

Université Batna 2

Année universitaire : 2022-2023

Faculté des Sciences de la nature et de la vie

Département d'Ecologie et environnement

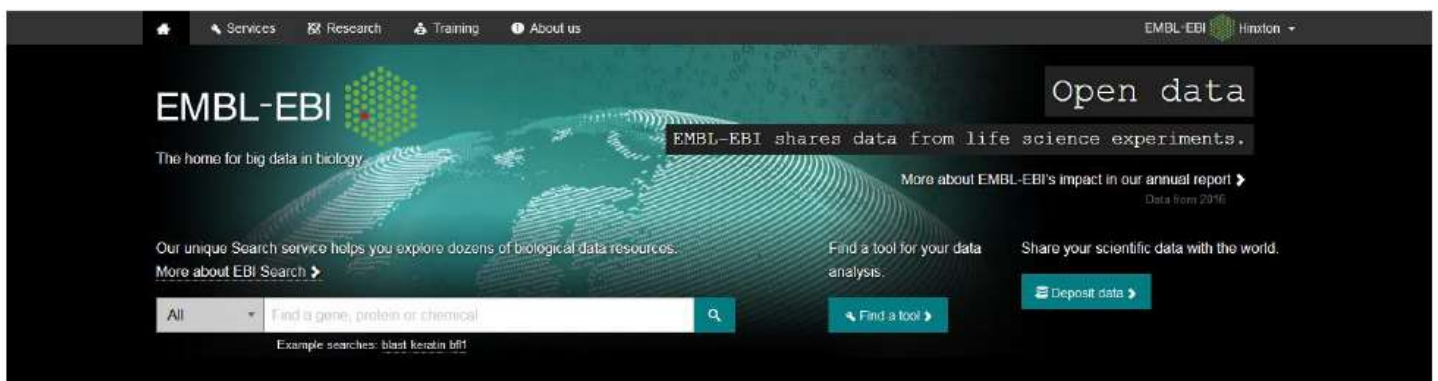
## Travail Pratique 1 : Les banques de séquences biologiques

### (Exemple : EMBL)

Réalisons le TP suivant sur Internet pour connaître la **séquence primaire d'ARNm** du **gène de réponse à l'auxine** de la plante *Oryza sativa*, en interrogeant la base de données **EMBL**.

Suivre les étapes suivantes :

1. Allez sur le moteur de recherche google et glissez l'adresse de EMBL : <https://www.ebi.ac.uk/>



#### Explore EMBL-EBI and our mission

The European Bioinformatics Institute (EMBL-EBI) shares data from life science experiments, performs basic research in computational biology and offers an extensive user training programme, supporting researchers in academia and industry. We are part of EMBL, Europe's flagship laboratory for the life sciences. [More about EMBL-EBI and our impact](#)

##### Services

We provide freely available data and bioinformatics services to all facets of the scientific community >

##### Research

We contribute to the advancement of biology through basic investigator-driven research >

##### Training

We provide advanced bioinformatics training to scientists at all levels >

##### Industry

We help disseminate cutting-edge technologies to industry >

##### ELIXIR

We support, as an ELIXIR node, the coordination of biological data provision throughout Europe >

2. Sur l'interface recherche de celle-ci tapez le mot **Auxin Response Factor** (toujours en anglais svp) puis cliquez sur le bouton search

Noté bien qu'EMBL propose les différentes banques de données qui ont été consultées afin de répondre à notre requête (à gauche de la page).

On tape " **Nucleotide Sequences** " pour rechercher la séquence nucléotidique (ADN ou ARN)

3. Cliquez sur " **Nucleotide Sequences** ".

**4. Cliquez sur le numéro d'accèsion de la séquence : [AJ306306](#)**

C'est le code d'entrée à cette fiche.

Vous avez le choix entre 2 types de format :

- **Le format TEXT (EMBL)**

- **Le format FASTA**

\* Lorsqu'on clique sur EMBL on obtient la page suivante :

```

ID   AJ306306; SV 1; linear; mRNA; STD; PLN; 2511 BP.
XX
AC   AJ306306;
XX
DT   09-OCT-2002 (Rel. 73, Created)
DT   09-OCT-2002 (Rel. 73, Last updated, Version 1)
XX
DE   Oryza sativa subsp. japonica mRNA for auxin response factor 1 (arf1 gene)
XX
KW   arf1 gene; auxin response factor 1.
XX
OS   Oryza sativa (rice)
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; BOP clade;
OC   Oryzoideae; Oryzeae; Oryzinae; Oryza.
XX
RN   [1]
RP   1-2511
RA   Waller F.;
RT   ;
RL   Submitted (08-FEB-2001) to the INSDC.
RL   Waller F., Institut fuer Biologie II, Albert-Ludwigs-Universitaet Freiburg,
RL   Schaezlestr.1, Freiburg, D-79104 , GERMANY.
XX
RN   [2]
RX   DOI; 10.1023/A:1019818110761.
RX   PUBMED; 12369618.
RA   Waller F., Furuya M., Nick P.;
RT   "OsARF1, an auxin response factor from rice, is auxin-regulated and
RT   classifies as a primary auxin responsive gene";
RL   Plant Mol. Biol. 50(3):415-425(2002).
XX
DR   MD5; dc65c8b1dd04ce3aa5b7ff4fa5dd7c02.
XX
FH   Key                Location/Qualifiers
FH
FT   source              1..2511
FT                       /organism="Oryza sativa"
FT                       /sub_species="japonica"
FT                       /cultivar="Nihonmasari"
FT                       /mol_type="mRNA"
FT                       /db_xref="taxon:4530"
FT   CDS                 1..2511
FT                       /gene="arf1"
FT                       /product="auxin response factor 1"
FT                       /function="putative transcription factor"
FT                       /db_xref="GOA:Q2R3F5"
FT                       /db_xref="InterPro:IPR000270"
FT                       /db_xref="InterPro:IPR003340"
FT                       /db_xref="InterPro:IPR010525"
FT                       /db_xref="InterPro:IPR015300"
FT                       /db_xref="InterPro:IPR033389"
FT                       /db_xref="UniProtKB/Swiss-Prot:Q2R3F5"

```

```

FT      /experiment="experimental evidence, no additional details
FT      recorded"
FT      /protein_id="CAC83756.1"
FT      /translation="MATAEVGGGGGGGGGGEDALFTELWSACAGPLVTVPRVGEKVFY
FT      FPQGHIEQVEASTNQVGEQRMQLYNLPWKILCEVMNVELKAEPDDEVYAQLTLLPESK
FT      QQEDNGSTEEEVPSAPAAGHVRPRVHSFCKTLTASDTSTHGGFVLRRADECLPPLDM
FT      SRQPPTQELVAKDLHGVEWRFRHIFRQPRRHLQSGWSVVFVSAKRLVAGDAFIFLRGE
FT      NGELRVGVRRAMRQQTNPSSVISSSHMLGLVLATAWHAVNTGTMFTVYKPRTPAEF
FT      VVPYDRYMESLKQNYSIGMRFKMRFEGEEAPEQRFTGTIVGMGSDPAGWPESKWRSLK
FT      VRWDEASSIPRPERSVPWQIEPAVSPPPVNPLVPRTRKRLRPNATALPADSSAIAKEAA
FT      TKVVVESEPNGTQRTFQTQENATPKSGFGNSSELESAQKSIMRPSGFDREKNNTPIQWK
FT      LGSDGRMQMSKPESYSEMLSGFQPPKDVQIQGFWSLPEQITAGHSNFWHTVNAQYQDQ
FT      QSNHNMFPSSWSFMPNTRLGLNKQNYSMIQEAGVLSQRPGNTKFGNGVYAALPGRGTE
FT      QYSGGWFGHMMPNSHMDDTQPRLIKPKPLVVAHGVDVQKAKGASCKLFGIHLDSPAKSEP
FT      LKSPSSVVYDGTPTPGATEWRRPDVTEVEKCSDPKAMKPLDTPQPDSPVEKPSQQA
FT      SRNMSCKSQGVSTRSCKKVHKQGIALGRSVDLTKFNGYEELIAELDDMFDFNGELKGPK
FT      KEWMVVYTDNEGDMMLVGDDPWIEFCDMVHKIFIYTREEVQRMNPGTLNSRSEDSHANS
FT      MERGSVGREMRGCLSTSSLNSENC"
XX
SQ

```

Sequence 2511 BP; 673 A; 584 C; 677 G; 577 T; 0 other;

```

atggcgacgg cggaggtcgg cggcggcgcc ggaggaggag gaggaggagg ggaggacgcg      60
ctgttcacgg agctgtggag cgcgtgcgcg gggccgctgg tgacggtgcc gagggttggg      120
gagaaggtgt tctacttccc gcagggggcac atcgagcagg tggaggcctc gaccaaccag      180
gtaggcgagc agcggatgca gttgtacaat ctcccatgga agatcctgtg tgaggttatg      240
aacgttgagc tgaaggccga gccagacacc gacgaggttt atgctcagct cactctgctt      300
cctgaatcga agcaacaaga agataacggc tctactgagg aggagggtgcc ttctgctcca      360
gcagctgggc atgttaggcc acgggtgcac tcattctgca agacattgac cgcctcggac      420
acgagcacac atggcggcct ctccggtgctg cgacgccacg cggacgaatg cctcccacca      480
ctggatatga gccgtcagcc tccaacacaa gagctggtgg ccaaggatct gcacggtgtg      540
gaatggcgct ttcgtcacat attcagaggt caaccacgaa gacacctttt gcagagtggc      600
tggagtgtgt tcgttagtgc caaacggcct gttgctgggg atgcctttat ctttctcaga      660
ggtgagaacg gggaaactgc tgttgagtc aggcgtgcaa tgaggcagca aactaatggt      720
ccatcttcag tgatatcaag ccacagcatg catcttgggg ttcttgccac agcatggcat      780
gctgttaaca cagggacat gttcactgtc tactacaagc ctaggacaag tccagccgag      840
tttgtggtcc cttacgatcg ctatatggaa tcaactgaaac aaaattactc cattgggatg      900
agatttaaga tgaggtttga aggtgaagag gctccagagc aaagattcac tgggactatt      960
gtcggaatgg gggattctga tccagctggc tggcctgaat caaaatggcg ctcccttaag      1020
gtgagatggg atgaagcttc ttccatacct cgccctgaaa gaggtttctcc ctggcaaata      1080
gaacctgtg taagccctcc tccctgtcaac ccaactaccg ttccaagaac caagaggctt      1140
cgtccgaatg taccgcttt gccggctgac tcttccgcta tagcaaaaga agctgctacg      1200
aaagtgtggt tcgagtctga accaaatggt acgcaaagga ccttccagac acaagagaat      1260
gcaaccccaa aaagtgttt cggcaatagc agcgagttag aaagtgctca gaaatcaatc      1320
atgcttccat ctggatttga tcgtgagaaa aataacactc ccatacagtg gaagctaggt      1380
tcagatggtc ggatgcagat gagcaagcct gagagttaca gtgagatgct atctggattt      1440
cagccacctc aagatgtaca aatcccacag ggtttctggt ctttacctga gcagattact      1500
gcaggccatt ctaacttctg gcacacagta aatgctcagt atcaagatca acagagcaat      1560
cacaatatgt tcccagctc atggctcttc atgcctccaa atactcgctc tggattaaac      1620
aaacaaaact attcaatgat tcaggaagct ggtgtggtgt ctcaaaggcc tggaaataca      1680
aagtttggga atggagttta tgctgcactg ccaggccgag gtactgagca atactcaggg      1740
ggttggtttg gccatatgat gccaaattcc cacatggatg atacacagcc acgcctgatc      1800
aagcctaac ctctggttgt tgccatggt gatgttcaga aagctaaagg tgcttcatgc      1860
aaactatttg gaattcacct tgacagccca gcgaaatctg aacctttgaa atctccatca      1920
agtgttgat atgatgggac gccacaaaca ccagggtgcta ctgaatggcg ccgaccagat      1980
gtaactgaag tagagaaatg ttctgatctc tctaaggcca tgaagccact tgatactccg      2040
caaccagact ctgttcccga gaagccttct tctcagcaag cttcacgaaa catgtcgtgc      2100
aatcacaag gtgtatcaac tagaagctgc aagaaggtcc acaagcaagg cattgcactt      2160
ggcaggtctg tggaccttac aaagtttaat ggctatgaag agttgattgc tgagctggat      2220
gacatgtttg atttcaatgg tgaactgaag ggtcctaaga aggagtggat ggttgtctac      2280
actgacaacg aaggtgacat gatgctagtt ggggatgatc cctggattga attctgcgac      2340
atggttcaca agatcttcat ctacacaaga gaagaggttc agcggatgaa tccgggaacc      2400
ctgaactcaa gatcagaaga tagtcatgct aattcaatgg aaaggggctc agtcggcaga      2460
gaaatgcgag gctgcttgct aacctcgtct cttaattctg agaactgcta a      2511

```

//

**3 – Lorsqu'on clique sur FASTA on obtient la page suivante :**

```
>ENA|AJ306306|AJ306306.1 Oryza sativa subsp. japonica mRNA for auxin response factor 1
(arf1 gene)
ATGGCGACGGCGGAGGTTCGGCGGGCGGGCGGGAGGAGGAGGAGGAGGAGGGGAGGACGCG
CTGTTTCACGGAGCTGTGGAGCGCGTGCAGCGGGGCCGCTGGTGCAGGTGCCGAGGGTTGGG
GAGAAGGTGTTCTACTTCCCGCAGGGGCACATCGAGCAGGTGGAGGCCTCGACCAACCAG
GTAGGCGAGCAGCGGATGCAGTTGTACAATCTCCCATGGAAGATCCTGTGTGAGGTTATG
AACGTTGAGCTGAAGGCCGAGCCAGACACCGACGAGGTTTATGCTCAGCTCACTCTGCTT
CCTGAATCGAAGCAACAAGAAGATAACGGCTCTACTGAGGAGGAGGTGCCTTCTGCTCCA
GCAGCTGGGCATGTTAGGCCACGGGTGCACTCATTCTGCAAGACATTGACCGCCTCGGAC
ACGAGCACACATGGCGGCTTCTCGGTGCTGCGACGCCACGCGGACGAATGCCTCCCACCA
CTGGATATGAGCCGTGAGCCTCCAACACAAGAGCTGGTGGCCAAGGATCTGCACGGTGTG
GAATGGCGCTTTCGTCACATATTCAGAGGTCAACCACGAAGACACCTTTTGCAGAGTGGC
TGGAGTGTGTTTCGTTAGTGCCAAACGGCTTGTGCTGGGGATGCCTTTATCTTTCTCAGA
GGTGAACGGGAACTGCGTGTGGAGTCAAGGCGTGAATGAGGCAGCAAACCTAATGTT
CCATCTTCAGTGATATCAAGCCACAGCATGCATCTTGGGGTTCTTGCCACAGCATGGCAT
GCTGTTAACACAGGGACCATGTTCACTGTCTACTACAAGCCTAGGACAAGTCCAGCCGAG
TTTGTGGTCCCTTACGATCGCTATATGGAATCACTGAAACAAAATTAATCCATTGGGATG
AGATTTAAGATGAGGTTTGAAGGTGAAGAGGCTCCAGAGCAAAGATTCAGTGGGACTATT
GTCGGAATGGGGGATTCTGATCCAGCTGGCTGGCCTGAATCAAAATGGCGCTCCCTTAAG
GTGAGATGGGATGAAGCTTCTCCATACCTCGCCCTGAAAGAGTTTCTCCCTGGCAAATA
GAACCTGCTGTAAGCCCTCTCTGTCAACCCACTACCAGTTCCAAGAACAAGAGGCTT
CGTCCGAATGCTGACCTTTGCCCAGTCACTCTTCCGCTATAGCAAAGAAGCTGCTGACG
AAAGTTGTGGTTCGAGTCTGAACCAAATGGTACGCAAAGGACCTTCCAGACACAAGAGAAT
GCAACCCCAAAAAGTGGTTTTCGGCAATAGCAGCGAGTTAGAAAGTGTGCAAGAAATCAATC
ATGCGTCCATCTGGATTTGATCGTGAGAAAATAAACACTCCCATACAGTGGAAGCTAGGT
TCAGATGGTTCGGATGCAGATGAGCAAGCCTGAGAGTTACAGTGAGATGCTATCTGGATTT
CAGCCACCTAAAGATGTACAAATCCACAGGGTTTTCTGGTCTTTACCTGAGCAGATTACT
GCAGGCCATTTCACTTCTGGCACACAGTAAATGCTCAGTATCAAGATCAACAGAGCAAT
CACAATATGTTCCCCAGCTCATGGTCTTTCATGCCTCCAAATACTCGCCTTGGATTAAC
AAACAAAACCTATTTCAATGATTCAGGAAGCTGGTGTGTTGTCTCAAAGGCCTGGAAATACA
AAGTTTGGGAATGGAGTTTATGCTGCACTGCCAGGCCGAGGTAAGTACTGAGCAATACTCAGGG
GGTTGGTTTTGGCCATATGATGCCAAATTTCCACATGGATGATACACAGCCACGCCTGATC
AAGCCTAAACCTCTGGTTGTTGCCCATGGTGATGTTTCAGAAAGCTAAAGGTGCTTTCATGC
AAACTATTTGGAATTCACCTTGACAGCCCAGCGAAATCTGAACCTTTGAAATCTCCATCA
AGTGTGTATATGATGGGACGCCACAAACACCAGGTGCTACTGAATGGCGCCGACCAGAT
GTAAGTGAAGTAGAGAAATGTTCTGATCCTTCTAAGGCCATGAAGCCACTTGATACTCCG
CAACCAGACTCTGTTCCCGAGAAGCCTTCTTCTCAGCAAGCTTACGAAACATGTCGTGC
AAATCACAAGGTGATCAACTAGAAGCTGCAAGAAGGTCCACAAGCAAGGCATTTGCACTT
GGCAGTCTGTGGACCTTACAAGTTAATGGCTATGAAGAGTTGATTGCTGAGCTGGAT
GACATGTTTGAATTTCAATGGTGAAGTGAAGGGTCTAAGAAGGAGTGGATGGTTGTCTAC
ACTGACAACGAAGGTGACATGATGCTAGTTGGGGATGATCCCTGGATTGAATTCTGCGAC
ATGGTTTACAAGATCTTTCATCTACACAAGAGAAGAGGTTTCAGCGGATGAATCCGGGAACC
CTGAACTCAAGATCAGAAGATAGTCATGCTAATTTCAATGGAAAGGGGCTCAGTCGGCAGA
GAAATGCGAGGCTGCTTGTCAACCTCGTCTCTTAATTTCTGAGAACTGCTAA
```

5- Allez sur le moteur de recherche google et taper GenBank :

6- sur la barre de recherche retaper le numéro d'accèsion de la séquence précédente : [AJ306306](#)

- Qu'est-ce que tu remarques ?

7- Allez sur le moteur de recherche google et taper DDBJ:

8- Cliquer sur [Séquences annotées/assemblées \(DDBJ\)](#), puis sur Search et sélectionner Getentry :

- Dans la barre identifiant retaper le numéro d'accèsion de la séquence précédente : [AJ306306](#)

- Qu'est-ce que tu remarques ?
- Que peut 'on dire sur ces trois banques ?

**Application individuelle :**

Il s'agit de vous informer sur la structure primaire d'une enzyme clé de la photosynthèse : la Rubisco.

Support pédagogique : <http://www.com.univ-mrs.fr/IRD/atollpol/glossaire/photos.htm>