

TP 4 : Alignement des séquences du gène ARNr16S sur NCBI.

1. Introduction:

Au cours de l'évolution naturelle, les mutations causent des erreurs au moment de la réplication de l'ADN car l'évolution se fait par mutations successives. Ces erreurs peuvent être :

- ✓ Des substitutions (changement ponctuel d'un nucléotide par un autre). On parle de transition ou de trans-version,
- ✓ Des insertions (ajout d'un ou plusieurs nucléotides),
- ✓ Des délétions (suppression d'une base ou d'un segment d'ADN).

2. Alignement :

processus par lequel deux (ou n) séquences sont comparées afin d'obtenir le plus de correspondances (identités ou substitutions conservatives) possibles entre les lettres qui les composent

L'objectif de l'alignement est de disposer les composants (nucléotides ou acides aminés) pour identifier les zones de concordance

3. Les étapes de l'identification moléculaire:

- ✓ Extraction,
- ✓ PCR
- ✓ Electrophorèse
- ✓ Séquençage et Analyse Bioinformatique

4. Programmes de comparaison avec les banques

Recherches de similitudes dans les banques de séquences, Pourquoi ?

- ✓ Savoir si ma séquence ressemble à d'autres déjà connues• Trouver toutes les séquences d'une même famille
- ✓ Rechercher toutes les séquences qui contiennent un motif donné

Les deux types de programme les plus utilisés par les biologistes sont les logiciels :

4.1. FASTA

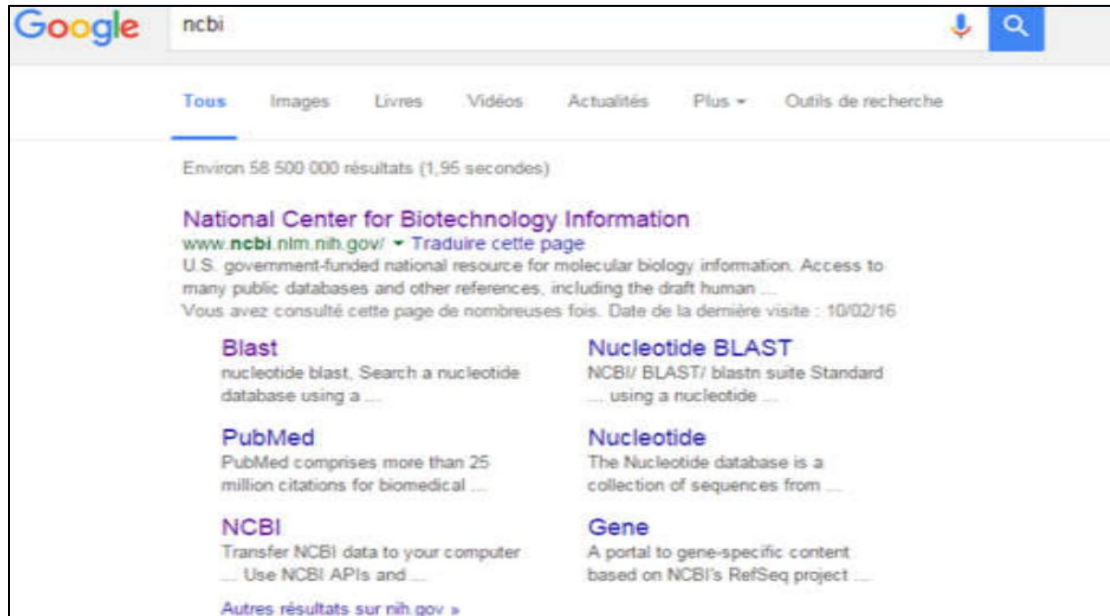
Le programme FASTA compare respectivement une séquence nucléique avec une base nucléique ou une séquence protéique avec une base protéique.

4.2. BLAST : Basic Local Alignment Search Tool

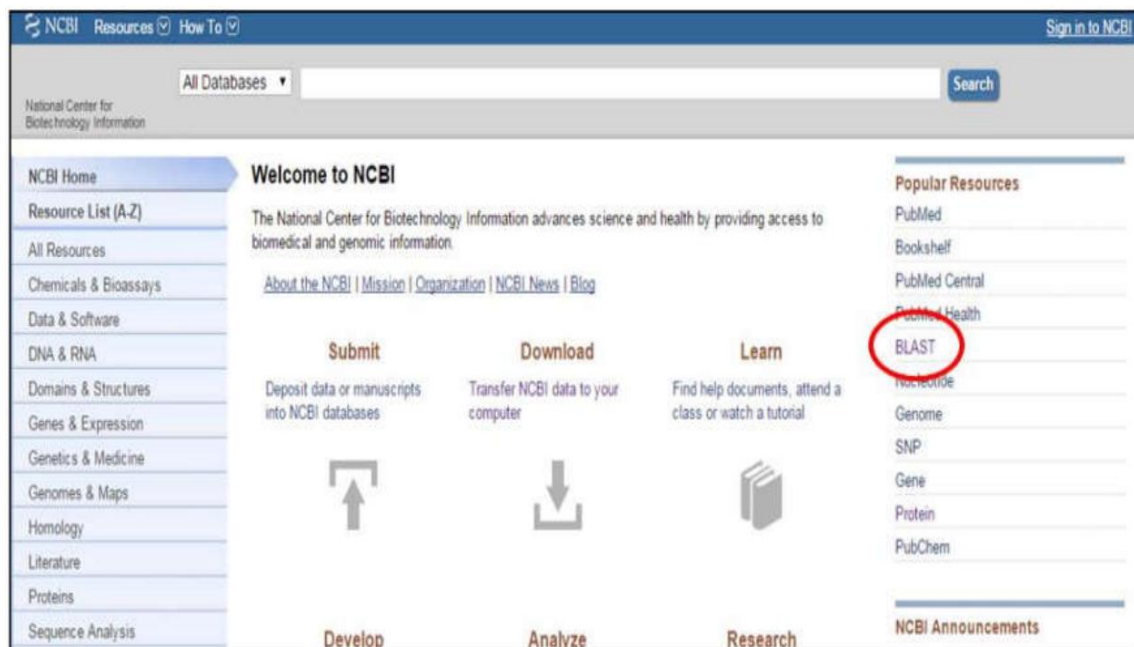
BLAST est l'outil de recherche basique d'alignement local. BLAST, cherche les bases de données des protéines et ADNs pour des séquences (sujets) qui ressemblent à notre séquence (requête) utilisée comme mot clé.

5.Méthode du travail

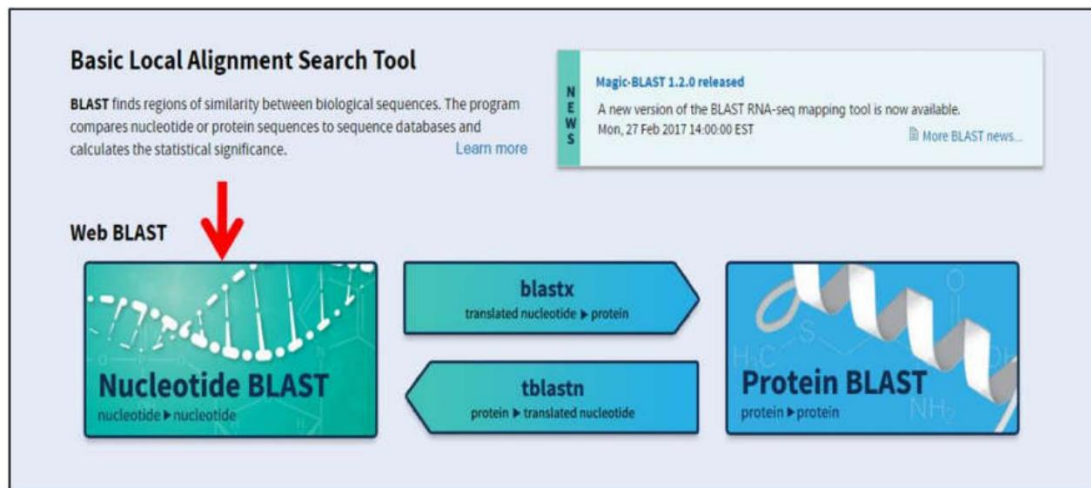
1.Ouverture du lien NCBI sur internet par l'utilisation du moteur de recherche Google



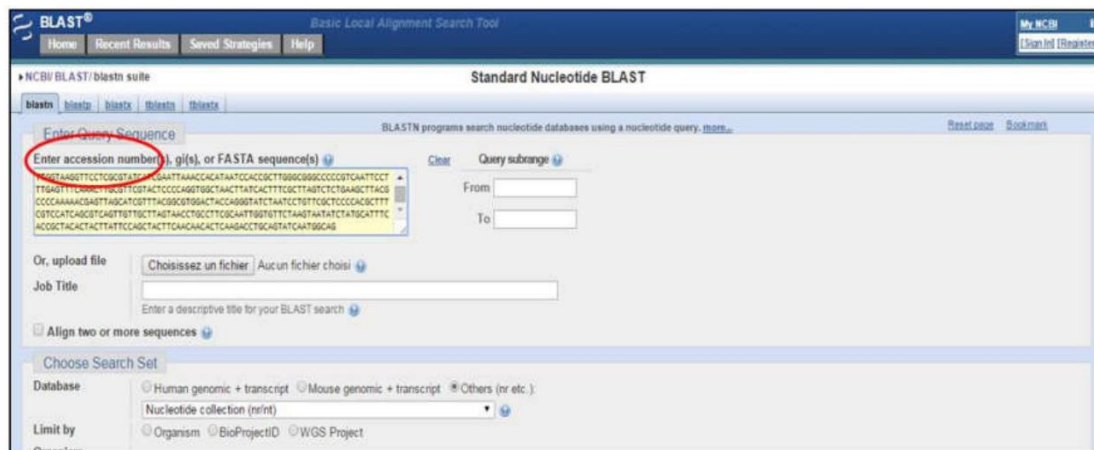
2.Choix du programme BLAST



3. Choix de l'outil nucleotide BLASTn



4. Insertion de la séquence ADN ou le Numéro d'Accès sur Gene Bank et activation de l'outil BLAST. (Code gene :HF678414.1).



5. Lecture de la liste des résultats de l'Alignement

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected: 0

Alignments [Download](#) [View](#) [Graphics](#) [Distance tree of results](#)

Description	Max score	Total score	Query cover	E value	Ident	Accession
Chryseobacterium indologenes partial 16S rRNA gene, isolate 6	1328	1328	100%	0.0	100%	HF678414.1
Chryseobacterium indologenes partial 16S rRNA gene, isolate 12	1323	1323	100%	0.0	99%	HF678418.1
Chryseobacterium indologenes partial 16S rRNA gene, isolate 3	1323	1323	100%	0.0	99%	HF678415.1
Bacterium 14S134 16S ribosomal RNA gene, partial sequence	1317	1317	100%	0.0	99%	KC734385.1
Bacterium 14S132 16S ribosomal RNA gene, partial sequence	1317	1317	100%	0.0	99%	KC734383.1
Chryseobacterium enrichment culture clone RA-M137 16S ribosomal RNA gene, partial sequence	1317	1317	100%	0.0	99%	JQ083171.1

6. Lecture du détail des résultats de l'Alignement

Download ▾ GenBank Graphics ▾ Next ▾ Previous ▾ Descriptions ▾

Chryseobacterium indologenes partial 16S rRNA gene, isolate 6
Sequence ID: **HF678414.1** Length: 719 Number of Matches: 1

Range 1: 1 to 719 [Sequence](#) [Statistics](#) [Text Match](#) [Previous Match](#)

Score	Expect	Identifiers	Caps	Strand
1328 bits(719)	0.0	719/719(100%)	0/719(0%)	Plus/Minus

Query 1 TSCCGATTACTAGCGATTCCAGCTTCATAGAGTCGAGTTCAGACTCCAATCGAACTG 60
 Sbjct 719 TSCCGATTACTAGCGATTCCAGCTTCATAGAGTCGAGTTCAGACTCCAATCGAACTG 60

Query 61 AGACCGCTTTCGAGATTTCATCAGATCCCTGTGTAGCTGCCCTGTACCCGCCATTG 120
 Sbjct 659 AGACCGCTTTCGAGATTTCATCAGATCCCTGTGTAGCTGCCCTGTACCCGCCATTG 600

Query 121 TATTACGTGTGTGGCCCAAGGCTTAAGGCGCTGATGATTTGACGTATCCACCTTCC 180
 Sbjct 599 TATTACGTGTGTGGCCCAAGGCTTAAGGCGCTGATGATTTGACGTATCCACCTTCC 540

Query 181 TCTCTACTTGCSTAGGCACTCTCAGTACAGTCCCACTTAATGATGCAACTAGTGA 240
 Sbjct 539 TCTCTACTTGCSTAGGCACTCTCAGTACAGTCCCACTTAATGATGCAACTAGTGA 480

Query 241 GGGGTTGCCTGCTTTCAGGACTTAACCTAACACCTCACGGCAGCTGACGACCA 300
 Sbjct 479 GGGGTTGCCTGCTTTCAGGACTTAACCTAACACCTCACGGCAGCTGACGACCA 420

Query 301 TCGAGCACCTTGAAAAATGTCGAGAAAAAGTCTATTCTAACCTGTCTATCCCATTT 360
 Sbjct 419 TCGAGCACCTTGAAAAATGTCGAGAAAAAGTCTATTCTAACCTGTCTATCCCATTT 360

7. Récolte des informations sur l'individu par le numéro d'accès sur Gene Bank :Auteur, affiliation, publication, séquence, etc.

Chryseobacterium indologenes partial 16S rRNA gene, isolate 6

GenBank: HF678414.1
[FASTA](#) [Graphics](#)

[Go to](#) (🔍)

LOCUS HF678414 719 bp DNA linear BCT 21-FEB-2013
 DEFINITION Chryseobacterium indologenes partial 16S rRNA gene, isolate 6.
 ACCESSION HF678414
 VERSION HF678414.1 GI:452084714
 KEYWORDS -
 SOURCE Chryseobacterium indologenes
 ORGANISM [Chryseobacterium indologenes](#)
 Bacteria; Bacteroidetes; Flavobacteriia; Flavobacteriales;
 Flavobacteriaceae; Chryseobacterium.

REFERENCE 1
 AUTHORS Boubendir,A.
 TITLE Analyse et prevalence du risque infectieux de listeria monocytogenes dans les laits crus recoltés dans deux regions a climat different (Zone semi-aride et le Nord-Est algeriens) : Modelisation spatiale de la diversite floristique
 JOURNAL Thesis (2012) Constantine 1 University, Algeria

REFERENCE 2 (bases 1 to 719)
 AUTHORS Hamidechi,A.
 TITLE Direct Submission
 JOURNAL Submitted (11-FEB-2013) Constantine University, Constantine, Route de Ain El-Bey, 25000, ALGERIA

FEATURES
 Location/Qualifiers
 source 1..719
 /organism="Chryseobacterium indologenes"
 /mol_type="genomic DNA"
 /isolate="6"
 /isolation_source="raw milk"
 /db_xref="taxon:253"

Customize view ▾

Analyze this sequence ▾
 Run BLAST
 Pick Primers
 Highlight Sequence Features
 Find in this Sequence

Related information ▾
 Taxonomy

LinkOut to external resources ▾
 Ribosomal Database Project II
 [Ribosomal Database Project II]
 SILVA SSU Database [SILVA]

Recent activity ▾
 Turn Off Clear
 Chryseobacterium indologenes partial 16S rRNA gene, isolate 6 Nucleotide
 Nucleotide Sequence (719 letters) BLAST