCRGMRRRCRGLVL : 180
Arabidopsi : LAYLPRAAQRRLRPWQLILCLGSSADRAVDSALAADLGLTRLVHVDTSRABEADTVMALLLGLLRRTHLLSRHALS----ASGWLGSVQPLCRGMRRRCRGMVL : 166
Zea : LAFLPRAAQRRLRPWQLLCLGSSADRAADAALADLGLRLVHVDANRABEAVDTVMALLLGLLRRTHLLSCHAS--SVPAAGWLGSVQPMCRGMRRRCRGLVL : 164
Oryza : LAFLPRAAQRRLRPWQLILCLGSSPDRAADAAVAELGLRLVHVDANRABEAVDTVMALLLGLLRRTHLLSRHASSYSAPPAGWLGSVQPLCRGMRRRCRGLVL : 175
Larix : LAYLPRAAQRRLQPWQLILCLGSSDKAVDSALASDLGLQLLHVDTARABEAVDTVMALLLGLLRHTHLLSKQGFSS--SSGWLGSVQPLCRGMRRCHGLVL : 168
liverwort : LAYLPRAAQRRLQPWQLILSLGCADKAVDSCLASDLGLQLLHVDSGRSEEVADTVMALLLGLLRRTPALAAQAGA---SAGWLGLALPAACRGMRRRCRQGVL : 199
```

```
Tobacco-S : VSHQPIEDV---AENSQKRLVNDSSRESPSQLQGSIVSQNSSGRSEVKRRSR--SGKKAKKRHRGKHSQHKVDD---SLAFBKES--TSHHEDGATMSGTDQGVSS : 459
Tobacco-T : VSHQPIEDV---AENSQKRLVNDSSRESPSQLQGSIVSQNSSGRSEVKRRSR--SGKKAKKRHRGKHSQHKVDD---SLAFBKES--TSHHEDGATMSGTDQGVSS : 459
Ipomoea : VGERYTEDVNLIASQTKIMSSREPPVQPGSIVLSQNVSRSEVKRRSR--SGKKAKKRHRGKHSQHKVDE---HLKFBKES--T SQNDDGAALSQTDQVLS : 481
Arabidopsi : TSRQGGEST-LTSTEIVRREASELKEKSLSPGQ-QHVSQNTAVKPEGRSSR--SGKKAKKRHSQQKYMOKTDG---SSGLNEES--TSRRDD-IAMSDTEEVLS : 463
Zea : FDTIDEISH-ITLSEKRAISHHKPQASGK-SVN---IGSRSEGRSSR--SGKKGKKRPAHRPQOKPDD---LSAVESDSNYSRRDDDTAMSSRDQVVS : 462
Oryza : -GQOTDESQ-LTLECDKRRALSHSEBPQASGQ-SQNRENVVPRSEGRSSR--SGKKGKKRPAHRKQOKRDEL--LSTLEGGSNYSRRMDDDTVTSGKDQVLS : 476
Larix : DGEQLGDESQTSPEFRQQAAYYQSQEQLRQSQDSSISQSQSGGQREVRHGR--SGKKGRKRPGRRKSQOKSDSV-VLLERENNCVALQRDDNGASVSGRDPVVS : 470
liverwort : DGEQRSGESQTAQDFRHKQMYFQSQDQRVHSQEVLDLSHGLVQRDSRNNHVKSGKKGKKRAGRKAQQQPSGTSIASERDATWLTLOREDRGNGTSSKDAVVN : 507
e e s e q s s e rsr SGKK 4KR r 4 qqk d e s s d s d v6s
```

```
Tobacco-S : SSSRFASPEDLR-GRKTP----IBSIQESSGE-QLPKQGMDSLRSKSAELLKDGYYIIALHARHHPALHVSQRQVKGCGG-WFLDIMSVDVTKRDPAAQFLVVFRSK : 555
Tobacco-T : SSSRFASPEDLR-GRKTP----IBSIQDSSGV-QLPKKGMDSLRSKSAELLKDGYYIIALHARHHPALHVSQRQVKGCGG-WFLDIMSVDVTKRDPAAQFLVVFRSK : 553
Ipomoea : SSSRFESSPEDVR-SRKTTP----IBFIQESSSE-KLLKSNMDSLRSKSGELLKDGYYIIALYARHHPALHVSQRQVQGGG-WFLDMSNITKRDPAQFLVVYRSK : 577
Arabidopsi : SSSRCASPEDSR-SRKTTP----LVMQESSPN-QLVMSSKKFIGKSSSELLKDGYYVVALYAKDLSGLHVSQRQTKNGG-WFLDTLSNVSKRDPAAQFLIAYRNRK : 559
Zea : SSSRFASPEDPKYKHKSL----SESPMEITSE-KKVPVL--LSRKYPDKLKDGFIVALRARDNSGYHVARQRVVGCGGWLDDVVSNAITNRDPAAQFLVTFKKNK : 558
Oryza : SSSRFASPEDCKTKLRS----AEFEMEISE-NKLTAG--LSIKPLERLKDGFVVALRTRDNSGFHVARERVAVCG-WYLLDVVSQATKRDPAQFLITFRNK : 551
Larix : SSSRFASPEDSRHRGGD----VGSTVDTASE-QSVTSKGPYQGTGDLKDLGLVLSLRAKDRAGYHVSQRQVGGG-WFLDIMSVDVTKRDPAAQFLVVFRSK : 557
liverwort : SNSRFASPEESKIKREEDGTLGVEVIVESSQEVQKVAVKGTNQAALDYLKEGQIVALARADGGYYVARQKGPGRG-WFLDIMSVDVTKRDPAAQFLVVFRNR : 689
SsRfASPEd 4 r e e e k LKdG 66aL a4 hv R24 ggG W LD 6s 3kRDPaAQF66 4 4
```

```
Tobacco-S : DTIGLRSEFAGGKLLQINRRMEEVFASHSFDVWESWTFEGT-MEECRLVNCRNPLAVLDVRVEVLAAVGE-DGITRWLD----- : 632
Tobacco-T : DTIGLRSEFAGGKLLQINRRMEEVFASHSFDVWESWTFEGT-MEECRLVNCRNPLAVLDVRVEVLAAVGE-DGITRWLD----- : 630
Ipomoea : DTIGLRSEFAGGKLLQINRRMEEVFASHSFDVWESWTFEGS-LEECRLVNCRNPLAAILDVRIEVLAAIGE-DGITRWLD----- : 654
Arabidopsi : DTVGLRSEFAGGKLLQINRRMEEVFASHSFDVWESWMSLEGS-LDECRLVNCRNSSAVLDVRVEILAMVGD-DGITRWLD----- : 636
Zea : DTMGLRSEFVAGGKLLQINRRMEEVFASHSFDVWESWMLDGSLLLEGSKLINCNRNPSAVLDICIEILAAPSEEDGVTRWLDSPRWGL : 643
Oryza : DTMGLRSEFVAGGKLLQVNRKTMELVFASSHFDVWESWTFEGSLLDCCKLVNRKIPSVVLEVIYIEILAAVSEEDGVTRWLD----- : 650
Larix : ERIGLRSLAAGKLLQINRRMEEVFASHNFDVWENMILEGSTLECLLRNTRSPSVALDVSIIEILAAVGEEDGIARWLS----- : 646
liverwort : DRIGLRSLAAGKLLQANKKLELVFNHTFDVWESWIVVEGSTELEDCLTNTSKFRGVSVDVSIIEILAAVGEEDGVTRWLS----- : 688
dt6GLRSf AGGKLLQ N4 6E vFash FDvWESw eG3 6e c L N 4 p Ld6 6E6LAa ge DG6tRW6d
```



```
>userseq29283 1237nt
+ ACTCCCAAG ACCATGACAA AACGCTAGCT TAATTACTTT GGAATGAGT AATGTATCTT TTTTGGAAAC
- TGAGGTTTC TGCTACTGTT TTGCGATCGA ATTAATGAAA CCTTACTTCA TTACATAGAA AAAATCTTTC

+ AATATCTAC TCCTACTCGA TTAAATTACT TATATCAAGT CAAACTCCTC AATATGCTCA ACTTTTAAAG
- TTTATAGATG AGGATGAGCT AATTTAATGA ATATAGTTCA GTTTGAGGAG TTATACGAGT TGAAAAATTC

+ TTTTAAAGG AATTTTAGST ATATATAGAA GTAGCAITAG TACCGAGTAC TTCAGTGGG TCTACAAAAA
- AAAAATTTCC TTAAAAATCCA TATATATCTT CATCGTAATC ATGGCTCATG AGTTTCACTC AGATGTTTTT

+ GTTGCAGAA TTTGAATTTG TGATAAATAA AATGGATGCA TATCTAGAAC TATTTACACA TTTTATATTT
- CAACTTCTT AAACTTAAAC ACTATTTAT TTACCTACST ATAGATCTTG ATAAATGTGT AAAATATAAA

+ GTACCAATGA ACAAAACTT CTCTCCGAT CTCCATCTAG TAACCAATTT TTTTTCSTCA TCAAAGGATA
- CATGGTACT TGTTTTTGAA GAGAGGCGTA GAGGTAGATC ATTGGTTAAA AAAAAGCAGT AGTTTCTCAT

+ AITAAITAIT TAITGCATTT AITACTTAGT AAATAAATAT ATGGSTTTTT GTAAGGATA AITAGCTTCC
- TAAITAAATA ATAACTGAAA TAATGAATCA TTTATTTATA TACCCAAAAA CAITTTCTCAT TAATCGAAGG

+ ATCCAAACT AITGGAAAT TTGAATTTAC AATAGGAATA ACTAAAAITA AGGATTTGTA AACTATTTTC
- TAGTGTTTGA TAACTTTTAA AACTTAAATG TTATCCTTAT TGATTTTAAAT TCTTAAACAT TTGATAAAGG

+ TAITAGATA CATTGAAAAA AATTAAAAAA TATAACAATA GCATATCCAA TATAITCCAA CATGGTAGAA
- ATAATTTAT GTAAACCTT TTAATTTTTT ATATTTTAT ATATAGGTT ATATAGGTT GTACCATCTT

+ AATAGAAAT ACTTTTTAAA AGCTATTGTA CTAATCTCA TGGTGAGTG CATAAAAAAT AAATTAGGAG
- TTATCTTTTA TGAAAAATTT TCGATAACAT GATTTAGAGT ACCACTCTAC GTATTTTTTA TTTAATCTCC

+ TGAACAAC TGTGAACITG AAATGAACAT AGGCAITGAT GACAGAAITA AAACGATAGA ACTTGCAATT
- ACTTTGTGTA ACACCTGAAAC TTTACTTGT TCCGTAACATA CTGCTTAAT TTTGCTATCT TGAACTTAA

+ AACACTCATA AITATAGAGC AAGAGCAGAT TTTTACTTTT AAAAAAACAA TATGTAAACA CAACACAAA
- TTGTGAGTAT TAATATCTGC TTCTCGTCTA AAAATGAAAA TTTTTTGT TTATCAITTT ATTTTGTTT

+ AAAAAAAT AITTCAGAAC CATTITGTTT ATCAGATAAA AITTTCCATT TTTTAAAAAC GTTTCAAAA
- TTTTTTATA TAAGTCTTGT GTAAAAACAA TAGTCTATTT TAAAAGTAA AAAATTTTGG CAAAGTTTTT

+ ATCTGTACAA ACGCAATAT ATTCAGTTT CACTTAAACA AAAAAAGAGA AAGGAAACA ATGAAAAACA
- TAGACATGTT TCCGTTATA AAGTTTTAAA GTTGAITTTT TTTTTTCT TTCTTTTGT TACTTTTTGT

+ AAAAGGAT CCAAAATAA ATCCAGGCAA TGATTCATC ACGTTGCCCA AGTATACTCC TGCAGATAG
- TTTTCTATA GTTTTTATTT TAGSTCCGTT ACTAAGGTAG TGCAACGGGT TCATATGAGG ACGTCTATTC

+ GGACTTGAT GGGAAATAG GTAATTAATA AAGGCTTTT TGAACAAAA GAGAAAGTCA CGGGCCAGT
- CTTGAACTAA CCTTTTATTC CATTAAITAT TTCCGGAA AATTTGTTTT CTCTTTCAGT GCCCCGCTCA

+ STGCACAAA AAAGTGACA GAGAGGACTA GGTCTCTCA CACTCTTTA GGCACAACT GACACTCAT
- CACTGTGTT TTTTCACTGT CTCTCTGAT CCCAGAGAT GTTGAAGAA CCGTGTTTGA CTGTGAGTA

+ TCTTTCTAA AITCCCGGA GAAAAAGCG GAAATGTTCA CCGCGTCT CCGCGCTCC TAATAGTTT
- ACGAAGATT TAGCGGCT CTTTTCTGC CTTTACAGT GCGCGCAGA GCGCGCAGG AATATCAAG

+ CGACAAAC STCAATCC AGATCTCTC TCATCCGGC MCAAC
- GCTTTTGTG CAGTATAGG TCTAGGAGG ATAGGCTTG TCGTTG
```

5. Promoters Sequence Analysis

T-Type



+ AE-box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
AE-box	<i>Arabidopsis thaliana</i>	65	+	8	AGAAACAA	part of a module for light response

+ Box 4

+ Box I

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
Box I	<i>Pisum sativum</i>	505	-	7	TTTCAAA	light responsive element
Box I	<i>Pisum sativum</i>	1019	-	7	TTTCAAA	light responsive element
Box I	<i>Pisum sativum</i>	861	+	7	TTTCAAA	light responsive element

+ CAAT-box

+ CGTCA-motif

+ G-Box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
G-Box	<i>Pisum sativum</i>	950	+	6	CACGTT	cis-acting regulatory element involved in light responsiveness

+ G-box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
G-box	<i>Zea mays</i>	950	+	6	CACGTT	cis-acting regulatory element involved in light responsiveness

+ GAG-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
GAG-motif	<i>Hordeum vulgare</i>	308	-	7	GGAGATG	part of a light responsive element

+ GCN4_motif

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
GCN4_motif	<i>Oryza sativa</i>	1052	-	7	TGTGTCA	cis-regulatory element involved in endosperm expression

+ Gap-box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
Gap-box	<i>Arabidopsis thaliana</i>	898	+	9.5	CAAATGAA (A/G)A	part of a light responsive element

+ I-box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
I-box	<i>Flaveria trinervia</i>	531	+	10	cCATATCCAAT	part of a light responsive element
I-box	<i>Triticum aestivum</i>	974	+	8	AGATAAGG	part of a light responsive element

+ Skn-1_motif

+ TATA-box

+ TATA-box

+ TCA-element

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
TCA-element	<i>Brassica oleracea</i>	1071	+	9	GAGAAGAATA	cis-acting element involved in salicylic acid responsiveness
TCA-element	<i>Nicotiana tabacum</i>	1142	-	9	CCATCTTTT	cis-acting element involved in salicylic acid responsiveness

+ TGACG-motif

+ Unnamed__13

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
Unnamed__13	<i>Zea mays</i>	957	+	9	TCCAAGTATA	

+ Unnamed__4

+ chs-CMA2a

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
chs-CMA2a	<i>Petroselinum crispum</i>	192	-	8	TCACTTGA	part of a light responsive element

+ circadian

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
circadian	<i>Lycopersicon esculentum</i>	497	-	6	CAANNNNATC	cis-acting regulatory element involved in circadian control
circadian	<i>Lycopersicon esculentum</i>	1197	+	6	CAANNNNATC	cis-acting regulatory element involved in circadian control

>userseq3768 1345nt

```

+ ATTTATAAAA GGAGATAGAA AGAAGCACAT TTCACCAAAA TTATTCTACA CACACGATCT TCATAAAAGT
- TAAATATTTT CCTCTATCTT TCCTCGTGA AAGTGTGTTT AATAAGATGT GTGTGCTAGA AGTATTTTCA

+ GGTGCATTG ATGTGCAACA AATCCGTTCA AGTGACAATC CAACAAATTT ATTCACTAAA TCITTTGCCAA
- CCACGTGTAAC TACACGTTGT TTAGGCAAGT TCACGTGTAAG GTTGTGTTAAA TAAGTGATT AGAAAACGTT

+ CTCCAACTTT TGAGAAGATG GTATACAAGA TTGGAATGCG GAGACTCAAA TATTTGAAT AAGATTTTTA
- GAGTTGAAA ACTCTTCTAC CATATGTCT AACTTACGC CTCGTAGTTT ATAAACITTA TTCTAAAAAT

+ TCAGGGGAG TAAAATAGT GATGCACTCT TTTTTCTTA CTAATGTTTT TTCTATGAG GTTTTTCTTA
- AGTCCCCCTC ATTTTATGCA CTACGTGAGA AAAAAAGAA GATTACAAA AAGGATACTC CAAAAGGAAT

+ TAAGTTTTT AATGAGGCAC ATACAATGC GTATTACTAA ATATGTGTAC TCTTTTTTCT TCATTAGGAT
- AITCCAAAAA TTACTCCGTG TATTGTTACG CATAATGATT TATACACATG AGAAAAAAGA AGTAATCCTA

+ TTTTCCACC TGATTTTTCT AGTAAGSTTT TAATGAGGCA CATTATCTTT TAATGAACAT CCAGGGGGA
- AAAAGGTTGC ACTAAAAAGA TCAITCCAAA ATTACTCCGT GTAATAGAAA AITACTTGTG GGTCCCCCT

+ GTGTTATAAA TATATTATAT TATGAATGTT CAITTAGTAC TCAGTTGTAA ATAACTTCC TAAGAGGCT
- CACAATATTT ATATAATATA ATACTTACAA GTAAATCATG AGTCAACAIT TATTTGAAAG AITTTCTTGA

+ TATCCATATG GGACTCCGCC GTAANAATGT TTAATCTATT AGTACTCTAT TGAAAAATAG AACTTATCA
- ATAGGTATAC CCTGAGGCGG CAITTATACA AATAGATAAA TCATGAGATA ACTTTTATTC TTTGAAATGT

+ CTTCACTACC CGGTTATGGA TAAACATTAC TCCCCTAGA AGATTATCCA TACCGGGTAT AATAAGCTTA
- GAAGTCATGG GCCAATACCT ATTTGTAATG AGGGGCATCT TCTAATAGST ATGGCCCAT AATTTCGAAT

+ TCCATTCACT ACTCCGTTAT GGATAAACAT TGCTCTCACT AAAAAATTAT CCATATATGA TATAATGGCA
- AAGTAAGTCA TGAGGCAATA CCTATTGTA ACGAGAGTCA TTTTTAATA GGTATATACT ATATTACCST

+ GCITACACAG CAGCTTGCAG TAGCAGCTTA CACAGTAGCT TGCACTAGCA GCITACACAA CAGTTTGCAG
- CGAATGTGTC GTCGAACGTC ATCGTCGAAT GTGTCACTGA ACGTCACTGT CGAATGTGTT GTCAAAACGTC

+ TAGCAGCTTA CACAACAGCT TCTTTCTTTT TATAAATAGA AGAGATTICA GTTCATTATG TACATCAATT
- ATCGTCGAAT GTGTGTCGA AAGAAAGAAA ATATTATCT TCTCTAAGT CAAGTAATAC ATGTAGTTAA

+ TTAATTCGAA TAATATATCA GTTCTCTCT ATATTAGTCT TTACTTTATC GTCTTTTATT TTATAACAAA
- AATTAAGCTT ATTATATAGT CAAGAGGAGA TATAATCGA AATGAAATAG CAGAAAATAA AATATTGTTT

+ MCACATTTTT TCTCAGATA AAAATTTCCA TTTTGTGAAA ACCSTTCCG AAAATCTCTC TAAACGGCAA
- TGTGTA AAAAGTCTAT TTTAAAAAGT AAAACACTTT TGGCAAGGC TTTTATGACG AITTTGCCST

+ TATTTTCAAA TTTAAAGCA AAAAAGGGA AAAAAACAA ATGAGAAAAA TTTCCAAAT AAATCCAGGC
- ATAAAGTTTT AAATTTGGT TTTTCTCTT TCTTTTGTT TACTCTTTTT ATAGSTTTTA TTAAGTCCG

+ GATGATTCCA TCACSTGCC CAAGTATATT CCTGCAGATT AAGGACTTGA TTCAGAAATA AGSTAATTAA
- CTACTAAGST AGTGCACCGG GTTCATATAA GGAAGTCTAA TCCCTGAACT AAGTCTTTAT TCCATTAAIT

+ TAAAGCTTT TTTGAAACAA AAGAGAAAGT CACGGGCCG GTGTGCACA AAAAAAGTGA CAGAGGAGC
- AITTCGAAA AACTTTGTT TTCTTTTCA GTGCCCCGC CACACTGTGT TTTTTTCACT GTCTCTCTG

```

```

+ TAAAGCTTT TTTGAAACAA AAGAGAAAGT CACGGGCCG GTGTGCACA AAAAAAGTGA CAGAGGAGC
- AITTCGAAA AACTTTGTT TTCTTTTCA GTGCCCCGC CACACTGTGT TTTTTTCACT GTCTCTCTG

+ TAGGCTCTCT CACAACCTCT TAGCAAAAAA ACTGACACTT CATTGCCTTC TAAATTCGCC GGAGAAAAAG
- ATCCAGAGA GTGTTGAGAA ATCGTTTTT TACTGTGGA GTAACGGAAG ATTTAAGCGG CCTCTTTTC

+ AAGGAAATGC TCGCCGCTT TCTCCGCTT CCTTAATCAT TTCSAGCAAC AACGTCATCA TCCAGATCTC
- TCCCTTACG AGCGGCGAA AGAGGCCCA GEAATTAETA AAGCTCGTTG ITGCACTAGT AAGTCTAGAG

+ CTTGSMCAG AAC
- GAGCCTGTC TTTG

```

S-Type



Motifs Found

AAGAA-motif

ABRE

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
ABRE	<i>Arabidopsis thaliana</i>	226	+	6	TACGTG	cis-acting element involved in the abscisic acid responsiveness
ABRE	<i>Arabidopsis thaliana</i>	357	+	6	CACGTG	cis-acting element involved in the abscisic acid responsiveness

AE-box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
AEBox	<i>Arabidopsis thaliana</i>	549	+	8	AGAAACTT	part of a module for light response

ARE

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
ARE	<i>Zea mays</i>	995	-	6	TGGTTT	cis-acting regulatory element essential for the anaerobic induction

Box 4

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
Box4	<i>Petroselinum crispum</i>	1116	+	6	ATTAAT	part of a conserved DNA module involved in light responsiveness

Box I

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
Box.I	<i>Pisum sativum</i>	195	-	7	TTTCAAA	light responsive element
Box.I	<i>Pisum sativum</i>	1151	-	7	TTTCAAA	light responsive element
Box.I	<i>Pisum sativum</i>	984	+	7	TTTCAAA	light responsive element

CAAT-box

CATT-motif

CGTCA-motif

+ CGTCA-motif

+ ERE

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
ERE	<i>Dianthus caryophyllus</i>	193	-	8	ATTTCAAA	ethylene-responsive element

+ G-Box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
G-Box	<i>Antirrhinum majus</i>	226	-	6	CACGTA	cis-acting regulatory element involved in light responsiveness
G-Box	<i>Pisum sativum</i>	1062	+	6	CACGTT	cis-acting regulatory element involved in light responsiveness
G-Box	<i>Pisum sativum</i>	357	+	6	CACGTG	cis-acting regulatory element involved in light responsiveness

+ G-box

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
G-box	<i>Daucus carota</i>	226	+	6	TACGTG	cis-acting regulatory element involved in light responsiveness
G-box	<i>Arabidopsis thaliana</i>	357	+	6	CACGTG	cis-acting regulatory element involved in light responsiveness
G-box	<i>Brassica napus</i>	356	-	7	CACGTGG	cis-acting regulatory element involved in light responsiveness
G-box	<i>Zea mays</i>	1062	+	6	CACGTT	cis-acting regulatory element involved in light responsiveness

+ GA-motif

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
GA-motif	<i>Arabidopsis thaliana</i>	521	-	8	ATAGATAA	part of a light responsive element

+ GCN4_motif

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
GCN4_motif	<i>Oryza sativa</i>	1164	-	7	TGTGTCA	cis-regulatory element involved in endosperm expression

+ LTR

+ MBS

+ Skn-1_motif

+ TATA-box

+ TATCCAT/C-motif

+ TC-rich repeats

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
TC-rich repeats	<i>Nicotiana tabacum</i>	242	+	9	GTTTTCTTAC	cis-acting element involved in defense and stress responsiveness
TC-rich repeats	<i>Nicotiana tabacum</i>	1009	-	9	GTTTTCTTAC	cis-acting element involved in defense and stress responsiveness
TC-rich repeats	<i>Nicotiana tabacum</i>	334	+	9	ATTTTCTTCA	cis-acting element involved in defense and stress responsiveness

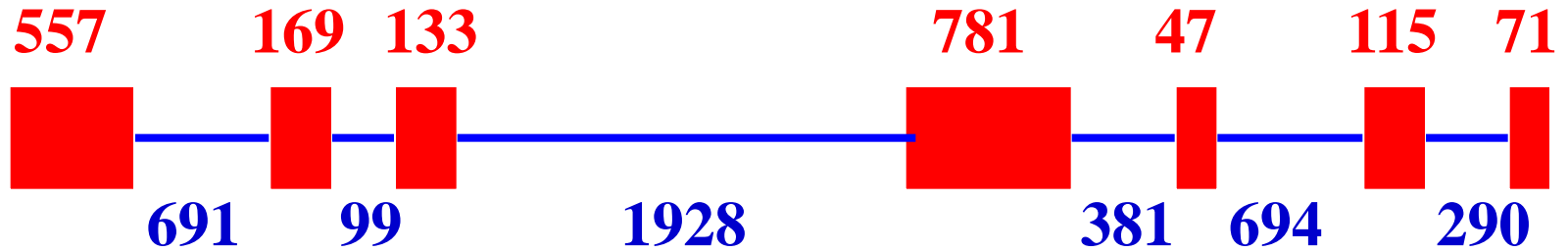
+ TCA-element

Site Name	Organism	Position	Strand	Matrix score.	sequence	function
TCA-element	<i>Brassica oleracea</i>	1023	+	9	GAGAAGAATA	cis-acting element involved in salicylic acid responsiveness
TCA-element	<i>Nicotiana tabacum</i>	1255	-	9	CCATCTTTTT	cis-acting element involved in salicylic acid responsiveness

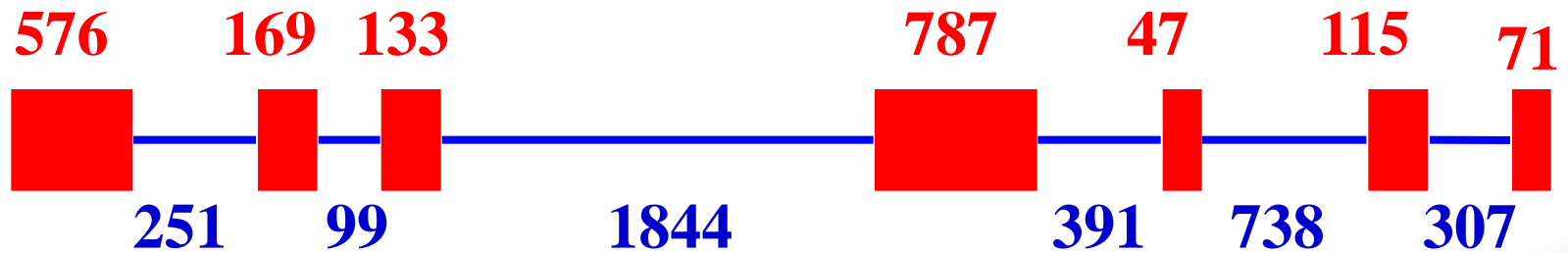


6. Gene Structure Analysis

T-Type



S-Type



Thank you!

