****

* **NCBI BLAST**

<http://blast.ncbi.nlm.nih.gov/Blast.cgi>

* ***Jurassic Park* DNA sequence**

<http://www.ncbi.nlm.nih.gov/Class/MLACourse/Modules/BLAST/q_jurassicparkDNA.html>

* ***The Lost World* DNA sequence**

<http://www.ncbi.nlm.nih.gov/Class/MLACourse/Modules/BLAST/q_lostworldDNA.html>

* **sequence for NP\_247556**

<http://www.ncbi.nlm.nih.gov/protein/15668757?report=fasta>

* **PHI-BLAST (cAMP protein kinase) sequence**

<http://www.ncbi.nlm.nih.gov/Class/NAWBIS/Modules/Similarity/Data/cAMP_Protein_Kinase.html>

* **PHI-BLAST pattern**

[LIVMF]-G-E-x-[GAS]-[LIVM]-x(5,11)-R-[STAQ]-A-x-[LIVMA]-x-[STACV]

* **Human ALU repeat sequence:**

aggccgaggccggcggatcgcctgaggtcgggagtttgagaccagcctgaccaacatggagaaaccctgtctctactaaa

gatacaaaattagccgggcgtgg

* ***C. Elegance* Protein:**

>C.elegans protein

MFHPGMTSQPSTSNQMYYDPLYGAEQIVQCNPMDYHQANILCGMQYFNNSHNRYPLLPQMPPQFTNDHPY

DFPNVPTISTLDEASSFNGFLIPSQPSSYNNNNISCVFTPTPCTSSQASSQPPPTPTVNPTPIPPNAGAV

LTTAMDSCQQISHVLQCYQQGGEDSDFVRKAIESLVKKLKDKRIELDALITAVTSNGKQPTGCVTIQRSL

DGRLQVAGRKGVPHVVYARIWRWPKVSKNELVKLVQCQTSSDHPDNICINPYHYERVVSNRITSADQSLH

VENSPMKSEYLGDAGVIDSCSDWPNTPPDNNFNGGFAPDQPQLVTPIISDIPIDLNQIYVPTPPQLLDNW

CSIIYYELDTPIGETFKVSARDHGKVIVDGGMDPHGENEGRLCLGALSNVHRTEASEKARIHIGRGVELT

AHADGNISITSNCKIFVRSGYLDYTHGSEYSSKAHRFTPNESSFTVFDIRWAYMQMLRRSRSSNEAVRAQ

AAAVAGYAPMSVMPAIMPDSGVDRMRRDFCTIAISFVKAWGDVYQRKTIKETPCWIEVTLHRPLQILDQL

LKNSSQFGSS

* **Compare two sequences**

[Sequence 1](http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?WebEnv=1NGKUdxnyn2ngCKQ1QQK5dH3QujpI09o0BwF0A1sg9UlZ1mBOPHq1KLqqjbJce8dWp95W3vhFQAwmVbT0ng7Q6y%40264668BF8F8DEAF0_0089SID&db=protein&qty=1&c_start=1&list_uids=157632&uids=&dopt=fasta&dispmax=5&sendto=&from=beg)

[Sequence2](http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?WebEnv=1La8ccN8ag6yvRynuF4MaXM8hfqLdM1hY7Dvtw-m7qo_hTMZK1Wgjx4ObndDuxPQhlkCkbQrwqfjweUjUpYxxy3%40264668BF8F8DEAF0_0089SID&db=protein&qty=1&c_start=1&list_uids=156933&uids=&dopt=fasta&dispmax=5&sendto=&from=beg)

* **BLAST2Seq**

<http://blast.ncbi.nlm.nih.gov/bl2seq/wblast2.cgi>

* **LALIGN**

<http://www.ch.embnet.org/software/LALIGN_form.html>

* **EBI Sequence similarity search tools**

[FASTA search](http://www.ebi.ac.uk/Tools/fasta/)

[Smith and Waterman algorithm- MP Search](http://www.ebi.ac.uk/Tools/MPsrch/)

* **WU-BLAST**

<http://www.proweb.org/Tools/WU-blast.html>

* **BLINK**

<http://www.ncbi.nlm.nih.gov/sutils/blink.cgi?mode=query>

Sample Sequences :

>Human gi|29725609|ref|NP\_005219.2| epidermal growth factor receptor isoform a [Homo sapiens]

MRPSGTAGAALLALLAALCPASRALEEKKVCQGTSNKLTQLGTFEDHFLSLQRMFNNCEVVLGNLEITYV

QRNYDLSFLKTIQEVAGYVLIALNTVERIPLENLQIIRGNMYYENSYALAVLSNYDANKTGLKELPMRNL

QEILHGAVRFSNNPALCNVESIQWRDIVSSDFLSNMSMDFQNHLGSCQKCDPSCPNGSCWGAGEENCQKL

TKIICAQQCSGRCRGKSPSDCCHNQCAAGCTGPRESDCLVCRKFRDEATCKDTCPPLMLYNPTTYQMDVN

PEGKYSFGATCVKKCPRNYVVTDHGSCVRACGADSYEMEEDGVRKCKKCEGPCRKVCNGIGIGEFKDSLS

INATNIKHFKNCTSISGDLHILPVAFRGDSFTHTPPLDPQELDILKTVKEITGFLLIQAWPENRTDLHAF

ENLEIIRGRTKQHGQFSLAVVSLNITSLGLRSLKEISDGDVIISGNKNLCYANTINWKKLFGTSGQKTKI

ISNRGENSCKATGQVCHALCSPEGCWGPEPRDCVSCRNVSRGRECVDKCNLLEGEPREFVENSECIQCHP

ECLPQAMNITCTGRGPDNCIQCAHYIDGPHCVKTCPAGVMGENNTLVWKYADAGHVCHLCHPNCTYGCTG

PGLEGCPTNGPKIPSIATGMVGALLLLLVVALGIGLFMRRRHIVRKRTLRRLLQERELVEPLTPSGEAPN

QALLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSPKANKEILDEAYVMASVD

NPHVCRLLGICLTSTVQLITQLMPFGCLLDYVREHKDNIGSQYLLNWCVQIAKGMNYLEDRRLVHRDLAA

RNVLVKTPQHVKITDFGLAKLLGAEEKEYHAEGGKVPIKWMALESILHRIYTHQSDVWSYGVTVWELMTF

GSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMVKCWMIDADSRPKFRELIIEFSKMARDPQRYLV

IQGDERMHLPSPTDSNFYRALMDEEDMDDVVDADEYLIPQQGFFSSPSTSRTPLLSSLSATSNNSTVACI

DRNGLQSCPIKEDSFLQRYSSDPTGALTEDSIDDTFLPVPEYINQSVPKRPAGSVQNPVYHNQPLNPAPS

RDPHYQDPHSTAVGNPEYLNTVQPTCVNSTFDSPAHWAQKGSHQISLDNPDYQQDFFPKEAKPNGIFKGS

TAENAEYLRVAPQSSEFIGA

>Mouse gi|46560582|ref|NP\_997538.1| epidermal growth factor receptor isoform 1 [Mus musculus]

MRPSGTARTTLLVLLTALCAAGGALEEKKVCQGTSNRLTQLGTFEDHFLSLQRMYNNCEVVLGNLEITYV

QRNYDLSFLKTIQEVAGYVLIALNTVERIPLENLQIIRGNALYENTYALAILSNYGTNRTGLRELPMRNL

QEILIGAVRFSNNPILCNMDTIQWRDIVQNVFMSNMSMDLQSHPSSCPKCDPSCPNGSCWGGGEENCQKL

TKIICAQQCSHRCRGRSPSDCCHNQCAAGCTGPRESDCLVCQKFQDEATCKDTCPPLMLYNPTTYQMDVN

PEGKYSFGATCVKKCPRNYVVTDHGSCVRACGPDYYEVEEDGIRKCKKCDGPCRKVCNGIGIGEFKDTLS

INATNIKHFKYCTAISGDLHILPVAFKGDSFTRTPPLDPRELEILKTVKEITGFLLIQAWPDNWTDLHAF

ENLEIIRGRTKQHGQFSLAVVGLNITSLGLRSLKEISDGDVIISGNRNLCYANTINWKKLFGTPNQKTKI

MNNRAEKDCKAVNHVCNPLCSSEGCWGPEPRDCVSCQNVSRGRECVEKCNILEGEPREFVENSECIQCHP

ECLPQAMNITCTGRGPDNCIQCAHYIDGPHCVKTCPAGIMGENNTLVWKYADANNVCHLCHANCTYGCAG

PGLQGCEVWPSGPKIPSIATGIVGGLLFIVVVALGIGLFMRRRHIVRKRTLRRLLQERELVEPLTPSGEA

PNQAHLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSPKANKEILDEAYVMAS

VDNPHVCRLLGICLTSTVQLITQLMPYGCLLDYVREHKDNIGSQYLLNWCVQIAKGMNYLEDRRLVHRDL

AARNVLVKTPQHVKITDFGLAKLLGAEEKEYHAEGGKVPIKWMALESILHRIYTHQSDVWSYGVTVWELM

TFGSKPYDGIPASDISSILEKGERLPQPPICTIDVYMIMVKCWMIDADSRPKFRELILEFSKMARDPQRY

LVIQGDERMHLPSPTDSNFYRALMDEEDMEDVVDADEYLIPQQGFFNSPSTSRTPLLSSLSATSNNSTVA

CINRNGSCRVKEDAFLQRYSSDPTGAVTEDNIDDAFLPVPEYVNQSVPKRPAGSVQNPVYHNQPLHPAPG

RDLHYQNPHSNAVGNPEYLNTAQPTCLSSGFNSPALWIQKGSHQMSLDNPDYQQDFFPKETKPNGIFKGP

TAENAEYLRVAPPSSEFIGA

>Chimp gi|114613403|ref|XP\_001156495.1| PREDICTED: epidermal growth factor receptor isoform 5 [Pan troglodytes]

MRPSGTAGAALLALLAALCPASRALEEKKVCQGTSNKLTQLGTFEDHFLSLQRMFNNCEVVLGNLEITYV

QRNYDLSFLKTIQEVAGYVLIALNTVERIPLENLQIIRGNMYYENSYALAVLSNYDANKTGLKELPMRNL

QEILHGAVRFSNNPALCNVESIQWRDIVSSDFLSNMSMDFQNHLGSCQKCDPSCPNGSCWGAGEENCQKL

TKIICAQQCSGRCRGKSPSDCCHNQCAAGCTGPRESDCLVCRKFRDEATCKDTCPPLMLYNPTTYQMDVN

PEGKYSFGATCVKKCPRNYVVTDHGSCVRACGADSYEMEEDGVRKCKKCEGPCRKVCNGIGIGEFKDTLS

INATNIKHFKNCTSISGDLHILPVAFRGDSFTHTPPLDPQELDILKTVKEITGFLLIQAWPENRTDLHAF

ENLEIIRGRTKQHGQFSLAVVSLNITSLGLRSLKEISDGDVIISGNKNLCYANTINWKKLFGTSGQKTKI

ISNRGENSCKATGQVCHALCSPEGCWGPEPRDCVSCRNVSRGRECVDKCNILEGEPREFVENSECIQCHP

ECLPQAMNITCTGRGPDNCIQCAHYIDGPHCVKTCPAGVMGENNTLVWKYADAGHVCHLCHPNCTYGCFG

PHSHASMPRPWNLSATSPFLVPPLHSRPKIPSIATGMVGALLLLLVVALGIGLFMRRRHIVRKRTLRRLL

QERELVEPLTPSGEAPNQALLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSP

KANKEILDEAYVMASVDNPHVCRLLGICLTSTVQLITQLMPFGCLLDYVREHKDNIGSQYLLNWCVQIAK

GMNYLEDRRLVHRDLAARNVLVKTPQHVKITDFGLAKLLGAEEKEYHAEGGKVPIKWITLESILHRIYTH

QKKVWSYGVTVWELMTFGSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMVKCWMIDADSRPKFRE

LIIEFSKMARDPQRYLVIQGDERMHLPSPTDSNFYRALMDEEDMDDVVDADEYLIPQQGFFSSPSTSRTP

LLSSLSATSNNSTMACIDRNGLQSCPIKEDSFLQRYSSDPTGALTEDSIDDTFLPVPEYINQSVPKRPAG

SVQNPVYHNQPLNPTPSRDPHYQDPHSTAVGNPEYLNTVQPTCVNSTFDSPAHWAQKGSHQISLDNPDYQ

QDFFPKEAKPNGIFKGSTAENAEYLRVAPQSSEFIGA

>Rat gi|25742617|ref|NP\_113695.1| epidermal growth factor receptor [Rattus norvegicus]

MRPSGTARTKLLLLLAALCAAGGALEEKKVCQGTSNRLTQLGTFEDHFLSLQRMFNNCEVVLGNLEITYV

QRNYDLSFLKTIQEVAGYVLIALNTVERIPLENLQIIRGNALYENTYALAVLSNYGTNKTGLRELPMRNL

QEILIGAVRFSNNPILCNMETIQWRDIVQDVFLSNMSMDVQRHLTGCPKCDPSCPNGSCWGRGEENCQKL

TKIICAQQCSRRCRGRSPSDCCHNQCAAGCTGPRESDCLVCHRFRDEATCKDTCPPLMLYNPTTYQMDVN

PEGKYSFGATCVKKCPRNYVVTDHGSCVRACGPDYYEVEEDGVSKCKKCDGPCRKVCNGIGIGEFKDTLS

INATNIKHFKYCTAISGDLHILPVAFKGDSFTRTPPLDPRELEILKTVKEITGFLLIQAWPENWTDLHAF

ENLEIIRGRTKQHGQFSLAVVGLNITSLGLRSLKEISDGDVIISGNRNLCYANTINWKKLFGTPNQKTKI

MNNRAEKDCKATNHVCNPLCSSEGCWGPEPTDCVSCQNVSRGRECVDKCNILEGEPREFVENSECIQCHP

ECLPQTMNITCTGRGPDNCIKCAHYVDGPHCVKTCPSGIMGENNTLVWKFADANNVCHLCHANCTYGCAG

PGLKGCQQPEGPKIPSIATGIVGGLLFIVVVALGIGLFMRRRQLVRKRTLRRLLQERELVEPLTPSGEAP

NQAHLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSPKANKEILDEAYVMASV

DNPHVCRLLGICLTSTVQLITQLMPYGCLLDYVREHKDNIGSQYLLNWCVQIAKGMNYLEDRRLVHRDLA

ARNVLVKTPQHVKITDFGLAKLLGAEEKEYHAEGGKVPIKWMALESILHRIYTHQSDVWSYGVTVWELMT

FGSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMVKCWMIDADSRPKFRELILEFSKMARDPQRYL

VIQGDERMHLPSPTDSNFYRALMEEEDMEDVVDADEYLIPQQGFFNSPSTSRTPLLSSLSANSNSSTVAC

INRNGSCRVKEDAFLQRYSSDPTSVLTEDNIDDTFLPVPEYINQSVPKRPAGSVQNPVYHNQPLHPAPGR

DLHYQNPHSNAVSNPEYLNTAQPTCLSSGFDSSALWIQKGSHQMSLDNPDYQQDFFPKEAKPNGIFKGPT

AENAEYLRVAPPSSEFSGA

>Chicken gi|45382797|ref|NP\_990828.1| epidermal growth factor receptor [Gallus gallus]

MGVRSPLSASGPRGAAVLVLLLLLLGRVALCSAVEEKKVCQGTNNKLTQLGHVEDHFTSLQRMYNNCEVV

LSNLEITYVEHNRDLTFLKTIQEVAGYVLIALNMVDVIPLENLQIIRGNVLYDNSFALAVLSNYHMNKTQ

GLRELPMKRLSEILNGGVKISNNPKLCNMDTVLWNDIIDTSRKPLTVLDFASNLSSCPKCHPNCTEDHCW

GAGEQNCQTLTKVICAQQCSGRCRGKVPSDCCHNQCAAGCTGPRESDCLACRKFRDDATCKDTCPPLVLY

NPTTYQMDVNPEGKYSFGATCVRECPHNYVVTDHGSCVRSCNTDTYEVEENGVRKCKKCDGLCSKVCNGI

GIGELKGILSINATNIDSFKNCTKINGDVSILPVAFLGDAFTKTLPLDPKKLDVFRTVKEISGFLLIQAW

PDNATDLYAFENLEIIRGRTKQHGQYSLAVVNLKIQSLGLRSLKEISDGDIAIMKNKNLCYADTMNWRSL

FATQSQKTKIIQNRNKNDCSKSVCFPAFAKAHNEMEE

>C elegans gi|25146697|ref|NP\_495962.2| LEThal family member (let-23) [Caenorhabditis elegans]

MRYPPSIGSILLIIPIFLTFFGNSNAQLWKRCVSPQDCLCSGTTNGISRYGTGNILEDLETMYRGCRRVY

GNLEITWIEANEIKKWRESTNSTVDPKNEDSPLKSINFFDNLEEIRGSLIIYRANIQKISFPRLRVIYGD

EVFHDNALYIHKNDKVHEVVMRELRVIRNGSVTIQDNPKMCYIGDKIDWKELLYDPDVQKVETTNSHQHC

YQNGKSMAKCHESCNDKCWGSGDNDCQRVYRSVCPKSCSQCFYSNSTSSYECCDSACLGGCTGHGPKNCI

ACSKYELDGICIETCPSRKIFNHKTGRLVFNPDGRYQNGNHCVKECPPELLIENDVCVRHCSDGHHYDAT

KDVRECEKCRSSSCPKICTVDGHLTNETLKNLEGCEQIDGHLIIEHAFTYEQLKVLETVKIVSEYITIVQ

QNFYDLKFLKNLQIIEGRKLHNVRWALAIYQCDDLEELSLNSLKLIKTGAVLIMKNHRLCYVSKIDWSSI

ITSKGKDNKPSLAIAENRDSKLCETEQRVCDKNCNKRGCWGKEPEDCLECKTWKSVGTCVEKCDTKGFLR

NQTSMKCERCSPECETCNGLGELDCLTCRHKTLYNSDFGNRMECVHDCPVSHFPTQKNVCEKCHPTCYDN

GCTGPDSNLGYGGCKQCKYAVKYENDTIFCLQSSGMNNVCVENDLPNYYISTYDTEGVIETHCEKCSISC

KTCSSAGRNVVQNKCVCKHVEYQPNPSERICMDQCPVNSFMVPDTNNTVCKKCHHECDQNYHCANGQSTG

CQKCKNFTVFKGDIAQCVSECPKNLPFSNPANGECLDYDIASRQRKTRMVIIGSVLFGFAVMFLFILLVY

WRCQRIGKKLKIAEMVDMPELTPIDASVRPNMSRICLIPSSELQTKLDKKLGAGAFGTVFAGIYYPKRAK

NVKIPVAIKVFQTDQSQTDEMLEEATNMFRLRHDNLLKIIGFCMHDDGLKIVTIYRPLGNLQNFLKLHKE

NLGAREQVLYCYQIASGMQYLEKQRVVHRDLATRNVLVKKFNHVEITDFGLSKILKHDADSITIKSGKVA

IKWLAIEIFSKHCYTHASDVWAFGVTCWEIITFGQSPYQGMSTDSIHNFLKDGNRLSQPPNCSQDLYQEL

LRCWMADPKSRPGFEILYERFKEFCKVPQLFLENSNKISESDLSAEERFQTERIREMFDGNIDPQMYFDQ

GSLPSMPSSPTSMATFTIPHGDLMNRMQSVNSSRYKTEPFDYGSTAQEDNSYLIPKTKEVQQSAVLYTAV

TNEDGQTELSPSNGDYYNQPNTPSSSSGYYNEPHLKTKKPETSEEAEAVQYENEEVSQKETCL