

أدوات ال-BLAST

-BLASTn (nucleotide blast) - مُحرك بحث يجد
سجلات تسلسلات نوكلونيتيدات مشابهة لتسلسل نوكلونيتيدات
مُعطى (تسلسل استعمال).



-BLASTn - أدوات ال-BLAST، Basic Local Alignment Search Tool، تُستعمل للبحث عن تسلسلات مُتشابهة .
BLASTn هو مُحرك بحث يجد في مخازن تسلسلات
نوكلونيتيدات سجلات مشابهة لتسلسل نوكلونيتيدات مُعطى.
في مُحركات البحث مثل Google أو Entrez الاستعمال
مُكوّن من كلمات مُفتاح، أما في مُحرك البحث BLASTn
الاستعمال مُكوّن من تسلسل نوكلونيتيدات. يُقارن مُحرك
البحث تسلسل الاستعمال مع تسلسلات موجودة في مخزن
المعلومات ويعرض سجلات ذات تسلسل مُشابه لتسلسل
الاستعمال. يُعرض تسلسل الاستعمال مُقابل تسلسل السجل
المُشابه، موضع مُقابل موضع، في طريقة تُسمّى "تراصف
تسلسلات" (Sequence Alignment- העמדת רצפים).
(. فيما يلي عدّة حالات نستعمل فيها مُحرك البحث BLASTn
لإيجاد تسلسلات نوكلونيتيدات مُشابهة لتسلسل استعمال:

-BLASTp (protein blast) - مُحرك بحث يجد سجلات
تسلسلات أحماض أمينية مُشابهة لتسلسل أحماض أمينية
مُعطى (تسلسل استعمال).



-BLASTp - أدوات ال-BLAST، Basic Local Alignment Search Tool، تُستعمل للبحث عن تسلسلات مُتشابهة.
BLASTp هو مُحرك بحث يجد في مخازن تسلسلات
أحماض أمينية سجلات لتسلسلات مُشابهة لتسلسل أحماض
أمينية مُعطى. في مُحركات البحث مثل Google أو
Entrez الاستعمال مُكوّن من كلمات مُفتاح، أما في مُحرك
البحث BLASTp الاستعمال مُكوّن من تسلسل أحماض
أمينية. يُقارن مُحرك البحث تسلسل الاستعمال مع تسلسلات
موجودة في مخازن معلومات تسلسلات الأحماض الأمينية
ويعرض سجلات ذات تسلسل مُشابه لتسلسل الاستعمال.
يُعرض تسلسل الاستعمال مُقابل تسلسل السجل المُشابه له ،
موضع مُقابل موضع في طريقة تُسمّى "تراصف تسلسلات"
(Sequence Alignment- העמדת רצפים). فيما يلي
عدة حالات نستعمل فيها مُحرك البحث BLASTp لإيجاد
تسلسلات أحماض أمينية مُشابهة لتسلسل الاستعمال:

- مقارنة تسلسل بروتين لا نعرف مبناه أو فعاليته،
مع مخزن تسلسلات من الأحماض الأمينية بإمكانها
أن تُكشف عن بروتينات بُحثت في الماضي
وتسلسلها مُشابه لتسلسل الاستعمال. تشابه في
التسلسل يُمكنه أن يدلّ على تشابه مُمكن في مبنى
أو في فعالية البروتين.

- مقارنة تسلسل جينٍ فعاليته غير معروفة مُقابل
مخزن تسلسلات نوكلونيتيدات بإمكانها أن تعثر
على جينات بُحثت في الماضي، تسلسلها مُشابه
لتسلسل الاستعمال. هذا التشابه في التسلسل بإمكانه
أن يدلّ على تشابه في المبنى أو في الوظيفة لنتائج
هذا الجين أي في البروتين المُشفّر من هذا الجين.

• مُقارنة تسلسل جين من كائن آخر مع مخزن تسلسلات النوكليوتيدات بإمكانها أن تكشف عن جينات هومولوجية موجودة في كائنات أخرى. مثلاً وجد باحث جين يسبب الصلح لدى الفئران. هل يتواجد جين هومولوجي في الإنسان أيضاً؟ مقارنة الجين من الفأر مع تسلسلات الجينات في الجينوم البشري تُجيب عن هذا السؤال.

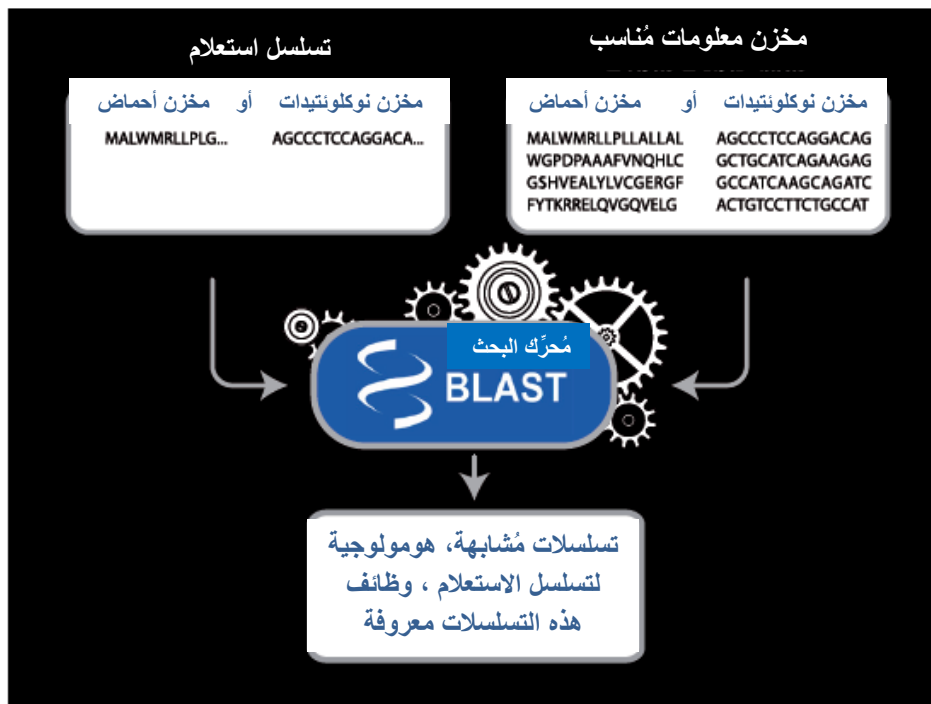
• مُقارنة تسلسل بروتين من كائن آخر مع مخزن أحماض أمينية بإمكانها أن تكشف عن بروتينات هومولوجية في كائنات أخرى. مثلاً اكتشف باحث بروتين ذا فعالية إنزيمية معينة في بكتيريا. هل البروتين الهومولوجي موجود أيضاً في الثدييات؟ مقارنة تسلسل البروتين الموجود في البكتيريا مع تسلسلات البروتينات الموجودة في مخازن المعلومات تُجيب عن هذا السؤال .

بإمكانك الوصول إلى الأدوات بالضغط على الرابط

<http://blast.ncbi.nlm.nih.gov/Blast.cgi>

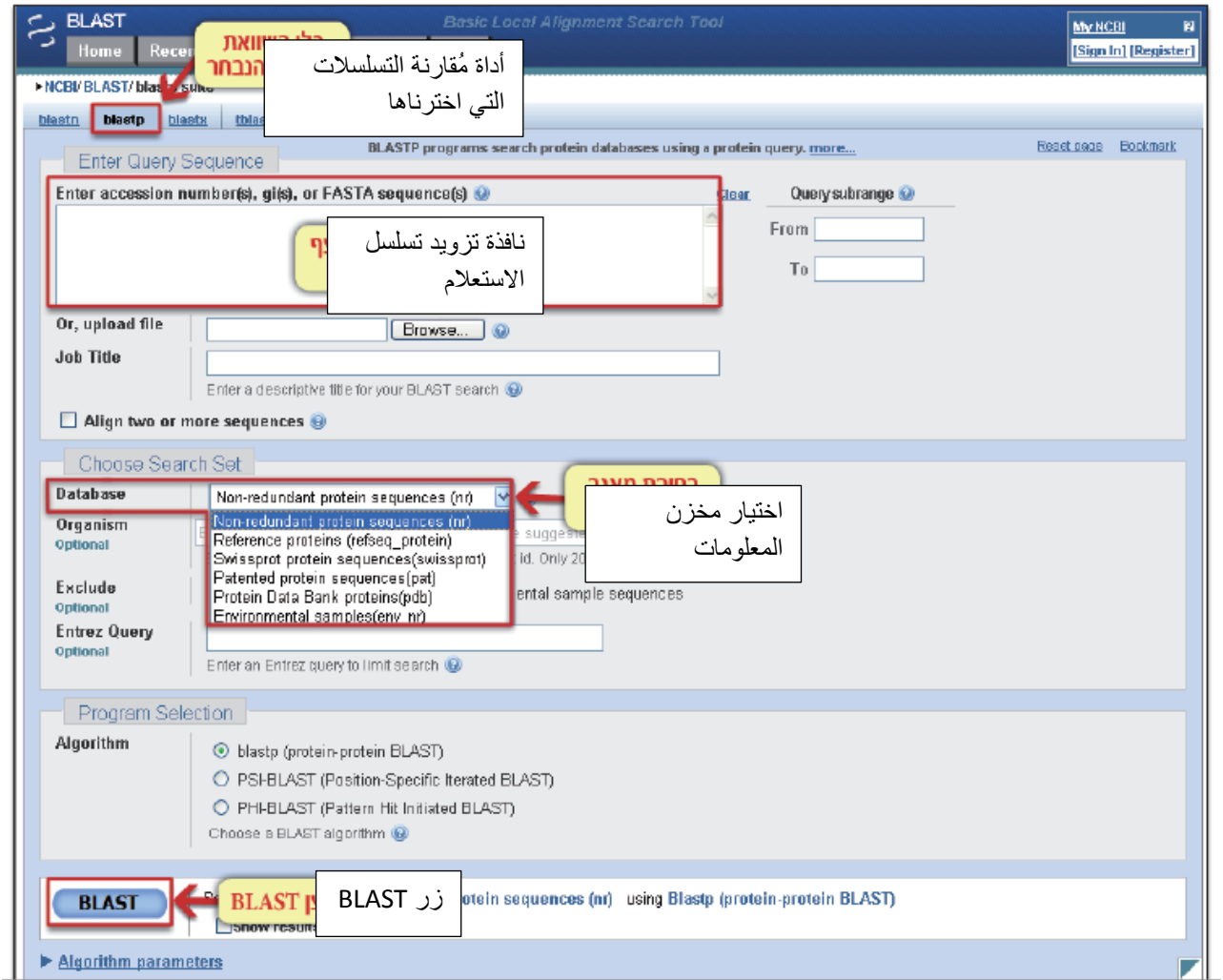


أهلاً بكم في الجولة الإرشادية للأداة **BLAST (Basic Local Alignment Search Tool)**، الأداة التي تستعمل مُحركات البحث عن التسلسلات . عندما نريد فحص إذا كان تسلسل مُعيّن من النوكليوتيدات هو جين مُشفر لبروتين، أو إذا كان لتسلسل أحماض أمينية وظيفة مُعيّنة أو ما هو مصدر التسلسل (من أي كائن عُزل) – نستعين بِمُحرك البحث BLAST لتحقيق هذا الهدف. بِمُساعدة مُحرك البحث BLAST نستطيع أن نجد في مخازن المعلومات تسلسلات مُشابهة، هومولوجية، لتسلسل الاستعلام الذي بحوزتنا. بسبب كون مبنى ووظيفة تسلسلات النتيجة الموجودة في مخزن المعلومات معروفة ومُرتبة في سبجلات، نستطيع أن نتعلّم منها عن مبنى ووظيفة تسلسل الاستعلام.



قائمة أدوات ال BLAST

هذه هي قائمة أدوات BLAST. في البداية نختار مُحرك البحث المُناسب: عندما يكون الاستعلام مكوّن من تسلسل نوكليوتيدات - نختار الأداة nucleotide blast؛ وعندما يكون الاستعلام تسلسلاً من الأحماض الأمينية- نختار protein blast.



صفحة النتيجة: المعطيات التقنيّة والوصف البياني

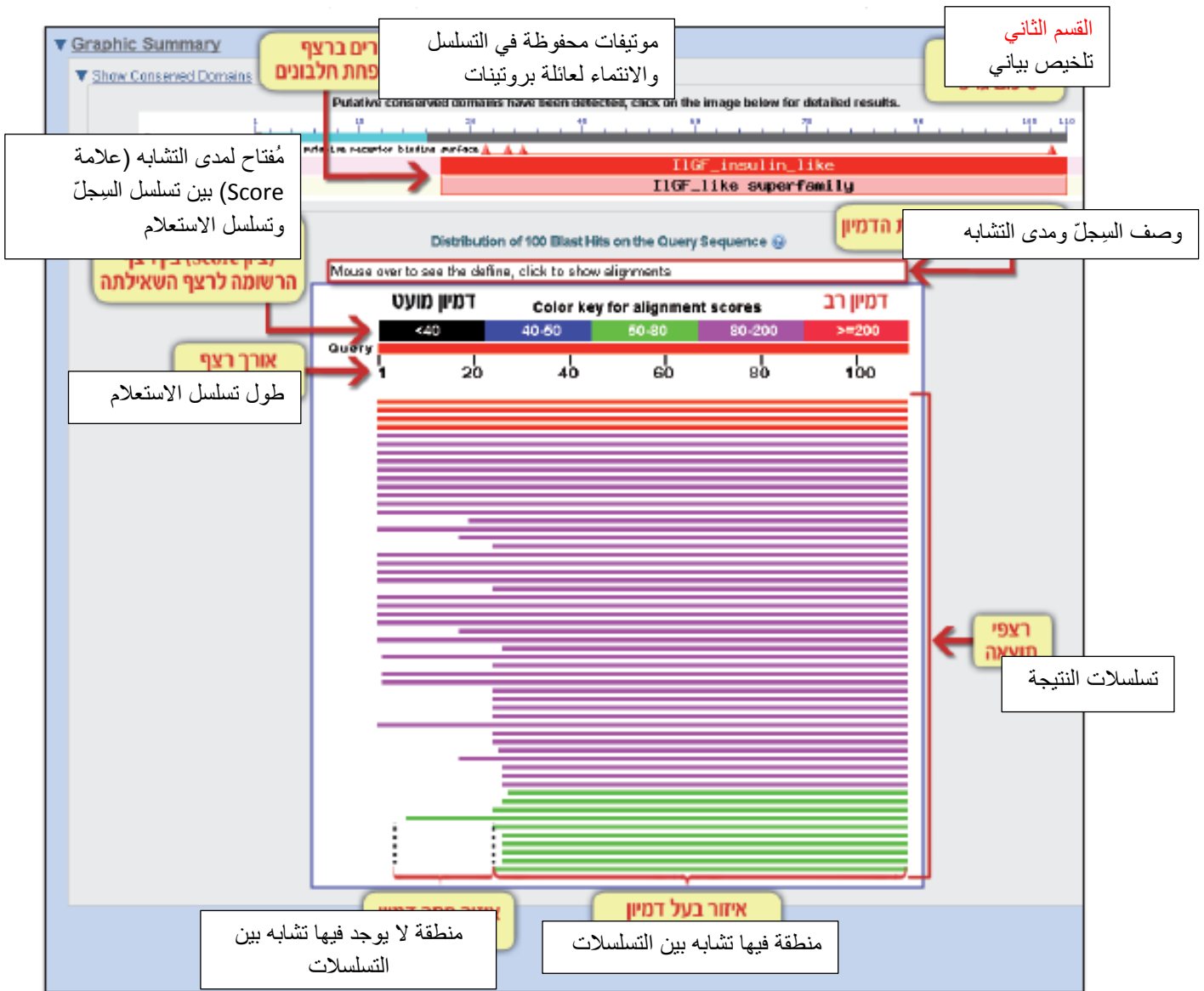
صفحة النتيجة مُقسّمة لعدّة أقسام.

يعرض القسم الأول مُعطيات تقنيّة عن البحث: اسم التسلسل، نوع التسلسل، طول التسلسل، في أي مخزن معطيات يتم البحث وهكذا.



صفحة النتيجة: الوصف البياني

يظهر في القسم الثاني من صفحة النتيجة وصف بياني للنتائج التي حصلنا عليها. السطر الأعلى يعرض تدرج (مقياس) لطول تسلسل الاستعلام، ويتغير هذا التدرج بحسب طول التسلسل. تُعرض تحته التسلسلات الأكثر شبيهاً بتسلسل الاستعلام والموجودة في مخزن المعلومات، بترتيب تنازلي بحسب مدى التشابه مع تسلسل الاستعلام: في البداية تُعرض التسلسلات الأكثر شبيهاً بتسلسل الاستعلام وتليها التسلسلات التي تشبه التسلسل بنسبة أقل. يتم تحديد لون كل خط بحسب مدى التشابه بينه وبين تسلسل الاستعلام، يتم حساب مدى التشابه وعرضه في العلامة Score. في صفحة النتائج الموجودة أمامنا حصلنا في التسلسلات الأربعة الأولى على علامة تشابه أعلى من 200، ولذلك لون التسلسلات الأربعة الأولى أحمر. علامة التشابه للتسلسلات التالية يتراوح بين 80-200، لذلك لونها زهري وهكذا. مكان التسلسلات بحسب مسطرة التدرج التابعة لتسلسل الاستعلام يدل على المنطقة التي يتواجد فيها التشابه بين تسلسل النتيجة وتسلسل الاستعلام. إذا وضعنا سهم الفأرة فوق أحد الخطوط، ستظهر معلومات تُعرّف عن التسلسل في نافذة وصف السجل. هكذا نستطيع أن نعرف أن التسلسل الأكثر شبيهاً بتسلسل الاستعلام هو بروتين الانسولين، وعلامة التشابه هي 223. الضغط على الخط ينقلنا إلى تراصف التسلسلات، وسنشرح عن ذلك لاحقاً.



صفحة النتيجة: وصف السجلات

يظهر في القسم الثالث من صفحة النتيجة وصف للسجلات. يُعرض لكل سجل: كود التعرف (الضغط على رابط كود التعرف يفتح السجل الموجود في مخزن المعلومات)، اسم السجل، علامة التشابه مع تسلسل الاستعلام، وغيرها.

تسلسلات النتائج

القسم الثالث
وصف السجلات

links to other resources: [U](#) UniGene [G](#) GED [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

is producing significant alignments:

| Accession | Description | Max score | Total score | Query coverage | E value |
|--------------------------|---|-----------|-------------|----------------|---------|
| P01308.1 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 223 | 223 | 100% | 5e-43 |
| P01308.2 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 222 | 222 | 100% | 5e-43 |
| P30410.1 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 221 | 221 | 100% | 1e-42 |
| P30406.1 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 219 | 219 | 100% | 3e-42 |
| P30407.1 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 198 | 198 | 100% | 2e-40 |
| P01323.1 | RecName: Full=Insulin-2; Contains: RecName: Full=Insulin-2 B chain; | 172 | 172 | 100% | 6e-41 |
| P01313.2 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 172 | 172 | 100% | 5e-43 |
| P07972.1 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 171 | 171 | 100% | 1e-42 |
| P01322.1 | RecName: Full=Insulin-1; Contains: RecName: Full=Insulin-1 B chain; | 171 | 171 | 100% | 1e-42 |
| P01321.1 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 170 | 170 | 100% | 3e-42 |
| P01325.1 | RecName: Full=Insulin-2; Contains: RecName: Full=Insulin-2 B chain; | 169 | 169 | 100% | 3e-42 |
| Q02907.1 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 165 | 165 | 100% | 6e-41 |
| P06306.2 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 163 | 163 | 100% | 2e-40 |
| Q9LX13.1 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 161 | 161 | 100% | 1e-39 |
| P01311.2 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 154 | 154 | 82% | 1e-37 |
| P01325.1 | RecName: Full=Insulin-1; Contains: RecName: Full=Insulin-1 B chain; | 154 | 154 | 100% | 1e-37 |
| P21563.1 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 142 | 142 | 84% | 5e-34 |
| P01327.2 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 133 | 133 | 70% | 2e-31 |
| P01315.2 | RecName: Full=Insulin; Contains: RecName: Full=Insulin B chain; Cor | 131 | 131 | 100% | 1e-30 |

كود تعرف السجل

اسم ووصف محتوى السجل

معايير لقياس مدى التشابه مع تسلسل الاستعلام

اسم الجين الأكثر شبيهاً بتسلسل الاستعلام وعلامة التشابه

صفحة النتيجة: تراصف التسلسلات

تُعرض في القسم الرابع من صفحة النتيجة مقارنة التسلسلات. يعرض السطر الأول تفاصيل السجل الموجود في مخزن المعلومات، بعدها تُعرض علامة التشابه في الحقل Score، يُشار في الحقل Identities إلى عدد الأحماض الأمينية المتطابقة بين تسلسل الاستعلام وتسلسل النتيجة الموجود في مخزن المعلومات. في المثال الموجود أمامنا يوجد تطابق بنسبة 82%، وهي تُشكل 91 حامض أميني مُتطابق من بين الـ 110 أحماض أمينية التي تكوّن التسلسل في هذه المقارنة. في الحقل Positives يُشار إلى عدد الأحماض الأمينية المُتشابهة، وليس بالضرورة المُتطابقة بين تسلسل الاستعلام وتسلسل النتيجة الموجود في المخزن. الأحماض الأمينية المُتشابهة هي أحماض أمينية ذات صفات كيميائية وفيزيائية مُتشابهة. مثلاً ليسين وأرجنين هي أحماض أمينية مُتشابهة لأنّ كلاهما مشحونتان بشحنة موجبة. في الحقل Gaps يُشار إلى عدد الفراغات التي أُدخلت إلى تراصف التسلسلات الواحد مقابل الآخر، أي مواضع فيها مُقابل حامض أميني مُعيّن في أحد التسلسلين لا يوجد حامض أميني في التسلسل الآخر.

تُعرض بعد ذلك مقارنة التسلسلات: يُعرض في السطر العلوي تسلسل الاستعلام ويُشار إليه بـ Query، أما السطر السفلي فيعرض التسلسل المُماثل الموجود في مخزن المعلومات، ويُشار له بـ Sbjct. تُعرض في السطر الأوسط نتيجة المقارنة بين التسلسلين: إذا تواجد في موضع مُعيّن نفس الحامض الأميني في التسلسلين، يظهر رمز هذا الحامض في السطر الأوسط (سطر المقارنة). إذا كانت الأحماض الأمينية في موضع مُعيّن مُتشابهة ولكنها غير مُتطابقة في التسلسلين، تظهر الإشارة + في سطر المقارنة، أما إذا كانت الأحماض الأمينية مُختلفة الواحد عن الآخر ولا يوجد تشابه في صفات هذه الأحماض، لا تظهر أي إشارة في سطر المقارنة. تُعرض في بداية التسلسلات وفي آخرها أرقام المواضع في كل تسلسل. هكذا مثلاً يُمكن أن نرى أنّ التراصف الموجود أمامنا يعرض تراصف من الحامض الأميني الأول وحتى الموضع 110 في تسلسل الاستعلام (Query) وأيضاً في تسلسل النتيجة الموجود في مخزن المعلومات (Sbjct).

The screenshot shows a sequence alignment tool interface with the following components and annotations:

- Header:** "Alignment" and "Select All" options.
- Query:** "Query 1" with sequence "MALNMEILLPLALLL...".
- Sbjct:** "Sbjct 1" with sequence "LQVQVELGQGFSA...".
- Statistics:** "Score = 223 bits (569)", "Expect = 2e-58", "Method: Compositional matrix adjust.", "Identities = 110/110 (100%)", "Positives = 110/110 (100%)", "Gaps = 0/110 (0%)".
- Annotations:**
 - ربط للسجل في مخزن المعلومات:** Points to the "Query" field.
 - رقم الموضع في تسلسل الاستعلام:** Points to the "Query 1" label.
 - رقم الموضع في تسلسل النتيجة:** Points to the "Sbjct 1" label.
 - علامة التشابه مع تسلسل الاستعلام:** Points to the "MALNMEILLPLALLL" sequence.
 - عدد ونسبة الفراغات التي أُدخلت في تراصف التسلسلات:** Points to the "Gaps = 0/110 (0%)" statistic.
 - عدد ونسبة الأحماض الأمينية المتطابقة أو المتشابهة بين تسلسل الاستعلام وتسلسل النتيجة:** Points to the "Identities = 110/110 (100%) Positives = 110/110 (100%)" statistics.
 - نتائج المقارنة:** Points to the alignment lines.
 - أحماض أمينية متطابقة:** Points to the aligned residues.
 - أحماض أمينية مختلفة:** Points to the '+' symbols.
 - أحماض أمينية مُتَشابهة:** Points to the aligned residues.
 - نقص في أحماض أمينية:** Points to the '-' symbols.
 - عدد ونسبة الأحماض الأمينية المتطابقة بين تسلسل الاستعلام وتسلسل النتيجة:** Points to the "Identities" statistic.

تلخيص

تعلمنا كيف بإمكاننا أن نبحث عن تسلسلات هومولوجية لتسلسل استعمال بواسطة الأداة BLAST

إذا كان لدينا تسلسل نوكلوتيدات – نستعمل الأداة nucleotide blast؛ أمّا إذا كان لدينا تسلسل من الأحماض الأمينية – نستعمل الأداة protein blast. عملية البحث، كذلك صفحة النتيجة مُتشابهتان جدًّا في الأدوات.

أهمية مقارنة التسلسلات في البحث البيوإنفورماتي كبيرة جدًّا، و تُسيطر استعمالاتها على مجالات عديدة مثل إيجاد بروتينات مُتشابهة تنتمي لنفس العائلة البروتينية، البحث في التطور والارتقاء (مחקر ابولوزيوني) الذي يُقارن بين تسلسلات بروتينات أو جينات من كائنات مُختلفة وغيرها. هذا فقط قسم صغير من الأمثلة العديدة التي يؤدي فيها البحث ومُقارنة التسلسلات بواسطة الأداة BLAST إلى تطوير البحث البيولوجي.

انتم مدعوون لاختبار الأداة، نتمنى أن تستمتعوا بوقتكم. بالنجاح!