

Computational Analysis of Unknown DNA Sequences Using Integrated Bioinformatics Tools

Eman Sajjad

BS-vII



NCBI BLASTn alignment

DNA ANALYSIS

Descriptions Graphic Summary Alignments Taxonomy

Alignment view Pairwise CDS feature Restore defaults

3 sequences selected

[Download](#) [GenBank](#) [Graphics](#)

PREDICTED: Solanum pennellii transcription factor ILR3-like (LOC107024201), transcript variant X1, Sequence ID: XM_015225123.2 Length: 1266 Number of Matches: 1

Range 1: 1 to 1266 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
2338 bits(1266)	0.0	1266/1266(100%)	0/1266(0%)	Plus/Plus
Query 1	AGAAAAGCTCAAATAGAAAGTTACAGTAGTTGTCATTCCAGTTGCAGTTCCCTTCTCA	60		
Sbjct 1	AGAAAAGCTCAAATAGAAAGTTACAGTAGTTGTCATTCCAGTTGCAGTTCCCTTCTCA	60		
Query 61	GCAACAAATAGTATAATAGCAATGAAAACGATGATTGATGAATGATAGCTACTTGCTAG	120		
Sbjct 61	GCAACAAATAGTATAATAGCAATGAAAACGATGATTGATGAATGATAGCTACTTGCTAG	120		

Descriptions Graphic Summary Alignments Taxonomy

Sequences producing significant alignments Download

select all 3 sequences selected GenBank Graphics Distance tree of results MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
PREDICTED: Solanum pennellii transcript...	Solanum...	2338	2338	100%	0.0	100.00%	1266	XM_015225123.2
PREDICTED: Solanum stenotomum transc...	Solanum...	2023	2023	97%	0.0	96.27%	1233	XM_049558004.1
PREDICTED: Solanum tuberosum transcri...	Solanum...	2019	2019	98%	0.0	95.91%	1280	XM_006360518.2

RNA Fold Image

• RNA is thermodynamically stable, but is structurally heterogeneous
• Collective contribution of multiple alternative conformations determines stability
• Biologically important for RNA function, interactions, or regulation.

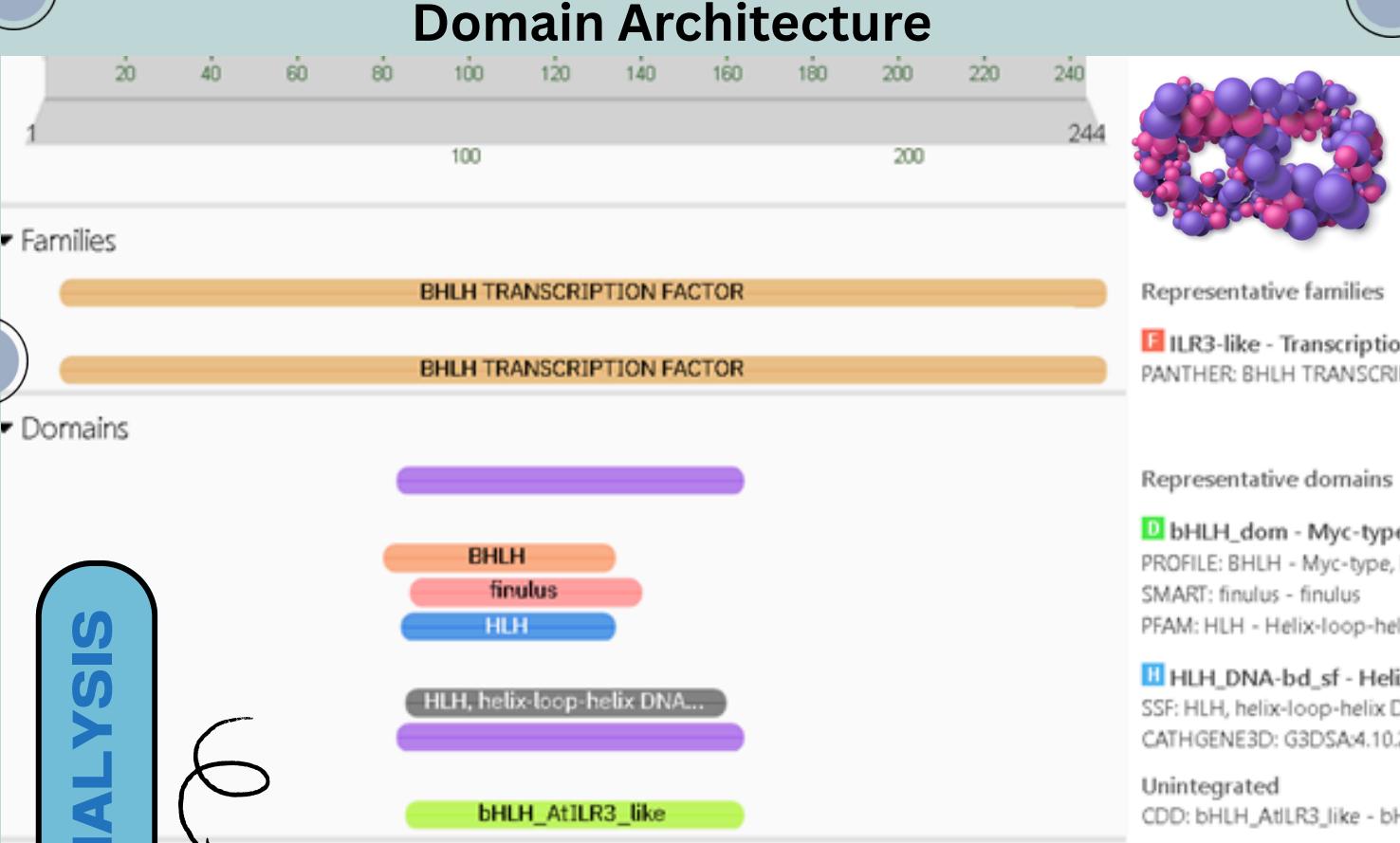
RNA ANALYSIS

S' cap S' UTR Start codon Coding sequence (ORF3) Stop codon 3' UTR Poly(A) tail

Annotated mRNA drawing



Among 6 identified ORFs, ORF3 located on the forward strand (+3 frame) was the longest, spanning 735 nucleotides and encoding a 244 amino acid protein, so was selected as the most likely coding region. It contained a canonical start codon (ATG) and a valid stop codon with no internal interruptions.

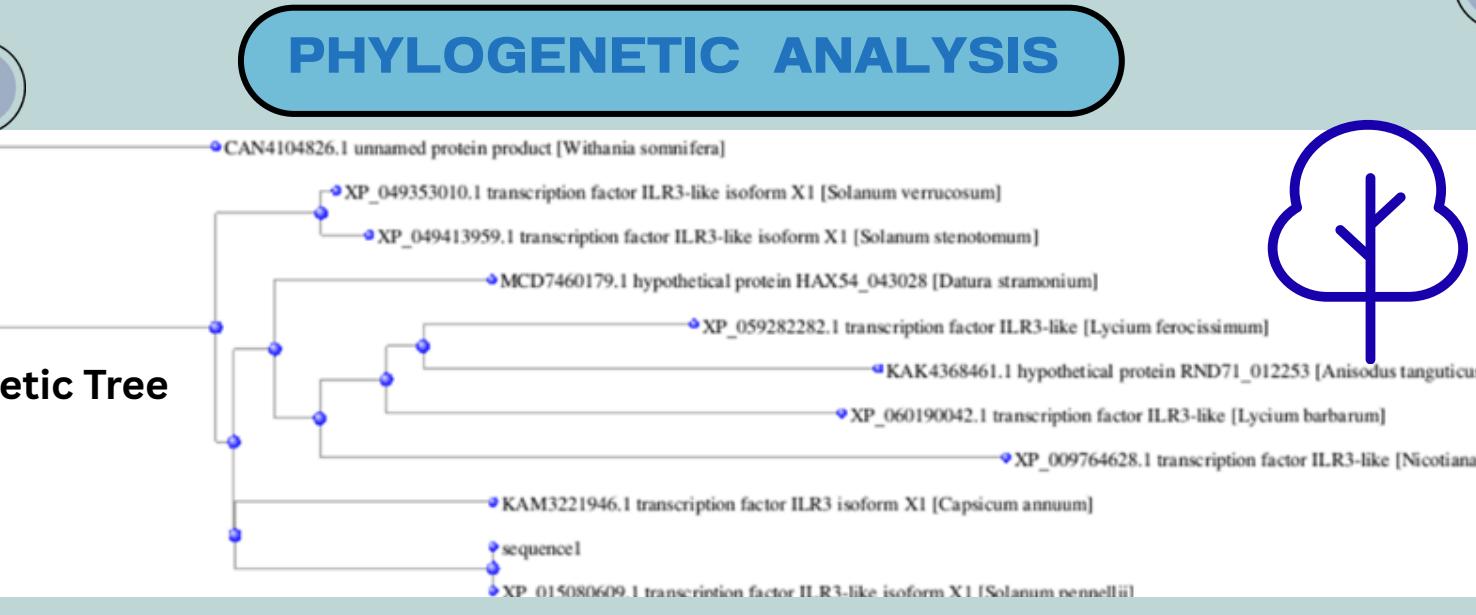


- Functional prediction**
- bHLH transcription factor > DNA binding, dimerization
 - Regulatory role in gene expression
 - Involved in Developmental and stress-responsive pathways

PROTEIN ANALYSIS

Protein Properties

Molecular weight	27377.98 > moderate size
Theoretical pI	5.65 > acidic
Instability index	53.66 > unstable in vitro
Aliphatic index	72.83 > moderate thermostability



ILR3-like transcription factor is evolutionarily conserved, supporting its importance in plant growth, development, and stress response pathways.

Alignments Select All Re-align Mouse over the sequence identifier for sequence title

View Format: Compact Conservation Setting: 2 Bits

Query_10001	NEVDSS-GNPM6LFYELMTDITSAASVTVAEFPSPATIDFSWPAQTIYASNLITETDYTFADSEVSKEARRKRLKE	79
Query_10002	NEVDSS-GNPM6LFYELMTDITSAASVTVAEFPSPATIDFSWPAQTIYASNLITETDYTFADSEVSKEARRKRLKE	79
Query_10003	NEVDSS-GNPM6LFYELMTDITSAASVTVAEFPSPATIDFSWPAQTIYASNLITETDYTFADSEVSKEARRKRLKE	79
Query_10004	NEVDSS-GNPM6LFYELMTDITSAASVTVAEFPSPATIDFSWPAQTIYASNLITETDYTFADSEVSKEARRKRLKE	79
Query_10005	NEVDSS-GNPM6LFYELMTDITSAASVTVAEFPSPATIDFSWPAQTIYASNLITETDYTFADSEVSKEARRKRLKE	79
Query_10006	NEVDSS-GNPM6LFYELMTDITSAASVTVAEFPSPATIDFSWPAQTIYASNLITETDYTFADSEVSKEARRKRLKE	79
Query_10007	NEVDSS-GNPM6LFYELMTDITSAASVTVAEFPSPATIDFSWPAQTIYASNLITETDYTFADSEVSKEARRKRLKE	79
Query_10008	NEVDSS-GNPM6LFYELMTDITSAASVTVAEFPSPATIDFSWPAQTIYASNLITETDYTFADSEVSKEARRKRLKE	79
Query_10009	NEVDSS-GNPM6LFYELMTDITSAASVTVAEFPSPATIDFSWPAQTIYASNLITETDYTFADSEVSKEARRKRLKE	79
Query_10010	NEVDSS-GNPM6LFYELMTDITSAASVTVAEFPSPATIDFSWPAQTIYASNLITETDYTFADSEVSKEARRKRLKE	80
Query_10011	NEVDSS-GNPM6LFYELMTDITSAASVTVAEFPSPATIDFSWPAQTIYASNLITETDYTFADSEVSKEARRKRLKE	75
Query_10001	CCSSPISKACREKLRRDRNRFLERFLAASSLVLDPGRPPKTEKVALSQAQRILZELRTERTOKLKESHEEQIKEIKLAEW	159
Query_10002	CCSSPISKACREKLRRDRNRFLERFLAASSLVLDPGRPPKTEKVALSQAQRILZELRTERTOKLKESHEEQIKEIKLAEW	159
Query_10003	CCSSPISKACREKLRRDRNRFLERFLAASSLVLDPGRPPKTEKVALSQAQRILZELRTERTOKLKESHEEQIKEIKLAEW	159
Query_10004	CCSSPISKACREKLRRDRNRFLERFLAASSLVLDPGRPPKTEKVALSQAQRILZELRTERTOKLKESHEEQIKEIKLAEW	159
Query_10005	CCSSPISKACREKLRRDRNRFLERFLAASSLVLDPGRPPKTEKVALSQAQRILZELRTERTOKLKESHEEQIKEIKLAEW	159
Query_10006	CCSSPISKACREKLRRDRNRFLERFLAASSLVLDPGRPPKTEKVALSQAQRILZELRTERTOKLKESHEEQIKEIKLAEW	159
Query_10007	CCSSPISKACREKLRRDRNRFLERFLAASSLVLDPGRPPKTEKVALSQAQRILZELRTERTOKLKESHEEQIKEIKLAEW	159
Query_10008	CCSSPISKACREKLRRDRNRFLERFLAASSLVLDPGRPPKTEKVALSQAQRILZELRTERTOKLKESHEEQIKEIKLAEW	159
Query_10009	CCSSPISKACREKLRRDRNRFLERFLAASSLVLDPGRPPKTEKVALSQAQRILZELRTERTOKLKESHEEQIKEIKLAEW	159
Query_10010	CCSSPISKACREKLRRDRNRFLERFLAASSLVLDPGRPPKTEKVALSQAQRILZELRTERTOKLKESHEEQIKEIKLAEW	160
Query_10011	CCSSPISKACREKLRRDRNRFLERFLAASSLVLDPGRPPKTEKVALSQAQRILZELRTERTOKLKESHEEQIKEIKLAEW	155

SUBCELLULAR LOCALIZATION PREDICTION

prediction for 176666846422687 is nuc (k=23)

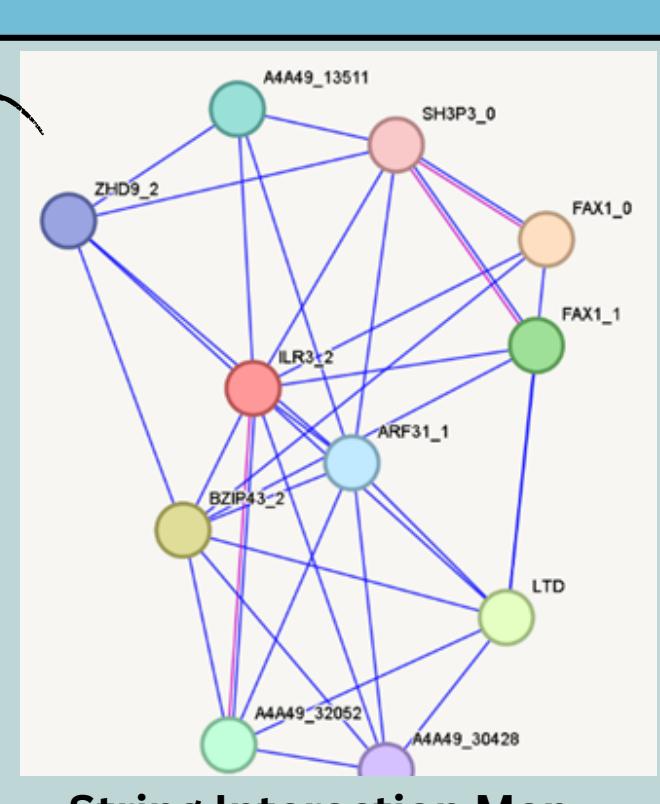
PSORT TargetP

Protein type	Other	Signal peptide	Mitochondrial transfer peptide
Likelihood	1	0	0

- Predicted Localization to nucleus (PSORT)
- absence of targeting peptides (TargetP), and
- the presence of a bHLH transcription factor domain

STRING PROTEIN -PROTEIN INTERACTION ANALYSIS

- a hub protein
- key transcription factor
- functional importance in plant physiology
- multiple interacting partners



Writing for Different Audiences



PEERS / FRIENDS

Who They Are:

- People your age, friends, classmates.

What They Know:

- Similar background knowledge, current slang/interests.

Writer's Choices:

- Use casual language.
- Use humor or personal anecdotes.
- Can refer to shared experiences.
- Keep sentences engaging, maybe shorter.

ADULTS (FORMAL)

Who They Are:

- Teachers, principals, parents, other adults in authority.

What They Know:

- Educated, expect formal language, may know topic well.

Writer's Choices:

- Use formal and respectful language.
- Use correct grammar, spelling, and punctuation.
- Provide strong evidence and clear explanations.
- Be organized and logical.

YOUNGER CHILDREN:

Who They Are:

- Elementary school kids, younger siblings.

What They Know:

- Limited vocabulary, less background knowledge.

Writer's Choices:

- Use simple words and short sentences.
- Explain ideas very clearly.
- Use lots of repetition or simple examples.
- Include bright colors or fun pictures (if applicable).
- Focus on core concepts.

GENERAL PUBLIC

Who They Are:

- A diverse group of people

What They Know:

- Knows little to some, varying backgrounds

Writer's Choices:

- Define all specialized terms
- Use clear + accessible language
- Provide strong context