

犏牛DMTR7基因的生物 信息学分析

G12小组：赵 翔 马永华
 宋兴超 吴晓云

研究背景

牦牛是分布在青藏高原及其毗邻高山地区的重要畜种，是青藏高原居民赖以生存的生活资料。



野牦牛



家牦牛

犏牛是牦牛和黄牛杂交产生的后代，其具有很强的杂种优势，但是F1代公犏牛因不能正常产生精子，导致无法将杂种优势固定下来，这成为牦牛杂交改良的一大难题，犏牛雄性不育的分子机理至今仍不明确。



公犏牛

组织形态学观察发现犏牛初级精母细胞很少，无次级精母细胞，附睾内未观测到精子，而黄牛、牦牛睾丸组织中可见各级生精细胞，附睾内可见发育良好的精子，可见F1代犏牛雄性不育主要是由精子发生过程中减数分裂障碍所引起的。

雄性哺乳动物XY染色体的联会配对依赖于一种特殊的染色体结构XY body而完成。XY body在雄性生殖细胞减数分裂过程中发挥着重要作用。DMRT7蛋白选择性的集中于粗线期精母细胞XY body上。

DMRT7基因缺陷小鼠生精细胞缺乏，生精过程被阻滞在减数分裂粗线期，表现出性染色体联会调控缺陷，小鼠表现出雄性不育，雌性正常，说明**DMRT7**基因在减数分裂粗线期和双线期之间雄性特有的性染色体联会过程中发挥重要作用。

F1代犏牛雄性生殖细胞减数分裂停滞在粗线期阶段，表现出的染色体联会障碍表型与DMRT7基因缺陷小鼠表型非常相似。

假设：

1、犏牛DMRT7蛋白质中是否存在影响其功能的变异位点。

2、犏牛DMRT7基因表达量的变化是否影响其功能。

密码子偏好性

#Codon	AA	Fraction	Frequency	Number
GCA	A	0.121	10.782	4
GCC	A	0.364	32.345	12
GCG	A	0.030	2.695	1
GCI	A	0.485	43.127	16
IGC	C	0.615	21.563	8
IGI	C	0.385	13.477	5
GAC	D	0.800	10.782	4
GAI	D	0.200	2.695	1
GAA	E	0.429	16.173	6
GAG	E	0.571	21.563	8
IIC	F	0.500	5.391	2
III	F	0.500	5.391	2
GGA	G	0.375	32.345	12
GGC	G	0.281	24.259	9
GGG	G	0.250	21.563	8

牦牛

#Codon	AA	Fraction	Frequency	Number
GCA	A	0.121	10.782	4
GCC	A	0.364	32.345	12
GCG	A	0.030	2.695	1
GCI	A	0.485	43.127	16
IGC	C	0.615	21.563	8
IGI	C	0.385	13.477	5
GAC	D	0.800	10.782	4
GAI	D	0.200	2.695	1
GAA	E	0.429	16.173	6
GAG	E	0.571	21.563	8
IIC	F	0.500	5.391	2
III	F	0.500	5.391	2
GGA	G	0.375	32.345	12
GGC	G	0.281	24.259	9
GGG	G	0.250	21.563	8

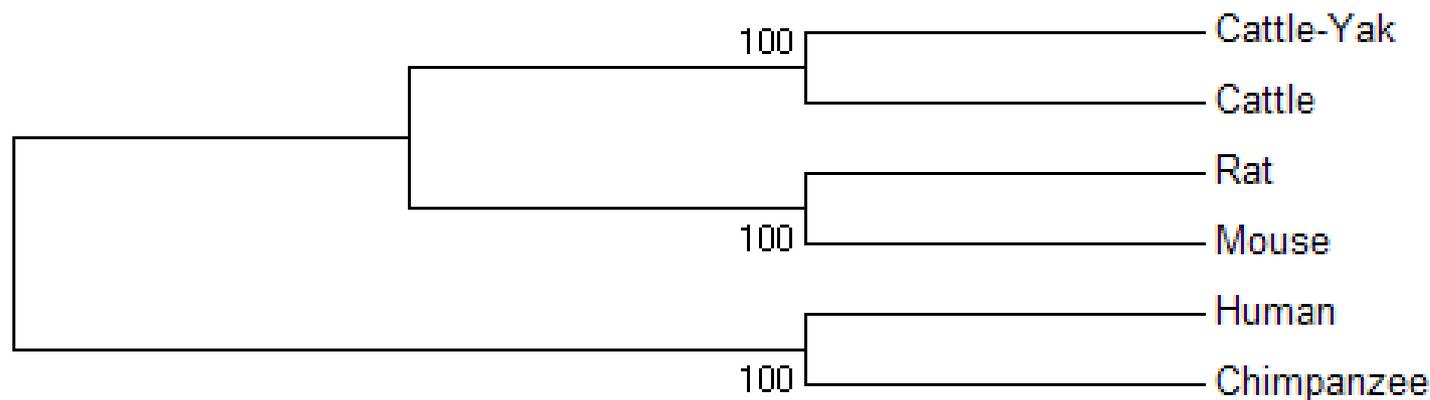
犏牛

用GENSCAN将犏牛DMRT7 mRNA序列翻译成氨基酸序列

```
>DMRT7                               Cattle-yak                               370_aa
MDPSEMPAVHHCPSSDSATGGETRAPQGMELIPRRAVS
RSPTCARCRNHGVT AHLKGHKRLCLFQACECHKCVLIL
ERRRVMAAQVALRRQQEAQLKRHLTQGLIRGAAPPRAP
SRVKKGVTRPGVHSGKENIAPQPQIPHVVPLALTPPGK
ENSRGPLLSRPPEALPLPWTPMPPGPWASGHWLPPG
LSMPTPVVCRLLCQEP AIPLHPFPGFDPGTTLRLPTHGP
LPTCTGSHPILTAPLSGESQGPSTLPRTCSTLILQPCGTP
DPLLQAPGPSRLTWTSSASSEWQLQREAAEALVGLK
DSSQAPRLTPGPANPAWISLLHPCGPAAAGGRGFQPV
GPSLRPSPAPSVALHI GRLGSISLLS
```

同源性比较

犏牛DMRT7蛋白与黄牛、大鼠、小鼠、人和猩猩DMRT7蛋白的同源性分别为：99%、84%、83%、80%和79%。



Cattle-Yak	MDPSEMEAVVHHCPSDSATGGETRAPQGMELIPRRVSRSP	40
Cattle	MDPSEMEAVVHHCPSDSATGGETRAPQGMELIPRRVSRSP	40
Rat	MDPSETEALHHCSPDSSPADAEVPCSTELIPRRFVSRSP	40
Mouse	MDPSETEALHHCSPDSSPADAEVPCSTELIPRRFVSRSP	40
Human	MEPSDMEAGYHCPIDSAFWDETRDPCSTELIPRRVSRSP	40
Chimpanzee	MEPSDMEAGYHCPIDSAFWDETRDPCSTELIPRRVSRSP	40
Consensus	m ps a hc ds e r pq eliprr srsp	

Cattle-Yak	TCARCRNHGVTAHLKGHKRLCLFQACECHKVLIILERRRV	80
Cattle	TCARCRNHGVTAHLKGHKRLCLFQACECHKVLIILERRRV	80
Rat	TCARCRNHGVTAHLKGHKRLCLFQACECHKVLIILERRRV	80
Mouse	TCARCRNHGVTAHLKGHKRLCLFQACECHKVLIILERRRV	80
Human	TCARCRNHGVTAHLKGHKRLCLFQACECHKVLIILERRRV	80
Chimpanzee	TCARCRNHGVTAHLKGHKRLCLFQACECHKVLIILERRRV	80
Consensus	tcarcnrnhgvtahlkghkrlclfqacechkc lilerrrv	

Cattle-Yak	MAAÇVALRRQCEAQLKRHIITCGIIRGAAPPRAFSRVKKEV	120
Cattle	MAAÇVALRRQCEAQLKRHIITCGIMRGAAPPRAFSRVKKEV	120
Rat	MAAÇVALRRQCEAQLKRHIAQGIMKGAAPLKAFLRVKKEG	120
Mouse	MAAÇVALRRQCEAQLKRHIAQGIMKGAATPLKAFLRVKKEGA	120
Human	MAAÇVALRRQCEAQLKKHIMR...RGEASPKAANHFRRGT	117
Chimpanzee	MAAÇVALRRQCEAQLKKDIMR...RGEASPKAANHFRRGT	117
Consensus	maaqvalrrqceaqlk l g ap g	

Cattle-Yak	IRPGVESGKENIAPQPQIIPHVVPLALTPPGKENSRGPLL	160
Cattle	TQPGVESGKENIAPQPQIIPHVVPLALTPPGKENSRGPLL	160
Rat	IRPGVESGKENIAPQPQNPHGAVPLVLTTPPGKEN.YGPLL	159
Mouse	IRPGVESGKENIAPQPQS PHGAVPLVLTTPPGKEN.YGPLL	159
Human	TQPGVESGKENIAPQPQT PHGAVLLAFTTPPGKNS.CGPLL	156
Chimpanzee	TQPGVESGKENIAPQPQT PHGAVLLAFTTPPGKNS.CGPLL	156
Consensus	p sgkeniapqpq ph v l tppgk gpll	

Cattle-Yak	LSRPPEALPLFWTEVPPGPWASGHWLPFGLSMFTFVVCRL	200
Cattle	LSRPPEALPLFWTEVPPGPWASGHWLPFGLSMFTFVVCRL	200
Rat	LSRPPEALPLFWTEVPPGPWASGHWLPFGLSMFTFVVCRL	199
Mouse	LSRPPEALPLFWTEVPPGPWASGHWLPFGLSMFTFVVCRL	199
Human	LSHPPEASPLSWTEVPPGPWASGHWLPFGLSMFTFVVCRL	196
Chimpanzee	LSHPPEASPLSWTEVPPGPWASGHWLPFGLSMFTFVVCRL	196
Consensus	ls rpea pl wte vppgpw asghwlpfg lsmftfvvcrl	

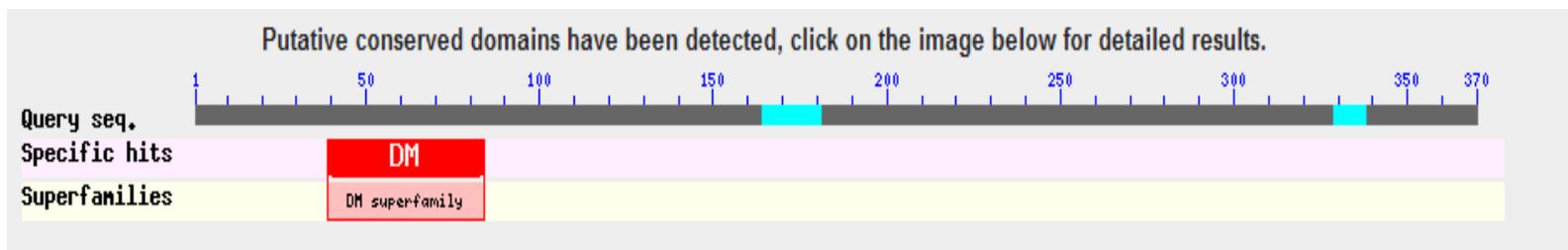
Cattle-Yak	LCQEPALPLHPPFGFDPGTTLRLPTHGPLPTCTGSHPILT	240
Cattle	LCQEPALPLHPPFGFDPGTTLRLPTHGPLPTCTGSHPILT	240
Rat	LCQEPAVPLHPPFGFDPGTSRLRPTHGALPTCPGSRSVIT	239
Mouse	LCQEPAVPLHPPFGFDPGTSRLRPTHGTLPTCPGSRSVIT	239
Human	LYQEPAVSLPFPFGFDPGTSLLQLPTHGPFITCPGSHPVIT	236
Chimpanzee	LYQEPAVSLPFPFGFDPGTSLLQLPTHGPFITCPGSHPVIT	236
Consensus	l qepa l pfpfgfdpgt l lpthg tc gs lt	
Cattle-Yak	APLSGESQGGSTLPRFTCSILLIQFCGTPDPLLLQPQAPGP	280
Cattle	APLSGESQGGSTLPRFTCSILLIQFCGTPDPLLLQPQAPGP	280
Rat	APLSGEPQGGPNLPRFTCSILLIQSCGTPDSLILLQPQAPGA	279
Mouse	APLSGEPQGGPNLPRFTCSILLIQSCGTPDSLILLQPQAPGA	279
Human	APLSGEPQGGPSSQPRTHSTILLIQFCGTPDPLQLQPQASGA	276
Chimpanzee	APLSGEPQGGPSSQPRTHSTILLIQFCGTPDPLQLQPQASGA	276
Consensus	aplsge qgp p t stillq cgtpd l lqpqa g	
Cattle-Yak	SRITWTSASSEWQLQREAAEAIVGLKSSQAPRITEG.PA	319
Cattle	SRITWTSASSEWXLQREAAEAIVGLKSSQAPRITEG.PA	319
Rat	SCIARTSGPSEWQLQREAAEAIVGLKSSQAPRITESVFP	319
Mouse	SCIARTSGPSEWQLQREAAEAIVGLKSSQAPRITESVFP	319
Human	SCIARTSGPSEWQLQREAAEAIVGLKSSQAPRVTESVFP	316
Chimpanzee	SCIARTSGPSEWQLQREAAEAIVGLKSSQAPRVTESVFP	316
Consensus	s l ts se lq eaaealvglkdssqapr tp p	
Cattle-Yak	NPAWISLIHPCGPPAAPGGRGFQIVGFSIRPSPAPSVLH	359
Cattle	NPAWISLIHPCGPPAAPGGRGFQIVGFSIRPSPAPSVLH	359
Rat	NPAWISLIHPCGPPAPPGGRGFQIVGEPVIRPSPGSSVSLH	359
Mouse	NPAWISLIHPCGPPAPPGGRGFQIVGEPVIRPSPGSSVSLH	359
Human	NPAWISLIHPCGPPAPAGGRGFQIVGECIRPSPAPSVLH	356
Chimpanzee	NPAWISLIHPCGPPAPAGGRGFQIVGECIQPSPAPSVLH	356
Consensus	npawisl hpcgppa ggrgfqpvgp l psp sv lh	
Cattle-Yak	IGRLGSISLL	369
Cattle	IGRLGSISLL	369
Rat	IGRLGSISLL	369
Mouse	IGRLGSISLL	369
Human	IGRLGSISLL	366
Chimpanzee	IGRLGSISLL	366
Consensus	igrlgsisll	

疏水性分析

```
181 ASGHWLPPGLSMPTPVVCRLLCQEPAIPLHFFPGFDPGTTLRLPTHGPLFTCTGSHPILT 240 Cattle-Yak
181 APGHWLPPGLSMPTPVVCRLLCQEPAIPLHFFPGFDPGTTLRLPTHGPLFTCTGSHPILT 240 Cattle
180 GPGHWLPPGLSMPPPVCRLLCQEPAVPLHFFPGFDPGTSLRLPTHGALFTCPGSRSVLT 239 Rat
180 GPGHWLPPGLSMPPPVCRLLCQEPAVPLHFFPGFDPGTSLRLPTHGTLFTCPGSRSVLT 239 Mouse
177 VPGHWLPPGFSMPPPVCRLLYQEPAVSLPFFPGFDPGTSLQLPTHGPFTTTPGSHPVLT 236 Human
177 VPGHWLPPGFSMPPPVCRLLYQEPAVSLPFFPGFDPGTSLQLPTHGPFTTTPGSHPVLT 236 Chimpanzee
*****:*** ***** *****: * *****:*****:***** : ** **: :**
```

结构域预测

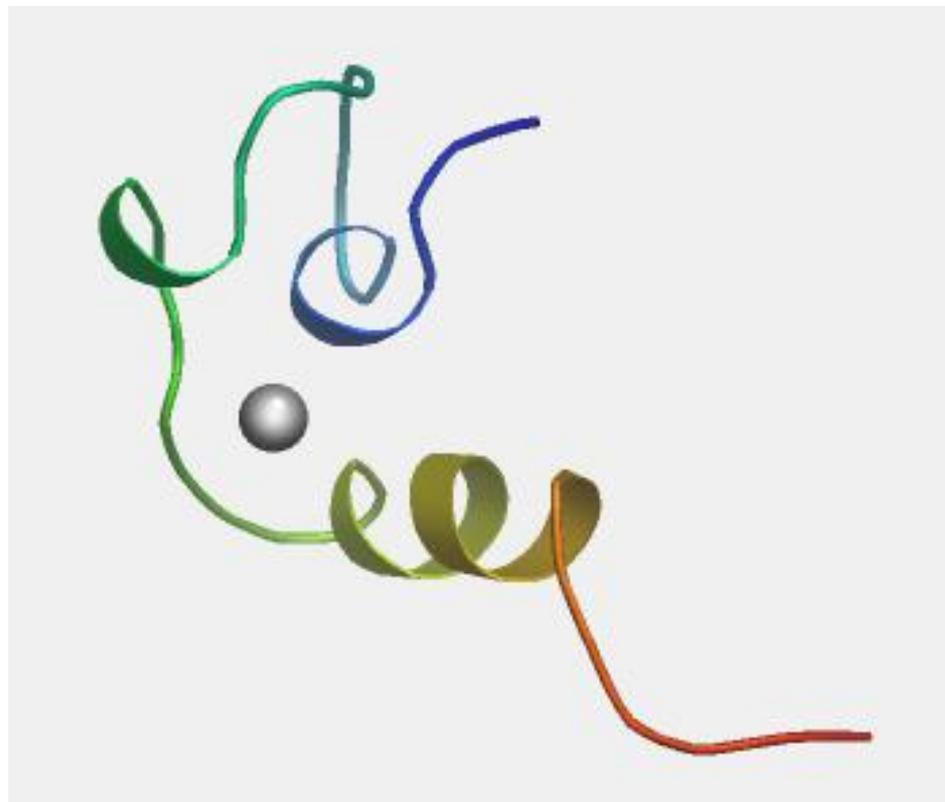
犏牛DMRT7蛋白含有DM超家族的结构域。



亚细胞定位

Subcellular location	Cytoplasmic	nuclear	Cytoskeletal	Plasma membrane	Mitochondrial
Cattle-Yak	45.3%	39.1%	8.7%	4.3%	4.3%
Cattle	34.8%	47.8%	8.7%	4.3%	4.3%

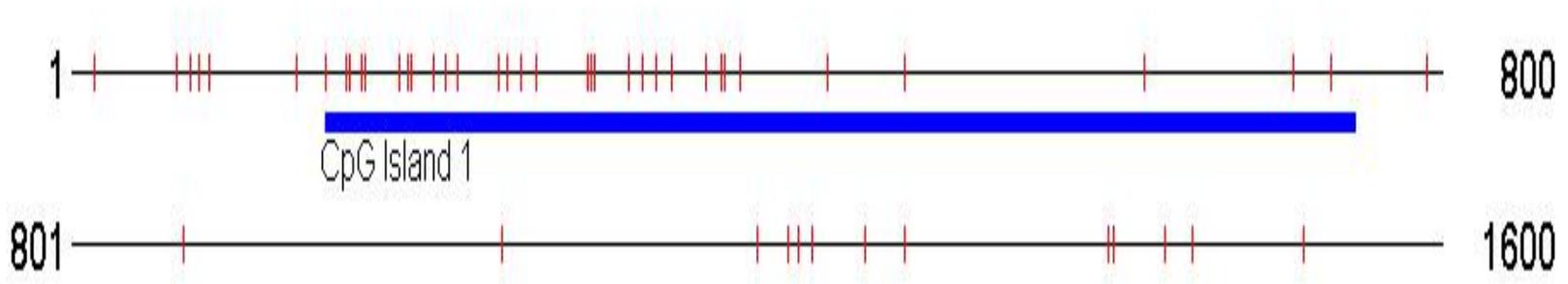
三级结构预测



犏牛DMRT7蛋白质中变异位点可能会对其功能产生影响，后续将克隆牦牛DMTR7基因的mRNA序列，通过比对进行验证。

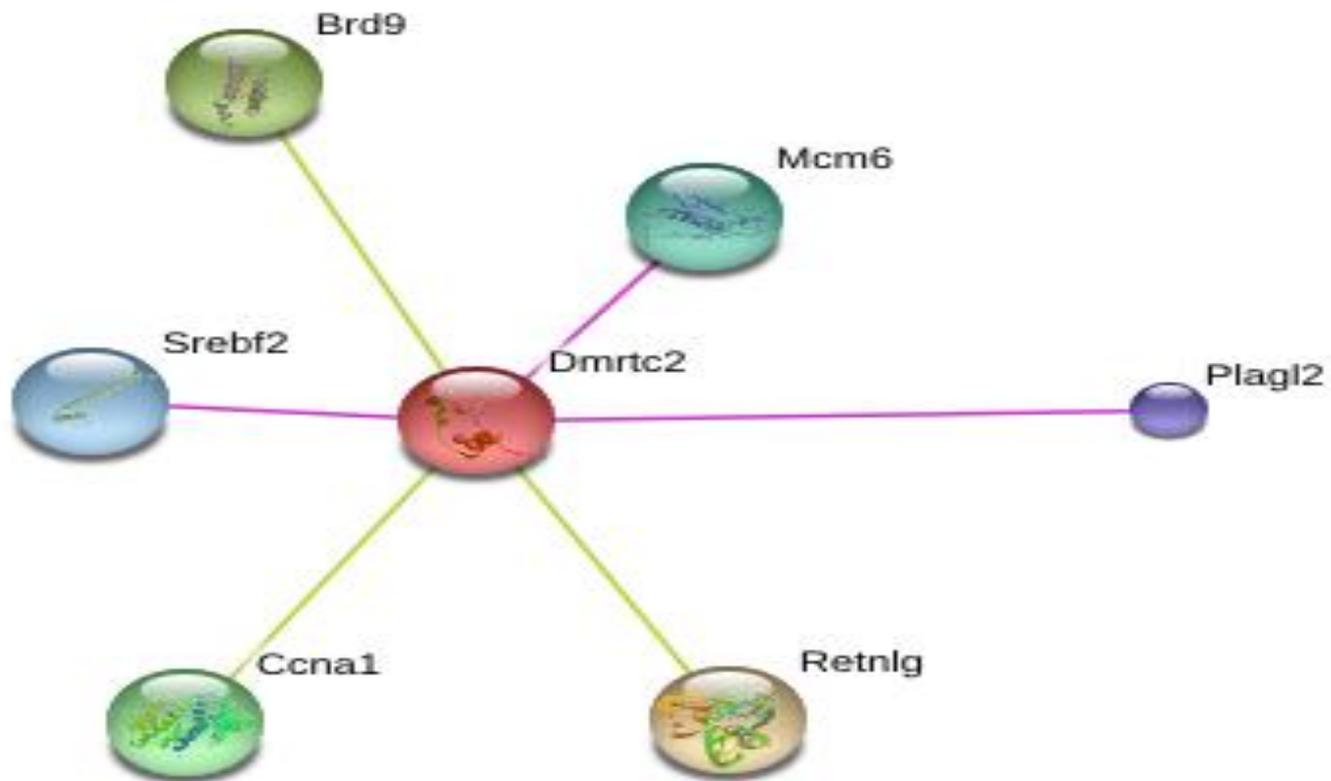
CpG岛预测

用 CpG Island Searcher 软件分析黄牛 DMRT7 基因启动子区的 CpG 岛，发现 CpG 岛在 148-750bp 间，GC=55%。



用Proscan软件分析DMRT7基因启动子区的转录因子结合位点，发现DMRT7基因启动子区存在SP 1、S P 2、AP-2、CREB、STAT5等多个转录因子结合位点。其中SP 1、CREB等位点为甲基化敏感位点，CREB为生殖和胚胎发育相关的转录因子结合位点。

蛋白互作分析



下一步将分析犏牛、牦牛和黄牛**DMRT7**基因表达量的差异和该基因启动子区的甲基化程度。

同时也将分析**DMRT7**互作基因的特征及表达量。

参考文献

1. Kim S, Namekawa SH, Niswander LM, Ward JO, Lee JT, Bardwell VJ, Zarkower D. A mammal-specific Doublesex homolog associates with male sex chromatin and is required for male meiosis. *PLoS Genet*, 2007, 3(4): e62.
2. Hou Y, Zhou X, Liu J, Yuan J, Cheng H, Zhou R. Nuclear factor-Y (NF-Y) regulates transcription of mouse *Dmrt7* gene by binding to tandem CCAAT boxes in its proximal promoter. *Int J Biol Sci*, 2010, 6(7): 655-64.
3. Kawamata M, Nishimori K. Mice deficient in *Dmrt7* show infertility with spermatogenic arrest at pachytene stage. *FEBS Lett*. 2006;580(27):6442-6446.



Thank you