

ANNEXE 2: Le barcoding ou comment identifier une espèce à partir de son ADN

Dans certaines conditions, il est impossible d'identifier des individus par une reconnaissance morphologique (espèces trop similaires entre elles ; individus incomplets). L'identification peut alors s'effectuer à partir de fragments d'ADN récoltés dans l'environnement étudié.

Le logiciel BLAST (Basic Local Alignment Search Tool) permet de comparer les fragments d'ADN non identifiés et récoltés dans un environnement avec des bases de données internationales. Cette comparaison peut permettre ainsi de déterminer une espèce à partir d'une séquence d'ADN.

Protocole à suivre :

1. Ouvrir le logiciel BLAST (accès via internet en tapant « blast.ncbi » dans le moteur de recherche)
2. Cliquer « Nucléotide Blast »
3. Recopier une des séquences notées sur la page 2, puis copier-la dans la fenêtre « Enter accession number »
4. Lancer la recherche en cliquant sur BLAST (en bas de page) et **attendre quelques secondes**.

Plusieurs séquences peuvent être proposées comme identiques ou proches de la séquence inconnue à identifier. L'identification de l'espèce à laquelle appartient la séquence inconnue est possible si le pourcentage d'identité (notée Per. Ident) est supérieur à 85%. Si cette valeur est de 100 %, cela signifie que la séquence est complètement identique à celle de l'espèce proposée par le logiciel. Cela indique que l'identification de l'origine de la séquence d'ADN est certaine. Dans certains cas, l'espèce n'est pas connue.

5. A partir du nom d'espèce proposé (colonne « scientific name ») par le logiciel, légèrer les images de la fiche.
6. Renouveler ces différentes étapes pour déterminer les autres espèces

Séquences d'ADN à utiliser pour l'identification des espèces

Séquence A

```
ATGAGTAGTGTATACGATTGGTTCCAGGAACGTCTTGAGATTCAGGCGATTGCTGATGATATTACCAGCAAATATGTTCCCCACA
CGTAAATATTTTTATTGCTTGGCGGGATCACTTTAACTTGTTTTATTGTACAAGTTGCAACAGGTTTCGCAATGACATTCTACTAC
CGTCTACGGTTACAGAAGCATTTCATCGGTTGAATATTTAATGACTGAAGTGAACCTTCGGATGGTTAATTAGATCTGTGCACCG
TTGGTCAGCGAGTATGATGGTTTTAAACATGATTCTTCATGTTTGTGCGAGTATATTTAACAGGTGGTTTTAAAAGCCACGTGAATT
AACATGGGTAACCTGGTGGTTTTATTAGCATCAGTAACTGTATCATTGGTGTAACTGGGTATTCATTACCTGGGATCAAGTAGGCTA
CTGGGCTTGTAAAGATTGTAACAGGTGTACCAGATGCCGTACCAATTGTTGGTGGTTTAAATTGTACAAGTTTTACGTGGTGGTGTA
GTGTAGGTCAATCAACATTGACGCGTTTCTATAGTGCCCATACTTTTCGTATTACCCGTTGTAGCAGCAGTTTTAATGCTAACGCACT
TCGTAATGATTCGTAAGCAAGGGATTCAGGGCCTCTATAA
```

Séquence B

```
CGCATATTTGCGGATATACATCACATTTATGAGATTTAAATTTGACAAGAAAGGGGAAGTGTGTAAGAGGCGCGCCGCACTCAGG
CGAAGCCGTGAAAGGAGTGAAACGAAAAAAGCCCCGTAGGGGCTAATCAGTGTTTACAATATCGAAAAGTATTAATGCGAA
TTGGCAAGTATACAGCGTAAAATAATCAAGGATGTGATAGCGGCAGTTTCTACCGTTTCGCTTTCCCCGTCTCTAATCTAATGAC
CTTACCCGACGTTTCGATTAATATTTAACTTTCCGCCCCACCAATTATTGATTGATATAATCATGCTATTCTTTAGGTGAATCAA
ACTTTGTCGCACTTGGCACCAATACAATTTAACCCCGCTTCATGCGGAAGCGGTACGGGGTCAAGTTCGTAATGGTCGCCCAT
TGGTCATGCATATACTCGATGATATCGTCTGCGTCGTCTTCATGTAATTAGTGATGAGATACAAAACGCCGTTAATTTGATTTTA
CATATCTAGCCATTACTGAACCCTCTCCAATTCTGTAACAGTAACTAAGTTTACTTTAGCACCCAACGATAGCGGTTAAAGCGTAGT
TCTCCAATCTGGCTCTGGCATCACCACGAACGCGCCAGCTTCCAATCAACCATTTGAGCTTTAAACGACGCAGTGTCCGCAA
CTGCGTGAATCATCGTAACTGATTCTCGCCGTGGTTTGAATTTGAAACATTCTAGCCATTTTGAAGTCCAAGTAAAAAGCGA
TGTTAGCAACGTATAGATTCCAAACACTGATTTCGCGCACCAATGGATTCATATACAGTGATATCAACAAAAGGAGCAACAATAGG
AGCGAAACCGTTTCGAGTTATTTAGCTTTTTCGTAACCATCCGCAATTCCGCCTTGCGCACCAACTGTAAGCCCAATATTGTTGT
CGAGCCTGTAATCGCGCCAAGTATAAACGATTGAGAG
```

Séquence C

CGCATTCTCTTTTTGCCCTCGAGCACCCCCAGCGCTGTGCCTACGTTTTAGACCTCAAATTGACGAACACCAAGCCGATCAA
TTCGGCAGTGGATTCTCCAGGCCAGTCAACTCCCTCGGCAGGGGTCTGCCAGCAATAGTCGCGAACACTGCAGCTGGAATTTGA
GCGTAAGCGCCTGGGACTCGAAGCCCTACTACAGGGCAGCACGGGGTCCATTCTGCCGAGTGAAATCGTTGCAGTGGCAACGT
TAGCTGGC

Séquence D

ACACAGTTGGTACAAGAGATTGCTTCAAACGCATTCCGCATCATCCACACGACAATTACAACAATGGTCAACACTGAGCAGTTG
GAAGGCGCCACGGCCGCTATTAAGAGTTGATTGCAGCCAAAGCGTGTGGTCCAATTCTGATCAGGCTGGCGTGGCACGATGCT
GGCACTTACGATGACAGTATTGGCGCGGCCGCTTGGCCAAAGTGCGGAGGTGCGAACGGCTCAATTCGCTTTGATCCTGAGATC
CTTACGCGCTAACGCAGGGCTGAAGAAGCTCTCATTCTGCTGGAGCCGATTAAGCTCAATCCCCGAGGTTGGCTATGCCG
ACTTGTTCAGCTGGCCTCCGCCACAGCCGTTGAGGTCATGGGTGGGCCACCATCCCGATGAAGTATGGCCGCAAGGATGCGA
CTGGCCCCGACATGTGTCACCCAGAGGGAAACCTTCCAGCTGGTGTCTGCCCTTGGCCCACTGGTGGTGACGCTGCTGGACACC
TTCGCGCTGTTTTCCACCGCATGGGCCTCTCTGATCAGGATATCGTCGCGCTGTCCGGTGCCATTGCGTTGGTCGTGCCATGCC
AGCCGCTCTGGCTTGTGCCACAAGGCTGAGACCAAGTACACAGCTGCTGGCGCATGTCTATGGGCACAGCAGCTACTGGTGGT
GCCTCTGGACCCCTGAGTGGACAAAGTTCGACAACTCATACTTCCAAGTGGTGAAGGATCCCAAGGACGAGGAGCTGCTTGCC
CTGGAGACAGACTGTTCTGTTCAAAGATCCCGAATTCCTGAAGTATGCTGAGAAGTATGCAGAGGACCAGGACGCCTTCTTT
GCTGACTACGCCGTGTCTCATGCAAACTGTCCGAGTTGGGTGTGGCATGGGAGGCTTAGCTTGCCC

Séquence E

ATGGCAAGCCTGCGAAAGACTCACCCACTTATAAAAATCGCCAACGACGCACTAGTCGACCTCCCGGCCCCCTCGAACATCTCG
GTCTGATGAAACTTCGGATCGTTCTTGGCCTCTGTCTAGCGGCCAGATTCTGACAGGGTTGTTCTTAGCCATGCACTACACCTC
TGACATCGCAACCGCCTTCTTCTGTTGCCACATTTGTCGTGACGTAAACTACGGATGACTGATTCGAAGTATGCACGCAAATG
GAGCATCTTTTTCTTTATTTGATTTACGCCACATTGGGCGAGGGCTCTATTATGGCTCCTATCTCTACAAGGAAACATGAAACA
TTGGAGTTGTCTCTTCTTTTGGTCATGATAACTGCCTTTGTTGTTATGTCTTACCATGAGGCCAAATGTCCTTCTGGGGGGCCA
CGGTCACTAACTACTGTCCGCCGTCCCCTACGTGGGGGTGCGCTGGTAGAGTGAATCTGAGGGGGTTTCTCTGTAGATAA
TGCGACGCTGACCCGGTTTTTCGCGTTCACCTTCTTGTTCCTTCTGATCGCGGGGCCACAGTTCTTACCTGCTCTTCTCTCC
ACGAGACTGGCTCAAACAACCCCGCAGGGCTCAACTCGGACGCGGACAAGATCTCCTTCCACCCCTACTTCTCGTACAAAGACC
TCCTGGGATTTGCAGTGATGTTGCTAACCCCTTACCTACTAGCCCTTTTCTCCCCAATCTTTTAGGAGACCCAGACAATTTACC
CCTGCCAACCCGATGGTTACCCACACACATTAAGCCGGAGTGATACTTCTGTTTGCCTATGCCATTCTCCGATCAATCCCGAA
CAAAGTAGGGGGAGTTTTGGCCCTCTATTCTATTTTGGTGCTCATGGTCGTCCTTACACCTCAAACAGCGAGGCC
TAACATCCGACCTATCACGCAGTCTTGTCTGAACCCTTGTGGCAGACGTCATCATTCTGACGTGAATTGGGGGAATGCCAGTC
GAGCATCCTTACGTATCATTGGGCAAGTGGCGTCAGTTATTTACTTCTCACTTTTCTAGTCCTCGCCCCATTGGCTGGGTGGC
AGAAAACAAAGTGCTTGATTGAAAAT