

Université Batna 2

Année universitaire : 2022-2023

Faculté des Sciences de la nature et de la vie

Département d'Ecologie et environnement

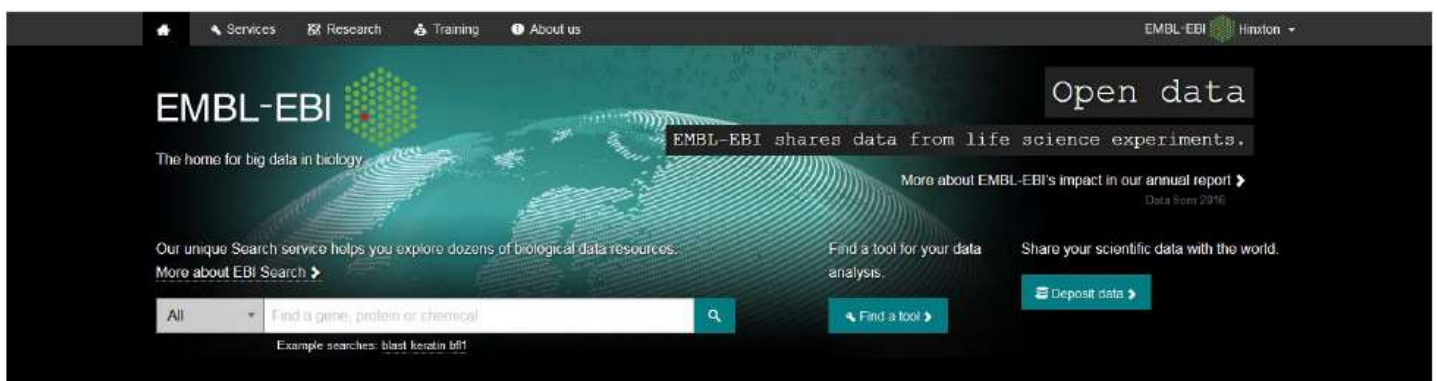
Travail Pratique 1 : Les banques de séquences biologiques

(Exemple : EMBL)

Réalisons le TP suivant sur Internet pour connaître la **séquence primaire d'ARNm** du gène qui code pour une enzyme dite **PPO (polyphenol oxydase)** de la plante *Prunus armeniaca*, 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 **polyphenol oxydase** (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 : [AF020786](#)

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   AF020786; SV 1; linear; mRNA; STD; PLN; 2070 BP.
XX
AC   AF020786;
XX
DT   02-JUL-1998 (Rel. 56, Created)
DT   15-APR-2005 (Rel. 83, Last updated, Version 4)
XX
DE   Prunus armeniaca polyphenol oxidase precursor, mRNA, nuclear gene encoding
DE   chloroplast protein, complete cds.
XX
KW   .
XX
OS   Prunus armeniaca (apricot)
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae; Pentapetalae;
OC   rosids; fabids; Rosales; Rosaceae; Amygdaloideae; Amygdaleae; Prunus.
XX
RN   [1]
RP   1-2070
RX   DOI; 10.1104/pp.119.4.1261.
RX   PUBMED; 10198084.
RA   Chevalier T., de Rigal D., Mbeguie-A-Mbeguie D., Gaillard F.,
RA   Richard-Forget F., Fils-Lycaon B.R.;
RT   "Molecular cloning and characterization of apricot fruit polyphenol
RT   oxidase";
RL   Plant Physiol. 119(4):1261-1270(1999).
XX
RN   [2]
RP   1-2070
RA   Chevalier T., Mbeguie-A-Mbeguie D., Fils-Lycaon B.;
RT   ;
RL   Submitted (25-AUG-1997) to the INSDC.
RL   Technologie des Produits Vegetaux, INRA, Site Agroparc, Domaine Saint Paul,
RL   Avignon 84914 cedex, France
XX
DR   MD5; a6caa9a6ccb2d004e4c703cf6964eef5.
DR   EuropePMC; PMC196857; 12960376.
XX
FH   Key                Location/Qualifiers
FH
FT   source              1..2070
FT                       /organism="Prunus armeniaca"
FT                       /strain="Bergeron"
FT                       /mol_type="mRNA"
FT                       /dev_stage="green fruit"
FT                       /clone="pAPRI101G"
FT                       /tissue_type="mesocarp; exocarp"
FT                       /db_xref="taxon:36596"
FT   transit_peptide    4..306
FT   CDS                 4..1797
FT                       /codon_start=1
FT                       /product="polyphenol oxidase precursor"
FT                       /function="catalyze the oxidation of mono- and O-diphenols

```

```

FT          to O-diquinones"
FT          /note="catalytic activity: 2 catechol + O(2) = 2
FT          1,2-benzoquinone + 2 H(2)O; cofactor: contains two copper
FT          atoms (by similarity); belongs to the tyrosinase family;
FT          located within the thylacoid lumen"
FT          /db_xref="GOA:O81103"
FT          /db_xref="InterPro:IPR002227"
FT          /db_xref="InterPro:IPR008922"
FT          /db_xref="InterPro:IPR016213"
FT          /db_xref="InterPro:IPR022739"
FT          /db_xref="InterPro:IPR022740"
FT          /db_xref="UniProtKB/Swiss-Prot:O81103"
FT          /protein_id="AAC28935.1"
FT          /translation="MATAPSPTTMGTYSLLISTNSFSTFLPNKSQLSLSGKSKHYVARR
FT          SSISCKATNNNNSNNQNEQQEESRLLGKLDRRNILIGLGGLYGATTLDRKPFPAFADPI
FT          APPDLTTCKPAEITPGGSETVPCPPVTTKIKTFKPDLSIPLRTSPAAHQVTDEYLAKF
FT          KKAQAAMRALPDDDPMSMVQQAQKVHCAYCNGAYPQVGFTDNDIQVHFSWLFPPFHRMYL
FT          YFYERILGKLIDDPTFALPYWNWSDSPVGFPIPDITYTDTSSPLYDQYRNADHQPPVLVDL
FT          SYGGKDDDDVDEQTRIDENLAIMYRQMVSGAKTDLFFGHAYRAGNLNTGKYPGTIENMP
FT          HNNIHIWVGDPQSQTHQEDMGNFYSAGRDLFYAHHANVDRMWNWIKTLGGKRKIDITDID
FT          WLDAEFLFYDENAELELVVRVVRDLSLEPEKQLRYNYEPVSLPWLFTKPTARKTKNKTAKV
FT          AATQLTSKFPATLVEVTTVEVARPKPRKRSKKEKVDEEELLIKIDIEFEGTEAVKFDVF
FT          INDDAESLSRRDKSEFAGSFVHVPQGKTTKAKTKTNLKLGITDLLEDLGAEDDSSVLVT
FT          LVPRVSNPITIGGFKIEYSS"
FT mat_peptide 307..1794
FT          /product="polyphenol oxidase"
XX
SQ

```

```

Sequence 2070 BP; 564 A; 546 C; 504 G; 456 T; 0 other;
gccatggcca ctgcaccctc tcccacaacc atgggcactt actcttctct catctccacc      60
aattccttct caactttcct cccaaacaaa tcccaacttt cttatctgga aaaagcaag      120
cactacgttg cagcaggtc atcaatctca tgcaaagcca caataataa caatagcaat      180
aatcaaaatg aacaacaaga agaatcttct cgtttggttg gaaaattgga caggagaaat      240
atcctcattg gcctcggcgg cttgtatggc gccaccaccc tcgaccggaa accattcgcc      300
ttcgccgacc cgatagcccc accagactta accacgtgca agccggccga aatcaccccc      360
ggaggaagtg aaaccgtccc ctgctgccct cccgtgacga caaagatcaa aaccttcaa      420
ccagacctgt cgattccgct ggggacgagc cccgccgccc accaagtcac ggacgagtag      480
ttggccaagt tcaagaaagc ccaagctgct atgagagcct taccgatga cgatccacgt      540
agcatggtgc aacaagctaa agttcactgt gcctactgta acggtgcata tccacaggta      600
aggtttacgg acaatgacat ccaagttcac ttttcatggc tcttcttcc cttccaccgc      660
gtgtacttgt acttctacga gagaatctc ggcaagctca ttgatgacc aactttcgct      720
cttcataact ggaactggga ctctccagta ggcttcccaa ttccagacat ttatactgac      780
accagctccc cactctacga ccagtacaga aacgccgacc accagccgcc ggttctggtg      840
gacctcagct acggtgggaa agacgatgac gtggacgaac agacaagaat agacgagaat      900
ctggccatca tgtaccgtca gatggtttcc ggtgccccaa ccccgatct gtttttcggc      960
cagcgtaca gggccgggaa cctaaacacg ggaaaatacc ccggcaccat cgagaacatg     1020
cctcacaaca acatccatat atgggtcggc gacccaagcc agaccacca agaggacatg     1080
ggcaacttct actccgcccg tcgtgatccc ctgttttacg cccaccatgc caacgtggac     1140
cgcatgtgga acatctggaa aacccttggg ggcaagcgca aggacatcac cgacaccgat     1200
tggctcgacg ccgagtttct gttctacgac gagaatgctg agcttgtgag agtcaaagtt     1260
cgggacagcc tcgaaccaga gaagcagctt cgatataatt acgaacctgt gagtcttccg     1320
tggctgttca caaaaccaac cgctcgtaaa acgaagaaca agacgaaagc taaggtggct     1380
gcgacgcagc tgacgtcgaa gttcccggcc acgttggttg aggtgacgac ggttgaggtg     1440
gcgaggccga agccgaggaa gaggagcaag aaggagaagg ttgatgagga ggagctgctg     1500
atcatcaagg acattgagtt cgagggtacc gaggcgggtg agtttgatgt gtttattaac     1560
gacgacgctg agtcgctgag tcggagggac aagtcggagt ttgctgggag cttcgtgcac     1620
gtgccacaag ggaagacgac gaaggcgaag acgaagacga acctgaaatt ggggataacg     1680
gactgtttgg aggacttggg agcggaggat gacagcagtg tgctgggtgac tttggtgcct     1740
agggtttcga actcgcgat caccattggg gggttcaaga tcgagtattc ttcttgatct     1800
caattaagta tcgtggccct taatgaatca ataaagtggg ttcttgtggt aatgtaataa     1860
tgtctgtgag gaattgatgg cctgcgttgc cccgctttta gttcatctga ttttctatca     1920
gtgtttggat tatggagcca tgcaattaca taagctggcc tttccatgta attaataatc     1980
caccgtttac gtggatggaa tatggatggt gtctacttgt tcatcaataa agccacgttg     2040
tttgatttaa aaaaaaaaaa aaaaaaaaaa

```

//

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

```
>ENA|AF020786|AF020786.1 Prunus armeniaca polyphenol oxidase precursor, mRNA, nuclear
gene encoding chloroplast protein, complete cds.
GCCATGGCCACTGCACCCTCTCCCACAACCATGGGCACTTACTCTTCTCTCATCTCCACC
AATTCCTTCTCAACTTTCCTCCCAAACAATCCCAACTTTCCTTATCTGGAAAAAGCAAG
CACTACGTTGCACGCAGGTCAATCTCATGCAAAGCCACAAATAATAACAATAGCAAT
AATCAAAATGAACAACAAGAAGAATCTTCTCGTTTTGTTGGGAAAATTGGACAGGAGAAAT
ATCCTCATTGGCCTCGGCGGCTTGTATGGCGCCACCACCCTCGACCGGAAACCATTGCGC
TTCGCCGACCCGATAGCCCCACCAGACTTAACCACGTGCAAGCCGGCCGAAATCACCCCC
GGAGGAAGTGAACCCGTCCCCTGCTGCCCTCCCCTGACGACAAAGATCAAAACCTTCAA
CCAGACCTGTTCGATTCCGCTGCGGACGAGCCCCGCGCCACCAAGTCACGGACGAGTAC
TTGGCCAAGTTCAGAAAGCCCAAGCTGCCATGAGAGCCTTACCCGATGACGATCCACGT
AGCATGGTGCAACAAGCTAAAGTTCCTGTGCCTACTGTAACGGTGCATATCCACAGGTA
GGGTTTACGGACAATGACATCCAAGTTCACCTTTTCATGGCTCTTCTTTCCCTTCCACCGC
ATGTACTTGTACTTCTACGAGAGAATCCTCGGCAAGCTCATTGATGACCCAACCTTTCGCT
CTTCCATACTGGAAGTGGGACTCTCCAGTAGGCTTCCCAATTCCAGACATTTATACTGAC
ACCAGCTCCCCACTCTACGACCAGTACAGAAACGCCGACCACCAGCCGCGGTTCTGGTG
GACCTCAGCTACGGTGGGAAAGACGATGACGTGGACGAACAGACAAGAATAGACGAGAAT
CTGGCCATCATGTACCGTCAGATGGTTTTCCGGTGCCAAAACCCCGATCTGTTTTTCCGGC
CACGCGTACAGGGCCGGAAACCTAAACACGGGAAAATACCCCGGCACCATCGAGAACATG
CCTCACAACAACATCCATATATGGGTGGGCGACCCAAGCCAGACCCACCAAGAGGACATG
GGCAACTTCTACTCCGCGGTCTGTATCCCCTGTTTTACGCCACCATGCCAACGTGGAC
CGCATGTGGAACATCTGGAAACCTTGGAGGCAAGCGCAAGGACATCACCGACCCGAT
TGGCTCGACGCCGAGTTTCTGTTCTACGACGAGAATGCTGAGCTTGTGCGAGTCAAAGTT
CGGGACAGCCTCGAACCCAGAGAAGCAGCTTCGATATAATTACGAACCTGTGAGTCTTCCG
TGGCTGTTTCAAAAACCAACCGCTCGTAAAACGAAGAACAAGACGAAAGCTAAGGTGGCT
GCGACGCAGCTGACGTCAAGTTCGCGCCACGTTGGTTGAGGTGACGACGGTTGAGGTG
GCGAGGCCGAAGCCGAGGAAGAGGAGCAAGAAGGAGAAGGTTGATGAGGAGGAGCTGCTG
ATCATCAAGGACATTGAGTTCGAGGGTACCGAGGCGGTGAAGTTTGTGTTTATTAAC
GACGACGCTGAGTCGCTGAGTCGGAGGGACAAGTCGGAGTTTGTGGGAGCTTCGTGCAC
GTGCCACAAGGGAAGACGACGAAGGCGAAGACGAAGACGAACCTGAAATTGGGGATAACG
GACTTGTGGAGGACTTGGGAGCGGAGGATGACAGCAGTGTGCTGGTGACTTTGGTGCCT
AGGGTTTTCGAACTCGCCGATCACCATTGGTGGGTTCAAGATCGAGTATTCTTCTTGATCT
CAATTAAGTATCGTGGCCCTTAATGAATCAATAAAGTGGGTTCTTGTGTTAATGTAATAA
TGTCTGTGCGGAATTGATGGCCTGCGTTGCCCGCTTTTAGTTTCATCTGATTTTCTATCA
GTGTTTGGATTATGGAGCCATGCAATTACATAAGCTGGCCTTTCCATGTAATTAATATTC
CACCGTTTACGTGGATGGAATATGGATGTTGTCTACTTGTTCATCAATAAAGCCACGTTG
TTTGGATTAATAAAAAAAAAAAAAAAAAAAAAA
```

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

6- sur la barre de recherche retaper le numéro d'accension de la séquence précédente : [AF020786](#)

- 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'accension de la séquence précédente : [AF020786](#)

- 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>