

MANUAL

Step by step procedure for analysis of protein sequence(s)

1. Take the protein sequence in fasta format (can also use multiple fasta), shown below the image

```
>sp|Q7XR88|STRK1_ORYSJ Salt tolerance receptor-like cytoplasmic kinase 1 OS=Oryza sativa subsp. japonica OX=39947 GN=STRK1 PE=1 SV=2
MFTGCGGLFACVRRCDGGDVRKRGEGAGMSRVAADPAGVEEGSCKNVAAASARQLAWAD
VESVTGGGFSRVIGHGGFSTVYLASLSSRLGAVKVCSSERLHRAFRQELVLLSLRHP
HIVRLLYCDEDEGVLFVEYAPNGDLHERLSECVAGGVASVLEFWARRVIAIAFOVAMAL
EYLHESRHPAVIHGDIKASNVLLDANMNAKLCDFGFAHVGFSAIVGCRPSARAVMGSPGY
VDFPLIRSGVATKKSDVYSFGVLLLELVGTGKEAVCRDTGRRLTAAVGPMLESGKVADVVD
RRLGGEGHDGAEAAVMAELAMQCIGDSFGLRPSMADVVRALQEKTSALASAVGSRDLRKM
F
>sp|Q5Z9S0|PYL9_ORYSJ Absciscic acid receptor PYL9 OS=Oryza sativa subsp. japonica OX=39947 GN=PYL9 PE=1 SV=1
MEAHVERALREGLTEEERAALEPAVMAHHTFEPSTITATTAATCTSLVTORVAAFEVRAV
WPIVRSFGNPNQRYKHFFVRTCALAAGDGASVGSVREVTVVSGLPASTSTERLEMLDDDRHI
ISFVVGQGHRLNRYRSVTSVTEFQPPAAGGPPAPPYCVVVEVYVDVDPGNTAEDTRMF
TDTVVKLNLMQAAVAEDSSASRRRD
>sp|Q1QUN20|MPKS_ORYSJ Mitogen-activated protein kinase 5 OS=Oryza sativa subsp. japonica OX=39947 GN=MPKS PE=1 SV=1
MDGAPVLEFRPTHTIGGRYLLVDIFGNRFEVNNKYOPPIMPIGRGAGVGVCSMNFETRE
MYAIAKIANAFNMMDAKRTLREIKLLRHLDHENIIIGIRDVIPPPIQAFNDVYIATELM
DIDLHIIIRSNQELSEHCQYFLYQILRGLKYIHSANVHRDLKPSNLLNANCDLKICD
FGLARPSSESDMTEYVTRWYRAPELLNSTDYSAIDVWVSGCIFMELINRQPLFFGR
DHMQMRLITEVIGTPTDDELGFIRNEDARKYMRHLPOYPRRTFASMFPRVQPAALDLIE
RMLTFNPLQRTVEEALDHPYLERLHDIADEPICLEFFSFDFEQKALNEDQMQLIFNEA
IEMNNIRY
>sp|Q6EN42|PYL3_ORYSJ Absciscic acid receptor PYL3 OS=Oryza sativa subsp. japonica OX=39947 GN=PYL3 PE=1 SV=1
MVEVGGGAAEAAAGRRRLADERCDLRAAETEYVRRFHRHEPRDHQCSSAVAKHIKAPVH
LVNSLVRFDQPLQKFPFVSRCEMKNIEIGSVREVNVKSGLPATRSTERLELLDDNEHI
LSRVFVGGDHRLKNYSILTVHPEVIDGRPTLVIESFVVDVFEQNTKDETCTYFVEALLK
CNLKSIAEYSERLVVKQDTEFLDR
>sp|Q6K4R2|CIPK3_ORYSJ CBL-interacting protein kinase 31 OS=Oryza sativa subsp. japonica OX=39947 GN=CIPK3 PE=1 SV=1
MYRAKRAALSPKVRKRGVYELGRITIGGTFKVRFAKNTENDEPVAIKILDKEKVKQHR
LVEQIRREICTMKLVKHPNVVRLFEVMSGKARIFIVLEYVTGGELFEIATNGRLKEEA
RKYFOQLINAVDYCHSRGVYHRLKLENLLDASGNLKVSDFGLSALTEQVQKADGLLHTT
CGTPNVYAVEVIEDRGYDGAADIWSCGVILYVLLAGFLPFEDDNIILYKKISEAQFTC
PSWFSTGAKRLITRLDPNPTTRITISQILEDPNFKKGYKFPVDFEKYETSFDDVDAAFG
DSEDRHVKEETEQQGTSMAPELISLQALNLDMLFEAKKEYKRETRFTSCQCKPKEITK
IEEAAKFLGFDIQKKNYKMRMENLKAGRKGNLNVATEVQVAFSLHVLEKKAKGDTLEF
QKFYRTLSTQKDVVWKCDEGEVENGAAA
>sp|P53684|CDPK7_ORYSJ Calcium-dependent protein kinase 7 OS=Oryza sativa subsp. japonica OX=39947 GN=CDPK7 PE=2 SV=2
MGNQCQNGTLGSDYHNRFPREHAGVYVQGDYLDLKKFDDTWPEVNNFKPTAASILRRGL
DPTISNVLGRITADLREHYIIRGLGQGFSTTLCTEINTGCEYACKTIFKRKLITKED
```

2. Click on the Analysis for analyzing the protein sequences in fasta file, shown below

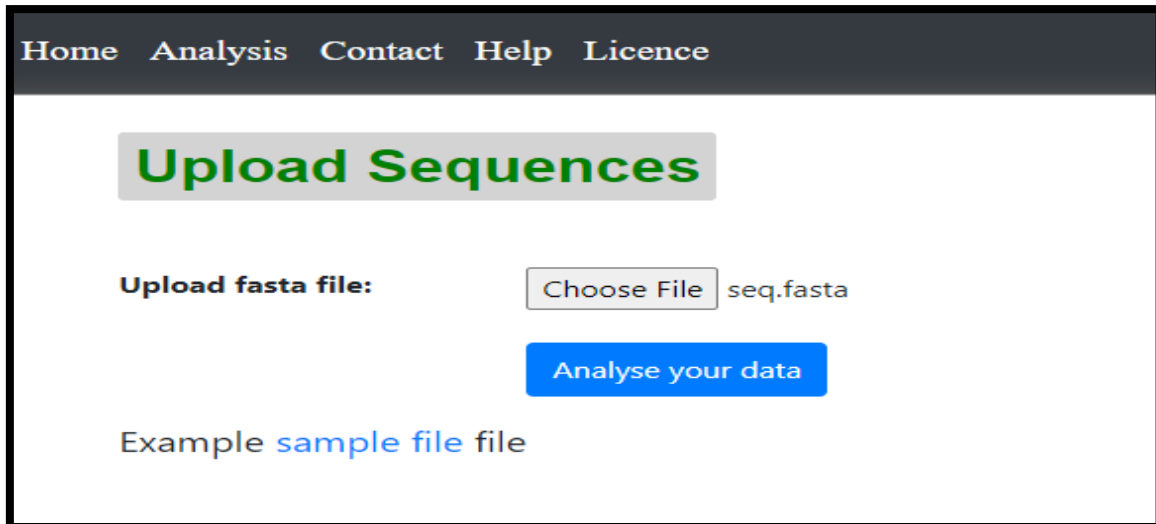
DeepAProt: Crop specific biotic stress protein classification tool using Deep Learning

Home Analysis Contact Help Licence

Welcome to DeepAProt

Protein sequences are the basic information of a living organism which are made of long chain of amino acids. Long chain of primary protein sequence are linear which form peptide chains of protein. primary protein sequences binds together with hydrogen bonds to form a secondary structure. Secondary protein structure are broadly alpha-helix, beta-sheet and coil. Alpha-helix are form by hydrogen-bond in oxygen atom of carbonyl group of an amino acid and hydrogen atom in amine group. In beta-sheet, hydrogen bonds are formed in oxygen atom of carboxyl group of one strand and hydrogen atom of amine group of another strand. And coils are the bending which are found in both N-terminal and C-terminal. Tertiary structure of a protein are stable which contain dimensions, atomic coordinates. Stability of a tertiary structure is achieved by thermostability, kinetic trap and metastability. Further these tertiary structure gives a final finishing known as quaternary structure which gives a basic form of life. Quaternary structure consists of multiple folded protein subunits in a multi-subunit complex. Finding these structures need high level machinery, lots of money and time. So keeping in mind, this tool has been developed.

3. After clicking on the Analysis tab, upload fasta file and then click on analysis your data, shown in below image



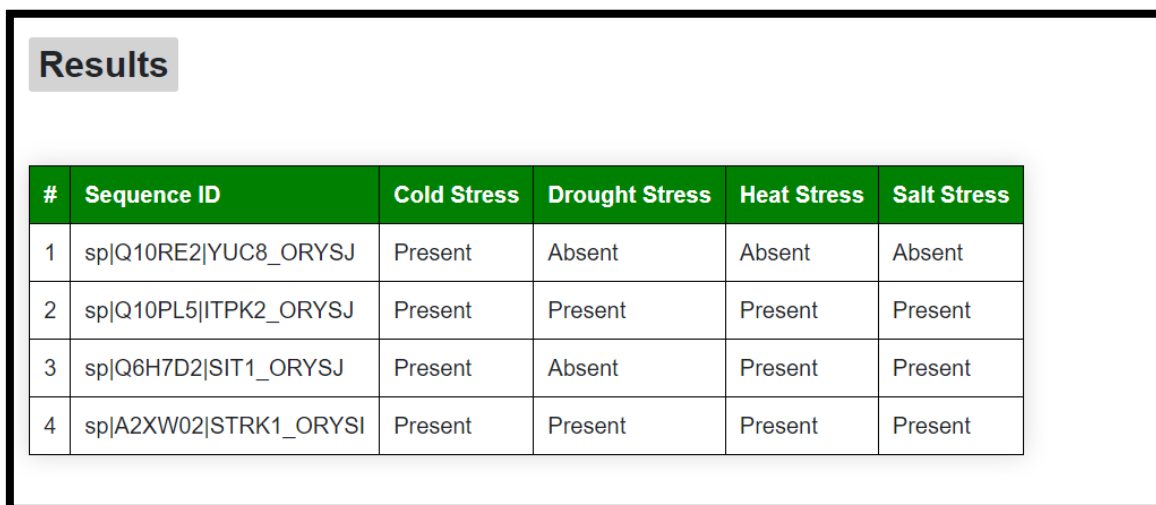
Home Analysis Contact Help Licence

Upload Sequences

Upload fasta file: seq.fasta

Example [sample file](#) file

4. Finally, the result will be displayed on the screen, shown in the image below



Results

#	Sequence ID	Cold Stress	Drought Stress	Heat Stress	Salt Stress
1	sp Q10RE2 YUC8_ORYSJ	Present	Absent	Absent	Absent
2	sp Q10PL5 ITPK2_ORYSJ	Present	Present	Present	Present
3	sp Q6H7D2 SIT1_ORYSJ	Present	Absent	Present	Present
4	sp A2XW02 STRK1_ORYSI	Present	Present	Present	Present
