



# ***In silico* identification of Nuclear factor Y subunit B genes potential drought tolerance in Poplar genome and performance in *P. euphratica***

Dong-hui Yan, Hecheng Zhang, Sha Tang,  
Hongshuang Ma, Xilin Xia and Weilun Yin

College of Biological Sciences and Biotechnology, Beijing Forestry University, Beijing, 100083

Institute of Forest Ecology, Environment and protection, Chinese Academy of Forestry, Beijing, 100091

October 28, 2008, Beijing

# Nuclear Factor Y Heterotrimeric Protein (NF-Y protein)

## Characterization:

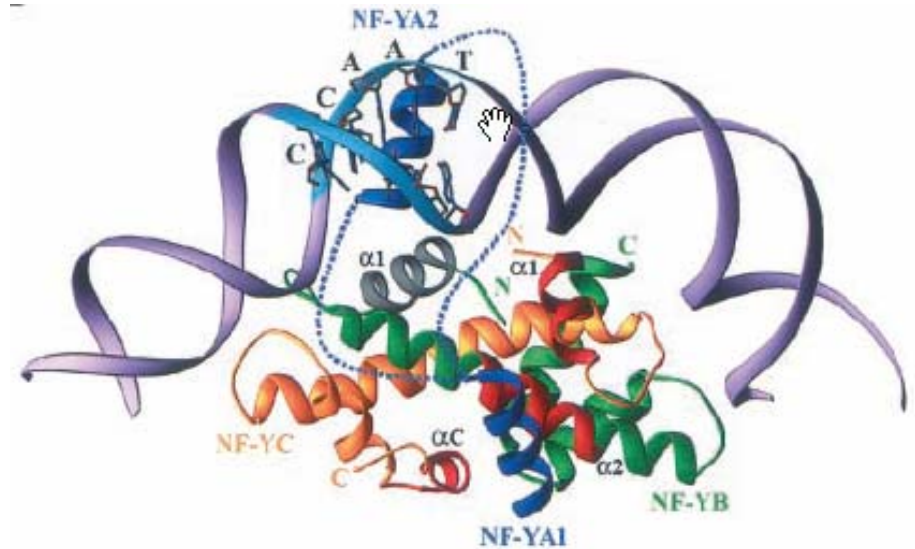
Nuclear transcription factor Y subunit A, subunit B and subunit C ( NF-YA,NF-YB and NF-YC )(in animal and plant), or HAP2,HAP3 and HAP5 (in yeast) or CBF-A, CBF-B,CBF-C( in animal)

## Function:

Transcription role in gene expression by binding CCAAT regulatory element.

## More genes to one subunit in plant

Each subunit encoded by a single gene in animal and fungi, by multiply genes in plant



Romier *et al.*, 2003, J. Bio. Chem.

**Multiply genes for one subunit protein implicate multiply function roles?** (Nelson et al., 2007, Shiu et al. 2005, Edwards et al. 1998)

**Knows three genes responsive to drought in NF-YB subunit family**

*AtNF-YB1 Arabidopsis thaliana* (Nelson, et al., 2007)

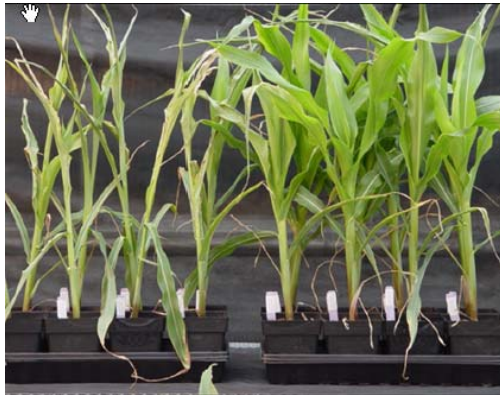
*ZmNF-YB2 Zea mays* (Nelson, et al., 2007)

*TaNF-YB2 Triticum aestivum* (Stephenson, et al., 2007)

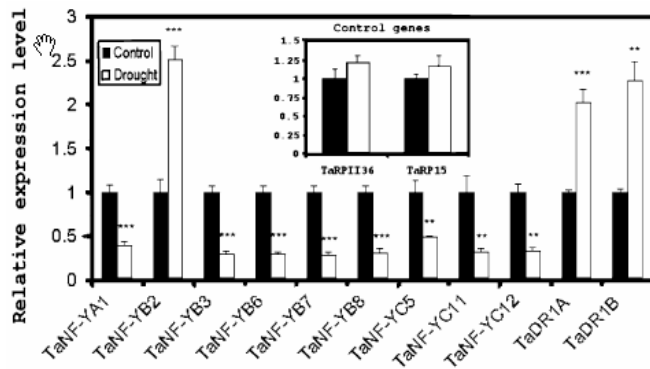
**one gene in NF-YA subunit family**

*AtNF-YA5 Arabidopsis thaliana* (Li et al., 2008)

## Plant NF-Y subunit B genes confer drought tolerance



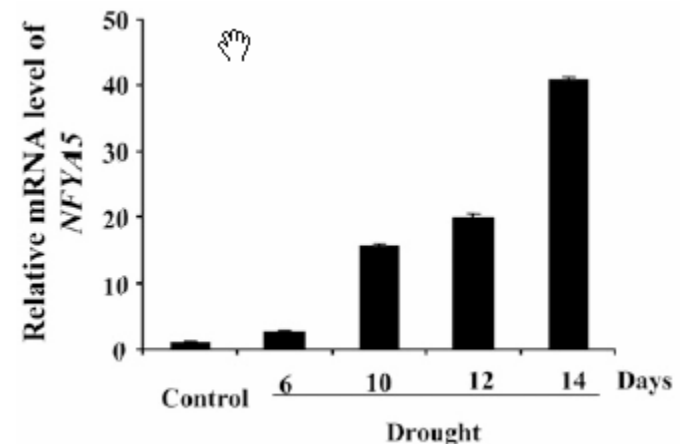
*ZmNF-YB2*. Nelson *et al.*, 2007



*TaNF-YB2*. Stephenson *et al.*, 2007



*AtNF-YB1*. Nelson *et al.*, 2007



*AtNF-YA5*. Li *et al.*, 2008



*Field performance of maize modified by ZmNF-YB2. Nelson et al., 2007*

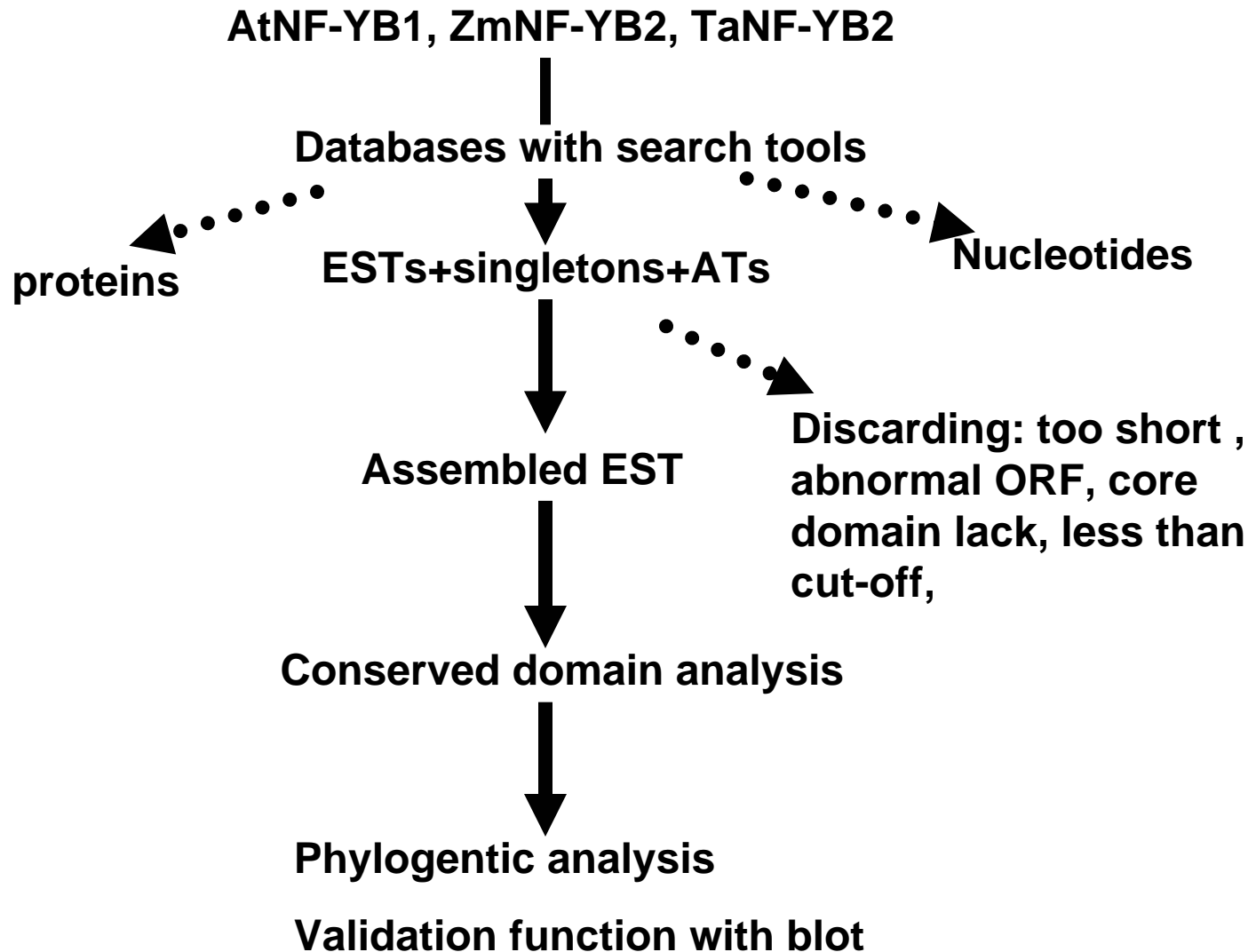
## Address two questions:

How about NF-Y genes,, especially subunit B genes, in drought condition?

More NF-Y genes in trees than in herbaceous plants?

Work on *Populus trichocarpa* and *P. euphratica*

# Procedure for genome analysis of NF-YB subunit in Poplar



# Databases and Tools

Used DataBase for ESTs, genes,  
transcriptors:

**GenBank: Protein,Nucleotide,EST**

<http://www.ncbi.nlm.nih.gov/Genbank/>

**PlantTFDB: Peking university**  
<http://planttfdb.cbi.pku.edu.cn/>

**PlantGDB: Iowa**  
<http://www.plantgdb.org/>

**TIGR: Plant Genomics Claig Venter**  
<http://www.tigr.org/tdb/e2k1/ath1/>

**JGI: Populus trichocarpa v1.1**  
[http://genome.jgi-psf.org/Poptr1\\_1/Poptr1\\_1.home.html](http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html)

**PopulusDB**  
<http://www.populus.db.umu.se/>

**Tools:**

tBLASTn, BLASTp, BLASTn

Clustalx

pfam

<http://www.sanger.ac.uk/Software/Pfam/>

Consensus Logo  
([http:// weblogo.berkeley.edu/](http://weblogo.berkeley.edu/))

Phylip

# Results

**Arabidopsis transcription factor subunit NF-YB1 in all populus ESTs**  
**Distribution of 104 Blast Hits on the Query Sequence**

**Arabidopsis transcription factor subunit NF-YB1 in trichocarpa ESTs**  
**Distribution of 30 Blast Hits on the Query Sequence**

**Arabidopsis transcription factor subunit NF-YB1 in euphratica ESTs**  
**Distribution of 10 Blast Hits on the Query Sequence**

**The same results with ZmNF-YB2, TaNF-YB2**



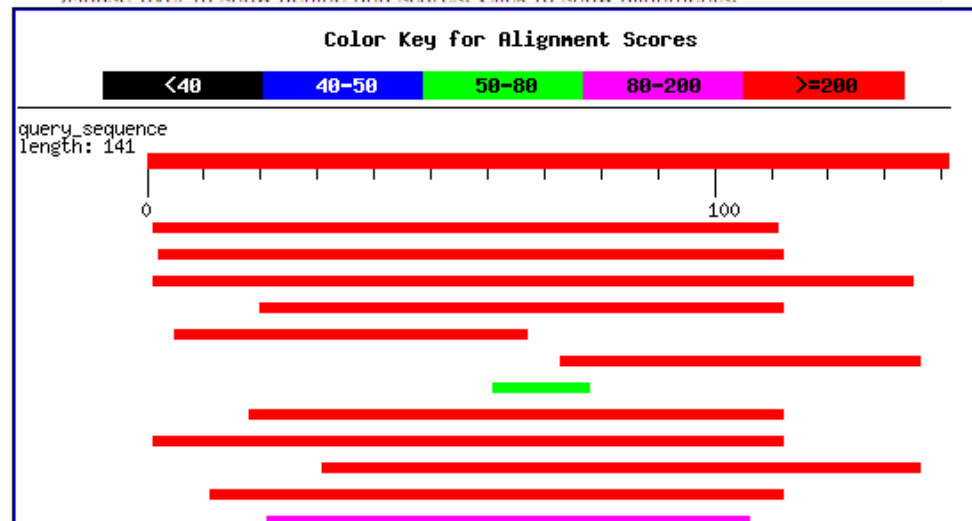
## TIGR Plant Transcript Assemblies



<a href="#">CV225747</a>	Populus trichocarpa Transcription factor NF-Y CC...	+1	<a href="#">495</a>	1.0e-47	1
<a href="#">CV243355</a>	Populus trichocarpa Nuclear transcription factor...	+3	<a href="#">472</a>	3.5e-45	1
<a href="#">TA17904_3694</a>	Populus trichocarpa Putative transcription f...	+1	<a href="#">444</a>	1.7e-42	1
<a href="#">CV243504</a>	Populus trichocarpa Nuclear transcription factor...	+2	<a href="#">445</a>	2.2e-42	1
<a href="#">TA19827_3694</a>	Populus trichocarpa Nuclear transcription fa...	+3	<a href="#">255</a>	4.3e-42	2
<a href="#">TA19882_3694</a>	Populus trichocarpa Putative transcription f...	+2	<a href="#">421</a>	7.3e-40	1
<a href="#">TA19543_3694</a>	Populus trichocarpa Nuclear transcription fa...	+3	<a href="#">416</a>	2.8e-39	1
<a href="#">BU871529</a>	Populus trichocarpa Transcription factor NF-Y CC...	+1	<a href="#">417</a>	3.0e-39	1
<a href="#">CV226095</a>	Populus trichocarpa Nuclear transcription factor...	+1	<a href="#">391</a>	1.1e-36	1
<a href="#">TA19400_3694</a>	Populus trichocarpa Repressor protein [Glyci...	+3	<a href="#">169</a>	7.5e-13	1
<a href="#">TA17806_3694</a>	Populus trichocarpa Hypothetical protein OJ1...	+2	<a href="#">127</a>	2.7e-08	1

### Distribution of Blast Hits on the Query Sequence

Mouse-over to show define and scores. Click to show alignments.







gw1.XIII.914.1	SFEAQMME	ATERISEV	GEASDQME	NRKIVGDDI	CWALISLGETD	DVAELIVRYL
gw1.133.33.1	SFEAQMME	ATERVSEV	GEASDQME	NRKIVGDDI	CWALISLGETD	DHAEMVRYL
gw1.XIII.909.1	SFEAQMME	ATERISEV	SEASNDQME	NRKIVGDDV	CWALISLGETD	DVADIVRYL
fgenes4_pg.C_scaffold_132000044	SFEAEMME	ASERISEV	GEASEDQME	KRKIVGDDV	CWANGALGETD	DVAGLIRYL
gw1.XIV.2649.1	SFEAEMME	CVSERISEV	GEASDQME	KRKIVGDDI	CWALASLGETD	DVSELRRL
gw1.XVI.590.1	SDDAMEIE	CVSERISEI	SEANERQRE	QRKIVAEIV	LWANSKLGTD	DVIELTIVL
gw1.VI.1909.1	SDDAMEIE	CVSEYISEI	SEANERQRE	QRKIVAEIV	LWANSKLGTD	DVIELTIVL
gw1.XVI.597.1	SDDAMEVE	CVSERISEI	GEASDQRE	KRKIVGDDL	LWAMITLGETE	DVVELMIVL
gw1.VI.2407.1	SDDAMEVE	CVSERISEI	GEASDQRE	KRKIVGDDL	LWAMITLGETE	DVAELMIVL
gw1.I.4755.1	SFEAMEVE	CVSERISEI	GEASDQRE	KRKIVGDDL	LWAMITLGETE	EVVELMIVL
gw1.XIV.3795.1	SDDAMEVE	CVSERISEI	GEASDQRE	KRKIVGDDL	LWAMITLGETE	EVVELMIVL
gw1.V.3062.1	SFEAMEVE	CVSERISEI	GEASDQRE	KRKIVGDDL	LWAMITLGETE	NVVGLMIVL
gw1.VII.2124.1	SFEAMEVE	CVSERISEI	GEASDQRE	KRKIVGDDL	LWAMITLGETE	NVVGLMIVL
grail3.0049010802	AMDAMDVE	CVSERISEV	SEASDQME	KRKIVGDDL	LWAMAYLGETE	DVIELMIVL
grail3.0022035801	AMDAMDVE	CVSERISEV	SEASDQME	KRKIVGDDL	LWAMAYLGETE	DVIDLMIVL
eugene3.00160779	AMDAMEVE	CVSERISEI	SEASDQRE	KRKIVGDDL	LWAMAYLGETE	DVIDLMIVL
gw1.V.2524.1	SDDAMEVE	CVSERISEV	GEASDQRE	KRKIVGDDI	IWAMITLGETE	DVVALMIVL
gw1.VII.1509.1	XDDAMEVE	CV---ISEV	GEASDQRE	KRXIVGDDI	IWAVITLGETE	DFIALQFXL
estExt_fgenes4_kg.C_LG_IX0001	HDDALLFE	SARITIHVLS	ATANDIQRES	KRQIMADIV	FKALEIETP	EVVGLIVSL
Consensus/60%	t++t+-slp-	slt-altals	t-tt-+sp+-	+++slst--l	latbsslta-	-al-sl+lal

# Result



gw1.XIII.914.1	RYREI--E-	-----	-----	-----	-----	-----
gw1.133.33.1	RYREA--E-	-----	-----	-----	-----	-----
gw1.XIII.909.1	RYREA--ER	-----	-----	-----	-----	-----
fgenes4_pg.C_scaffold_132000044	QRYREI--EG	DANQEKPAN	ANNNSIADQE	KEAPSSSSSS	YRNNQGIRM-	-----
gw1.XIV.2649.1	YRYREV--EG	ERAS-	-----	-----	-----	-----
gw1.XVI.590.1	RYREI--EG	EGSMRGEPL	MK-----	-----	-----	-----
gw1.VI.1909.1	RYREI--EG	ESSMRCEPL	VK-----	-----	-----	-----
gw1.XVI.597.1	QRYREI--EG	ETAAMGIVR	QGDQRDGTAG	DGGVVNSGNP	GGGFGGGGGG	NMYVGMQSSM
gw1.VI.2407.1	QRYREI--EG	ETAAMGT--	-----	-----	-----	-----
gw1.I.4755.1	QRYREI--EG	ET-	-----	-----	-----	-----
gw1.XIV.3795.1	QRYREI--EG	ET--SSMG--R	PAGEKDGPGG	GSGGSGAAGG	GGGANSTGGG	-----
gw1.V.3062.1	RYREI--EG	ENSMARQED	Q-----	-----	-----	-----
gw1.VII.2124.1	RYREI--EG	ENSMARQE-	-----	-----	-----	-----
grail3.0049010802	ARYREI--LWQG	DAKGSARGGD	GSSKRDVGG	LPGQNAQFAF	QGSMN----Y	TSPQVQGQHM
grail3.0022035801	ARYREI--LWQG	DAKGSARGGD	GSSKREAVGG	LPAQNAQFAL	QGSMN----Y	ISPQGQGQHM
eugene3.00160779	RYREI--EG	DIKGSAKTGD	TSAKKDIH--	-PGPNAQISH	QGSFSQGVSY	GNSNSQAPHM
gw1.V.2524.1	RYREI--EG	E-	-----	-----	-----	-----
gw1.VII.1509.1	RYREI--EG	ET-	-----	-----	-----	-----
estExt_fgenes4_kg.C_LG_IX0001	SEPRRN--GG	KIVGSAQNKE	VQKKRKIGGP	SKKSGGKSTP	KKGDIGEK--	-----
Consensus/60%	p+a+-b...-t	++..t.....	.....	.....	.....	.....

# Hits from *P. trichocarpa*

<a href="#">gb AC210129.1</a>	Populus trichocarpa clone POP009-F11, complete...	<a href="#">179</a>	9e-46
<a href="#">gb AC214028.1</a>	Populus trichocarpa clone POP025-N21, complete...	<a href="#">176</a>	1e-44
<a href="#">gb EF146061.1</a>	Populus trichocarpa clone WS0115_E10 unknown mRNA	<a href="#">58.9</a>	3e-09
<a href="#">gb CV240829.1</a>	WS02510.B21_H09 PT-MB-N-A-15 Populus trichocar...	<a href="#">225</a>	6e-60
<a href="#">gb CV225747.1</a>	WS0162.B21_E13 PT-DX-A-7 Populus trichocarpa c...	<a href="#">203</a>	3e-56
<a href="#">gb CV243355.1</a>	WS02518.B21_K05 PT-MB-N-A-15 Populus trichocar...	<a href="#">209</a>	5e-55
<a href="#">gb CV243504.1</a>	WS0252.B21_A22 PT-MB-N-A-15 Populus trichocarp...	<a href="#">196</a>	5e-51
<a href="#">gb BU871529.1</a>	Q031E12 Populus flower cDNA library Populus tr...	<a href="#">188</a>	1e-48
<a href="#">gb DT483220.1</a>	WS02521.B21_P21 PT-MB-N-A-15 Populus trichocar...	<a href="#">185</a>	8e-48
<a href="#">gb DT478153.1</a>	WS02521.BR_P21 PT-MB-N-A-15 Populus trichocarp...	<a href="#">185</a>	8e-48
<a href="#">gb DT478398.1</a>	WS02522.BR_K19 PT-MB-N-A-15 Populus trichocarp...	<a href="#">184</a>	1e-47
<a href="#">gb DT490508.1</a>	WS02545.B21_C13 PT-MB-N-A-15 Populus trichocar...	<a href="#">180</a>	3e-46
<a href="#">gb CV225383.1</a>	WS0161.B21_D19 PT-DX-A-7 Populus trichocarpa c...	<a href="#">179</a>	4e-46
<a href="#">gb DN491129.1</a>	UM63TC12.3pR Populus female catkins cDNA libra...	<a href="#">175</a>	1e-44
<a href="#">gb BU880488.1</a>	UM49TG09 Populus flower cDNA library Populus t...	<a href="#">175</a>	1e-44
<a href="#">gb DN501137.1</a>	UM63TC12.5pR Populus female catkins cDNA libra...	<a href="#">174</a>	2e-44
<a href="#">gb BU881483.1</a>	UM63TC12 Populus flower cDNA library Populus t...	<a href="#">171</a>	2e-43
<a href="#">gb CV226095.1</a>	WS0163.B21_E08 PT-DX-A-7 Populus trichocarpa c...	<a href="#">165</a>	1e-41
<a href="#">gb DT490620.1</a>	WS02545.B21_H11 PT-MB-N-A-15 Populus trichocar...	<a href="#">97.4</a>	2e-25
<a href="#">gb DT476620.1</a>	WS01230.B21_D17 PT-GT-FL-A-3 Populus trichocar...	<a href="#">73.2</a>	5e-14
<a href="#">gb DT473941.1</a>	WS01230.BR_D17 PT-GT-FL-A-3 Populus trichocarp...	<a href="#">73.2</a>	5e-14
<a href="#">gb CV245214.1</a>	WS0256.B21_M17 PT-MB-N-A-15 Populus trichocarp...	<a href="#">73.2</a>	5e-14
<a href="#">gb DT498409.1</a>	WS0115.BR_E10 PT-P-FL-A-2 Populus trichocarpa ...	<a href="#">58.9</a>	1e-09
<a href="#">gb DT495778.1</a>	WS01121.BR_H08 PT-P-FL-A-2 Populus trichocarpa...	<a href="#">58.9</a>	1e-09
<a href="#">gb DT483467.1</a>	WS02522.B21_K19 PT-MB-N-A-15 Populus trichocar...	<a href="#">55.1</a>	2e-08
<a href="#">gb CV250926.1</a>	WS0115.B21_E10 PT-P-FL-A-2 Populus trichocarpa...	<a href="#">55.1</a>	2e-08
<a href="#">gb CV229161.1</a>	WS01912.B21_G15 PT-DX-N-A-10 Populus trichocar...	<a href="#">54.7</a>	2e-08
<a href="#">gb CV243110.1</a>	WS02516.B21.1_O21 PT-MB-N-A-15 Populus trichoc...	<a href="#">52.8</a>	8e-08
<a href="#">gb CV248643.1</a>	WS01121.B21_H08 PT-P-FL-A-2 Populus trichocarp...	<a href="#">51.6</a>	1e-07

# Some translated protein sequences in trichocarpa's CDS

```

>BUB71529|cds13..417|trichocarpa
MKKALPANGKIAKDAKDTUQECUSEFISFUTSEASDKCQKEKRKTINGDDLLWAMATLGFEDYIEPLKUYLARYRELEGDAGKSARGGGDSSKRDAUGGLPGQNAQFAFQGSMMYTSPPAQGGHM
ILPSMPGNE
>CU225747|cds16..390|trichocarpa
MADNPTSPAAGSHESGGEQSPRSGVREQDRYLP IANISRIMKKALPANGKIAKDAKDTUQECUSEFISFUTSEASDKCQKEKRKTINGDDLLWAMATLGFEDYIEPLKUYLARYREVISLFSCL
>CU226095|cds79..630|trichocarpa
MKKSLPANAKISKEAKETUQECUSEFISFITGEASDKCQKEKRKTINGDDLLWAMSTLGFENYUGSLKUYLNKYRDETEGENSHARQEDQLSPTNHGSLGTHETIKUNGSPSLSTGMDLQSFNAG
LYSLGTEUTAKSYQNTFRFLGYQESSMUGDFNMNKRIRDHGDRNGNIAMAGQLHHGUGW
>CU243355|cds93..623|trichocarpa
MAAEAPASPGGSHESGDQSPRSNSNUREQDRFLPIANISRIMKKALPANGKIAKDAKETUQECUSEFISFITSEASDKCQKEKRKTINGDDLLWAMATLGFEDYIDPLKIYLSRYREMEGDTKG
SAKAGDTSAKKDUIHPGPNQKWSFMLSYCHFLFLUNQRTSLSLFYSANF
>CU243504|cds47..451|trichocarpa
MKKALPANGKIAKDAKETUQECUSEFISFITSEASDKCQKEKRKTINGDDLLWAMATLGFEDYIDPLKIYLSRYREMEGDTKGFAKTGDTFAKKDIPPGPNAQIFHQGFFSQGUSYGNNSNFQAPP
MMUPMQSNE
>TA17806_3694|cds32..469|trichocarpa
METEKUVAEAEELPKAIVRRUUKDTL SRCSPEDSDINUHKDALLAFSESARIFIHYLSATANDICKESKRQTMNADDVFKALEDIEFPEFUGPLKUSLSEFKRKNGGKKUGSAQNKEUQKKRKTG
GPSKSGGKSTPKKGDTEK
>TA17904_3694|cds139..711|trichocarpa
MADSDNESGGHNAUSELSAKEQDRFLPIANUSRIMKKALPANAKISKDAKETUQECUSEFISFITGEASDKCQKEKRKTINGDDLLWAMTTLGFEEYUEPLKIYLYQYREMEGEKSSMGRPAGEK
DGPGGGSGGSAAGGGGGANSTGGGGGFGGGGGQMYGGMMMHGHHQHUVYGGAGGFHHQGNRS
>TA19400_3694|cds156..629|trichocarpa
MEPMDIUGKSKEDASLPKATHTKIIEKMLPPDURVARDTQDLLIECCUEFINLUSSESNEUCSREDKRTI APEHULKALQULGFGEYIEDUYTAYEQHKLETMHDSLKGGGKSTGAAMTEEEAA
AAQQRMLDEARARHNGUUSAPKQPEANQSLES
>TA19543_3694|cds18-569|trichocarpa
MADSDNESGGQNNSTNYSETSSREQDKLLPIANUSRIMKKALPANAKISKDGKETUQECUSEFISFITGEASDKCQKEKRKTINGDDLLWAMTTLGFEDYAEPLKIYLYQKRFRETEGERTAAMGT
URQGDQRDGAUVNSGDSGGGFGGGGGNUYGGMQSSMMMHGHHPLHQGNMYGSASGRPK
>TA19827_3694|cds114..221|trichocarpa
MKKALPANGKIAKDAKDTUQECUSEFISFUTSEGE
>TA19882_3694|cds8..601|trichocarpa
MADSDNESGEQNNSTNYSTETSPREQDRLLPIANUSRIMKKALPANAKISKDAKETUQECUSEFISFITGEASDKCQKEKRKTINGDDLLWAMTTLGFEDYUEPLKIYLYQKRFREMEGEKTAAMG
IURQGDQRDGTAGDGGVUNI GNPPGGFGGGGGGNMYUGMQSSMAMMHGHHYQHRHPGNMYGPGGGGASSGRPS
>eugene3.00160779
MAAEAPASPGGSHESGDQSPRSNSNUREQDRFLPIANISRIMKKALPANGKIAKDAKETUQECUSEFIS
FITSEASDKCQKEKRKTINGDDLLWAMATLGFEDYIDPLKIYLSRYREMEGDTKGSAKTGDTSAKKD IHP
GPNAQISHQGSFSQGUSYGNNSQAPHMMUPMQSNE*
>grail3.0049070802
MADNPTSPAAGSHESGGEQSPRSGVREQDRYLP IANISRIMKKALPANGKIAKDAKDTUQECUSEFISFU
TSEASDKCQKEKRKTINGDDLLWAMATLGFEDYIEPLKUYLARYREQLVQGDAGKSARGGGDSSKRDAUG
GLPGQNAQFAFQGSMMYTSPPAQGGHMILPSMPGNE*
>grail3.0022035801
MADNPTSPAAGSHESGGEQSPRSGVREQDRYLP IANISRIMKKALPANGKIAKDAKDTUQECUSEFISFU
TSEASDKCQKEKRKTINGDDLLWAMATLGFEDYIDPLKUYLARYREQLVQGDAGKSARGGGDSSKREAVG
GLPAQNAQFALQGSMMYISPQGGQHMILPSMQNE*

```

## Translated protein sequences from 3 euphratica's ESTs

```
>AJ775556|cds166-339|euphratica  
MLPPDURVARDAQDLLIECCAEFINLUSSESNEPCSREDKRTIAPEHULKALQALGFG  
>AJ768375|cds379..594|euphratica  
MADSDNDSGGQNPTSTNEULFPREMDRFLPUANUSRIMKKALPANAKISKDAKETUQECUSEFISFITGEAS  
>TA3229_75702|cds172..753|euphratica  
MADSDNESGGHNAUSELSAKEQDRFLPIANUSRIMKKALPANAKISKDAKETUQECUSEFISFITGEASDKCQREKRKTING  
DDLLWAMTTLGFEEYVEPLKIYLQKYREMEGEKSSMGRPAGEKDGPGGGSGGSGAAGGGGGANSTGGGGGGFNGWVGARDUM  
RDDDDGSSPSRTRURWLUGFTTKUIQGRY
```

# Populus trichocarpa v1.1

Found 18 results					
<input type="button" value="Check All"/>		Displaying results 1 to 18		<input type="button" value="Uncheck All"/>	
	<u>Relevance</u>	<u>Type</u>	<u>Assigned Name</u>	<u>Detail</u>	<u>Links</u>
<input type="checkbox"/>	1.0	transcript	GW1.XVI.590.1	CCAAT-binding factor, subunit A (HAP3)	P T G
<input type="checkbox"/>	1.0	transcript	GW1.VII.1509.1	CCAAT-binding factor, subunit A (HAP3)	P T G
<input type="checkbox"/>	0.8	transcript	GW1.XIV.2649.1	CCAAT-binding factor, subunit A (HAP3)	P T G
<input type="checkbox"/>	0.8	transcript	GW1.I.4755.1	CCAAT-binding factor, subunit A (HAP3)	P T G
<input type="checkbox"/>	0.8	transcript	GW1.VI.1909.1	CCAAT-binding factor, subunit A (HAP3)	P T G
<input type="checkbox"/>	0.8	transcript	GW1.VII.2124.1	CCAAT-binding factor, subunit A (HAP3)	P T G
<input type="checkbox"/>	0.8	transcript	GW1.XVI.597.1	CCAAT-binding factor, subunit A (HAP3)	P T G
<input type="checkbox"/>	0.8	transcript	GRAIL3.0022035801	CCAAT-binding factor, subunit A (HAP3)	P T G
<input type="checkbox"/>	0.8	transcript	GW1.V.3062.1	CCAAT-binding factor, subunit A (HAP3)	P T G

# Each peptide sequence from assembled EST against pfam for confirming core domain:



[HOME](#) | [SEARCH](#) | [BROWSE](#) | [FTP](#) | [HELP](#)



## Sequence search results

We found **1** Pfam-A match to your search sequence. You did not choose to search for Pfam-B matches.

The Pfam graphic below shows the arrangement of the domains on your search sequence. Clicking on any of the domains will take you to a page of information about that domain.



Below showing the details of the domains that were found. Rows containing significant hits are **highlighted**. Hits which do not start and end at the end points of the matching HMM are also **highlighted**.

For Pfam-A hits we show the alignments between your search sequence and the matching HMM. You can show individual alignments by clicking on the "Show" button in each row of the result table, or you can show all alignments using the links above the table.

You can bookmark this page and return to it later, but please note that old results will be removed after **one week**.

[Return](#) to the search form to look for Pfam domains on a new sequence.

## Pfam-A Matches

[Show](#) or [hide](#) all alignments.

Pfam-A	Description	Entry type	Sequence		HMM		Bits score	E-value	Alignment mode	Show/hide alignment
			Start	End	From	To				
<a href="#">CBFD_NFYB_HMF</a>	Histone-like transcription factor (CBF/NFY) and archaeal histone	Domain	24	89	1	66	126.8	6.2e-35	ls	<a href="#">Hide</a>
#HMM	+-->aeLPiArvkrIMKsdLpdagkiskdAkeliasecveePiefiaseAaeickkekRRTikaehiklAv<--									
#MATCH	+ LPiA++ rIMK++Lp++gki kdAk++++ecv+eFi+fi+seA+++c+kekRRT++++++A+									
#SEQ	RYLPiANISrIMKKALPPNGKIGKDARDTVQECVSEPISEFITSEASDRKQREKRRRTVNGDDLLWAM 89									

>AJ775556|cds166-339|euphratica

MLPPDVRVARDAQDLLIECCAEFINLVSSESNEPCSREDKRTIAPEHVLKA  
LQALGFG

## Sequence search results

We found **1** Pfam-A match to your search sequence. You did not choose to search for Pfam-B matches.

The Pfam graphic below shows the arrangement of the domains on your search sequence. Clicking on any of the domains will take you to a page of information about that domain.



Below showing the details of the domains that were found. Rows containing significant hits are **highlighted**. Hits which do not start and end at the end points of the matching HMM are also **highlighted**.

For Pfam-A hits we show the alignments between your search sequence and the matching HMM. You can show individual alignments by clicking on the "Show" button in each row of the result table, or you can show all alignments using the links above the table.

You can bookmark this page and return to it later, but please note that old results will be removed after **one week**.

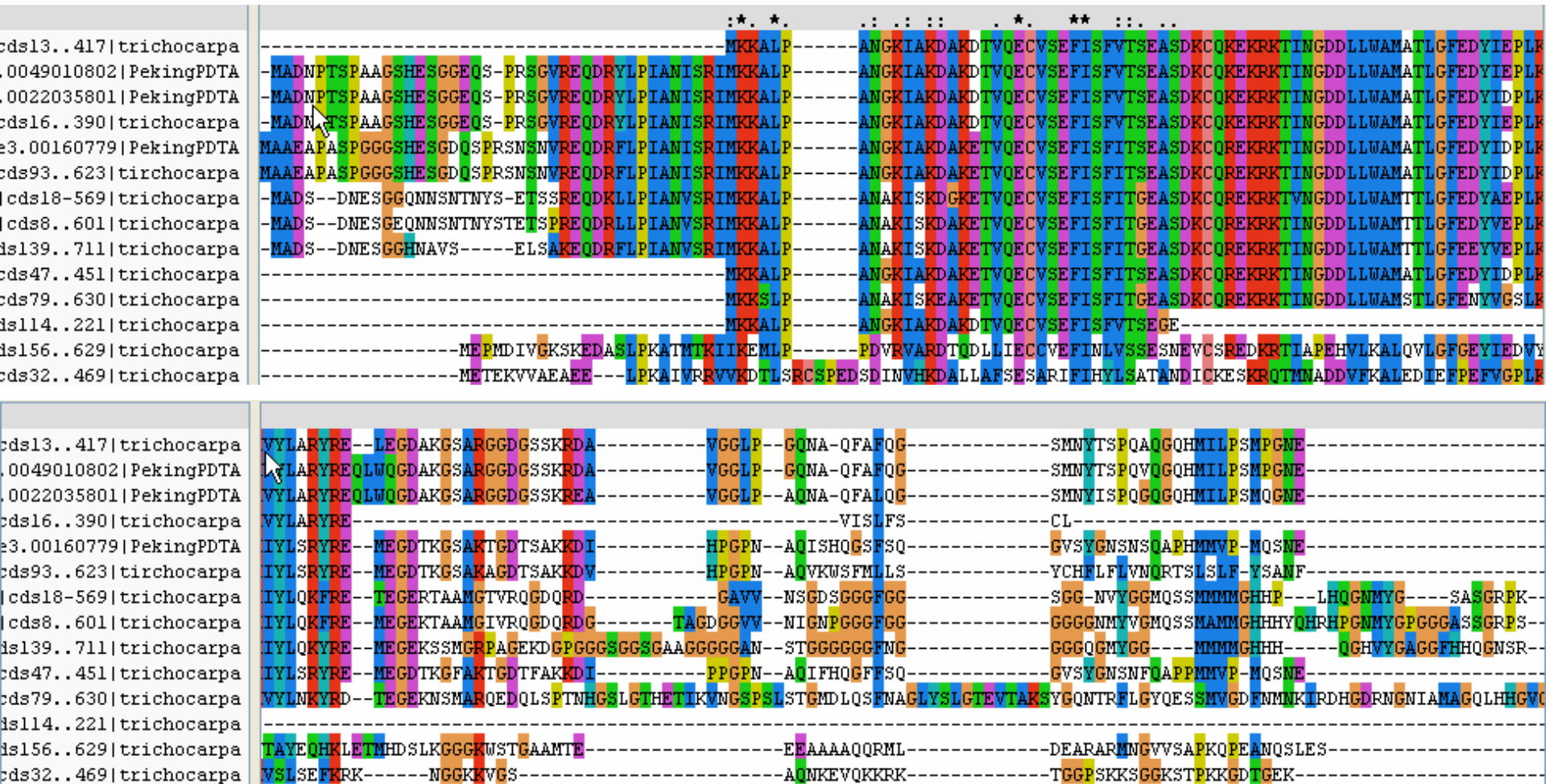
[Return](#) to the search form to look for Pfam domains on a new sequence.

## Pfam-A Matches

[Show](#) or [hide](#) all alignments.

Pfam-A	Description	Entry type	Sequence		HMM		Bits score	E-value	Alignment mode	Show/hide alignment
			Start	End	From	To				
<a href="#">CBFD_NFYB_HME</a>	Histone-like transcription factor (CBF/NFY) and archaeal histone	Domain	1	52	1	66	48.6	2.2e-11	ls	<a href="#">Show</a>

## Clustalx alignment in *P.trichocarpa* homologues to NF-YB1



# Alignment after appending *AtNF-YB1*, *ZmNF-YB2* and *TaNF-YB2*

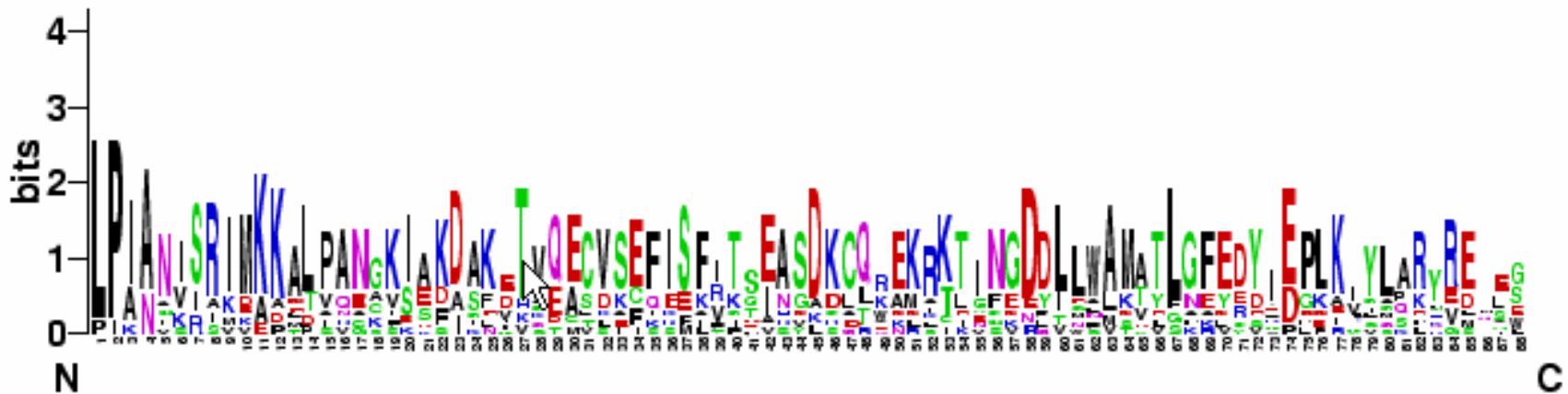
		.* ..	.: ::	. *	**	:: ..
cds18-569 trichocarpa	ETSSREQDKLLPIANVSRIMKKALP	-----	ANAKISKDGKETVQECVSEFISFITGEASDKCQREKRKTVNGDDLWAMITLGFEDYAEPLKIYLLQKFR	---	TEGERTAAMGTV	
cds8..601 trichocarpa	ETSPREQDRLLPIANVSRIMKKALP	-----	ANAKISKDAKETVQECVSEFISFITGEASDKCQREKRKTINGDDLWAMITLGFEDYVEPLKIYLLQKFR	---	MEGEKTAAMGIW	
sls139..711 trichocarpa	ELSAKEQDRFLPIANVSRIMKKALP	-----	ANAKISKDAKETVQECVSEFISFITGEASDKCQREKRKTINGDDLWAMITLGFEEYVEPLKIYLLQKYR	---	MEGEKSSMGRPA	
0049010802 PekingPDTA	RSGVREQDRYLPANISRIMKKALP	-----	ANGKIAKDAKDIVQECVSEFISFVTSEASDKCQKEKRKTINGDDLWAMATLGFEDYIEPLKVYLARYR	---	QLWQGDAGKSARGG	
0022035801 PekingPDTA	RSGVREQDRYLPANISRIMKKALP	-----	ANGKIAKDAKDIVQECVSEFISFVTSEASDKCQKEKRKTINGDDLWAMATLGFEDYIDPLKVYLARYR	---	QLWQGDAGKSARGG	
cds16..390 trichocarpa	RSGVREQDRYLPANISRIMKKALP	-----	ANGKIAKDAKDIVQECVSEFISFVTSEASDKCQKEKRKTINGDDLWAMATLGFEDYIEPLKVYLARYR	---	-----	
TaNF-YB2 aa	FGGVREQDRFLPIANISRIMKKAIPAN	-----	GKIAKDAKETVQECVSEFISFITSEASDKCQREKRKTINGDDLWAMATLGFEEYIEPLKVYLQKYR	---	TEGDSKLAGKSC	
ZmFN-YB2 aa	-GSVREQDRFLPIANISRIMKKAIPANGKTIIPANGKIAKDAKETVQECVSEFISFITSEASDKCQREKRKTINGDDLWAMATLGFEDYIEPLKVYLQKYR	---	-----	---	MEGDSKLTAKSS	
gi 18404885 AtNF-YB1	-GSVREQDRYLPANISRIMKKALPPN	-----	GKIGKDAKDIVQECVSEFISFITSEASDKCQKEKRKTVNGDDLWAMATLGFEDYLEPLKIYLLARYR	---	LEGDNKSGKNSC	
e3.00160779 PekingPDTA	MSNVREQDRFLPIANISRIMKKALP	-----	ANGKIAKDAKETVQECVSEFISFITSEASDKCQREKRKTINGDDLWAMATLGFEDYIDPLKIYLLSRYR	---	MEGDTKGSANTG	
cds47..451 trichocarpa	-----MKKALP	-----	ANGKIAKDAKETVQECVSEFISFITSEASDKCQREKRKTINGDDLWAMATLGFEDYIDPLKIYLLSRYR	---	MEGDTKGFANTG	
cds93..623 trichocarpa	MSNVREQDRFLPIANISRIMKKALP	-----	ANGKIAKDAKETVQECVSEFISFITSEASDKCQREKRKTINGDDLWAMATLGFEDYIDPLKIYLLSRYR	---	MEGDTKGSAKAG	
cds13..417 trichocarpa	-----MKKALP	-----	ANGKIAKDAKDIVQECVSEFISFVTSEASDKCQKEKRKTINGDDLWAMATLGFEDYIEPLKVYLARYR	---	LEGDAKGSARGG	
sls114..221 trichocarpa	-----MKKALP	-----	ANGKIAKDAKDIVQECVSEFISFVTSEGE	-----	-----	
cds79..630 trichocarpa	-----MKSLLP	-----	ANAKISKEAKETVQECVSEFISFITGEASDKCQREKRKTINGDDLWAMSTLGFENYVGSLEKVVYLNKYRD	---	TEGEKNSMARQE	
sls156..629 trichocarpa	IVGKSKEDASLPKATMTIKIIEMLP	-----	PDVVRWARDTDLLIECCVEFINLVSSESNEVCSSREDKRTIAPHEVLKALQVLGFGEYIEDVYTAYE	---	QHKLETMHDSLKGGGKWS	
cds32..469 trichocarpa	VVAEAEELPKAIWRVVKDILSRCS	-----	PEDSDINVHKDALLAFSESARIFIHYLSATANDICKESKRQTMNADDVFKALIEDIEFPEFVGPLKVSLSLSEFTRK	-----	NGGKIVC	

# Alignment after appending euphratica homologues

cds18-569 trichocarpa	--ETSSREQDKLLPIANVSRIMKKALP-----ANAKISKDGKETVQECVSEFISFITGEASDKCQREKRKTVNGDDLLWAMTTLGFEDYAEPLKIYLQKFRE--TEGERTAAMG
cds8..601 trichocarpa	--ETSPREQDRLLPIANVSRIMKKALP-----ANAKISKDAKETVQECVSEFISFITGEASDKCQREKRKTINGDDLLWAMTTLGFEDYVEPLKIYLQKFRE--MEGEKTAAMG
cds172..753 euphratica	--ELSAKEQDRFLPIANVSRIMKKALP-----ANAKISKDAKETVQECVSEFISFITGEASDKCQREKRKTINGDDLLWAMTTLGFEEYVEPLKIYLQKYRE--MEGEKSSMGR
ds139..711 trichocarpa	--ELSAKEQDRFLPIANVSRIMKKALP-----ANAKISKDAKETVQECVSEFISFITGEASDKCQREKRKTINGDDLLWAMTTLGFEEYVEPLKIYLQKYRE--MEGEKSSMGR
3.00160779 PekingPDTA	RSNSNVREQDRFLPIANISRIMKKALP-----ANGKIADAKETVQECVSEFISFITSEASDKCQREKRKTINGDDLLWAMATLGFEDYIDPLKIYLSRYRE--MEGDTKGSAR
cds93..623 tirchocarpa	RSNSNVREQDRFLPIANISRIMKKALP-----ANGKIADAKETVQECVSEFISFITSEASDKCQREKRKTINGDDLLWAMATLGFEDYIDPLKIYLSRYRE--MEGDTKGSAR
.0049010802 PekingPDTA	-PRSGVREQDRYLPANISRIMKKALP-----ANGKIADAKDITVQECVSEFISFVTSEASDKCQKEKRKTINGDDLLWAMATLGFEDYIEPLKVYLARYREQLWQDAKGSAR
.0022035801 PekingPDTA	-PRSGVREQDRYLPANISRIMKKALP-----ANGKIADAKDITVQECVSEFISFVTSEASDKCQKEKRKTINGDDLLWAMATLGFEDYIDPLKVYLARYREQLWQDAKGSAR
cds16..390 trichocarpa	-PRSGVREQDRYLPANISRIMKKALP-----ANGKIADAKDITVQECVSEFISFVTSEASDKCQKEKRKTINGDDLLWAMATLGFEDYIEPLKVYLARYRE-----
TaNf-YB2 aa	GGFGGVREQDRFLPIANISRIMKKAIPAN-----GKIADAKETVQECVSEFISFITSEASDKCQREKRKTINGDDLLWAMATLGFEEYIEPLKVYLQKYRE--TEGDSKLAGK
gi 18404885 AtNF-YB1	---GSVREQDRYLPANISRIMKKALPPN-----GKIGKDARDITVQECVSEFISFITSEASDKCQKEKRKTVNGDDLLWAMATLGFEDYLEPLKIYLARYRE--LEGDMKGSCK
ZmFN-YB2 aa	5---GSVREQDRFLPIANISRIMKKAIPANGKTIIPANGKIADAKETVQECVSEFISFITSEASDKCQREKRKTINGDDLLWAMATLGFEDYIEPLKVYLQKYRE--MEGDSKLTAK
cds379..594 euphratica	--VLFPREMDRFLPVANVSRIMKKALP-----ANAKISKDAKETVQECVSEFISFITGEAS-----
cds79..630 trichocarpa	-----MKSLLP-----ANAKISKEAKETVQECVSEFISFITGEASDKCQREKRKTINGDDLLWAMSTLGFENYVGSLLKVYLNKYRD--TEGEKNSMAR
cds47..451 trichocarpa	-----MKKALP-----ANGKIADAKETVQECVSEFISFITSEASDKCQREKRKTINGDDLLWAMATLGFEDYIDPLKIYLSRYRE--MEGDTKGFAR
cds13..417 trichocarpa	-----MKKALP-----ANGKIADAKDITVQECVSEFISFVTSEASDKCQKEKRKTINGDDLLWAMATLGFEDYIEPLKVYLARYRE--LEGDAKGSAR
ds114..221 trichocarpa	-----MKKALP-----ANGKIADAKDITVQECVSEFISFVTSEGE-----
ds156..629 trichocarpa	--IVGKSKEDASLPKATWTKIIEKMLP-----PDVVRVARDITDILLIECCVEFINLVSSSENEVCSREDKRTIAPHEVLKALQVLGFGYIEDWYTAAYE QHKLEIMHDSLKGGGK
cds166-339 euphratica	-----MLP-----PDVVRVARDADILLIECCAEFINLVSSSENEPCSREDKRTIAPHEVLKALQALGFG-----
cds32..469 trichocarpa	--VVAEAEF---LPKAIWRRVVKDITLRSCL-PEDSDINVHKDALLAFSE SARIFIHYSATANDICKESKROTMMADDVFKALEIDIEFPEFVGPLKVSLSSEFKRK-----NGGKK



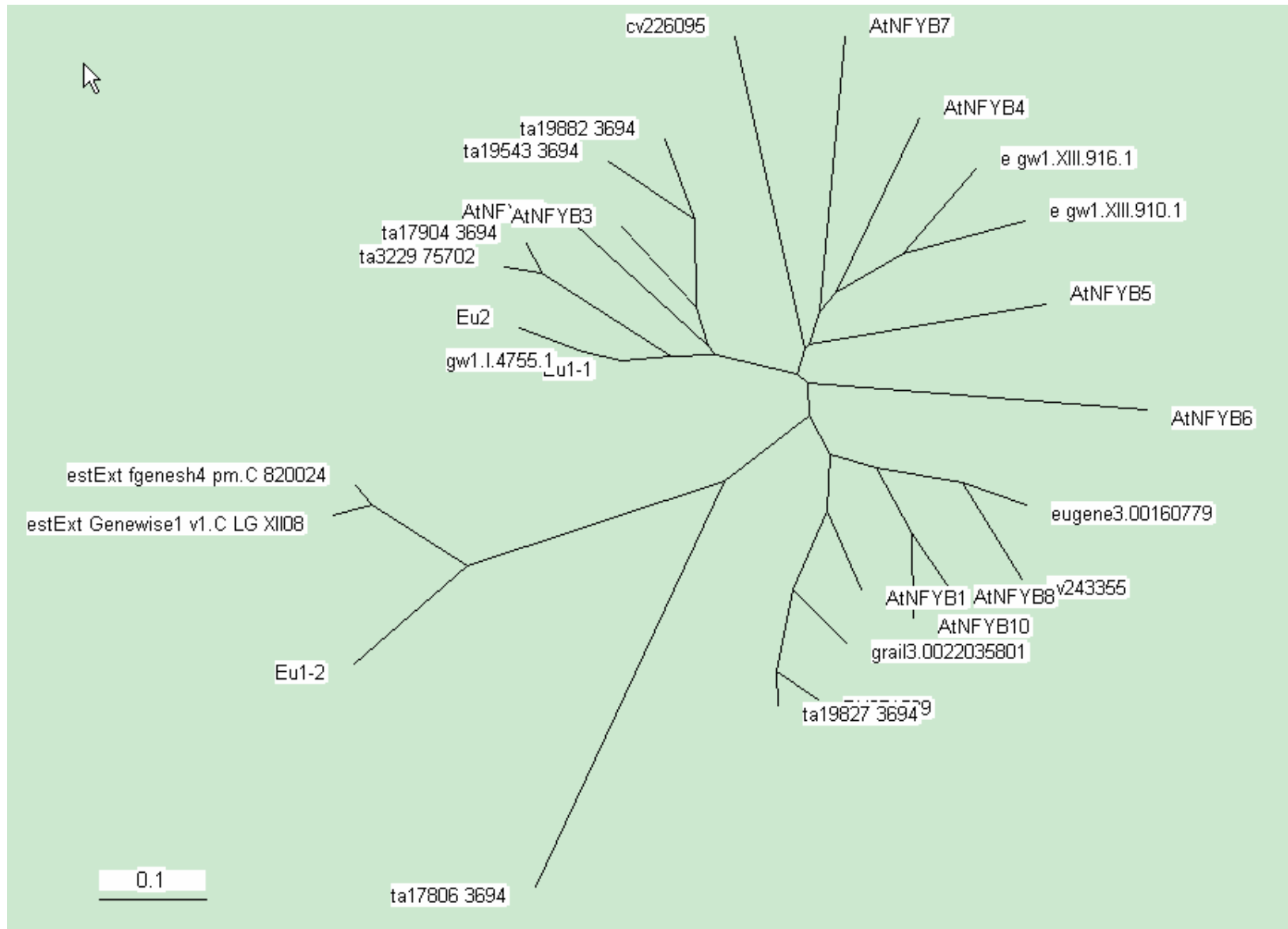
# Analysis on NF-YB motif in identified sequences in poplar

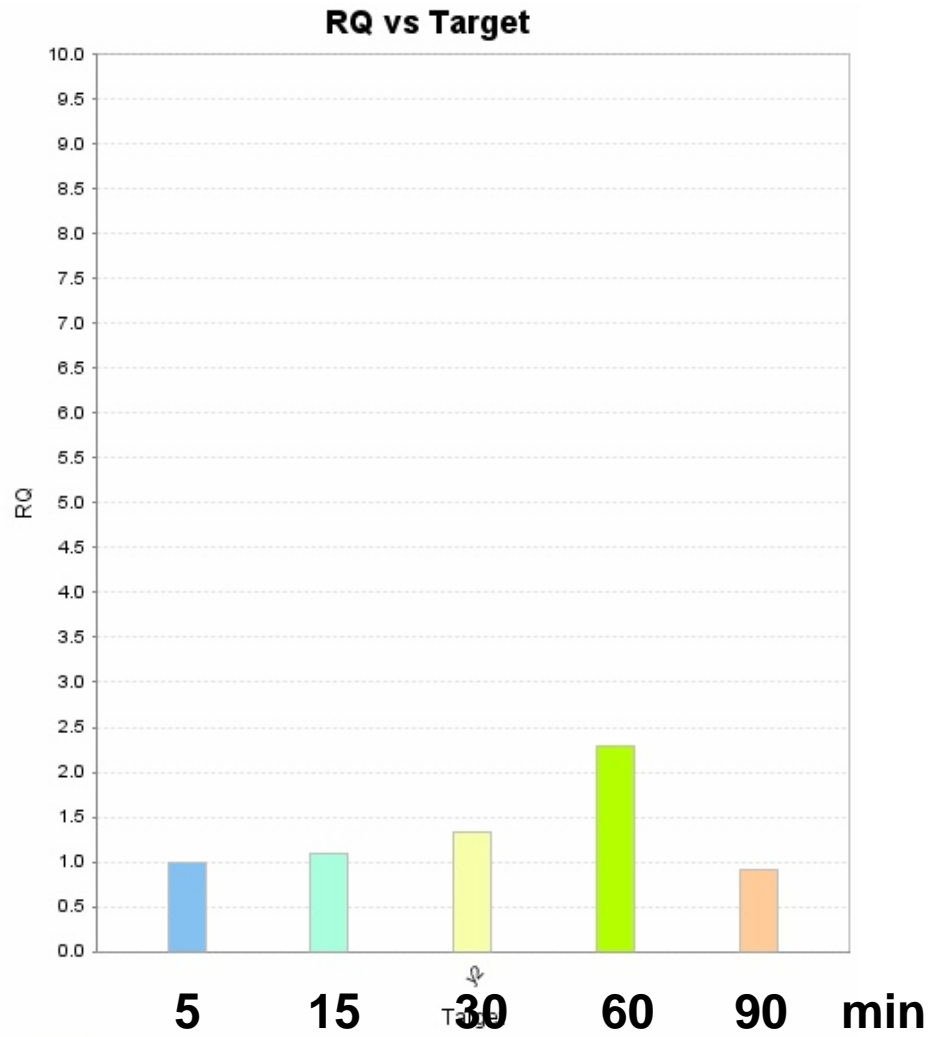


LPiANISrIMKKALpANGKIAkdAKETVQecvSeFiSfITseASDKcQrekRKTINGD DLLWAMATLGFEDYIEPLKIYIARYRE

The length of the NF-YB subunit core region is 90 residues in an average. (Maity and de Crombrughe, 1992). The NF-YB subunits in poplar have 87 amino acids in length.

# Phylogenetic analysis





*P. euphratica*

PEG 6000 20%

Sequence:  
CV226095

## conclusion

We retrieved a total 18 ESTs and transcripts in *P. trichocarpa* (15) and *P. euphratica* (3) molecular sequence databases by computational analysis. And also confirmed that identified sequences were NF-YB homologues

Poplar tree has more NF-YB subunit family genes than herbaceous plant ( Arabidopsis 8, wheat 11)

There exist NF-YB genes involved in response to drought stress for poplar trees. We found a NF-YB gene (naming it afterward) with EST sequence ID, CV226095, associate with *P. euphratica* response to stress by PEG so far.

Thanks!