SUPPLEMENTARY MATERIAL 2

Taxon abbreviations/Database:

Co: *Capsaspora owczarzaki* (BROAD INSTITUTE-Origins of Multicellularity)

Mb: *Monosiga brevicollis* (JGI)

Ac: *Acanthamoeba castellanii* (own assembly)

Am: *Allomyces macrogynus* (own assembly)

Ap: *Amoebidium parasiticum* (NCBI scaffolds)

Tt: Thecamonas trahens (own assembly)

Sp: *Spyzellomyces punctatus* (BROAD INSTITUTE-Origins of Multicellularity)

Rel/NFkappaβ

```
>Co_NFkappaB GU985464 (predicted CAOG_01632)
MDLSELSGWDPNLSLQEHTANLLAMDDSTMAAMILHSDGLSLFRDMGLYNSVSTSLDISGIPKFPPPPQQ
PQLAQAPPRRGHNQSSSSDSHSTPSPGSVLFSPSPASQDMSLQSPELGLVGLTGNSRLASTSEADDILLA
NILGHPVRSVSANTSMVGLPDDLGFSPTTGMDLTITATSPATADSSASATAFPAASPIASPSVSTSSGPV
TVGGTSAALAALTPDRINRLLSAAEAAAEGAATLVEDLLMVTEEPAQFARFRYMSEQRERSLAGENSFPT
LMVNPKYARVVPEMALVTAVLVTKMPDPHTGRQQKHWHHLGGIPAAPLEGPQRIARFDNIAVIMDKANNK
DKDKSKAPVRSKDDQRCVRIMFELVFVSGNTQFYGRAISQPIYNAKLAITKISHSSGPVTGGNEVIMLCS
KIRKGVTGVRMTDPTQWSVQAPSGSAWELNPQTLKADCNVPGANLFFHHQYAVVLTLPPYHTQTITAPVT
VRISILDTDDETESQYVEYTYLPAEAAVRNAELAARKRRRDDSMRDFMDRFDGSDGGNGSGSGRGNNGGH
DGSDANNNGRGGGGGSSSSKGGDEPFNFNSLIPMHQHKLHQLALSTVRAVQGFAASGDARYLLALHRQLL
AAPNENGDSPLHTAVAQGNLRSTMALLPLLAAEDLQSVNDMGETVLHSAVIEKRAAIARLLLVAGADLGQ
SNARNFNRNSLHYLARHGDRATAMAVFGVFGSAQAPPANTNTPAQAPAGETKPKPADLRLLARIQAQAIK
ALLACELETGATPAHLAIRGGHWHVFEACAKLAASAPIPKAAGSLLSMVAEKSSGHSLLHSCVLANNEQA
VRLLINLGASGNARDFGKNTPLHLAARQGHIGIAALLVEAGATLSLNAVSQTPLDVLTSEGSGLSRDQLR
ALVAVLRGELKYADMRGRPTLRMPTHAELHSTAAALTSASPGAVSLADFYAGKKASRSPAPLGASSSLLS
STGASAAGASAPTIAAVHAASATPVERTSMNNDDDYVLLEKDAPYPVEQQPHGKRNKHSHHRFTRSSHGS
QDKDELKKDKDDPKKEKEPKELSKFTLKEAFVDGTNFWELTRKFAGKKKMASASTGEMEPLSPERPLSPT
NAGSGAASPFNQAKEQVSPGAVPPTGLEKLVNKLMDASEATLSSQPAEAVTPEQKLAEKLEKLGLAPAST
TSAPPPHPKVAALNAQSVEDARKTSTHALYSVD
```

Runx

```
>Co_Runx1 GU985462
MSLTATAAATSPISPASAVSSSTAGGAGLVGPPRRRVLQSSRSSSFSRTTSTSSEDVTANSEQDFPFVSS
IVSTTHPQVLCSNLPEHWRCNKSLPAPFVVYAQVNVPDDTEVTVSAGNDEHAIAEMRNFATVMSNNTATF
SDLRFMGRSGRGKRLTVSITIHTTPTPIVAQLVEVIKMTVDGPRDPRRRRPGSEMIASASEDDLSDASGF
ASDLGSGTMSAEAFGSTASTGSPLSGSSQQLAPGQTPPAQPNSAGTGGSFKSHRTRSLSTTLRRCQSAPY
MAPEFRRRTGSGTLLPHDKPSSPRASSNNNDSLDPRPRSAFPSDLADSEMQRFRSVSMGSSADSMAEDVF
SPLESPRSPSVAALYHLQHSGVSSPSPDHSLLSSMESLLAPLTTPAALSFNAADVVLSSGAMNLVPDFTL
VSLSVYEAPEGSACVATMRVGSAGQVMTDVGVMFGNAIPTLDSVNLGIDGTVQCMFKVPSRSECGASLQR
GNRLEVVGFYKLADRVCCSSNSLGFAFKDVEGAITERQLNAVLHKLHLKNQQTQLAHSEPLQKQVLGGGA
PRNLLRQLRPLEDNLFARSQAVLQQCGMQHTLNPIDLPIGPDGQISVSNVQSSGLTALHVAAEFGWNKFV
TQLVTAGASINQRDNFGNTALDWAFMSDQHVTCQLLQSAGGLFNMFAQPKVTDDLQASFQQFTLSSPVLG
QLPSPTTRIASSRSPRSPHLGSPLNFGGLSQNTSPPGQVLSEIPEAQSSAQQSYQTQLQFLQQLQHHQLV
EQEQQQKQQQLLVQLMQQQQPQLQLQQQAKQQAAHHNLVQHLLSMDMSASASGSSSRAVSTSSLLSNASA
VTSAPSMLPLDFTTTSDAHLYSEPDLLALDLPSPNPSTLSMGELHF
```

```
>Co_Runx2 GU985463
MSIAAASPTSYTSPSSGLRRTTRLRSNPHHSGNEDEDDGDYSAGKGRAKRAARPASGRGDDSDEEGDVDH
SATVSQTDNPYIFVVGLPKHWRANKALPATFRIGIHPGYKVANGTQVILHAKNDEVGQAQIKGGMTVIQD
```

```
NAALFTDLRFVSRSTSRSGRGKRFDLLISILCEPPMYATVMEALKITADGPRVPRFHKEGKSPNLGVLKA
PDTTQDSPMKPAFAPSQQRQQPMPQQQQAPPPQAQLLQQTAPQLGLQLQHPQQQQQQQQQQQPQQPQQQQ
QHPLSLNNMYALPHNPANTTPVRQRTVPPPLASSAGSAGNMAGAAGIAMYGTQGNFLTSPGFFAGEFTPG
FGSAITGQWLGNPSGAGVTRGTPLPSPGLFGPQVPLQTPQGGQGMYYSTGHIPMPPQSGQMLQPPPVQQP
QPHHLQQPQQQQQPQQQQQQQQQQQQQQQPQSQPHHPQQPGMHQFQPPMPPAFNVGGGHPQQQQVNSQHLS
HPPHHQQPGQHGTPLPTPMSATLHFNFPASARETAMLEESANVSRCSMIVETQLTGLSQRAQNALNACTH
SMLSIGSCNEALIPLNEALWNLEVREFNRLKEQTVCFAMAV
```

T-box

```
>Co_Bra-like GU985459
MTSNHFSSTNVATSAPIVVPTPSTSDVTLSRASSSAAAAAAALDSSDTHRYIKHVAYHNQPSAELEQPNV
ALESSKLWAKFADATNEMIVTRTGRNMFPVLRFDITNLDPNTLYAIMLDIAPADGIRWKYVSGEWWVGGK
SDAPTGPTMYVHPDSPKSGKYWMQRPISFSRLKLTNRPNTKGNIVLSSLHKYQPQLHIVSVPSMNAPTHG
QPTSVHTFPETAFIAATAYQNDYIKDLKIKNNPFAKAFLNPRDRPRSSRGDDDDDDDNDDNDDNDADEHE
EDSGSDGSLPGAGGREGGRPVRTRKPPSPFKGGSSTTSGTTRTRVPSGSRSSSSSLPSSSASELHAIQGL
RPSDLSAMPNSMHVQPPSQQQLSSMQYVSPQRVFYAADSAPPPYAGQPQQLQQGMPYGSNPTTTTSVAFP
AAGGEGMNGGYYQQPVSLVQSMQPGQPQSMQPIQQQPIQQQQPIQQQQQQQQQLGQYAAQTNVLPYGQPQ
MVDRRVFYEQQQPQLQQQQQLQQLQPLQQQQQQLQPLQRQPPTSSSEIGSSLDGVMTLPLASNTMPHQSM
QMGSWSQPPSGGYYSMSGQPVSVSYSAPTMQPMTTFAQRPLMQQISQQTPSQQQQHLPQVVFTSQGGGGH
VTFTGALPAQNHQQQMAFSHVQPIPYQSMAHPHAQQLQQQQQQQQHHQPLPPHQHQPQYSTLPPLGMVPQ
QLESQALLWLQQQPGPVNAPTQGLKRAALEAFNASESDPGAPWTKREHIDAARVE

>Co_Dtbox GU985460
MGERRITRSLSGSIGGGSAGAHAAFGFSMDPGIGIGSGLGFAAPVAATDANAASMTTSASPGRAAGQQQQ
QGQQQQQGQQQQQREPRLSARNLALPDDRSALAGTSASSMRSAASTASSGAHAASLGALRRGGASGHSLA
GRRGGGNNNHDDTNNNRSSSGDATADDADHPLAVTRSNGSAIQVSLRGEALWKRFSENTTEMILTRNGRC
LFPVPKIDVGGLQPNALYSIVLEFVPANGARWKFDLVDQEWRPLVGPSRPITQASYTHPDSPALGSVWNA
SKINFNKLKITNVPNTGDHVVLNSFQMYKMVVRIDQLNPATNQPTSTVFSETLPYSSFIAVTVYQGLAMR
ELKKSTNPHSGIQHHQSQKELLLAAAAAAATTGTAKSKGNKRRPTGAAVASADSGSTDEEPQFREAKDDR
AKRIRLDTQRQSPPRNGNNNGNGNNNSNSNSNSNSNGLLQQSDLISGSQFDFLQNELNLFNSAFLQSGNS
INPSASTFSLFKNAAAAALAETQAPSTSFAQPSDSTITGGNPSRMLLHNADSLRSGSSNSVLGIIPSSSA
FALAAAAAHSHQLGLLSLPPSSNAPSNALPTTNGAVAAVKQEAESTTTGSAPAASSSSAAAAASAAAAQT
SPQAQEKARLLGQRRHGGRLSGAQHRGDTLAGPDNSMFAAPNSTLFGGFASFEVPPAPSAHMRTTPSPLQ
LQNNDASPPVMSSHPMRTRRSITSQDGLALHQPHAPQHQAQYQQQQQQQQLQQQQQLQQQQQQPLDQPQQ
TQPWRSQQNQLTHVFSASAQSFPSLLATFGAGGLAGSASVNASSFSLADSQPLAQWLSSFADPNALPFSR
ESVMHLSDLKGSLSGGLSSLGLGRRDVPMATQFGFGPIGTQFGVAPMPQPTLQQLQVRQQQFQQEQQQQQ
QQQQQRQLQEFQPQAPYQQFGHTQQQQQQPFYNQQMQTPVPAAQPFYQSAPAPQQFGQESLSNSYSARGG
FPGGFAAQPSASNLFASFSGFGGLDPSMTLSANSSGGVAQFASGRRVTRSMSGSSGHPFPSMESSAAPAT
SSQTYPPAAAAQGLDQDQNASGAMLTTPMGSFASNNNNNMLRPSFSSEGLPLPPPSTIGIELVGKDLWEQ
FHQLGTEMILTKKGRCLFPHPVIRLTGLQPDAKYTVILELSPVDSTRWKYANAEWQPNGRIALPLLPTLC
THPDSPATGEAWMAKEVSFVRVRVTNTFNETGNVLLNTFQKYVLHFHVVQVDSVTHAPRVVETRRFPETE
FVAVTVYQNVNIRKLKKQTNPHATRDFDTEVEEDES

>Co_Tbox3 GU985461
MGSPLRLISMPGVHIALEKEELWRTFFRHGNEMILTKKGRRLFPTPSFTITGLPEHRSFSIVINFVPASK
VAHRYVGEAWAEVPTSKSLDCNVVTQFVHNSSPALGSVWVEHPFMFDFKITNNEITSTTSDAVLLHALHK
YIFCVQVFEPLAEDESSMLEVDLEAPAGANRMHLPGHRLVHSQLFEETQFIAVTAYCNPNIRALKISHNP
FARAFTDRARGDIETARALTQALENQDDSSAGEASPIPVAMVGPQRAPVRRSSFGGADMLSPATTDDGTP
LRPQFLPLCFEPERSRSAALSSAPTPPTPIPTRITARRNRTVSEMDAVDSTASLSPAHSLASRNHTPSPM
IQQGSTAFPFTTLPSASPKRARLDHSQGLRSITSLGGGARKLSFESMYSSDFSDLLLDAQRMFSNANARN
SGHGGLTTSASCSSMASTNSRSAGLPHLSHTYFPGSSADSTPLPTPGLGHSGTQGTQPTARPRVMSISQN
RANWDAFQHATQSANLDSANTSTVEELEGLIGAVARSAKEAAIEMMTKMHATSPQFRSSEEYAHQVKLAM
NDAKIQIVVCDLAVFTPTTLQRLRANPVSADLSLMVVVPPNTLPSLIDSILASGVEDVLYEPLPPALLKQ
HLVRCIERRSWHAYM

>Sp_Tbox SPPG_06782
MHTTASSHTAVPKRKSDDDDSHESSYLIKAKTRRLSNNLMSSQSGNDQQVERRPLRQGGMRTLTVPPSLS
PHGSRQESTMRNANGPTILLEDAKLWAQFHSVGNEMIITKGGRCLFPTLRFRPICLDPQALYSIAIDIVQ
TVPHKFKFKNGHWAPVSLDAGSGGESEDMLVPTCRAILHSSATQSGAFWIQQGISFAKIKLTNRCPSSLR
SPSPGSSPDYPSVPEGYFYLNSFHQYQPRVHLIKHSAEGQSVTTYLFEETKFIAVTHYQSEKVNSLKKNY
```

```
NPHAKGFKDYETKLAALRIRSGSVASTTSSGTPTTTTNDAIDIPTFRISVNERSGRAVVKSLPATVRQSR
GVIRTTGNRYQIEEDEDESPYDNGQEDDQVEEDEVECSSTADTSDSDLSDSDDESTYSTTSTTSLSQVSS
DAASPSQENQPPQSSNSSSPLQLLALFCSYILDSGNLALNPAEIVTHSETTEVQRFQTLVDAYEIECKRL
REVRNMLMSSDRECAPAMRMI

>Ap_Tbox_partial manually annotated
QNVHLRQLELWTRLHQLRNEMIISSKGRQLFPPIELSISDVDPEQKYNVWLAFECEDGSGWRWVEHSGXX
XXXXXFYLLPGGPYEGRQISTDGISFNHCKITNKPNNKAAIHLESFHSYLPVGYLVPVHENSPSQNPPVR
NMTNFIGV
```

Churchill

```
>Co_Churchill CAOG_05736 reannotated
MCRNCLKEDTPARGTTCLDTGAYLVNFKGCAQCQSFEFPREQDRKVDEDDETGEETVTFTHVCKQCNHVI
AEHNYTFEIEDGYQEYTMECQLCGTADDTASVLPDDPRKAQTLF

>Ac_Churchill1 manually annotated
MCRRCLVHVCPQRDRTCLSSGSYLLNFKGCGECGKMEFIKIAELEKNETETDETVTYQHQCSECSHKIAD
HEYTFTIDGEYQEYAMNCVLCGYGEDRVSIMPDDPEKIAQYDMY

>Ac_Churchill2_partial manually annotated
RARRRARWSHLIIRHRITSKLKSVAANTRDAAAKALAKKTLKLVKKFALNGCGECGKMEFIKIAELEKNE
TETDETVTYQHQCSECSHKIADHEYTFTIDGEYQEYAMNCVLCGYGEDRVSIMPDDPEKIAQYDMY

>Tt_Churchill AMSG_09514
MCKDCLRSVCPDRGSMALDTGAYFANYTGCAECSVFRLAVRDKTADEDDEGNETVSFTLQ
RLSTDVCGACEHVIALHEFSCIVDDGRHEYSMSCLLCGMAEDSRAIGVYDAGAASALT
```

P53

```
>Co_p53 CAOG_00511
MDKLDATNLFELLGPSSSSSSSRRALPSRTQVATAGAGAGAGAGAAAGADADSKGSHDPADLNSTLLLGD
SRPASLRTSTSLLQSIFPPNLGTTPTQLFAEFGETTGTSLLNMSGLRGFNTPSSSSSHHSHHSHHSNSSN
SSFLGRPRTLPVQKPLSILQGSAPEPFVEPLTIGEPIYASPNSLGEAGFLLSVDVSNARHSAISSAYSYS
EALGTLFTNFDVGVPFVFRVAKPPPTVHPLHIRATLRYKQMQFMKEPVRRCPLHLSIDSDLHLLRACDQD
TVYSVDYHGRASIAVPFTPTMQPLVPVLINTLKDAHVPLVTRHPSSTHGEASRSNPYVCWSMTWFLKFMC
YTSCTGGMNRRATEIVFTLEDSQGLIYGAQALDFRTCASPSRDRKQLEKSPPPIPVGSIIPPEVVTAHQG
SALADPGTTASTPTRRSAQPVVPSPVVAASPTPAPGLLSSGRKAGAIPVTGVTPRPLVVIPAASSAQTAP
TAPAAAIAAPSPALTEAHAAPTTLLRDRSITAIAKRRHADDDDEIFTISIRGRDNFVLVSKIVEGLNALA
QRTSAAVASKLERTTTPPVVEDDPSAPPALSALDTLAQAAAGHVRAPDEAAPATHERKRARLADDDTKAG
RATEASGESISTWLGSFGLKQYAVNFAKNGYTNTPQILNLCEDDLKSIVRVKAHRVRIAAAIDELADAQA
EPTSTPPMSSDGDVTVVRMTIRRTTSSTAK
```

Stat

```
>Co_STAT CAOG_04636
MAEAAATPTRWSSISTHPEATAIYGRDFPLELRHYFAQWIEDQPWATMDPDNPEHYESIKDRVQTLGEML
HAKISQLAASGSEDKFLLQMRLQTIAADATFDENPTVLVRILKTCMERELIVWAKLHGLESSGLEASGMQ
LLDGMSAISALAALPQSAPEAHERDDISFLMMGPSSTSRSSVPVALNSSSSLLGSLPHVEGGARALAVEV
SDASGAAAHQLAECVRIAAHKAQATLHHLVGDTQHVDNELKRLQQMQEYFVIQYQEFCQLQAVLQRSPRS
DAARGQIELVKRRKQEVGTTLSTEANNIIALMQSLNRQLQANMGDIAELQLTVLPDALQAWQRAQRLESD
DADQLDSVQATCEMIAEALWRSREQTRALLLLQTQLPFDRKSYHQPVQQSGSHTPFVAQLNSHSPHHARN
SSSHHHMDTATPASSTSPVNSSPQATLSPRPGKSRSSSLSASASPASGLSSMSISSGSNAGGMSSNSSAA
AMDTEGATNTAASISSNGMVSLEDLHIRATQMLASLIGATFVVETQPPQALKTQSKFSAQVRLLVGRTLN
LHMSPPEVLTTIISEAQAKQVAQQPAVTVTDGKRASQPTSIPEDFNAGDILNCRKAMECHQPSGALRVIF
KNLSLKKIKRGSNKDETVTDEKFALLFQSTFSVGGGELLVAVRTLSLPCVVIVHNNQQHSAEATILWDNS
FATMDRAPFEVPEVVSWPDMLTALQHHFSLATQQPLSQRDLEYLTWKASDLIINESITWKRLTKEALPDR
TFTFWDWFFGAEDVIKRHMLETWQDGLIMGFCNKQQAHDLLINCVPGTFLLRFSDSEIGGLTVAWITEDE
RGMRQVFNLHPWFAKDFAIRALADRIHDLPQLQFLFPDTPKDAVFGRHYSVETAHVTSNNPDYVRSSIAA
```

VIPGSVPSFTMPIVNDQFLMKADDSSSSMMHPTSSNLLDLVGMGVGMGVGDADAGQMSSLDWSTFLESLA
EPIPQLQHDGI

>Tt_STAT1 AMSG_03678
MAKIAEKINSLALSPNQDAYILKIRLEEIAVALEAKYASDPLELVRVILYVLGHEREILETGGTKTKFAM
HSSSFSSTSYGMDVSKLNDASLNAEIQKLQEHTAETEAALAQLLQDQEYLVIWHQDWDELNARAQAASGA
QAKALAAECHTLRGNIMQAQEGILRTRREILEAHVGAITHSEQVLSILLAKLKQWKAVQRKQTMDESFLN
DVLNHMQLQFESLADVYWNIRHQLRQHEVQREQLDCGDENESAVFDQLSTQLHAQLETLLRASFVIVEQP
PQVMKTSTKFSSAARILVGGKLNMHLSPPEVNLQLVNEPQATALVNDDELIGNSGDLLNGKQMMAYSEST
RTLQATFKNLSLKKIRRHSARKEVVTEEKLALIFVTSVQLGDLTYNLWTMSVPVVVIVHNNQLVRATATV
FWDNSFGKTDRALWEVPDAVRWPKFAAGLNHFFTLANGRPLTPDALNFLGSKLLEPGEDTVEWRTFSKDN
LRNRSFTFWEWFYAIVELVRKHLAGPWIDGLVVGFVSKSAASEALMHQPRGTFLMRFSDSELGGVSVSWN
VVDGATGTSKVYNLQPWFEKDLAIRSLPDCIKDLDQLRILYPGTPKDDAFAKYYSEEKVVVGDGDYVRAT
INAAIIGDEDATGTLDSIATSSTLSTMDFNTGSLEYDSLAAITGFPMSSYMGDLAAVDYNSYGAPEAITF
PSMIDIEAPATMDE

>Tt_STAT2 AMSG_02036
MDGETLWSYIQDNPEVVAVYAESNPHFPKAPALARGLFDALTEVLQHLETVAASVPHARFKLNRLTQIGR
TLREEYAGNDARLLEALRVCMMAEAAVLEGAEGEGGVGGGSAGVGASSRAPQAMEDGGQSPREGEDVVSG
RAGPLLRQYNGLAHEARGSEDELRSVQELQEEFRVQYLQNEDRIAALAQGSGPNNPQARQQLHALYQSLA
QQSEQLIGMRRSLRSALAGEANMLRTLQATLGEMLGEWREKQFRKAVPEKILAAELRELRQLLMRQAAAL
FSLRERLEVLRFLRQQLPPAASADDDGEVSVPEAETLAAAADAMIENLVLGSWVVVEQPPSVVRRRVRTR
VSIELLIGASHVVSLPMVSVQLLDAVTGEPASAQLLNATKRAEATNDGRCVAVFSSLCIADIRHASLTKD
KLSAECMYRLAVSCEAGLAHGTYSLIAVSNPITLTVSRRQAALAAATMFWHEAFADSFEAITAPETVFWS
ALAGALEASFAASGRELGADVLAYFRSKVVSSVPGDSSDIVSFEQFAKADMGKGFSFWAWLFAAREVLES
SPHMRALWADSLVAGFVPKSGVGGMLSAYPSGSFLLRFADSLLGALSISYVWDDSRGRQVVHLQPWTLTE
FETISLADRIRQTEQLTLLYPTFQHRDEAFGAYYSPDEAVQAVDGYVTTGLKMTVEGGSESRQGGGESAF
SFPLDLGPVSPLGEASNPFYTGMEPAEPAAGYSGMLLSPSLGAPAFGGEGPSSEGGGRLPELSFNMGLEP
LRFDEAPSLGSPPLQGSASAEGGLFGFGGSRQRSLGGDELRDMLQ

>Mb_STAT JGI38850 reannotated (5’ truncated)
MTRLKQVTEQELVVDAQELLRSRQELKDRLAQCLPHVSDNVRDVAQTLQLWEQTHNAKQDSSASMTTPEF
TAARKRCETELVNTMLAYMDVLGRLLSSALVVISQPPSVLKTQTKFTTTCAFLISDALALFLDPPTVTVS
VINESQAIDVNERNQAGLSGLQRQDAVGESGGLQNNQKVLGSASADGVLQAAFKTLSLKKIRRSDKGESD
VAMWYVGFGWSRLHRDLRFVFFLTHPVTPSLRIKVLNLHKLSAPIVVVVHGNQLPSAEATIMWDRAFALP
ARDASPGTFVLRFSASMLGGLSVTWITVDEDGDRQPWQLKDFAIRNLPERILGALSCIWDKSFDSNCRRC
IASAHARHYRFSIASSVIGAVDLDDLTHLYPGRPKEDSFGPLITPVTKVEKKKASGYVMADIAAVIRQMN
TGLRFESRPQSSLDNMDGGGPAKRSVRRDTSRASHTSGRDGMDSRMSSAMSPATPMSAVGGAASSVPVGL
SASGMGATSHLSQPRHQEQQLFMPQQGQHQHQQQSLSQQPNMGYGGYAGAAMHSHPQSAGMQAQSTYVTL
PSNLGNIASLNPALASGNFDWANMDDFFMGLQQNVPAGAPPTTSSFEAFTSSLIGRGGMNQGQPTSSAPA
PGGGSTAV

bZIP

>Co1 CAOG_07161
MANAPSVKLVAPSLNNTSMLSMNTPSFASFALNNMADGAQTPELAKWLNDHSLNSEDILNVSVLPENTAN
NTNTAAMSSLLKSAVVTPHPHQQTAALMMHFTAKTSALSPPKTAAYDHHSSDSEESVATSSNSSTTNQRR
RRAQQDDNDELAAQSKRVRVQVSGTAAGNDDDDNDNDNEADSADDTDSASSVFEDAQEGFVSLMAANGIK
ASATAFNPAMMNHISAAQVSFCPAPITAGTVPFYSPPSSECDLGELPSLNAVLSQANSALASSASTISGV
IAGDSHSYIAYSSPQEYDNEPMAPRKRRRPATTEEERQSRRRERNRVAAVVCRQRKQRNEADLKDRVGML
TRDNAQLLSSATQLRQTIIVLKNALLRHHRTCTDPEMELAISSLDVATPASNVPLIAM

>Co3_CREB CAOG_04938
MSAHVNLLENRVSLGELDRTLASDSLERRASDSELMKRQSFLSLVDDLHKQEKAQSELNGEHTASRGVAE
PQAGHHFPPVPAHQQKLLDQQQPQQSLQQPWQPALDYTPTVVQYAPAVFRAQPPAPIPIHGQPPQFIMQV
LPSTPLTTAPPASATRGSTRSQQQQQQQQQQASAAAAATSQWHAQDDDHNDTSSDAESQHGFSTTRPADK
KQRRLQKNREAAKECRRKKKEYISTLEDRVKVLEQQNSALTEEVKRLQAALAQR

>Co4_PAR CAOG_08044

MLSFATAANPSSHVAGATQAAHAAASVAAHPSMHGLAPVVAPVPDLDAQDHSDFQLLLATPDIMGASVHS
HLADAANPLPACFVDLTAEPVLKKGRPMASPLASSFDAVAAAHTADQHQQQQQMADLQYLEHAFDSLDA
MLQHEVIPPVVAALPASPSSSSFASSPRLSARRSFAEIEESSTNMSSTSMLSSQTFDLSLYEKQAALVKG
PAPGSVAAVAAAAAAAAIAAGASPTTPARKPRVAVAPATQRTEKYLERRRKNNEAAKRCRDAKKLKEDMT
GLMADALSEENQQLRARVAFLEDEIEELKRMVLSRAVSTPSTA

>Co5_Oasis CAOG_01232
MAENSSPLPSAPDAFAPTVNWDALLQPQQHQPSFLPPLSFLTGGAPFGSSVAATAPCALPNPAMNGAHNT
LVNIPSIMDFVGDTFDVLEQQGTFRDDSFDAVIEQPFSPLPPLEDPEDILRQQGFAPHAYGFADEDDNDS
DADSAGSDSDSDDSKPKRGNKRSSAAARRLAPGAAQAAGVLLASAPRKSSGPRTLQDFMDELTEEERRGM
AAEKMDIKYGHPLTKAEERTLKKLRRKVRNKISAQESRKRKKEYMSGLEARVNVYTSANVSLRGRVNKLE
ERNRSLAVQVRKLRNLLGSVASAASTTVSTITGGLGANAGLFAGEDALLPTNVADVKFAAAESGVAPVKD
EFSSMSTAAAAAARRAPVGFGRVSNVSKGVCLMMVLLSFSMFGSLPLNTRGGAGLTGNATLDAAAPLLSG
SSSDMYRISRTLKSFESSNMVATVLANGNQQEALTAMQHVHLREPLSVVPPAASTATNDSQLTSTSNSLP
VSNEVTVALEVLRAQAPAKLAALLQTSDAKALAGGLLSQAAGASRLSAAAPASLGAAATHERAALARSPV
DALSG

>Co8_Atf6 CAOG_04023
MEQLGHHDNDAAAAAAAAMHHLECGGGGDDDPLKVEIDVGVFDSLTSDEDEHHHPSSHSNHSTSSSGNSG
VSQPWSVLDAASTLSPASLNDTDFLDLLDSTTTTAPTTTAAAAASMRAHVDVDVDDVGVGRQQADVALLM
GLDLHHHADDDAHNHQLMAWLDDSHHSATSSITGAPDDLLLMTAHHHHTHTQPHHQYAAADHHASLLGLD
LLQDALLHPSPTLVAEHYQPQQHQQPQQQVVKLNPASSGQYKRLASLPPLRPLQGGPVAPAPATTTPTMA
TTTTTTNLGSAYIHQQQQVVGVGAASPPSSSSGHSSAGEEAQFATANTAPAAMYQQAIHPEMDKKLQRLI
KNREAASQSRKRKKDQFDTLERDLNTIKTHNAALRSQVVALEQENAVLKADNERLRSFIDTTPELRVSYE
QQYGPLLGGSGSGMRGHGSQVGKMAASLSSRSGVAASVTVCAVLFFVGFAFLSPTSFPGAMLGGPRQRYD
MMVVNSERGVAHVGGRGLLAVDDQGWGLSVRESSAVGLLELPAAETRLSEWQGGPPARVGGQNTGLPLIT
NEQRALGGRGEPALITNEQRAIGGPALLTNQQPAIAGPALLANQQPAIAGPALLSNQQPAIASASSSSSS
SSSSSSVYAIEDNLSINMQQHLERIERASQQLQVLEQQRQEAIAVAFAAQLGLTSSDVAMLAPPTLDHVG
TFMEQLAIDRRVDTSYLVCADVRNIWPAMVETAKSGSRPRLSFLIPSTSVLNADGTVQDDVTVEDDIESA
QDSLPSSADAPQPASTTNATNAGAIKSLVQIDCEVLQTKLIPVRANATTAPAEPRRRAVAGGRRPYA

>Co9_C/EBP CAOG_02294
MDLPCRQQQKQQQQQQCACLLSVRGGRGSILNAWHSDSFIGSAAPFGTFGSTPFSLATAASQRAASILPP
SFDSVSSPLSTLSSPPSTLDACDDIPAPTPVAAFYLHQLLGSSSKERDVDRLLHHPDMLSYLAGHSAAPF
RRASTAATKRPSSEALASPPAQRPRLDDGSHSSAAAAAPPSSAVDWFEHASLRDILNDENLVLHLDAPSL
SMQQLDPSLASPLPVQTQPVAVRLDVPSSPPVLASIPLVPSATDLVTLADDHHLLLADHPRSLSVEGLLL
KSSPLVSSHDHNLAVAPLLLAGDSLPAGSWSLPVSNSHSDTSPSRALQSLLDDSSVEHALNSSSVANAHF
PDFAGLFEFDDLSPFQPPAAFFSEQSFSPASPLKESTPQPSPSAAAATPAAPAAPVVEQSPISKPRAASS
RSRATTQRAAASSNKPVKASAVAAVAAEKAAAAAPSASSDSDGAQELDAGDLPQRKLSRRERNNIAVRRC
RDKNREKSLAAKSQCETVAQENANLRVRIHSLEQEVSYLKSMLLSQVIPGMPLPSASR

>Co10_C/EBP CAOG_03823
MTIEPSFASQILNFPCTAPISAIADDTSAMMLMLEPSFVPQFRGDLNDFPSSDMLEFDAALLAFEAAAAS
EGLARVCKQSPCASDASSVSDSAVSDAADRASFASEPAWLSQTFDFSLHSNADVKKRSGRAPTHHTNTKN
NMSFIDESVSPRKRRAEEALRDPRSSGASQSLSGSPRVSPVASPAVHLSTDRQLEAGETLADYLDKRKKN
NDAVKKCRARKRMAVVATEEECQRLSGENASLRDRVGSLEAEVAYLKNLLISATAVAANLA

>Co13_Atf2 CAOG_06753
MTQRLHERRNAAAATTDAATASSHNNHHGHPHNHAHAHHHQSHHSAAAHGRHDMQADHPSPAVFTPPAFS
DHHAANHADRGWGNFGGPHLRSHGLTPGASGHHETSNLLSIPGFAGPLSTNSELELIQQISPFDASFAQV
LTDDGDRASRRSRRSKENGAAASPGVIHITPASPRPRSGRVRMEAHEPEQAADASSHETDDEPSAGPRRR
HSERNGRRQSPFTISTAPEQQLSARIITEADMEHSSNAAHPRRSARIAASPGLLLSTFQITEVITPSVLS
RVALNTPFLNEVHSGMLDLDNMPHEPITASIVHTTTPAQRVIELAKPVLLPGRQRRSLNGTRPTLSSPGV
DASSSSSPSLSASDATSSNLPVFVVTSATAREALPGTTSAAAASAAAASAASASHRNKPTHPHTAVVGPP
PITPGTFDRFISPLFQSGEAASVANTTSAAPSSSVKKSSAAKAVKPSKRSAAEADLDVNDSDADSTSEFT
ARPGATEEERRQLALARNRQAASRCRAKKKRWLSDLEQREQSIQMHNSALVLEVEELRMEIERLRQVLLK
HTDCAVTRALQASGGDTGMEVKPTIASASNRVAQIEQNAEIPTKPTVVKFALRTPAKHSQSTYKYAG

>Mb1_Oasis JGI32766
MAAPLSDVLSLPMLDVDVDMGAELEAVNDSELVVFTPGTLSTTFAALEQEGSVDGAHDALNMGDYDLEPG
TMAGASNRASTSQQDDTLLAGSMNLDGIAGIDDLVLEAELPGDSQTTQFHLDAATMNAHAALYSESAHTS

```
PTLEHEQASSVFSDEDSGHASDHSSGSEALSSAAGHFAVHLGEDEVDATVEAEVCVPRAGKRELAKRPFA
KRSKPNFYEDMTAEERRLVEREGFKLPKDRALTKAEEKELKKVRRKVKNKISAQDSRKRRKEYLSQLEDK
VKSAMTNNRSLKTRVSSLERQNSNLMEQINELHARLARATGGGSATAGTALMLVGLCFSLLFPADLVGSG
GLELGGPTPFQPFAPTTDSVKGFQARTLAFGPGPAGLMDLDGDGDADYVRHPGSHASPVPPASSAQSAAD
ERLARTVIDELGQAIAAMNTAPSTEASMEHLHTAFHNINRSAELDENTDEAPDLPTFTASEAVPMIAVGQ
ASAQA

>Mb2_Atf2 JGI37668
MALEEILDGGDLFSGLDLPDIDSIINNAASSSSEVPDDALWSNLKNNGKTIDAASSSSSSSKKGQVMKRKS
RASSVDPPIDERRLKHLERNRAAATRCRERKKQWLQQLQQKAATLTTSNRQMHEELKRLRDEVLNLKGNL
VQQQGPNALERAGIQLSKYSDPNFDFEAYTAQALAEQERQAAEQAANSKRAKSKKN

>Mb3_Atf6 JGI22289
MAPTAFDEFGSLEPQDQLFGADLNLGLNPSIFDPLMLEATVPFDDGAEGSLSLDDVTGTGVCEESPFPSL
SHDPAFLDAHATKPPRLDSLVSSALSPPFGEDCAPIGQDDHDGSIKLFDTTHELPTFADIVQSTSEGTMS
SAAASPLSTTNATSPLHHATAPTHTHAGSTQSQVQGVPIGLAPPTPPTAYPLSTTSARTTNTPLAARAEP
MKRKRRARKAIDPTALSPVELLEIKEKKERRMLKNRESASLSRKRKKEYLETLEHQLHDAQQQLGRAQHQ
IQQLQNDNHVLREQLANYHGFVNSQPSLQQQFAAFASQRPARTAAATVMLAALLCVAFVGAPGSNLLLRG
VEGPTSPMAVTGGAWHARTLQAAQKVPLLAHVHNAPSDTVYPLGLHHDHLTSSLRDGPPRAMPSNMLNPL
LKMHEHALEERMMLDGVRIPTPEPGHQSLVDPEEEELDGVPTLRIGHNQRDALETTVHRKDDTSYIFCTE
VKVISAQAASSRSRLSLIVPTTTLDDLDNEGEAAHLMQIDCDVVGTKLIATENATVVG

>Mb4_Oasis JGI30420
MDSLFANLHGANGDALSASDLMSAASDVWISDSASDLHFASDVFLEDADASNTLESVFGEPIACPATALS
HQEQQHNAAQLLVLPEQQQPQQPQQPQQQQQQQQPPQHASNASLLTATPLPITATAPGLLPASTASTTTS
GKRKCHATSHGVAPPPNPTKRPYTATPTLPQHNPSSARPLSQAPAQAYLAFPGQSLQLSHQQRQSQHRPQ
EHQPQQMAQQQQQQHIQRQQQHQQQQSHQQQQLQQQSQPQSQSAQQQLTELRPQAGSAYAAHPNTTPSLL
PPPHSGNNHASSSGAQQDLAPPNTLCPPLSSSAFASAKRQGLTPEGALAQFLTSITDEERALIISEGVDV
PLSGLITQAQSRELKRMRRKVKNKLSAKDSRRRRKEYVTQLEEENAQLRARLVTLHDQSMARQSMPATTS
SSSSTTTTTTTTTTTTSSASPSRAHSQPSSGRRPASRQHHQTYAMVMIMLALSSSTTRLGSQPLTGAPSP
NLSLNLAEKMGHGHNAVLASSPLSGSDCESVDDLYGDAGSELGLAENPLFSDLASNQDADLIEPLSTLLK
QGTAGLNMHVANLPVALQQQLAHVLSTAGPQFDSGAIEQAIVLAPKSNDNLWMQDTAPKIEPVSQ

>Mb6_Atf4/5 JGI31254
MTQPDPAYSVSTTAADFDAAIAASNLSSIAPTDAGLAQAYSTSSATGVNQNSSLFLDSPDTDALFNADNL
LQAATFPSTGYSLPGMSLPVPDADSFAEPFFAKFLTESDINAAFPAINPTNRPASSAAIGASTDLPLLPN
ASTASFPISTSRAVLPELHQSSALSAASGAPPLLGGAADLASPTLFAHLSDDGVGLRADALSQLLDSAPA
RPITRPPSHGHTTVSVPSGPVIAPTRPTAPAQLSPPGQASHTPDTGRARTAPPAGRSTAARPATSNVVSV
AATPHSSDTHEDDDDHAGSTSNPNKSAADRYRKKKREEFERLQHDTEAMKAENLELKTRLSKLRNEAEFL
ANMLQSAIQYSPNRLQGLLHMPARYVTQAMLPEGLELSTFLSKREVIAALLTQIEAPDVLTSD

>Mb7_Atf4/5 JGI32251
MALQEFDFSDWYAAAPQGSHDLSEMDFGSSLSSLDQSKANLTAELSGETTSKRAKTTDKKQLNKQAADRY
RRKKRQQFEELQSQSSELADENKALSVKCERLENEVAYLKDLLMTTIKTSQTAPATAQPSLDGLDATVDP
AVLEAVPPALHTLVKTLVNQRRSVMEARFAALEEKIAHLAAKCC
```

bHLH

```
>Co1_MLX/TF4 CAOG_06632
MPLSSSSSSSSSSSTIIPSSASSCKPRVLLPQPTITVAAPLQQLRLQAIPLHAIAQAAPSAGVCLVVPSS
NTQPHPEPPVSYLTPISERSSLALAVAPPLPKPPTTRGLDIPSHSAASPAGFNEGVRAQSPGAGVRDNRV
KLAHVASEKRRRDSIQVGFDELRTVVPDCSSCYSKASVLQRTLSYIRTFQRANGRVSGAVPTAPAATEDL
HKQQPPLHHHGAPNQVADALLQQQNAMIAVLRASPFLSEAQQNELVAQFTAKSSVDLWANMPNNAPAFAS
GALRPLTPACSLSTASMSDDDDEPATPSSLSD

>Co17_MLX/TF4 CAOG_00266
MLHAIDLKPHGLPPLQAMPMQPMQQAMTANGAANGAARKPRATAQKPLLAMAPMPMPMQQQQQPQQMMHF
QPQQLLLDDAMLLGDGMDVNDMMSFMNTRAAVDNDDDHDDGDDFDDEDDALAAAAPHGAASFDFASLMPQ
PLMPAASTAPFQTGTGLLPSGPGVDDLLASLMGTSSVMQPQTLLNPNAAPISLEQNANLINLIMSLGAAA
PPSAQAIDASAGANLAKPSLSAAAEIRRATRVEREKRRQKNINSGFDVIKELVPGCSDLKISRATVLQRC
ANYLTHAQQALGQYKVVNEHLKSENSRLVAENARLSALIEGMMKNSLSQFGITSQSFAAAAAAAAAAAAA
```

```
AAAANTTTSIAPGTSTLRPANPSSLSPASVDGSIVSRSSNRSSNSVSSASSSPSNDAAVHSPHGSSPSPS
SSTAFSGVSPKLRTLSPLLPHPTAPSVDLLIPPADLTEPGSNVYVKDSTRILLVLCVFFLIFPAVGFFSR
EDPHLIAGGNVPSFSSGRALSSVDAGLDGGAAIPTTTDDGFLSGIFSFVPLSVVYHILLCFRCLAAMIVA
RLILSSLEPIAYRSETMVSQAAVLRKAGETAYKEGRFEAAHEFFVQALGTLGNRYSFSSNWRSALNILIE
TGHHQLSKWKVGQIVDRLHRWATRTNTSSANHELARIYHGIHMLECFQHGREKAAIYSALKALNHAEAYE
HIYGATTPVLSQIYCSMAFIYDRIFPQFNYITRHYRNLVNRRILDLGEDHARGVQEILQLQRSEGHWCQI
NPMPGLFIVFILLLACTLCQFTAFFFSTHTYTDSALMHNAMVVPTLILFPVAVLTTMP

>Co18_MLX/TF4 CAOG_07309
MSGLAGSDEVLRERVAMMMDRSSSSTNLSAASSTFSSMDSMAVSDDDDDPDENDDRGGTTMMSSSSSSSS
NGNGTKPGNGTGNGAAAARIQQAAAEQRRRNSINSGFDELQRLVPVLQPKFQGHRFSKATVLAKTAAHVS
TIIRQSTAYSEEVNLLRQEAMALAIVNARYQMAIHAQQSASAQATDAMQLCEADTSFQRYEPALPHGNVT
SSIPAPIGWLPGQPQLSAPAPQVLVSTALGSVSPSSVQQISVTSTPLPTPDEAQAHLVKFQLFRHIIENL
FATFSPAVSLNSYDAMSASILAWVEEHCHPDRLRDVSLAAFCTVAGDYFVGPPST


>Co28_MLX/TF4 CAOG_07309
MNETNEIAVKQSLAHAQVPVPPPPPPSHPTVQLDESLARVFSHLSLEYDATVDAVVPRWKKFSRSNSYRG
QTLAASTASSSGAGTADDGRQNDLPSSSSALYGGVTGESEPRYTSNEAAQAKTAVYLEHIEKKRLDNAIW
RAWHLWRVHGITTPYVNFNVDRAALAQAAAAKHKNDVQVLRGQLARIAREYHSWRAFYKSDVYRQIQRRS
EEPTAPSTSLMPEVPAEPPAPIDMSTEPFSMSTTMDVEPSSGFALLESLASADEENLLKDLVENDIFADE
MLGEALNGLSDTLFSSLGPFAIMDAARANAPGNLDLSNYSTANKHSQVSRPLRGLQTIELELRIALQPTL
SSGPSVPPVTASYQSSSAPNSRRNSTTDVALSKLAHVAHEQKRRDTLRSGFDELTSLIQLPVAGHTSFSL
KPNQKLCKATVLSKTVEFIESVHRQNAAALAEANRLRQEIYALKMSISQQQAALLQSGANQAMLPTRPPD
PTNIMYREYVQATSTQNLKFWIFSSIADKLFATYSPIVRADTPDAVSDTMAMWLSNHCQIEQLREATLSA
VRAIGSALFPGTSSSVSTALPIMPKPR

>Co19_MITF CAOG_03375
MAGRGRGGRKSSSSTSMASLDGSTQQQQGFRFAGDPASAAAAVAAAVELQLGAGLSNTAAGNARGPFMVP
PAASGPASSLMLGMNMHSLQLGGGGGGLGAGAAAGLQNGELDAAALHGLDEFGAHAHQQLSASLPMSSML
DHSAVGADGLGSVDDLLLNFPSIGLQLTPNSAGRTTAPMRSMSMLGVAPGSAGPVVQATVLRDEQKRRPS
SSVPSTPLRSQMRLQLQQQQAMLGSRSRSSTLLAQPLPSAVKQEPFSEHEPFGNEHDASGGLDSQMDIAM
EDQQESGQSSLADSSRTSSISAGQPPNLASAALGVFAPSSFNGVQPAGAFPEAVTVPRQVLNVRSSLAFP
TAYHLHEKQRLQLQEFISTTMNQPGDAGMSLPPGLVGAEFGTAVSSTAASSVIGEEVSYDFDDGQDEAMS
SLESSMISQSGVPSGSGMFNFGAGMAGSAPSDYDPDDIQTLIKSRKKKDNHNAIERRRRYNINDRIVELG
SLLPNAEIDPKASKGSILKRSVDYIKYLQDINRSLSEKLAQTGTVVELPEFTSTGE

>Co27_USF CAOG_02389
MNPAQTSYFNPADATALLDMKPSRVPLAEHRKQMAELERLKIVMANSAHLMGAVSSAAAAASAPASSHHQ
GFMEVFANAIPAAAPLSARSAVSEEDESLDTFSNEDGTPANLMDDLLIDFDVNAFSSMNVDLGATTTLPA
AVPAAALPSAALPSAALPSAMPSSLPAAATVQTATANANATAAAVAAAFVAAASANAAAPGHSLPGPFGF
NPLAAAPMGAWAPGQFNPSVAAAAAVLGLAPMIDQSNMMRKQRRRDSHNAVERRRRDNINDRIIELSALV
PDNGGEKANKGETLRKAVDYIRYLHQCQARSFEQLKEATKQMQAMNAEIEALRKITGAAPSQQ

>Co8_SREBP CAOG_05901
MLSIQQQQQQQQQQQRIVAAATAAAVAALQQQHHALSPSSSSSTSSASSSSAVPSSSSMAGDAVLASATA
AAAAAAAVASGSAQHQHQQNQATSSSNAAPATATATAANLNLFNLMNMDEDDPLFTDFVLSNYYGFGDLD
DLSALTAHQQQQQQQQPQPQLGLFNAANANSIHLNSQFGVLGGLGALAGLPTSANPMLNAAAGLDVATVL
KQQQHFQQLQQLQLQQQQLQHMHHQVSRASSVSLDAMSASVDGSSTSGMLTSPGSPQSLHSHLSPAPSSA
SVHDDHHIAAIYNNAAAAAAGPSASASASAAVAAAASFAASPAGSAEQEAFVNAATAAARASANGLDPNA
AAAAYTATMAAHSSNNQQQQQQQQQQRAVASAQQAKSPNPSEGVSSPPSMTAAAAFGGLAHSIPAMTAA
AATSPKAASPLFSFGVLPTPTATGASAATVTGSAAAAAASSSSSKMAISRLPAAAAQTTKAVPKGAASGA
TTTPLMPSKTEKKSAHNVIERRYRNSINDRILELRDVVPVCREAAAMVAAGTNTVKIKLNKATILRKAIE
YIVFLQGSLERSRHEKATMQHALEMHGVDATQLLQTIEAELGAIKTEMAFSAADFDDNEELDEEFRDLPS
LSNLLESEDSNEDVSPSDSPTITAMTEAVKPAKRGRKPKAAAVAVGDGSTPDGRRKKAALVIAPASSGGL
LGGSTVGAMRDGSRMMLFVLMLVTLFVSPIDTLVDTNAPAEHLAVAARTLSSVEVAQLSNTIWMQGLLMT
AAWWFLRAVVVVVCLARVFFGEPVTSSKSLEYHRAQELQKAALANLANGDRAYAETNLIAALEILGRPQP
TSTPELYVAVAWQIIRQVLHRLWIGQWIDEFVASRNSNARVSFKAAAGCFHMLQMVQFMSSTSNSGHSTG
VSSSTTASDSDEQSSGNVSAKAPVSFRSSLHRLNVVLNAVNLAEASRGSVSPSFLIEVYSTAALQLKYAF
PRRFHFVTRQLLMLAKSVRADAGDLDPSLGWLFHTDGYAFISRGLWRASFARQAALSRTLSPLGLLGSMY
RTALLDRAFAVLVEQGNALRALDLMQEARQYALDAADVSAEWWAYIGLSSAHLRLGHTTEAIRAFQTAEH
```

```
LLSSGTMNEPVTQVQQALDQALRGSIALLMDDGVTAIRCTEMASALLRDDPALVVIRHRAAAASAFQQQQ
QQDSDGSSAQGISDTDSGSEDDTVSDLGQRRRSNGLALSTGSAPLQQLAASSSILPHSLTGQILACVHLM
DNRFALWQWLRASHLHPARVQELPLTTTAQRLSAGVNGDVDVLRCLSHAYPAAHSLAFLYQAMARTLAGG
SSQRTRAMFVRALAEARSATSPFAEACVLLKMGQLLVDTESQTNDAAAAGGAADRSDVPTVDERRHWLSA
ASHIFESLHATQELDLCRSLLKALLDSDK

>Mb2_SREBP 32649
MWRGCWWLGVGLSLIFSLSLSRARPDLNSNDSSSEAKLVAVLLCYHPGYLVSRAAVTLLDEISSSLPHLH
ATPEDDMHHLLHDPSEPLMHFNANDVATTLASTSMPTTGMPVPMMGASLIPEHHISAAVDHRVFAGKRTN
AHSYAAHDDVHPAAKRAAPVAGVDTNNHPIITHGFPDIGPVVSQTLSAPPAPQPMVNPNSPELKPAVDDL
SSSSVPENSDAASSLLPPVVPPPPPETFKKTVIKRQGPLKTAMDKKMAHNAVERRYRNSINDRITHLRRL
LPPTWTTGPKVRTNKAAILQKVTAYLHFLIENNQELHREVTSLREVLAAHGLADAGNAVPTVSTNPTANV
LEDLLAFDSAYDEAYPPTEPSTPSTVPAPAAAAMAMLNGTRRVQTGNMSPPSDVDSGHATSPNAPSSPGS
VRSISPGAPGAPATATRMFMVACVAGFTFFSHSLSAGFGGAVEGPAMAMGPGRVLQSVQASGAGVNTGGD
ESLGLWSTALFWCMQVVSLLLLAVVIVIRDPVTESESAKIKAAHEHERKTAVARMLQHHETASWHADKAL
ECLNQPQLGASRMQLGLANHVLRQALHTSRLGLWAERYFASSHGPAAAEFLAVLARTYHHRHQVMLDTRK
DAVVTAASLECALAAVNAAEAAGSQLSPVELGRFYAALAVELVLQRGGSSFLAGMYLDKAKAVAACAAAT
ASLAREACRTAVGPEQADEEDAESDHDVSLDNDGEGDPMTDPMANMEECEFAWMFEPQGRTYLLDGTWLT
VSALGARKPGTLLSLSVAFRHHLLQQGMQDLVAGRDLALATSRLLELGKHAHVANHAHDHWWALATMVPM
HWRQGQLREARSKLVQAERLCPKRDLLQESLYAVLRAQQALLDGDHALCWLALQRASCCYARLPPAHPVF
APALDEMARFLVLRQVMGVRVALYRLRSYLRRLPEAEVDALRERDGSVVTFDVMLTALQNDVDLATAFAQ
SLPYSQPQALLLRGICRCLAGARTEPTEQLFQQSWTLAEQLEQPYDVALAKLHMASCLSSSLPKALLRTQ
LEEAMHLFERLDAVDELNSCRKLTKLLDSR

>Co5_MAX-1 CAOG_05983
MSDSEDADTSSGEDVPQQDTSNMTPAAAAAIAALASASAKRSSGKIEKASKKKKPPRDADKRHSHNLLER
KRRDHIKERFTSLKDVVPTLDEEQNPSRSAILHEATVYINKMREATFSEEHQVRHLQTANQSLAKEVQQL
EAEINKR

>Co26_MAX-2 CAOG_01784
MEENTISQGDEYDSDDMEDVQPDASGRHVKRAHHNALERKRRDHIKEKFNELRDTVPSIAGDKASRSLIL
NRATEFIVTMKQRNTAHEAEIDAIRKQNETLRKQILDLENGHS

>Co6_MYC-1 CAOG_06158
MSTDALLASHSEIERALAQDAALLDYLSDDLASHHAVSAPGSPTPVAYDWTSFDSAPASPAHPTSHVEAS
PQAKKAPVAVASMLLERRGLSLDGATNLRMFVGHQNACLQASASVSASCNPSPTPMMDDLFSGEMVAALS
AYNTPSATPRHSPLHFRESADGNAFDVVPRSLSAHTSSASSDSDDDMISATMASHSSIDFPTLFISDTPS
VGGAAHANAVLRPSALTLSPDSRPASPLPADHAVDTALSALLHSRLNNYDLVVKMDAEDDDDELVGVFDD
EQAGFALPILNPAVSAATAAALRAAAASAPCSPFYDSGDAEDDEEDEADDEIDVETVDEYVPKRPLRKSA
ANSFVATQRRLQRGSASLPVSPASSPARRSIKRKADDDMYMPTGSRVSVRLQAAASLPSSRSASPLPFPS
SASYASTSAWAAKRSCTDVSVLLAATAAVTAAASGDPLPSELMNALGVNPDSGKPITAVAVAAAAITAAV
GGPVSEFDMKRNIHNVLERKRRNDLKDSFTELREVIPEMAEDERTPKAKLLKEQHALEERLRQLRSRKSR
A

>Co21_MYC-2 CAOG_03433
MTAIANRDDIPSGGALDDLGDLDHLHRADHLPDDFLDAAFEEFSAYSSSFSVSAAPSATASPYCSPAPPP
QPKRNLALSRTHQSRPLTHSASAAAAATASANAQRPPPIGDPATPKHAAITSSSTTSEGQTADHHHNSSP
HGVAMLAALDDALSQDAHASELASAALAAADAGDHVLLTAAQLDAFSHHPEDHTLPVDDTALRSISAGMD
ALANLPDDPAAAAIQAWESHADLFDPQLLDSHDDALWAGLLLNDDMSDHMHLGSVAHLDGANAEAAAAAA
AMDLSDDLAFMQHHDSEAHLSGALAALANSDPSNQQHHHHQAYSSPAQPSGLVLTLKTKKQHSKSTPSTP
VTTPTRKSTRLTRITAPTASPSKAQRKRDAAAAAALARANTLSGSVGDDDMDEDVEELEGFEDTETRLTK
VEPALAQAQALRAQQQSFGSYSAAGTPLASPSMAHRGGAATAAHLGVPLAPQFGSSFGTPQRPSRPAHDE
DDEDEEIDVVGDESDDDTAAQLPQRRPADSAANGRGPVRSNLVIPQPAFASPAAPIFAQAASSYAAAAST
GLTPRIGTMNLLNFPTRTQSAPSSPLQRPSAVATAARKRKNELVTPSQLDMVLLSPTKRARSVENVAAAA
AAAAAAAATTAAAAAIAAGGMSTPGTPNPRSGFMTSATMSPMSTAPSSLNPSPSTSPVPFNEDDGAMEGD
ATDSVAKRRFHNVLERRRRDDLNKSFIMLQKKIPELVDTSAAAKVHILRKAAEYIRSLCDREVELEAAKA
RALREHDELVARLQALRGSR

>Co10_MAD/MXD-1 CAOG_03883
MAALLQAAELLLQREEAQHFPMQGSGGMPPPPAPSSNQRFAPGAPMPMASPQQPHHPMMMHAPHGYAPMD
PGYHAGYQHPTHPANRTARSQSNPPQPLAYSSLSGFHAAPPTRTMPSPQAPMPHGGAFQHGMYPVAPPAN
```

```
AHHYMPPGMTGSSPSGKSLGIPTGAAARRRRNSSMRESFSYDVYEDDEEVDEEEEDEDEDDTEDGYNDGD
DDFQVSSSARRTPPAPARSGSRSRTMSTSFSSGRRNSESGFQPSSYPGPFAAGHSNPATMSPASNSSGSS
GATSSGGLKNANSGSTGSLGPADASPGGSSGGRSGKKRSVKNREQHNMLEKNRRAHLKTCFEELQEAIPG
LDMAKPSTVAILQHAKQFIETLHYRQAQQQAEIDRLQQLNVWCCSKLDALGHPVPAAYLQQPMQGKAPQH
HMHPAGEENGHYYESHHHHHSHHHEHGHQMDMESSSTSSSTPSDKSTDVDVERLDGNSPEPPAHGTTRSG
RTSRGRREQ

>Co13_MAD/MXD-2 CAOG_07879
MSAAVLLLEAAKILAASKQDGSGSGSSDEEPNNHSSGNNNGNNGSQPTTSAATAAAAVAGANTGMTAIAP
VTGMAGMAPLNGSQQQQEAYNLPGQPPQPLRPPPHPAMPGSNTMQTTAMMMPPPPPLLPAQFARSAVPLF
DPYLHRKQQQLQQPQLPYDAAHALLPPPLQHDQYAGLGGMTGTRVRASASFSGFGSTANYNPNGPSAPDL
NNSGGARGIMVSGAGGAFPRMSPNELGGKRSRAASMSAGDSPASGGGFSPASSIGGGGSSTSGTQWVKNR
ETHNLLEKNRRAHLKDCFTALKDAIPEIANNPKTSTVAILQHAQNFIVKLQFMTDSVVANINAFQQHNQQ
AMALLQQHMSAEEYEQVVRHVASMAPDGPASTPAAGRGASNGRIAAKRSATATFSGSDSDEDKPEHVARK
ARQSHAGRVGGAAKPSNGRGGVAKDPTRAVVGDLMINPHQFALTTSTMTSKHQMDRSLANRRRRTQSFAA
DTTLSRQRERSVQLRAPSPAPSENSEAPSSPSSPLSTSFNVDIEGSSESEDSPQSTSEAGDAEQVRKKPR
RLEA

>Co3_ARNT-like-1 CAOG_06106
MLGSQRKRKASSSTSHPGSQGAAADEQARSRLRLGDSSDEEDTSDQEAMQVTEPAAVPGASNGGRGRGRG
RGRIVKKTQSYRDSHNAIEKRRRDRINASLRALARLVPECALLESASSDKGKLDKADILQYTVQHLERQQ
QSDPSSSLLLPSAKQPHAADTPASSPGLTTSSSTQGQPLTLNTMTYRDWCQLLIDRDRDLVHVFCLISGE
ILHVSHACLATLGYSTPAEVLQLNALQWIHPDDLDSVMAHVKTAMLELLSVQESPDAVDLARQQDGGVAG
IIFPASRSAFVDLWYRLRSKSGSYQQLQVRARLVLFPGSGAPMTGSAEQTTGHPSTGISPRSSATPSSGA
RSSPAETSSSTSYMHGPLTRAEPTLVLVALGRPMGVQQHKPDPQPGTGQPITAFTSRLSLFGIFEYVSSL
CKPLVGFGAEELIGQSIYLFCHPDDVLAFSQMLGRVLGNPNATDVVVVRFRFVTKEGPYVWMTTKMHVKL
NSVTQAPECIVCLNTPLLDANAATESGAMAAPASSRSKPLEQHIPTPSDHHLPGSASSMALSFGFDPHMT
QGLLQPTLSNNNNNQPSQSTSSSLLLAQAQALQASQLGNPTAIQTLQQQQQQQQQQPHHPHHLQHQQKVA
PSQFFNTPPNQIADTSKGQPAVHATVSQSSLMPQPFPTTGPPEHMLSGLSSFGLDVNFFEPNSALLDEAL
NLSVRNKELQMQLDKIGVAGLSGGLANGLQPTTSASGAPGFSPNPVDEIHRRQSLAKMTNRGPQRFIPSI
ADIDKFLVTPPPYGSYAANPGTVLPWVQSVQQQQQQQQPLSPFPQQQHQYQPSQPPQSLPQQLPQHLPQ
QLPQPGQQLPDQFASFPQASTAANPGPSAMRTPALRIQHLLGDDTGMAGMEPPMSQWA

>Co9_ARNT-like-2 CAOG_06970
MLQHQHSGAGASHLPLGLGDDLAAMVGYTTDNDMLLEAIEEHLAATTAPAAPVPGSLQQLMHGQPTAAAA
AAAAGWTADGFGPAARVQQQQQQQVQQHYHDYHYPQQQSQQPQQSQQQQQQQQQQQEVVSAPTATTATRV
TLHQQQQQDALWSTAPLSLAAVQHQHHYQHQHQHLQQQRQRRQQHGGHQSQTLSLDDALSTVVAAAAG
LSPPASALASVSASASAAFPSASAAGVMAARVASSEASVEASSADDSGLLAQPTTLLLDDDQPLQQQHQQ
QQQQHQPQPKRAVKGTTVSHRQIEKRRRDRVNEILANLQEMVPECRKSEGRLDKTSVLAMTLRYLKQLVG
GSSSSSGGGGPSSLSSMDPAADDSTDSLDGNIIDRSSEPGVSSPVSLALDTATPREDQLLTTVNELARMQ
SSPGTRADDAEWHQLLLDHVEDLIIVLNRDGRILQISRSCFPILGYRPAQLVGKFLFDLCPSPDSARVRA
GLRHRELDPIKAVDYAATAELGATVINSRLVPIKGDPVWFEFKLREMSRNGHALDCLVAVGRQCSALRVL
HFGSPLENSDEFSARHDLFGRFLYVSPNITRLLGYQPVDLMGASPYRFHHQLDVPRTLNAHRGLFDQPAT
QVAFRFITKKGNSFCWMISRCEFVRNPVTRDVEYILSTSRVLSSSEIYLLLAGKPTNTSDPITALPAPLT
SQEDAEVCKSVYIDDQRQVPLVSNADEVARLYLDEPTKNLSEKSPLSVGVYSSSAPHVFPDLSGSNSGGS
APEQRSSEYQTRAQPFTDAGTTNFGNPAVGTVVSDLMSSVLTLQLENQRLHEELAKQASSRSADAADQPS
TMAVVSDL

>Co15_GroupC-like-1 CAOG_08129
MCAATGPSALLASDSPVTLFLNDGLLSVATAIPFVIPAGTSIQSAASMCHQALFGTPSSAAKLFTDLGVE
LMPVSMSRTPLVIPANQALVVTGGEPFFPLTMHSQIYQDATGASVAAGHTLAGDSQATQDRLSGIYNESR
RVQEQQYRLLQSLFGQLEQAQTEQQQKQLDPLELASASQDFLDTAAPLVTHDFANVLDSAVRALYRNDGE
IGPDKLASHVSYAMHDLQLMGQPTSTLTFPSAAAAAADVAAAAAAAAAAVDAAAAATTTSSATSSSSSTA
AKPHTGAVATQQVSNSIEAIFNNMALPRLNVPVATREQLSVANWTSQPAPETMDMCTINRKLQVERYRDR
LNLGVADSVYRVNPPVRPGALDELSKAESPAASETESTGSAAHSTSTAAAASMAAPPAAADSAHHEEFSM
DMSSDEESGAFSMGSGAGAARGKKVKERGMSHKQAEQRARDRLRTSMQELVLSLPAPENVKNASRATVLK
KAAEFLAFTRRRNDTLKAENDRAKREIMELEAEQRRLQQQIQAANLMTVEITDKNHRYVFVDHMWEIVTG
YQRTEALGKLVRDVAGCPTCALMIERGLEIMHSIEKGIEWSGLMMSRRKDGRIFASLAHIVPVMDETGQV
IQNVCTRRHFHLLDPMQASSVCTFEKSLEAEVQSAIPLNVVSPEIVAADVSQAFLSAPIS
```

```
>Co31_GroupC-like-2 CAOG_06284
MAKVYVTLFRNDGLLEAAQARPFFVFDDWHIDEARSRCHERLYGTPNSKARLFTDLGVPLDDVAFIPAKQ
SVVVSSGEPFVHMTIRERIAAAQQQQQQQEQLHQLQQQQHHHQVVSATPVANNLPVTTFSASPLTTPFAD
EDAAALFVPSQGPLDVGALLPGEDNLIDTILNDMYAQNTAENADFGYPMEDVLNRLANNLGSYQNEATAF
AQRLFDHSQGHESQPVTEQAIPYMQARAQQQQQQQQQQQQPQKQPSQQSLMDGDDEVSTWTVEDAQLQNI
FDNIQFPTADSPVSVFNQQTSPIQFTTIPFSPQDASYNSTPRQQDSGVSFSLGNGLPTKTSASQTTTTRR
STEARSREKVRQSIAELVLSVPALASLQSPSKATIIRKATEYVNQTKNANNALIKENQQARAAIIQLQQE
KLRVQHSVQVSNIMTVEITDPDMRYIFVDHMWEIVLGYPKHEVIGRFLKEVVGCPTCPLMQKKAADIREN
IEKGSVWTGYILGKRKDNRPFACEATINPVTDATGKIVQYICTRRNFHLLSVKECQKVCEDVEKTPRLDI
GMNLTGGIALTEESLAALTVSSATGLEYEADTMNLCPLVAE
```

```
>Mb7_GroupC-like-1 26507
MTPRAAKRAKSVAANPSTLASTPSSSSSPPTAATSVPPLPAGGTGSDPDEYQDYRKRRHRSTESQRRQRI
NDATEQLRELFAIDGKAERALVLEKAVNKVQRLQDEMRRLAEQVQTLQQQQLASTPSQSLAPNSGIKQEE
EEPATATASRLSSPLTTASESDSQVDDTVTIQPSALYDSFDPCDLTVTTSAAAQELRAGEWHPLPNALSP
NPEYSSTDRGVTVCPVTGVFLHGDNRSPSTSRAPSPFHVSLEDRLGLDESDSFDVTQMELASSILAEARL
QLTRANVPDPTNLESSTRSRSGQYTNTNQQTVFDGPDLRRNPDAVSCPLVQALLRRLLSEPELAAICQTV
APIDALQQRLQTTATQMRRHLIMQTKEYAALKHLLRRSSTQIVLEELSQYQRMYHPNLELQSLSNPVALC
SVSSHGVMLDCNSKALKLLECRRDDVIGKLLWQLKINPDWRVRKTQSVELVTGACSTIITVERVRTLHTR
SLRWVRQTLLAQDCVPPSDIGDGVYFCLAMLQPLDEAPSIGPHMLYRIAQKASSSFFIVLSLLSLSLSLS
LSLSLSLSLFFSLQCRLVS
```

```
>Mb8_GroupC-like-2 26679
MDEVTCEARFNVGRFAMLPGPSHEFEKWRRSVCVCVCVCVCVCVLSINKYCWAFPCVCVCVLCVCVCCTA
CERRSGSLSLTTNDSFFLSLSLSSLSLSLSLSLSLSLSLSISLSLSLLRLRLTHSSGLMDARAPKRAH
VIARGGEGTTGGQTRSQGRSQARNVKDEDSPVPEDSADRRLQHRRKEAHRRQRINQALDELRELTHSAHS
DEKSTVLNRTVQLIRTLQEENQRLRAASTRPSEGVNPNPTTPAQPFERPALHHLPPSIQSTAPASQGVNL
LPNGLQQIHNGSQGHSHGSPRLFSGNTTPRSFADLESTLDLPFANLPLPLLEDLDLSDLNPGDTGDGVYR
HRPSLGAGASGIGAGAASGSASSANGGPAYQRPDADTIRGEILQFVSQREAPLDQFVLPYPLGLAYVSTR
GVVLDCNRAFLDLYELERAEVVGRFLDATNTLVDPATGIGELRALVQGQRDVVAMVRRLRPRRSQGIIWV
KETYRRHGHVPPNSFGRRLKYGFGTFERVEPPEGGMPLVLPEQLGNSRSRSD
```

Mef2

```
>Co_Mef2 CAOG00636 reannotated
MGRKRIQIKPIFDDRSKQVTFTKRKLGLMKKAYELSVLCECEIALIIFSNQNRLFQYASQDMDNILLKYT
EYSEPHESRTNEDLAAQIAKKEARARGDSSLRDDSDSDTDGEDGIQSLDQKRAATDQIRQDFTLQSQQQQ
HHSTISANRIQQLQATSNDALQQLHNFQQQQQQQQQQPFSLIPPAYVPHPETLNGQPNLVFAPAPLSASS
MPFGSALGSAGYFPSYNLGSAELANAQAGSTLPPHINLGAPPFYTPQFAQGGPAPGASATMPTPPLPPPP
VTTAKSRGKK
```

```
>Sp SPPG_03091
MGRKKIAIKRIQDERNRQVTFSKRKFGLMKKAYELSVLCGCEIGLIMFASNGKLFQFASTDMDKILLRYT
ETEPHESRTNDDIIKTIQKHGSGGGDGGDDSDDSAERDDNGLIESTPRDDRSMDEHIPSPRSLRLARRKS
DKDTSKKNNSSGKGGGGEGGDDVESNHGDQGREYLSPQPGTPTHDMRHPPPSGSFHMSHLIHPPYTAVYP
SFAPYPAGFVLATGHHHPPSDSEGSPSPYIHG
```

```
>Am1 manually annotated
MGRRKIKIELIPGERTRQATFIKRRMGIMKKAYELAELCGCQIGMLIFSSSGKLFEFSSTGDLDSILLRY
AEYTQPHEIRTSRHPVRELLTLGCRSSTRSPDSKMRTWARPPRTPTRILTTTTTWSRTCLCLPRIWLRLA
LSRPWLPRLRRNPRRMALLLSSGSSSKCSSSSSSSSSNSWSALLRRQQQRPAPCLGTVPPLTTVYQAKAC
STQAVQRLRHTSTRPRAIRAITRSRGQRLLRNICIRTRSTQRLCSIHRLGRLSLHPRLRRRRTLMGHRCT
LMVRCTLARSRSTIIILTTTRIRRRLWRLRPCILILLLRRTTTTHTLTWRRIPLTDRLRVLRIMHPTLPH
HRLTLGRTTRTTMRRGPRSPSCPYAPTDLPSRFIDPHPHATPGSPLSTMAAQMDPSRTVHSPVLKRSASD
MLDPAEAAALAPTTNEPRSASKRKPPPFASKQPFSLRLQIPTPTSAGPTPARAVDVLPSAGSTYSPHHIG
ASTTPLHPPHDHAPIAAASPTVIPASPHVQPPPQPPLPHPHPAAASSSAPRGSLHAHHAPPLSAHSPLAQ
GAPYARLRGSPPSGLAGTRPPSVPGEATVAPGDLGRHDLR
```

```
>Am2 manually annotated
MWADLRMLSPHAPVATLPGAPRLGVDRLSPRACLVKKFIDPVCRLADFLLKKRRRREGNGRAALAWLATT
TTATSLSTNNTRTAADGEYRQHERPPDHPQAAHPGRPNAAASRAGRHDPRILGAPPPTTAARQLAALHDH
```

```
TDDQPAPAAAAARATSLRTLRARSQLVRASIPIPLPAHPNNRRTNQRPDRQSIMGRKKVRIQPIADARPR
QVTFIKRKTGLLKKCFELSTLCSCDVAVLIFSATGKLTEYASNGDLDSVLLRYTEYTKPHEIKSNQDFLD
DGAADKSKSSAKSIRNDDDDDEAGGDSDHDGAAVSAAGPSGTANAGPAVTPLLAPNPQSIMPAPGTSGVM
HHVGHHPQAGPYPIQLQPQQQPQPPPPPTAQYQQQPFQTHGAFYGHVPQPQQQQPQAPPYGGGQDMTGPF
HAFGAAPVQHHPHQHHQQQHFQGTPMQHSQQPFGPPMFPNDATAAAAMYGATTGQVSVPPQLSYTVDPAT
GYVTAVDPAAAARGNANDQYATYQAGGMFHPQQQQQQHPQPFPMQPQQQQHPYAVQSHAYAQPHPTQQHH
PQQQGYLMQQMPHAQQQQQQQPLQPAVFSSGAPTQLAPAAEPTTGFLSYDQVPPPSSSGPMGLTNPLTQA
TQIVAPSPASTAAAAGSPALTAAIGSPAMTARAGSASVAPSPLVVGQKRARGDESEVDRESTSPQSAKRK
PHFARRAPFSLKVHIPGTAAGEGDTAAGAGAAVAPLTLEPNPNTPTTWSPSLPPPPPPMRSGSTGSNGSG
TGAATTTTTTTASRAPPAIAVRPSSSVAASRPGDPALPASAFPPSLPSPYPAASTAFRGGAQGSSAFLLP
PPSASATAPGPLSALLQRSTAAALPSPSAFVPDHRGSHLLGGASLLASPVFPSPLGVVPMPAISVSAAAS
ATGRADMQGSRYEAGEEKEEEGGESRVKVENEDKGGAGASEDDIEKTVQHEVHSGVVNRMWKSFVECEQF
GVLEDSSRPETRCYCLSRLAPMRHMLATILLPPAIRCASSTLMPYKLQIVNLKSVSTSFATVLSSTPANF
TMNGSSNDQPQSTVPAAAAPNSTNLPAEPAMATVAVIAPTDHGRLEVGNGLVADPLSDAPERPGLPPPSL
NPSMADLPEIDYAAIEAYCQIPEKDRVQHWSEGTRMVLLPRVDGGVIVTSENDRATTEGRALDVEGINRQ
LHVIGRHFRSLTEAPFTIQRVCELVLFPGNGQHKTLEKYLRAIEKVVGVSSDWPMPPANGMALDEPTPTT
S

>Am3 manually annotated
MRWDAGRVARVATPAGACPAIPCACRKSPVPARACTPPPTLRASFKSRRPARDAPLLLPPANNCTCLLLP
PVVRQPHSRSAPQELKSVPAPILVVATPQVLGPSSTADRVLASLASAPGFAGHSQPTPELLKSGNFKATR
LGALLIYSARFRQLIWQPLATSRRPIGFIASFKCVLLEKGVCACRRVLPVHLHQFWNWPKLRESACHGPA
RHHPFPALAVLAGGASSSCAVCLPATQFLTRGTCGRAAPPKRKYTIQPIRDERSRQVTFVKRKGGLLRKA
YELSVLCGCEVGVVVFSKSGRLTTFASDGDLDSVLLRFVDVAVPHETRSNLDFLIEDGHLPPGSTLDDAK
NAFIEQPGADNQPAPPAATAAAAVPPPIPLVATNAQGQPQDPTALPTPAPHAAPDAAHPTATDPASRPIL
MHPMPPGTYPAPPGWYGHGAPPAPFMHPAYDGHAAAAAAAAAAAAAAAARPYSDPMLPPTTAAPGSADAR
PTAAAARPETTSPDTTTPISVAPPFSAAPMDPAVAGYPPRYPGYGVMPPPPVPIHPYAAAHYPYPRYAMR
PDYYAPYPHLPMEYPPVPLAGAHAARSRPGTHPRPSPPPPVPVTSAAYPAAQPDPAACAAAASASTAPAA
SAPGPAAPVVGALTAAPAPVPSSGGATPLGTKRAWDGSDEAVHPPAKRVAAGTENTLVEGSQDTSFSLSA
AASAALAHHAAVVLTSMPGSTQAPTSTAPAASGPPSAPAAPEAAAPVHLPPISVPEDHAGAPPREWPGK
```

Fox

```
>Cow1 CAOG_00937
MSSSGGFVPSPGSSAAAAGAGAGAGAAASGLASSGSVPTTDELSELTNMLLLQQQQQSQSQQQQQQQQQQ
HYLLLLQQQQQQQQQQQQHYSSLLQHLGTANGAGAVASSASNSPGSPLLGTAVSQQQQQQQQQVQNSSAT
VQAELAKQRRAQQSLAIIQQRLEQDWERQKYIQQHQSHQQQQQQQQQQQQQHPAFAAQQQRQPQSFDDLS
TGLAESAANVSSFMAGFDLSELNNLNSVSSEAARSLLLQQTLQEEQNLQQMLLSQTLSLQQQYDQLQQQL
LLQQQQQQQQQQQQQQQQQQQQHQHQQQLAQDFHLDQAAFASADQLPPQLLYYSQNQPQQVFPQQAHPSQ
PQPALPRAVTPTELFSHSSMPTSATSLNGLNNGSMFGAAPGAAPSTGSAFTPLSQHPSLTSLNSASPNGS
GASTPTKGRSRRASRASVSKNSALDLIATSMLATIQSQQPHGQQQSQQQSQQQAQQQAQQQQAQQSQSAA
QSSQSASAPAQPHDQHAGHPLTPAADSDFVSSPANNDFDDILDGALEADYNEHLGSIDSPGSSGPGSGGK
PPFSYATLIARAIQSTPDHRMTLSNIYNWISENYPFYQTTSTGWKNSIRHNLSLNKCFKRVTREREDPGK
GAYWTLDETQTTDLSSRQSKQRRRSRSNSASKGPMPYNPDIKEGEIDTNYLPRKPAKLIISGNTDSEQSS
QNSSAINTPTMSSAILNGGGDFGFVDDLNLSFKSLHTHSVNLSGTLPERIQLPGDDVTDTAAMFSVGMPL
AADLSLGESTDFLNEQLNLLLAQQQQQQPHPSQQQQQQHHQQQQQQQQLQHQLLQQHQQQQLLQLQLQQH
QQQQQQQQQQQQQQQYQLQQQQQQQQQQQQQQLQHHYHQQQQQQQQQQQQQHPMMSPQFMAPSPGSLRTN
ASETGSLADLRNGHNRRNVEDRKNALASRSASAMPTLSSLGRGRSPLSPPQALSGLMQQMQFGPGEGHSG
FSNGAGDVDPQSLDPAVRRRISTGSLDSNAINQAHQLSFMSHPQMSPGAYSPQQTQLVGTPPGDSARLTS
GTSNSL

>Cow2 CAOG_05750
MSDSAAGAGAVAAPQQQQPPPPQQLYMQQQGGAPQPDYQQQPQQQQQPPYYMHGAPMHPLNYMQMGGYPQ
QQPQQPQQPPQQHQQSQDPSFTPQQPQQPSQQQQQQQQQQQQQQQPSQQQQQPSGDMSQQQQQQQQQQQQ
QQPMPQPAMYGQPQQYYPGQPPPPQQMQQPGPYMYAQPPPQAFAPMGYNMYPNPYGQPPPPPGMAMAGMV
RPGMPPNPPHMYPQQPPFVNGHPSYGMDHHQAPGPLFAPGGLPPPPPPPPPPPPPMPPMPPQVLGGPYMM
PPHPSAMPPHMPHPSDPMGAAPPGSGEGIPFEVVETPAFAKLQGPTIDYCIRSLSVTLGRKPSHNVDSVD
VDLGRSKFISRRHAKIEYNFTMRHFEISALGKNGLLVDNEFCNNASAPVPLESKAFIQIGDVGFYFLLPL
GTVSQNVKPTKKKLQVAVPGGDEDAPHSDLGGIQSLGSSASGTGNFSILPSYTPPVSKARAGAAAGSGAG
GDGVSPHSPTGTLSPNESDEEGHDNGLHPPQSLVDYKGDPSVKPPFSYATLIAQAINISADRRLTLNGIY
TYITEHFPYYKRVDNGWQNSIRHNLSLNPCFVRVPRPDSEPGKGAFWTIDPNLQYMFTDTNFKRRHNRAP
```

```
RRPVEWPAGSELSPRGSAKAQKTPRTPRGGNRSAPTTPTGAQLAAHQQQQQQQQQLGMQAMPPMPPHNGS
FDGSSGMLHPVGLVPTPPQPAPLTFHVHAGSKMQPHSAPSNNVSSGTESSPSRKHAADSTPTGDTKRARL
DSPEGKQPHADEAPASNANEAPAPAPAASS

>Cow3 CAOG_01766
MSGFPATAVMASALDPTNAAALEDLLNVVPTPHSDTKPPYPYAALIARAIHDSPERRLTLNEIYKWIEEN
FPYYKEQDHSWKNSIRHNLSLKKFFLKVAKAVDDKGKGAWWMLDYEQLESMDRIGRRRSASETGVRPSTR
MIPTSLLPMAASGVAFGGDAMLSPATPGSSAPYGMPGGDFAQSIGSPNKVATSTGFHPSPKRRTTSVSIM
PMSGMSGNGYFSTSAPTPQWMVPDASGVLAEENHSLGPLPVSAPNESPVRATARSRRANSVLDASSQSFD
DSGNMYAMMTLHPPQHSTDRNHHAQQQQQQQQQQQQQQQQQLQQQQQQQHQQQQQQQQQQQQYQLAQYQQ
LQQAQMQQHHGKFMSAPSTPAPSQFAPLLAGGGGVPVSHSGTATPHPYGFGVNNYSLASTATANATTFTD
DHSPDEHMFESGAFSLDHISGEDLEATIAAAGQLDPSTDLMSAIGDNLFEGSDLISDQPLSASMDGFSSQ
TLTDATLRSVSTGLAYSASSSGRSTPVNTPQVATTTGTRGRKSSLRLNLNAFGEGAASAALSASLSAPSA
ISPAAIPEQPSAAAHQALRGTGMAVTTAASAAAPAQQSSQLVKGSGNRTVVSRQHSGGVTSSRSMSTASD
NMLTGHAGYIDALDIGHTPSSFSGSSTSTNGNSSVSIGSLAVSPTSMSYPAVDFSANLQQYLPEADYMIG
DDELTGGVTDLDWNDIGNPADPDLLLLDARHTAHPAEQQHHYSSSYGQQQQQQQQQQQLLQQQHQQHHSA
GRSSFSSAGLPNKAQPAPPAMLSSPGRLQMASPASMPFSTPSHTSHSEHSESSSGLITPRMRHANDSFDA
GHSQFGNLHFATGSASEQLMLDHSDLPSSLLLADYESLSNLGALPNSGEMMANGLSDLDWGQ

>Cow4 CAOG_07434
MILRRGRTASTSSAIVASSSSSSSSSPKRPARASKTAAAAAVAAVVSGGLPRIVVSTEDASMMLPATPAN
SDSVFMGAANSAMVESTSAFLNALPGTNNAPQIAMKPRKRAGRKPSASTIPEGDEQQPETLAPFTLPLHM
QLQNSSLLSMFPFAQQQQQQQQPPLLFDAQFQQPQQQSAQPLLSQQYSLLAPHYQMMLMSSAAITGNNSL
LTHNPALYAPFGGSSSLPFAQSSTSIPMDDTTAGMGGMAPSCSSSAAASPAASDASVAAPIAPTSSSSSI
TTTTTSAAAAPATRSNNAHLLGAQYAPDGRRYNYFLSRGRKRVALEDGVDYRNCEIKPPMSYAALIEEAL
NSTPTRSLTLSMVYDYIKDAYLYYRNNSGSWQNSIRHNLSLHRRFERVSRDKPGKGDFWTVNDNCGATHH
KRRRRNKKSIAAALAAAAAAATAAAVAAAAAEGSNDPALHAAAEAAALLASQTAQAAALVTGDTQRKRSR
QMSTDDTPQQHSHADDSDESDEDEDEEADEDEDDEPAMDALSTTAATVGDDTLSVADTFTTETASVAGMD
QAQTSKRRRSSIKVKMPSQQQQQLWSEQFLALPVSHMPSSASALDMTGAPFASNGVMYPSPAAAFASGAA
AAPMPAPASSSSTLLPAGSTSPSGYHHRRTGSDSSAITSISSVSSMGQTAEEQYYAQQQQLALTRMNVAM
SVQPQPWAMQDASASDAGKPFEQSQAAPQQQQQGRTAQQQPRHQQLFARTGAAPIMPMSMSPFGSQPLSA
LSASGLFLPLSATAPLSSTSLGDWDALTAEKLLDSISNSVSQPVNGDLWFSNTSAEDLMAGPGLDMDLFA
LQ

>Ac1 manually annotated
MEGSLQQQQPHKLHPQPRHVASGSSGSTSSPELAAAQTEDDYYARGGEKKGRVKQVGRGRKRHLDQTQDD
DTVDYSYHHHHPRSASSTSTTSSLYAHPAKSPKRPRTFDELEERAAASDEKQQSAGSTDEDHLVVQQYQE
MPDRDEEEDEEASHHVEEEEVEEEDYDEVPHPPPRTWVAKAEWALGRLGGRGTLADIYATIEDAFPEEID
GKRNWKAAIRSCLSKNEKNKFFKPETMSAAKGDHVWCLAASTPRRATTTTTKRGRPAGRARKPPSRTQGR
RRSLADERKAVEEKKEKEEEVKFAVPVLPARREKKKEKEEERGRLETLLRWKQQLEQDISRLQEEVNEDG
LSADQEDFLSQLLA

>Ac2 manually annotated
MQHSAVREAQGEEGHWLMTHASQILLMAVSTFVYTLYKKVIDEFADSVVKSLKYADVADAAGANPNPIKR
AGDAILESGIEQLLASLQPDHIKQVCAVTHTFGGDLDQLREAIVAQYLASGGGSLARFIEECPAPLQSAF
QRVLRLKKNARDPVAHQITTGIVLLGTHHYFRQLPNLLKKEWCRDLGLKVSGTAAELEERLMSHFFEQWA
THTESYPDPAVEPPGSGSDGDQPLTWADKIELALQQHGGEATLTEIYCAIEAEFPQDIVDKPNWQPAIRS
CLSTNTSKGRFAVEETADGRKWSLSQG

>Ac3 manually annotated
MVSHDPRCQKATHVAAPSEHLVDDNVSRGGEKGQVKQVGRKRRLALSASSDSNSNYADEDYQPPASGDVD
YEEPAPKRSRHVEEADEKQSGCTDEWQVPSSPLLDEEDDEEDSSTSSSSENEEDNEDVIEVPHPPPRTWV
AKVEWALGRRGGKGTLADIYTTIEAAFPEEVDGKRNWKAAIRSCLSKNEKRKFYKPHAHGSSLAAGSGKD
HLWSLVVPREGGATRPAVRRGRPAGRRATTTRRMTTRHARLVEEQQQQQQPPQQHQQEKEKAQELVMVHP
QELMDELDDLPDVLVEQFVKEERMAQLLRQKELLEQDIHRLLEDNNADCADGLSPDQEDFLSHLLA
```

HMG box genes

```
>Co_HMGB CAOG_02944
MARGNKKDEKDPNKPKKPTTAFMYFSNAMRERVKTQNPGLKMTDIASVLGKLWGQLPEADKDKYQTMANS
DKERYAKAMDGYVAPVSTGGKSGKKSKDPNAPKRPPSAYICFANAVRPELRKTYPSDTMPAISTKIGELW
RQLTDDNKEPYNKQAEALKLKFQTEMAAYKGGKGVAADESDEEDDGEEEEEDE
```

>Co_SSRP-1 CAOG_07557
MSEGVQFNNISLLSKGKLFKGMVKMQEAGVAFKNQATGAVTLAASSDLRGFAWHRAATGYELAIQLPSSK
LRLLGFKEAHLTDIEQYVSSTFHQTVQTSDIATKGWTWGEPRFTGNGLAFYVDGKVAFELPLSDVSQTHM
KKNDITLEFKQDDPNDDDQQLVNMRFVVHPSAVLESEDPVEKARREEERKKRERIERDKRNAARARRNDL
RRRRREAEQSTIMQVDGEEPAPTQVSSDEEDSASEPEEAEEGVEGDEHELPVTRLFQRILGKVDVAHTAG
DVIVSLLDVNCQTPRGRYQMDFYPTMLTLHGQTHDFKIPFASISKTFVLPHPDQFRVFVVLALDPPVRQG
QTPYPFIVFLLQTEGREISVELNMTEAEIAEKNLQVSKLINKNVREGDAPKPNHFAITKEMSGGEVMILA
RLLSAMSVRLPIQPAAASSKDKAYSEQHGYKASYRASDGYLFPLENAFIFVHKPLTYIHYSDIKTVTFER
GSSILKTFAFTIVTHSGTGFTFNSIPKEEQRNLEQFCNAKARSKGFSVVADKPPARGAAAAAGDDSDDED
GAIHSDDEENDSYKRRMKKEGHAREIAGDDYDSEDEDEDFQANSDVEEVGEEFDENYTSSSDDSGSDSSD
SGSGSEDENGEPKVKKAKKHKKEHKPKPVRKAPKKSSSPTKKAAKDKNAPKKPMSSYMLWANENRAAFKA
KNPDANVMELGSILGNAWKELGESEKNSWAEKATEARKAYEITLAEYEQKKKERAAAGEPMDVDEEPAKP
KKKAAASTSSSSSKKFKSEEFIEDSD

>Co_SWI/SNF CAOG_00396
MLSFYTFALPAAESTSDEEAYSLLNQVNPVIQNNRQPPTMLRGSKLKHGKRVFGMYDLAPFPDDGYSWTQ
QEAMRDVKINGQLAMKVFETSHGVDDAAPHIEVTRKRFALLPPGGLTMEMLQLPLLQLVHYTYVSKGGVD
AAAMQRLQAVQQQQQLLQLQQLQHQKGSAAPLPGAAGVQIQQLGLIYNTLKQEIAKQQAAGNTPDPELLR
RLAMLEAQLTRHQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQSQTPAQPSQPQQQRAQQPSAAMQLQLHQ
QQLQLLQQQQQQQQQQQQQQQQQQQQQQLLQQQQQLIQQKAAMQRAQQAQASPQPQAAAQAQQQQQQALL
QQARNQAQLQQQLQLQQQLQQGLPPQQQEQLLKHLQLQQQQQLQQQQQQQQQQQQQQQQQQQQQQKAAAS
HSSPAQRTPAQAATSHSQAGRSSAASGQPAPDPRLVAQQQLLLRAPALNLPERPTPPTRPPMPQTKFTQE
RWDSVRKQHPEATPLQLEEIIAKEWAELSAEARSVREHEWDRALAQYDLGVAEHENSVAFKKWATAMKDA
EHQIIQWKRSVQLQLQQQAAQQAAGRPTQVAPAVVVDEESEPYSIRQVAASRYYRNHELMAEIVDTIPMD
ESEASSFLASGVSQHQSLFDEECQQMEEDAQAMESRAAARLQRQADDSENFLASLREIRHAGSLEDLNSI
YHKQATGDLPAPHVCGTWRMVQLADNPVERDHTAALVAGCSSKPHLAL

>Co_MAT-alpha-like-1 CAOG_01776
MNHQAFGHMGSISGPAVPGSEHLAMSANGPVLARIAANFSMSPTTEEQLRQMSLACKLATTQPNQPSQKP
LVVLRRIAAKPQSPTPVSEPTQETTTTKRRGPNGFIVFRTMNKHRYSHLPSRASSIELGKLWKQLSSDER
VTFEMEAQKRELLQLNTNKPARATKVRKPRLLYNNRLTAAGRRQAQQVAESCLPRSFLAHETLHADHEMF
PLDVTDASSLFPGSKACDPPILSISSSNSSTFSSTTSAALSLASDSSSTLAEPSPISDSDSSCSVDDWRN
TNLEAPLHPRPFVRLLGLSGQDDDCENISSWLLPPAETAPLPVLPAPHPKNLACQLAQRATFVSACHSPT
LVCQSPVPLPPPSIAIDMSWLDWAERDLAVRSVDSNQSSPTQ

>Co_MAT-alpha-like-2 CAOG_07895
MNHQAFGHMGSISGPAVPGSEHLAMSANGPVLARIAANFSMSPTTEEQLRQMSLACKLTTTQPNQPSQKP
LVVLRRIAAKPQSPTPVSEPTQETTTTKRRGPNGFIVFRTMNKHRYSHLPSRASSIELGKLWKQLSSDER
VTFEMEAQKRELLQLNTNKPARATKVRKPRLLYNNRLTAAGRRQAQQVAESCLPRSFLAHETLHADHEMF
PLDVTDASSLFPGSKACEPPILPISSSNSSTFSSTTSAALSLASDSSSTLAEPSPISDSDSSCSVVDDWR
NTNLEAPLHPRPFVRLLGLSGQDDDCENISSWLSPPAESAPLPILPAPHPKNLACQLAQRATFVVSACHS
PTLVCQSPVPLPPPSIAIDMSWLDWAERDLAVRSVDSNQSSPTQ

Homeobox genes

>Co1_TALE-1 CAOG_08013
MSHISMDTSTSIHSNQNGVNLATSMLIQAASHCGLFSSCGTNDRAALSNRWREPRGYKRSLAESNETETN
LAETMSTKCRIEPQSELQEAQAILDRLDNTNPWKGYIAAMAQQTGHVDLKLVRWADSNHFPWDASVCSTA
ARVGRLDVLQWARANGCPWDESICTIAAALGYRETLKWAQENGCPMNQPVCQAILRKGLFGNTSAPRPVS
PNSYVAQQIRNNTPASNSAVSQSARALVAGSSGSSAAASVSAGEGSMLIPAGQTELGGNKTSKARSDDRS
VRDSTPQLVAPERFDPDWYSKYAIDNPQCVRTAHLTKDERQNIKDGDFFIFDWEEVTRVRDGLPYHKGYT
CCNGLVVYKSVHPGGLYRTVAKRPPFLIVFYSRTSPSQKPSFSQARLAAEHPMITQAFATTTTIPFESRG
ESDNKRFRFDDDDDEPSSPGSSSTLAPELVD**NTPRRSNLRKASVAILKQWLLDHVSNPYPTDIEKDALAQ**
ATDLNVSQVNNWFINARRRILQPLLLGE

>Co2_TALE-2 CAOG_02804

MQPPGGDPGPTAAAPAATATATATATAAPAPAPATAQIDIDRRFASAAEPPPRPPFQPLVQYHHHHHDHH
HHPASASPQGQGQGQGQGLGLGLGQRPPGAPPPAMPPSTAPARPPGPGLSSSLSIGAPPPPPPLSAGGSL
PHLARGMAAVVSSSWSWMPPPSTSYSSSSSNHHPSSSVAQGSTHAARDANSSHGQQHYGVPGALHATAYQ
PQQPQQQQHQHQHQQQNAQQSQLSQHQQPHPLQQQQQQQQGQAPSSSAVVVVVAGSLAKSGSDDMAIDRS
TAPSASASAAPLAAAPSASTPSTGTAATPVISTSTVDGSSSSITSGSTAYTDQALREQQFLKMASVYTQQ
LREIGNALVQVDKTLAARSAVADSNLSMLLRACAEHIDTASAERENTTRASPAGAASDIPASSFGAERRN
LEASDTESDGTPSKRARLDNGWSATSAGNNNNNNSSNNNNNNNNNNPLPPPPPGMSAYPPPPHAMQAPPQ
PMQMAYASNYSGHPYYMPSPYPDQSMPYNYSYGMYPPPPHYYPSQYGPRSMPPPPPLHHQQHPYPYQQPI
MQHPVSQPVPPVHHTHAEYKPTHASANSVAAAATAAPTVVPAAPTSDGAHKTKPLPTDPPHTNKPGSSNE
SERVIEKPASSESTQPRDNVQHPPPQSGGQQQPSQAHHQNGSEAPVTPNTLPHASTFLDSPRGAGKSPRE
DGPNPDPSNGAKMEQLADPTQPVQLGDPMYSHFSVKQAKAEPKSSSEFHFTWSKDGQHTVLTRG**RSKQRP**
NLSREVVEVLKNWLFAHSSRPYPSDVEKTAMMAETGISLLQLNNWFINARRRLLQKPD

>Co3_PBC-like CAOG_05454
MELDILQELSELEADQLVLGLSFFESKADNEQDSVDIGQAALAIGDAVPAWAASGYSDFWDELSLALQTP
TNRVGFASHDVTAAQLDVCLAQERAACAASETPSLAEIRGLFAFLDMQLLKVAHRLSSLEYQLKRKVYTK
VIFLKSTFLD**SRKRRINLSREAQQVLNDWFLAHIEHPYPSESEKEQLADQTNLTMRQISTWFANKRNRQA**
QD

>Co4_non-TALE CAOG_05256
MSSTTQSGAFKPPAVKTQPT**SKRRRTQISEQQVRELEMLFDVDPWPSAEDKIALSRRLELSFQSVQVWFQ**
NRRARAKRQDEVSHARVYSVDSDDDDNNDNDNQSDKAVEGHR

>Co5_non-TALE CAOG_04648
MDGISMLLGAIQFNRAATPQTVSDQTLFVQHPQFQQQQLYPQQPQLQQPQHVLVQQLVQQQQQQQQQQQQ
QQQQQQQQQPMMHFMYPTQSTVLVPAPQMMQYYPAQTQQQHHQDSQISQQQPQQQQQQQQQQQPQQQQQV
VGYDAATQQYVSMVHQQQQQQQQQPSQMPQQEQLVQYVQYVQQPMYGSQTQFSVVNMQPAFVPAAAPHSF
QPQYAPAMPSGPAPAMAMVRPPQMLSQQPAGYLPPSPSPPTGFLTQPAVKSEPVSKSSSRYTPASSPPST
QEPTGAGAGAGAGAGASTTSSRTRHHSSGSKRRTSSGNDDASVAGGRS**KTRRRTTITLEQLGMLEEAFNR**
NNLPDCFQRTEVSRTTGLSERVIRIWFQNRRAKQRRSERNPDVSAAAAGSAGSARRVLAPSSTVCIPTSR
NGSGASRRHTGAGSRRVTANGMQATRGVSSTDYSPLQRDQDERYDDDEEVEDDEDEDEELGSSDDEDGFH
SRSSEGSASPASLSYTGSAQHRIPTDESQSAQQSTLVVS

>Co6_Paired-like CAOG_04156
MELLHPGEQATQSMDNGEHHRDPLLQQAVAFAGQGQSQHTAAVAAAAAVAAAAAVAVAARTDPVETSDAR
HHSASLRTAVGDDDEESIAQLMAAQGYYGAASGGNTSVSSGHNSGQISQYSDSFSIASSSTST**ASTNKRY**
FLSDQLELLESFYDQNKFPKPSDEEALAAKMDESRARIQQWFRNRRAKERRTQRLGDASTSSEPAPTCLD
LMHASLDEENASLRQQLRTLWNSFEFYVPESQQARLPIVSIPSGLLEDMAEGGAGGPSVVGGARPPSLLS
ATSLAPSSGFASARSTSSMEPLSSAPIPQPGGSSTFELRLPAWSSSSSASSLNSGLSPVVPSKRRHSGAA
DDASGSAVADAAFAQYSSPSLSVAAGATDSSSHDAPVSYSGETTGTAPAGSSSVGYRRRLSALTHPSHSS
VPTLTITVEHLGAESIHHGSAAATESLSASTALPEPDRTAVADSGDPNKRIRLDPLLTS

>Co7_non-TALE CAOG_04599
MGRKRRDESDEDEQEQPTAPAPAAATTDKGKGGNKKSHDKQSSSSSSSSSSSSSSGGGGKKSSASKKKSA
HGAEDDDEDASASTAATSTGAAAAAAAAEDDGKKAKGGKGGKKGKKNRGGDDDDDDWEAEMNKLKDQGEVD
LAAQSAAVAAASASASASAAAASAAKAAGDDQEQVETAQPAAAAAAAAQGKKGKDAAAAGGKNNNAKGGKKGK
KSRDDDDDWEAEMNQLREENEKAPAASAKATSKADEDVDQLASTLADASIADDKSKSKKKSSKKGSKHDG
GDDADDSTVSALATKMAATLSTADGGEEEEEDANPFTVVSKKSKASGKSSKKNRRAAEQDNDMDDMLVEN
NDDAPEEEPPVTESKPKKQSKKAAAEQEEQDAAAEAEPAEEKKDKKKKKKAAAKEAEEDQPADEAPVEEE
PAAEDTVKEKKKKSKKSKQDPEEEQPEDAELPVDDAGEAPVKEKKKKKKKHDADEDGESEATAAAATAEE
PAEVVAEEAARDKK**KKKSKSELDEKQLRRLNRAFNEDSRPDDEQLAAIAAKVGLSEEEVYSWFKAQRKLT**
RQSNREAAAAEEESQAVEEEAAPAAAAAAAAPEPVVEAESNDRKAKSAKKLRARERALANKEKGRSTTKD
RNEDDDDDADEAERKASKASETYGDDDEDEDGDKPQETEAAKSASKGKKPRKARADDDDDDDEGYSRGMY
NEDAQQDEENDEKEELTGREERKARKLREREAKIAAARGPAGISQFSVQQANLNSQSVSDNATDVKIDGF
SISAGGKDLFVNANLTIVPGRRYGLLGPNGRGKTTLLRHIANRELAIPKNIDVLYVEQEIEADGTSVVEA
VLRADTVRINLLEEEKRLLAELQVKSSTEATKRLQQVTDELKAIGAASAESRARRILAGLGFDRKMQRRP
TNSFSGGWRMRVSLARALFIEPTLLLLDEPTNHLDLNAVIWLDNYLSTWKKMLLVVSHDQDFLNSVCTDI
IHLEDRKLHYYRGNYFDFKKMHAQKVVEQQKQWEKQQKQLKALKAAGHSSKDALEKARRLQVAKGGKKPA
GKNVDDDDVGNKLELIEKPRDYIVNFHFVDPPQLPPPIIEVVDVGFNYPNCPKLFDDVNMGVTMDTRMCI
VGPNGVGKSTLLNLIIGELNPTEGEVRRNRKLRIGRYNQHFVDVLPMDISPTEYLMNTFNKVNEQCRAQL

GRFGLVSHAHTIPIRDLSGGQKARVVFAALVFQEPHILFFDEPTNHLDIESIDALVEAIEQFKGGVVVVS
HDARLITETECQLWVCDERDCIKFDGDFDDYRDLILAQLEEAERNAA

>Co8_non-TALE CAOG_05435
MPLQEGTAGWNGSRDGISGREQPANHANLVPRMTATTATTATTAVAAATAAAAVDCHSSSDGSALAHPLL
GTAAAGTLGPWLAYNLAAAAAVPQGLPLSHAAVVAALRAAAASASAAGTTTTATHATHSLTTTTTTTTPT
TANIPVGSGSTRAATQTADRRHSVPPPPPPPPVVVESPLFSYLATGSGAMAQQPRQLVTLPPPVQAAQNA
QATTSVQTPLLLSEAPLLSIMDASRLPGLPWAAKNSAPTRIHNPSKASMRAKRHAARDLALTPPAAAHRH
SRPGAVDTA**MRSFRLKKSSEQVRVLEAFFREFPKPRKYQVIALCNDTGLLHTEVRNWFRNRRLKDAKLR**H
LGTHPEPHSQPQSGASDNPTHLMAGDAQGAAQGAALLPSDPSSLPERLSSAPVQTTPSHAFGSAMAGCVA
QGSSASSVAAPAGADEFDDDDLEEFAKFADHPDSDFGQFSQYGTLTNSARNDNELYDYDEEAAQDDDDDD
NDDNDDDDEQEEDAEDAEDDSDEDEDEDDDSGQNHLADAATKAFSKHRSSISSSSTCSSTSVGPVASRLA
AIATKQRLCDIPGFGRKRSAFVSAALEFPAGSKPRLSMLKHLVRGLDKRDKMFLAELIAPPRRRMQGLSS
ESNISIIATQGWSDLEQRFVPSHNSLSAQPKPASSSEARSTSLAAGALDPDMVHGLCALFAGHRDYLVEV
ARRLGVAFALSNASSRLLQQHTQPETPTGQPTVTGSHIPPFAHNTDTVSSLFMASSLPSPSSMSLM

>Co9_LAG-1 CAOG_06194
MSLWSLLSQFDDYVATELAPPGSAFPRLRATNPDIYPHLRQLWIVLPIAVALFVARIVFERFFIRPLGVR
L**GISTKPTKRNEPNPALEAEYKKKKRMTPERAADLTKKTDKTPEYIMTWFHRRRNADKPSK**MVRFQEAVW
RLIYYTTAFVWSVYILSGVSFPLPLAMHFTNYPWLTDTDYCWIGYPEKQTLDPTIQWIYFIQLGFYMSLL
FSQFTDVKRKDFWEMFIHHVVTIFLVAFSYHANFIRIGTLVLLVHDVSDIFLEGAKAFNYLKYQKLCDAT
FVVFAIVFFVARLFVYPRYVLKSAFWDVRALLNTEPFFGLWFFNILLWILQALHVMWFITIFKMVISFAS
KGEVSGDDRSDSEAEEEDDTAHPAAPAPVKAIPAAKPVKH

CBP/p300

>Co_CBP CAOG04903
MAEPRSPDPEDLPNSKRMRMDDSVDGVDPVDDFSFASAPTGEYGASEEVNSLARSLNDGLPDLDDETDTG
TGGETAMNQDPDTAQGALAALDGVGHSELSNFLPSDGNGGASSESIFDSAATTAANQGDVGAVMSSANAD
SFDQHAGAGASSSLASPFVGAVSSPSIKPLAADSSTDLSTSMLSSFGPAGSDASSGVANTSASTLASAQN
ASGSAVTDRKKGIQQQLLLLIHSYRCTQSTDKCKLPHCAVIKNVLSHMSTCKNGRDCATPHCASSRQIIA
HYSTCKNPDCVICQPLKQPVVQNRMPPVKAETHAGGAPSLLTNGPGLTHPLSSINLSNTGSGAHAAGRAP
PPLQLPNGGNNNPASALTSGNSQLAFPSASSLTPAQVASLRMTNSLSVKVPPAMAKAEHIQSPSAAQQLL
QQQLQQQQQHQQQQQQQQQQQQQQQQQQQTSDRSSVGLQAEQAQLLRAQLNTLMQIPQASQMLNDHGIRK
SADVLEAVKNHTIERTMNTIAAALGITLPTSLPSMTPIGGASSSDSSHLSSAGSVLGNPGSSLGGPASIP
PPSSSVSGSGASGAPWANNPADPNPDPQWRIKFNNNHRLKLVQKLIEALRSCCKPDENQDRIRQVAEKVE
SRVFQMAPTQDKYYHLLAEKIYRMKKQPTHVSLNTMGVDMSAGGAAGGMGAAGPNAVATAAAMGAVAGAR
SGLPLNALQQQVAAILQQLAQYQLQFSQLREQQTLANQLVNSHRANGNPVPDDLMTVYNSIQQQLSQLVA
YQQTAAGMLRSHAATAAHMSGGSEIAAQAEAAATTTLNTIGNLASVVGAGGNLPSGNALGNGAQNFARPQ
GMVVKLSAAPVQAQPPVDPAIYATLTCMAANKTITSVTQNGPMYVKTMREVFLPIFEELCRDENAGPFMV
PVDPVALGILDYLTVIKRPMDFSTIRTRLDRPTDKHFYRDPLGFVDDMRLVFTNALTYNKKNSRVHKMAT
KLSDLFENRVEPALRKLGYCCALRLLQFSPPSLYCFGKCVSTIPRDTQYHTYENRIHYCARCFKEAPDPI
VGIDGEESMAKAQFSTATNNKLDDEPFNDCDECGRRAHVICALHNKHTDTRFLCPLCKKKSGKKEMRFTS
KNLPKCELSTHLEARISTHLRSVPDCGEVTIRVILVRDHQVNFKEQMKKKSVEEGGPTGLPYRHKAMFAF
QKQDGVEVCFFGMHVQEYGSDCPEPNARRVYVSYLDSVFWFRPQTHRTFVYHEILIGYLEYCRNLGYRYA
HIWACPPTPGDDYIFYCHPEDQKVPKPKRLQEWYRQMLTRARDDKVILDFCDIYSQCMTDDILTVRQLPY
FEGDYWPGVIEDSLTEAAKERTDTEEAGEEKGKKGSSKKAALKKSGSKAKGAVKVRALTRQNSDVCEKLY
QMMEKAKDGFFVATLRPPTDPLVDIKDPDPLVACALMEGREAFLDLARENHYEFSTLRRAKYASLMFLYH
LHNPDDSAGYHITCDSCQVSLSVADTRYRCTVCPEFDLCPKCQTAKPHEHPMEQVAFDASGDGGAGAAVG
AMGASGGGDGGSDTANRALAIQRCIESLVHACQCSDTECKNGSCVKMKKVVAHTRVCQVKAPVCNICKQL
VALCFYHAKHCQDDVCVVPFCAGIKQRLQRQQQQTRMKAKKIAQRRMLFQRMQQRSVQEASGIIPVETQQ
HADSHPSSAPSSASVAAAPQLNSMVTTPMQVVPNSTLLRNNTSTGGTPSPAPAMAQHATSNALGNVGMSP
MLTSNNPMLPNMLAGNLGATPNLVGSPALGTLGNGGGGAAVAAASSSVSTLSRNQIQQTSQMLEQARNE
FQNRLRNLQLQQQAFAEKKASMTPEDQARALAALTAERESLTLEKKNLLDRSTILSAAMLVAKQSSGQQL
SAPELQVIENATLLKAKMQQQQQQQKQQQQQQAASALANNPLLGLNNVQAQQNFQVQALQLVQQQQQQLQ
QQQQAKANLTPAQRIQQTQLNLAQLNEAARLHANNPNTVALIEARKAKLINDLRLLTQAGAMNPNAFAPS
TSSAMDVLSGNANRNFAGLLDNPADQQQQQHDQQLSQMVSGL

>Mb_CBP JGI9670
MREFCFGRPCNIDPWQARASRDSHCKAQVSRAQVSRAQVSRAQVSRARAFKDRVFKDRVFKDRVFKDRVF
KAQAWLHSAHRCSSLVARCKIILCKANPSKDNPFKDSLRPASRPSDSPRCKVYSILPCQYNQSLVEAYHL

```
TAYRVHPSLPWATLAWLASQCTAKVLFLAKAAMAYRQQQQHQQQMALQQQYRQQQQQQQLHQQQQLQQQQ
QLQQMQMPSTSMMPSTSMMPSTSMMPSTSMMPATSMMPATSMMPANGGINPTMSANQRNASLRRPGAPSN
LDLRAISNHSSPNLNPGALPPQQQQQQQLPGQPPGMGMQGQPQRYLPQGSQPMVLPGGVVGRGAPMTNLG
SRPGQAAPNGLQLPNGSIASDASGGMSSLAGGNLGHDQLGALNATSMSGGATGMAPAHSETESKAALSRR
RAIQQQLMLLIHAYKCKLPDKDDRNVPPNGQVCSVRYCKVMKNVLRHMATCRDGKQCSVPVKRPTVRFVI
PCESRLTESDHQRARREIKQSSEISSQQRGHLLPSEPIVLCLFKKQRSRPTIKLAQVLPPLRLRLLLLLL
LLVQATDNESYYQKLAEKIYKIKKAKQDRQKRQATSRQIKEWPRDELKHHFSPVLERLKGMSAAECFLRP
VDPIELNIPDYFEIIKKPMDLSTIEDKLEKGTYKDPWEFCDDMRLMFKNAWTYNPKNHVVYKFTNEVSSV
FEDTIDLVMKRLRFCCGKELTFNPQVLYCSVKQICPIRRDEEYMKYDANLNSAFNCCMACWKETEGEKST
FTISAELTPDQKEYVGLIACNGSRRERFQHDPRSSRGASPYAVPFVATFRIEVSRTHFTKAINNHIDYEE
FVTCAGCGRRHHQICVMYHASIHKEFICQACSKSNSATLSNPYPAHKLPRTKLSDYLEKRVNDCLHDETV
SSPPEARRVTVRVHSSRERGVVVKDNIRRYYADTEDPLPAEYPYRTKAIFVYQESEGHEVCFFGLHVQEY
GDDCPGPNRRRVYISYLDSVFFFEPSSLRTKVYHQLLIGYLQYMAALGYMYAHIWACPPAAGDDYIFHCH
PTEQKTPTPKRLSQWYAKMLKAAKEQNAIHDFYDMYQYVCSRDIVRGRDLPYFEGDYWVNRLESLVQEID
QADGNDDGGMLSLDKVGKKKSKKNTKGGKKKGGKGSKKQAADIDETLANRAWPEIEKNKSVFFVVQLHAD
PVNLPTFKADPDPDITCDFMDGREPFLQHCRENNHEFSTRRRALYSSMMMLYKLHTDESSGYTCNHCSTQ
ISSMDFRYHCTVCPDYDLCVKCHKEGVHKHEMKQLGLGMAPAVNADGQKMTRSQQSVETLRRALLALHHA
HQCKNPQCDTKACPQIKGMIEHMNTCNMRSSQCAQCKPLLTIVAHHAKVCNNSNCQIPLCRQCKARFAEN
QRMALMRQQTQARRMNVMEHMNSTPAAAAPPAPAAATDAPMTREALAIVLRHLPEKEQKHYQTAMLNQKK
IKEAMAQLQAKIKAAPRDSAASKTDETHLKKMHARYIELRAFILQLHKKGHALAVNAGALPQLATATPSV
AGAPRATPAKAEPVSSAAATQPAVSGQGRNMSEKQRQVMIYVTKLAAMPPAQAEVELKRLPPNQAELLRN
LIKARHEVRVGWSGIIRDDDRR
```

LSF/Grainyhead

```
>Co_LSF CAOG_06888
MKETSVDASQSRGADNDAALNLRALSSSLASASPAHAIAAAAGGGQTGDAQAEPMPDAAATAATATIEAD
PLAALVRHEKTPALPPSLRFVYELNAKQTRSVTELPVTYLDRQRQYRLTLQALGHDPLLDSASDQAQQHA
QEDELWRQAPSSSALPTSTSARPSSSSFASIPTPGLIDVKPEVFNAKAAASSMIDPADSTSSSSSALRSI
IYVHFHDHKDELRETELWNNWVSETSQTSAVEIASSMSNGIAEIKPVGPNAVSFSWSAARGAAIYIVLNC
LSTEVGASRNSKGVPLRLHIDTYRSSGDAPAVIGPNAARGPTTIQGMTLISSAFTMIKVYKKGAEKQFQN
DLRRNQGLPLSVTTILSPAKAVPLPGGGASFVLLPKSPVTSVATKPDARWPAAPARPLGLPMDSPFSSKL
ALEQQLRDSADPERFLLSPDSLSDFLPNAMDVLSPGTIDALSDGLIQHASLKDYLLDDAIRSGHAGSPHQ
GLLSREAHAARLPMNASSGLVEDRALSAFDVSSVSSSLPGPEGHRNGSSSRSAVESHKQFLAAEELGFES
ITASGNPSSGDFPATTSSGAQKTHEVTAAWPSVREVDITNSPPFLAGRRDDATFAKKMAANPATNFGELR
FSPGISPQFVFQTLPNSTTTWPLNPSFLPMVDTPGTLPGLDGGAVYMSGLSTSPSTRRTASRVSIKAPNG
AGFDSAFSPSQLQTWLLRQRFSASVTDGLQDFSGEEVAALSRADLREIVGDIPAIRLQNLIKGSDQRAVS
QSILKCIMVQAEGEQLYRRVYLQSTNTLESLSLGIATKFAVRLPLRRLCMTTRRLDFKVAGLAKRLKITA
QSPDVTDLPSGQPVRALASAEVEAAAAGLHIGAPTPETESDRSAVAMTGDAAESLLEALSLTPMIPSSES
HRPVPSDARSSAAPSTPAQKRRNDTPPSATSASEDDASMRSVVKGAPHTATASPLLGAKHPSTNDKLARV
ASVTDTLQDATTIGTNTVHDSDSRYSFSQMSRALPAASNQLQNSPNGAAGGVSTAFEETAASEPADAIFE
DDVWEIDEDVILHEMDDNSRYIVTFRINDKEPGSMDAILRPK
```

```
>Co_GRH CAOG_04897
MALRAPLHQPPTLSQHLGIMSDSVSTTASTPSVATRSGGLKVDSSKKNDLNPPQPSTAGSLVKGKSANAM
TTTTTAPVPAPLVLSADTDMPEGFDMLQSTPLATLSATTMASLLQENADTPNSMGLGTFFGMDDASQQSQ
LQEQLQLQRQQLMQMRPGSGIDDLDLHVDFGPGNSFSGQAFADIRRPSSVASFQSNPASLHHGSVHGIPP
SSNDLLYDMSNLPSHNYPVLPTEPSQSVFISNMLRMMLARPTTTNPHSTQKYSFILDAPTSIAQKLEEGT
LTYVNKGQAYAVTFEGMRGRRGSAGELPHTVKSIIHLVFHDEHDQKNERGLWEYWRWQQPPTLRAMEVDR
KTCSGLTDINELAFNAFEFKWNPREGGKIVVRVNCLSTEFSTQKGVKGMPLRIQIDTYENVTEPYETSMP
VSRDYCQIKVFRDKGSERKSKDEAKSAEKKLLKILKQHELKYGERASEYPVDFPFCPPFNATLLVPSSKL
WPKPALLPSKPQAPGGSTESGGTNTAGGTGSLSTSSGAGSSGTLTAAHTGAPRTAHPSNASSSSHDVSNF
LQPHSMVSPSSTPTSAPTSPGGNVLSGVHTGLTSKKRSASSASLVDHAAESHTVTSLLPVMDNGIEFMVP
RTKIVRVVSKKKEPVLTLYLRQEAERYFHSVYLDALTADDLRDKLADLLKLPAASVREILRRTAKGVLVQ
MDTKCVEHLENEDDFIVAVELQTDGWYRIVLRSGST
```