

Sequence searching on STN[®]

Hands-on-practice

Example sequences (copy & paste to a .txt query file)

Human cholinesterase (NCBI [AAA98113](#))

```
>gi|180484|gb|AAA98113.1| cholinesterase (EC 3.1.1.8) [Homo sapiens]
MHSKVTTIICIRFLFWFLLLCMLIGKSHTEDDII IATKNGKVRGMNLTVFGGTVTAFLGIPYAQPPLGRLR
FKKPQSLTKWSDIWNATKYANSCCQNIDQSFPGFHGSEMWNPNNTDLSEDCLYLNWVWPAPKPKNATVLIW
IYGGGFQTGTSSLHVYDYGKFLARVERVIVVSMNYRVGALGFLALPGNPEAPGNMGLFDQQLALQWVQKNI
AAFGGNPKSVTLFGESAGAASVSLHLLSPGSHSLFTRAILQSGSFNAPWAVTSLYEARNRTLNLAKLTGC
SRENETEIIKCLRNKDPQEILLNEAFVVPYGTPLSVNFGPTVDGDFLTDMPDILLELGQFKKTQILVGVN
KDEGTAFLVYGAPGFSKDNNSIITRKEFQEGLKIFFPGVSEFGKESILFHYTDWVDDQRPENYREALGDV
VGDFNFICPALEFTKKFSEWGNNAFFYYFEHRSSKLPWPEWGMVMHGYEIEFVFGPLPLERRDNYTKAEEI
LSRSIVKRWANFAKYGNPNETQNNSTSWPVFKSTEQKYLTLNTESTRIMTKLRAQQCRFWTSFFPKVLEM
TGNIDEAEWEWKAGFHRWNNYMMDWKNQFNNDYTSKKKESCVGL
```

BRCA1 protein (NCBI [NP_009225](#))

```
>gi|6552299|ref|NP_009225.1| breast cancer type 1 susceptibility protein
isoform 1 [Homo sapiens]
MDLSALRVEEVQNVINAMQKILECPICLELILIKEPVSTKCDHIFCKFCMLKLLNQQKGPSQCPLCKNDITK
RSLQESTRFSQLVEELLKIIICAFQLDTGLEAYANSYNFAKKENNSPEHLKDEVSI IQSMGYRNRKRLQSQ
EPENPSLQETSLSVQLSNLGTVRTLRTRKQRIQPQKTSVYIELGSDSSEDTVNKATYCSVGDQELLQITPQ
GTRDEISLDSAKKAACEFSETDVTNTEHHQPSNNDLNTTEKRAAERHPEKYQGSSVSNLHVEPCGTNTHA
SSLQHENSLLLLTKDRMVEKAFCNKSQPLGARSQHNRWAGSKETCNDRRTPSTEEKVDLADPLCER
KEWVKQKLPCSENPRDTEVPWITLNSSIQKVNEWFSRSDLELLGSDSDHGESESNKAVADVLDVLEVD
EYSGSSEKIDLLASDPHEALICKSERVHKSVESENIEDKIFGKTYRKKASLPNLSHVTEENLIIGAFVTEP
QIIQERPLTNKLRKRRTSGLHPEDFIKKADLAVQKTPEMINQGTNQTQNGQVMNITNSGHENKTKGD
SIQNEKNPNPIESLEKESAFKTKAEPISSSISNMELELNHNKAPKKNRRLRRKSSTRHIAHALELVSRN
LSPPNCTELQIDSCSSSEI KKKKYNQMPVHRNRLQLMEGKEPATGAKKSNKPNEQTSKRHDSDTFPEL
KLTNAPGSFTKCSNTSELKEFVNPSLPREEKEEKLETVKVSNNAEDPKDLMLSGERVLTQTERSVSSSIS
LVPGTDYGTQESISLLEVSTLGKAKTEPNKCVSQAAAFENPKGLIHGCSKDNRNDTEGFKYPLGHEVNHS
RETSIEMEESELDAQYLQNTFKVSKRQSFAPFSNPGNAEEECATFSAHSGSLKKQSPKVTFECEQKEENQ
GKNESNIKPVQTVNITAGFPVVGQKDKPVDNAKCSIKGGSRFCLSSQFRGNETGLITPNKHGLLQNPYRI
PPLFPKSFVKTKCKKNLLEENFEHSMSPEREMGNENIPSTVSTISRNNIRENVFKEASSNINEVGSS
TNEVGSSINEIGSSDENIQAELGRNRGPKLNAMLRLGLVQPEVYKQSLPGSNCKHPEIKKQEEYEEVVQTV
NTDFSPYILSDNLEQPMGSSSHASQVCSETPDDLDDGEIKEDTSAFAENDIKESSAVFSKSVQKGELSRSP
SPFTHTHLAQGYRRGAKKLESSEENLSSSEDEELPCFQHLLFGKVNNIQSTRHSTVATECLSKNTEENL
LSLKNLNDCSNQVILAKASQEHHLSEETKCSASLFSQCELEDLTANTNTQDPFLIGSSKQMRHQSES
QGVGLSDKELVSDDEERGTLLEENQEEQSMDSNLGEAASGCESETSVSEDCSGLSSQSDILTTQQRDTM
QHNLIKLQQEMAELEAVLEQHGSSQPSNSYPSIISDSSALEDLRNPEQSTSEKAVLTSQKSSEYPI SQNPE
GLSADKFEVSADSSTSKNKEPGVERSSPSKCPSLDDRWMHSCSGSLQNRNYPSEELIKVVDVEEQQLE
ESGPHDLTETSYPRLQDLEGTPLYESGISLFSDDPESDPSEDRAPE SARVGNIPSSTSALKVPQLKVAES
AQSPAAHTTDTAGYNAMESVSRKPELTASTERVNRMSMVVSGLTPEEFMLVYKFARKHHITLTLNLI
TEETHVVMKTD AEFVTCERTLKYFLGIAGGKVVVSYFWVTQSIKERKMLNEHDFEVRGDVVNGRNHQGP
RARESQRDKIFRGL EICCYGPFNTMPTDQLEWMVQLCGASVVKELSSFTLGTGVHPIVVVQPDAWTEDNG
FHAIGQMCEAPVVTREWVLD SVALYQCQELDITYLIPQIPHSHY
```

Sequence searching on STN[®]

TNF Alpha (NCBI [AAC03542](#))

```
>gi|2905634|gb|AAC03542.1| tumor necrosis factor alpha [Homo sapiens]
VRSSSRTPSDKPVAVHVVANPQAEGQLQWLNRRANALLANGVELRDNQLVVPSEGLYLIYSQVLFKGGQCP
STHVLLTHTISRIAVSYQTKVNLLSAIKSPCQRETPRGAEAKPWYEPYIYLGGVFQLEKGDRLSAEINRPD
YLDFAESGQVYFGIIAL
```

Eukaryotic translation elongation factor 1 gamma (NCBI [NP_001395](#))

```
>gi|4503481|ref|NP_001395.1| elongation factor 1-gamma [Homo sapiens]
MAAGTLYTYPENWRAFKALIAAQYSGAQVRVLSAPPHFHFQTNRTPEFLRKFPAAGKVPVAFEGDDGFCVF
ESNAIAYYVSNEELRGSTPEAAAQVVQWVSFADSDIVPPASTWVFPPTLGIMHHNKQATENAKEEVRRILG
LLDAYLKTTRTFLVGERVTLADITVVCTLLWLKQVLEPSFRQAFNPNTNRWFLTCINQPQFRAVLGEVKLC
EKMAQFDAQKFAETQPKKDTPRKEKGSREEKQKQPAERKEEKAAAPAPEEEMDECEQALAAEPKAKDPF
AHLPKSTFVLDEFKRKYSNEDTLSVALPYFWEHFDDKDGWSLWYSEYRFPEELTQTFMSCNLITGMFQRLD
KLRKNAFASVILFGTNNSSSISGVVWFRGQELAFPLSPDWQVDYESTWRKLDPGSEETQTLVREYFSWE
GAFQHVKGAFNQKIFK
```

Human chorionic somatomammotropin hormone 2 isoform 1 (NCBI [NP_066271](#))

```
>gi|10334861|ref|NP_066271.1| chorionic somatomammotropin hormone 2 isoform 1
[Homo sapiens]
MAAGSRTSLLLAFALLCLPWLQEAGAVQTVPLSRLFDHAMLQAHRAHQLAIDTYQEFEEETYIPKDQKYSF
LHDSQTSFCFSDSIPTPSNMEETQKSNLELLRISLLLLIESWLEPVRFRLSMFANNLVYDTSDDYHLL
KDLEEGIQTLMGRLEDGSRRTGQILKQTYSKFDTNSHNHDALLKNYGLLYCFRKDMDKVVETFLRMVQCRS
VEGSCGF
```

Human D-amino acid oxidase (NCBI [NP_001908](#))

```
>gi|148539837|ref|NP_001908.3| D-amino-acid oxidase [Homo sapiens]
MRVVVIGAGVIGLSTALCIHERYHSVLQPLDIKVVYADRFTPLTTTDDVAAGLWQPYLSDPNNPQEADWSQQ
TFDYLLSHVHSPNAENLGLFLISGYNLFHEAIPDPSWKDVTVLGFRKLTPRELDMFPDYGYGWFHTSLILE
GKNYLQWLTERLTERGVKFFQRKVESFEEVAREGADVIVNCTGVWAGALQRDPLLQPRGQIMKVDAPWM
KHFILTHDPERGIYNSPYIIPGTQTVTLGGIFQLGNWSELNNIQDHNTIWECCCRLEPTLKNARIIGERT
GFRPVRPQIRLEREQRLRTGPSNTEVIHNYGHGGYGLTIHWGCALEAAKLFGRILEEKLSRMPPSHL
```

Hepatitis C virus 5' region (NCBI [M58406](#))

```
>gi|329885|gb|M58406.1|HPCRNAA1 Hepatitis C virus 5' region
GCCAGCCCCCTGATGGGGGCGACACTCCACCATGAATCACTCCCCTGTGAGGAACACTACTGTCTTACGCA
GAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCTGTCAGCCTCCAGGACCCCCCTCCCGGGAGAGCCA
TAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGGACGACCGGGTCCTTTCTTGATCAACCCG
CTCAATGCCTGGAGATTTGGGCGTCCCCCGCAAGACTGCTAGCCGAGTAGTGTGGGTGCGGAAAGGCC
TTGTGGTACTGCCTGATAGGGTGTTCGCGAGTCCCCGGGAGGTCTCGTAGACCGTGCACC
```


Sequence searching on STN[®]

Thermus thermophilus DNA polymerase I, DNA, CDS region (121-2622) (DGENE AAQ12748)

```
ATGGAGGCGATGCTTCCGCTCTTTGAACCCAAAGGCCGGGTCTCCTGGTGGACGGCCAC
CACCTGGCCTACCGCACCTTCTTCGCCCTGAAGGGCCTCACCACGAGCCGGGGCGAACCG
GTGCAGGCGGTCTACGGCTTCGCCAAGAGCCTCCTCAAGGCCCTGAAGGAGGACGGGTAC
AAGGCCGTCTTCGTGGTCTTTGACGCCAAGCCCCCTCCTTCCGCCACGAGGCCTACGAG
GCCTACAAGGCGGGGAGGGCCCCGACCCCCGAGGACTTCCCCCGGCAGCTCGCCCTCATC
AAGGAGCTGGTGGACCTCCTGGGGTTTACCCGCTCGAGGTCCCCGGCTACGAGGCGGAC
GACGTTCTCGCCACCCTGGCCAAGAAGGCGGAAAAGGAGGGGTACGAGGTGCGCATCCTC
ACCGCCGACCGCGACCTTACCAACTCGTCTCCGACCGCGTCGCCGTCTCCACCCCGAG
GGCCACCTCATCACCCCGAGTGGCTTTGGGAGAAGTACGGCCTCAGGCCGGAGCAGTGG
GTGGACTTCCGCGCCCTCGTGGGGGACCCCTCCGACAACCTCCCCGGGGTCAAGGGCATC
GGGGAGAAGACCGCCCTCAAGCTCCTCAAGGAGTGGGAAGCCTGGAAAACCTCCTCAAG
AACCTGGACCGGGTAAAGCCAGAAAACGTCCGGGAGAAGATCAAGGCCACCTGGAAGAC
CTCAGGCTCTCCTTGGAGCTCTCCGGGTGCGCACCGACCTCCCCCTGGAGGTGGACCTC
GCCCAGGGCGGGAGCCCGACCGGGAGGGGCTTAGGGCCTTCTGGAGAGGCTGGAGTTC
GGCAGCCTCCTCCACGAGTTCGGCCTCCTGGAGGCCCGCCCCCTGGAGGAGGCCCCC
TGGCCCCCGCGGAAGGGCCTTTCGTGGGCTTCGTCTCTCCCGCCCCGAGCCATGTGG
GCGGAGCTTAAAGCCCTGGCCGCTGCAGGGACGGCCGGGTGCACCGGGCAGCAGACCCC
TTGGCGGGGCTAAAGGACCTCAAGGAGTCCGGGGCCTCCTCGCCAAGGACCTCGCCGTC
TTGGCCTCGAGGGAGGGGCTAGACCTCGTGCCTGGGACGACCCCATGCTCCTCGCCTAC
CTCCTGGACCCCTCCAACACCACCCCGAGGGGTGGCGCGGCGCTACGGGGGGAGTGG
ACGGAGGACGCCGCCACCGGGCCCTCCTCTCGGAGAGGCTCCATCGGAACCTCCTTAAG
CGCCTCGAGGGGGAGGAGAAGCTCCTTTGGCTCTACCACGAGGTGAAAAGCCCCCTCTCC
CGGGTCTGGCCACATGGAGGCCACCGGGTACGGCTGGACGTGGCCTACCTTCAGGCC
CTTTCCTGGAGCTTTCGGAGGAGATCCGCCGCTCGAGGAGGAGGTCTTCCGCTTGGCG
GGCCACCCCTTCAACCTCAACTCCCGGGACAGCTGGAAAGGGTGCTCTTTGACGAGCTT
AGGCTTCCCGCCTTGGGGAAGACGAAAAGACAGGCAAGCGCTCCACCAGCGCCGCGGTG
CTGGAGGCCCTACGGGAGGCCACCCCATCGTGGAGAAGATCCTCCAGCACCGGGAGCTC
ACCAAGCTCAAGAACACCTACGTGGACCCCTCCCAAGCCTCGTCCACCCGAGGACGGGC
CGCCTCCACACCCGCTTCAACCAGACGGCCACGGCCACGGGGAGGCTTAGTAGCTCCGAC
CCCAACCTGCAGAACATCCCCGTCCGCACCCCTTGGGCCAGAGGATCCGCCGGGCCTTC
GTGGCCGAGGCGGGTTGGGCGTTGGTGGCCCTGGACTATAGCCAGATAGAGCTCCGCGTC
CTCGCCACCTCTCCGGGACGAAAACCTGATCAGGGTCTTCCAGGAGGGGAAGGACATC
CACACCCAGACCGCAAGCTGGATGTTTCGGCGTCCCCCGGAGGCCGTGGACCCCTGATG
CGCCGGGCGGCCAAGACGGTGAACCTTCGGCGTCTCTACGGCATGTCCGCCCATAGGCTC
TCCCAGGAGCTTGCCATCCCCTACGAGGAGGCGGTGGCCTTTATAGAGCGCTACTTCCAA
AGCTTCCCCAAGGTGCGGGCCTGGATAGAAAAGACCCTGGAGGAGGGGAGGAAGCGGGGC
TACGTGGAAACCCTCTTCGGAAGAAGGCGCTACGTGCCCGACCTCAACGCCCGGGTGAAG
AGCGTCAGGGAGGCCGCGGAGCGCATGGCCTTCAACATGCCCGTCCAGGGCACCGCGCC
GACCTCATGAAGCTCGCCATGGTGAAGCTCTTCCCCCGCTCCGGGAGATGGGGCCCCGC
ATGCTCCTCCAGGTCCACGACGAGCTCCTCCTGGAGGCCCCCCAAGCGCGGGCCGAGGAG
GTGGCGGCTTTGGCCAAGGAGGCCATGGAGAAGGCCTATCCCCTCGCCGTGCCCTGGAG
GTGGAGGTGGGGATGGGGGAGGACTGGCTTTCGCCAAGGGT
```