Entrez- مُحرِّك بحث يبحث في قواعد البيانات في الموقع NCBI (**N**ational **C**enter for **B**iotechnology **I**nformation)، "المركز القومي للمعلومات البيوتكنولوجيّة" بالعبرية "המרכז הלאומי למידע ביוטכנולוגי". بواسطة استعلام بحث مُكوّن من كلمات بحث أو كلمات مُفتاح، تقوم الأداة بإيجاد السِجلّات الموجودة في قواعد البيانات والتي تتلاءم مع شروط البحث.

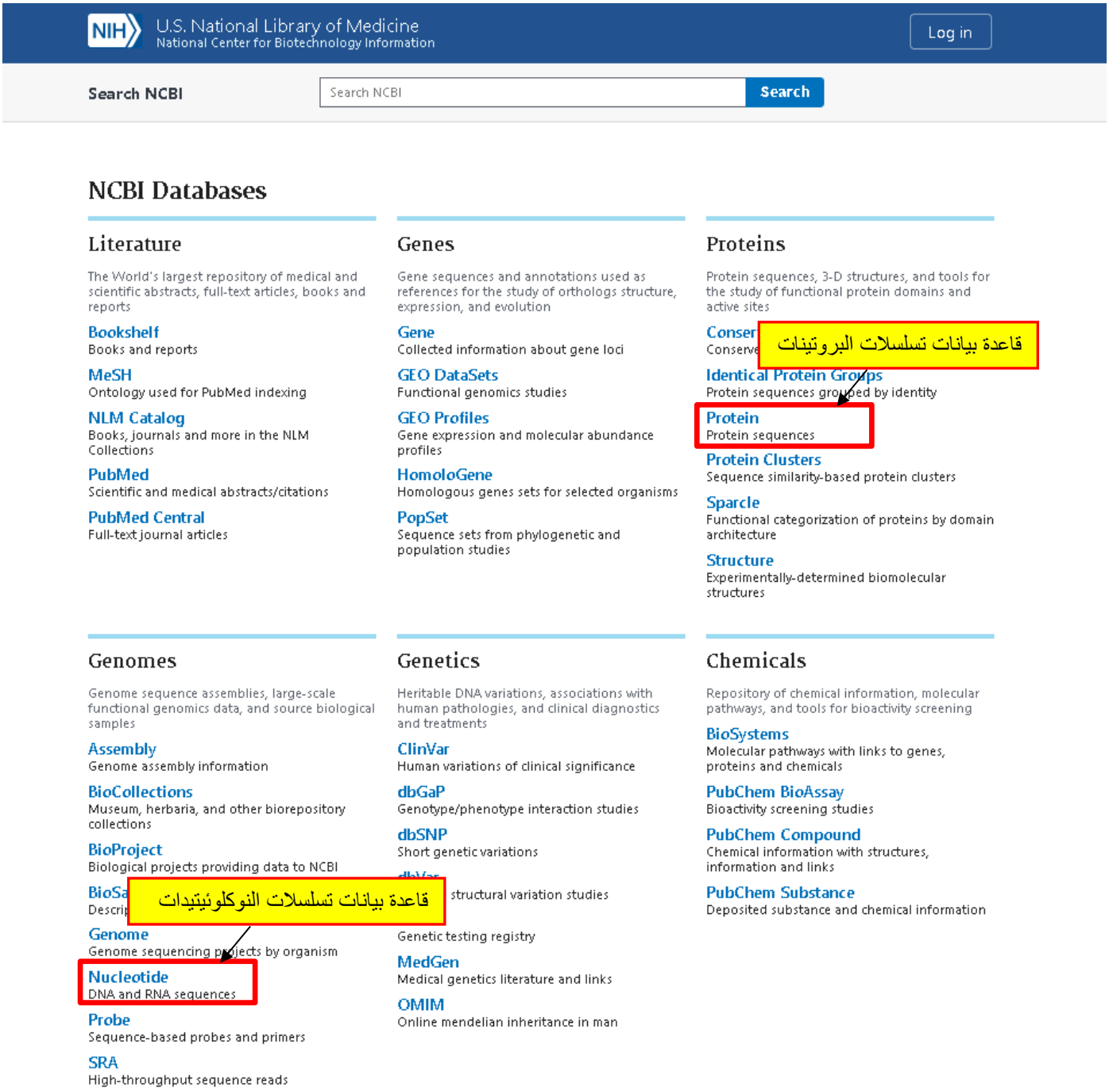
Entrez- مُحرِّك بحث يُمكِّن من إيجاد معلومات (פרטי מידע) موجودة على شكل سِجلّات (רשומות) في قواعد بيانات بمجال الصحة، البيولوجيا والبيوتكنولوجيا عن طريق استعمال كلمات بحث، التي تُسمّى أيضًا كلمات مُفتاح. يضمّ البحث الذكي أيضًا - بالإضافة إلى كلمات البحث – قيود مُناسبة، مثلًا التقييد لقاعدة البيانات المُناسبة لنوع الجُزيء المطلوب، تقييد لحقل البحث وغيرها. مصدر الاسم Entrez من اللغة الفرنسيّة، ومعناه -"زوّد" بالعبرية "הזן". يزوّد المُستخدِم مُحرِّك البحث بكلمات مُفتاح، وعندها يقوم مُحرِّك البحث بالبحث عن سِجلّات تحتوي على كلمات المُفتاح في كل قواعد البيانات التابعة له. يجب الحرص على اختيار كلمات مُفتاح مُناسبة وتقييد البحث بذكاء للحصول على نتائج صحيحة ومحدّدة.

من ضمن قواعد البيانات التي يتمّ البحث فيها يُمكن أنْ نجد قواعد بيانات لتسلسلات نوكلوئتيدات، قواعد بيانات تسلسلات أو مباني بروتينات، خرائط جينوميّة، قواعد بيانات المقالات العلميّة وغيرها. تحتوي قواعد البيانات على عدد كبير من السِجلّات، لذلك استخدام كلمات مُفتاح وقيود مُناسبة للبحث هو مرحلة ضروريّة لإيجاد المعلومات المطلوبة. مرحلة البحث وإيجاد السِجلّات المُناسبة في قواعد البيانات بواسطة محرّك البحث هي في الغالب المرحلة الأوليّة والضروريّة في أبحاث عديدة.

Entrez- يُمكن الوصول إلى الأداة في العنوان: https://www.ncbi.nlm.nih.gov/search/

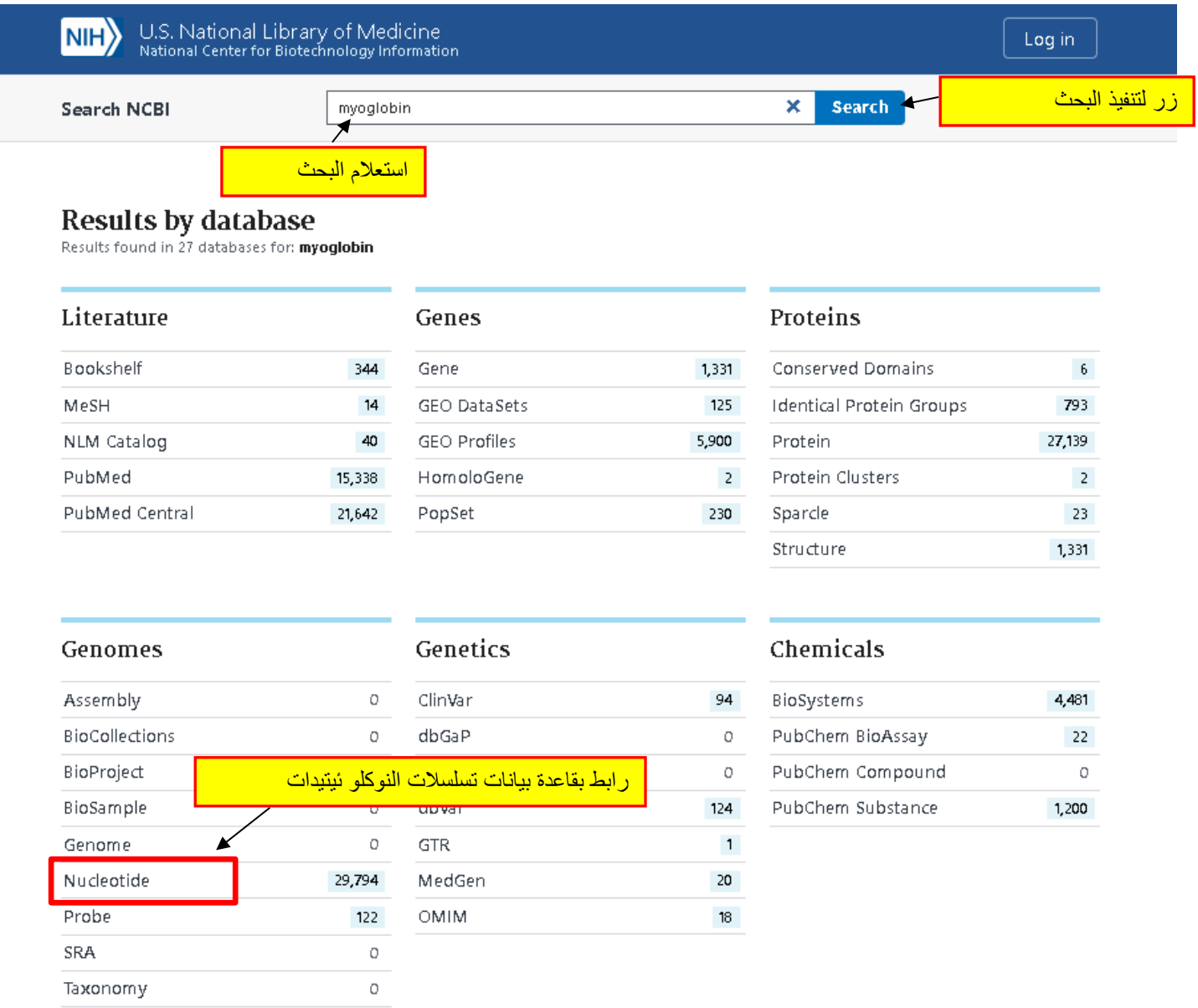
أهلًا بكم في الجولة الإرشاديّة لمُحرِّك البحث Entrez في الموقع NCBI. Entrez هو مُحرِّك بحث يُمكِّن من البحث عن تفاصيل معلومات، تُسمّى باسم سِجلّات، في قواعد بيانات بمجال الصحة، البيولوجيا والبيوتكنولوجيا. يتمّ البحث الذكيّ بواسطة استعمال كلمات مُفتاح ، بشكل مُشابه لمُحرِّك البحث Google، وتحديد قواعد البيانات المُناسبة للبحث.

يُمكن الوصول إلى مُحرِّك البحث من خلال صفحة البحث التابعة لموقع NCBI (الرابط مُرفق في الصفحة السابقة). عنما نرغب بالبحث عن معلومات تخص بحث مُعيّن، علينا في البداية اختيار كلمات المُفتاح المُناسبة لبحثنا، كذلك يمكننا تحديد قاعدة البيانات المناسبة لنوع المعلومات التي نبحث عنها. تُعرض في صفحة البحث التابعة للأداة قائمة قواعد البيانات التي يُمكننا البحث فيها والتي يقوم بصيانتها موقع NCBI. مثلًا، قاعدة البيانات Protein تحتوي على تسلسلات بروتينات، قاعدة البيانات Nucleotide تحتوي على تسلسلات نوكلوئتيدات كجُزيئات RNA رسول (mRNA)، جينات وغيرها.



البحث في قواعد البيانات

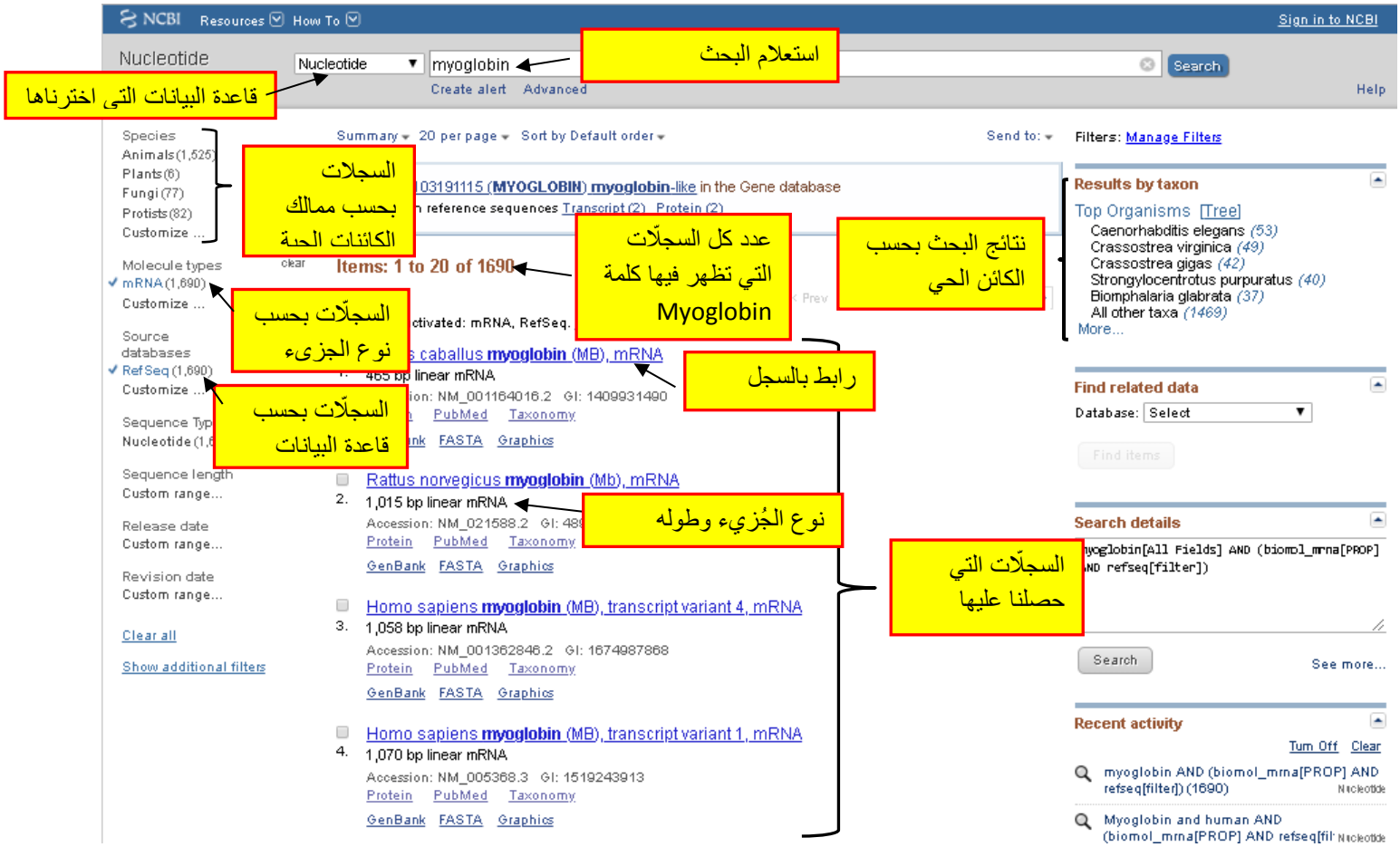
نرغب في المثال الموجود أمامنا بالبحث عن تسلسل الجين المُشفِّر للميوغلوبين، ناقل الأكسجين عند الانسان. علينا كتابة استعلام بحث مكوّن من كلمات مفتاح داخل نافذة البحث. اختيار كلمات المُفتاح الصحيحة هو أمر ضروريّ من أجل الحصول على نتائج دقيقة. في المرحلة الأولى نكتب كلمة Myoglobin في نافذة البحث ونضغط على Search.



نتائج البحث

المعلومات الموجودة في قواعد البيانات مُرتبة على شكل سجلّات. نتيجة البحث الذي قُمنا بتنفيذه تضم عدد السجلّات الموجودة في قواعد البيانات المُختلفة والتي تحتوي على استعلام البحث، وتضم روابط لفتح هذه السجلّات. مثلًا، في قاعدة البيانات Nucleotide، نجد سجلّات لتسلسلات نوكلوئتيدات، جينات، جُزيئات RNA وغيرها. يتواجد كل تسلسل في سجلّ مُنفصل.

في هذا المثال، يُمكن أن نرى في صورة الشاشة المعروضة، عدد السجلّات التي تحوي داخلها كلمة Myoglobin في مكان ما داخل السجلّ. مثلًا عدد السجلّات الموجودة في قاعدة البيانات Nucleotide هو 29,794 سجلّ. الضغط على قاعدة البيانات المطلوبة ينقلنا إلى صفحة نتيجة تحوي روابط تصلنا بهذه السجلّات. بما أننا نبحث عن تسلسل mRNA، نضغط على قاعدة بيانات النوكلوئيتيدات.



صفحة النتيجة تضمّ قائمة سجلّات. لكل سجلّ يظهر وصف قصير للسجلّ وكود تعرّف خاص به. الروابط الموجودة على يسار الصفحة مُقسمة إلى مجموعات مُختلفة. مثلًا إذا ضغطنا على الرابط mRNA نحصل على جميع السجلّات التي تصف جُزيئات RNA رسول والتي تظهر فيها كلمة ميوغلوبين. يُمكن أيضًا التمعّن في السجلّات الموجودة في ممالك مُختلفة في عالم الأحياء، مثلًا مملكة الـBacteria، أو في قواعد بيانات مُختلفة وغيرها. من الجهة اليُمنى لصفحة النتيجة تظهر روابط لسجلات تابعة لكائنات مُختلفة توجد لديها سجلات كثيرة تضمّ كلمة البحث المطلوبة.

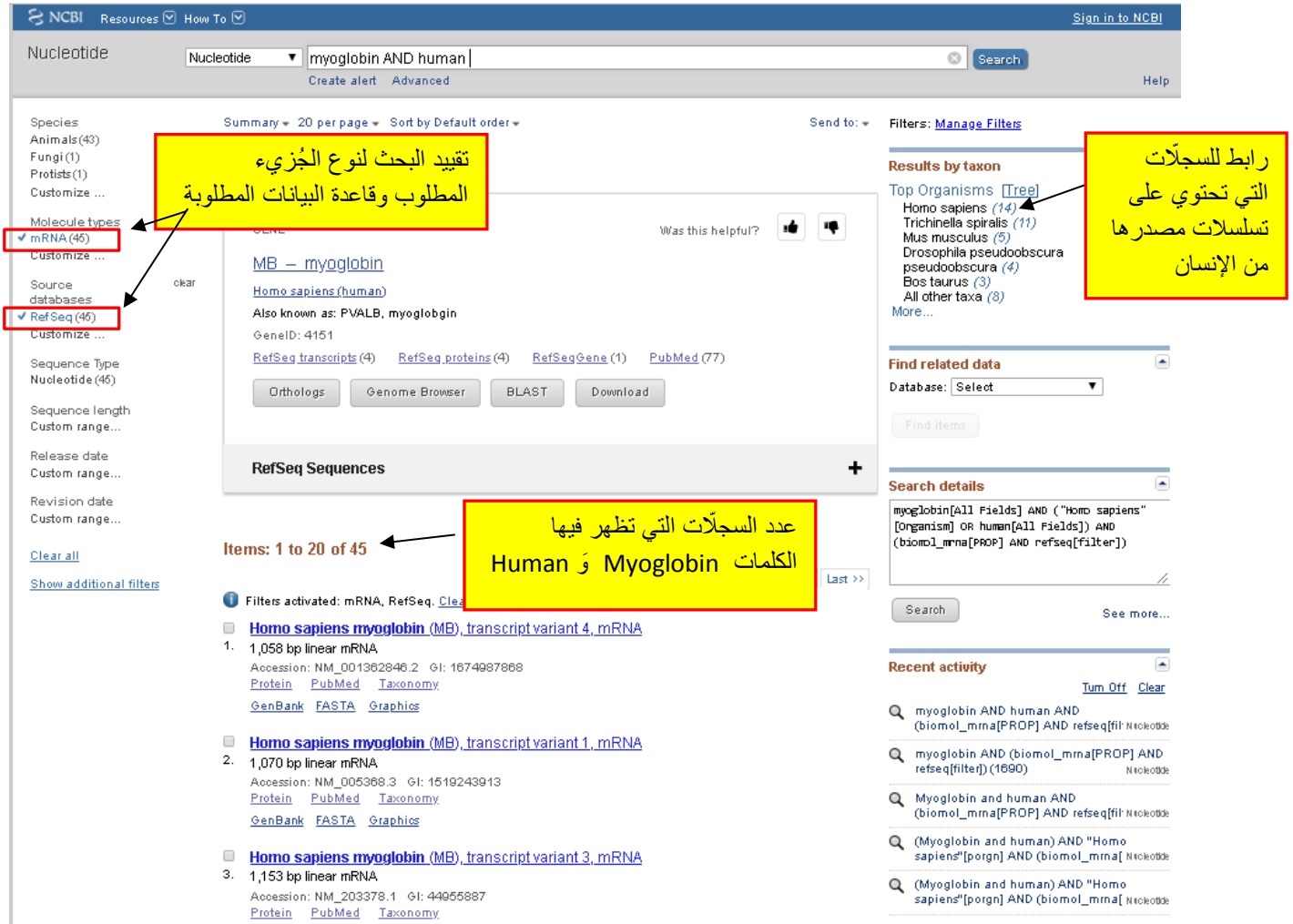
البحث الذكي- دمج عدَّة كلمات بحث

عدد النتائج التي حصلنا عليها كبير جدًّا، ولا تصف كل السِجلّات التي حصلنا عليها جين الميوغلوبين الموجود لدى الإنسان. الآن نحدّد البحث عن السجلات التابعة للكائن الحي الصحيح، في هذا المثال، الإنسان. لهذا الهدف نُضيف لاستعلام البحث كلمة Human، علينا إضافة كلمة AND بين كلمتي المُفتاح.

فيما يلي قواعد أساسيّة للبحث السليم بواسطة كلمات الربط التي تُكتب بأحرف كبيرة:

شروط البحث	الأمر
يحتوي السِجلّ بالتأكيد على كلمتي البحث: كلمة أ' **و** كلمة ب'	كلمة أ' **AND** كلمة ب'
يحتوي السِجلّ على الأقل على واحدة من كلمتي البحث، لكن ليس بالضرورة كلتاهما: كلمة أ' **أو** كلمة ب'	كلمة أ' **OR** كلمة ب'

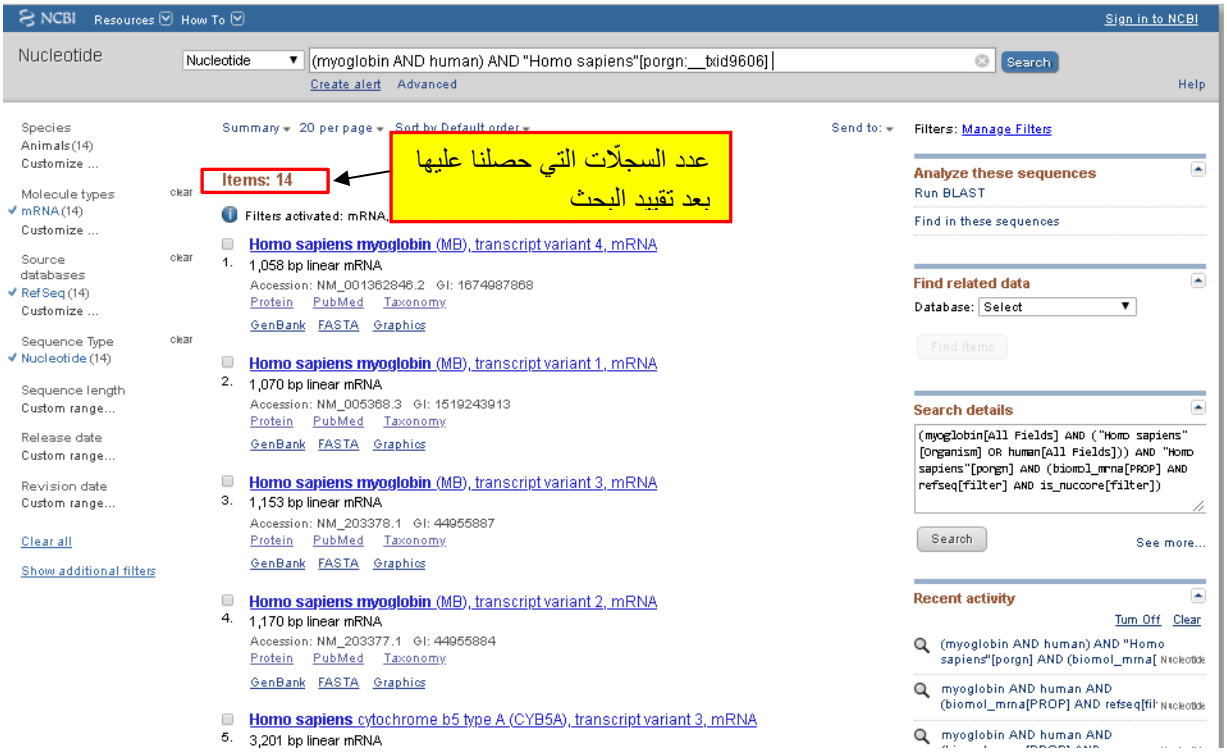
كما يُمكن أن نلاحظ في الصورة التالية فإن القيود التي قُمنا بوضعها أدت إلى تقليص كبير بعدد السجلّات، جميع السجلات التي حصلنا عليها تحتوي على كلمتي المفتاح. من أجل اختيار السجلّات التي تحتوي على تسلسلات تابعة للإنسان Homo sapiens، نختار هذه السجلّات من خلال الضغط على الرابط المناسب في الجهة اليُمنى للشاشة.



البحث الذكي – قيود إضافيّة البحث

بإمكاننا القيام بالبحث بشكل أدق عن طريق استعمال الإمكانيّات الموجودة في الجهة اليُسرى لصفحة النتيجة. بالإمكان تقييد البحث لنوع الجُزيء، لمكان الجُزيء في الخلية، لنوع قاعدة البيانات وغيرها.

في هذه المثال سنبحث عن جُزيئات RNA رسول فقط. نُلاحظ في الحقل Molecule types أنَّ أداة البحث اختارت بشكل تلقائي الإمكانيّة mRNA، يُمكن تغيير ذلك حسب متطلبات البحث. إضافة إلى ذلك يمكن تقييد نوع قاعدة البيانات التي يتم البحث فيها. تختلف قواعد البيانات المُختلفة فيما بينها بصفات عديدة. قاعدة البيانات RefSeq هي قاعدة موثوق بها للبحث عن سِجلّات نوكلوئتيدات، لذلك ستختار الأداة في حقل Source Database الرابط الذي يُظهر السجلّات الموجودة في قاعدة البيانات RefSeq. في حال عدم اختيار الأداة القيود الصحيحة والمطلوبة بشكل تقائي يمكن تغييرها أو اختيار القيود المناسبة لبحثنا.



بعد تقييد البحث عن السجلّات بحسب الكائن الحي المطلوب، نوع الجزيء وقاعدة البيانات، تقلص عدد السجلّات التي تُلائم شروط البحث بشكل كبير جدًا. الآن يُمكننا بسهولة الدخول إلى كل سجلّ على حدة ودراسة المعلومات الموجودة فيه بدقة.

انتبهوا، إذا عرفنا كود تعرُّف السِجلّ بإمكاننا إدخاله مُباشرة إلى نافذة البحث.

تعلَّمنا في هذا القسم، كيف يتمّ تنفيذ بحث ناجع في قاعدة بيانات مُعيّنة باستعمال كلمات بحث مُناسبة. لاحظنا أنَّ البحث البسيط يُعطي عادة عددًا كبيرًا من النتائج، لذلك يجب علينا تقييد البحث باستخدام قيود مُختلفة.

تحليل السِجلّ

نختار سِجلًّا مُعيّنًا ونتعرّف على طريقة استخلاص المعلومات اللازمة منه. الضغط على رابط السِجلّ يؤدّي إلى فتحه. السِجلّات ذات مبنى ثابت يضُمّ أسماء الحقول في الجهة اليُسرى للصفحة وتفاصيل المعلومات في الجهة اليُمنى للصفحة. الحقول الموجودة في قاعدة بيانات معيّنة ثابتة في كلّ السِجلّات، وتتغيّر فقط تفاصيل المعلومات المُرافقة لها من سِجلّ لآخر.

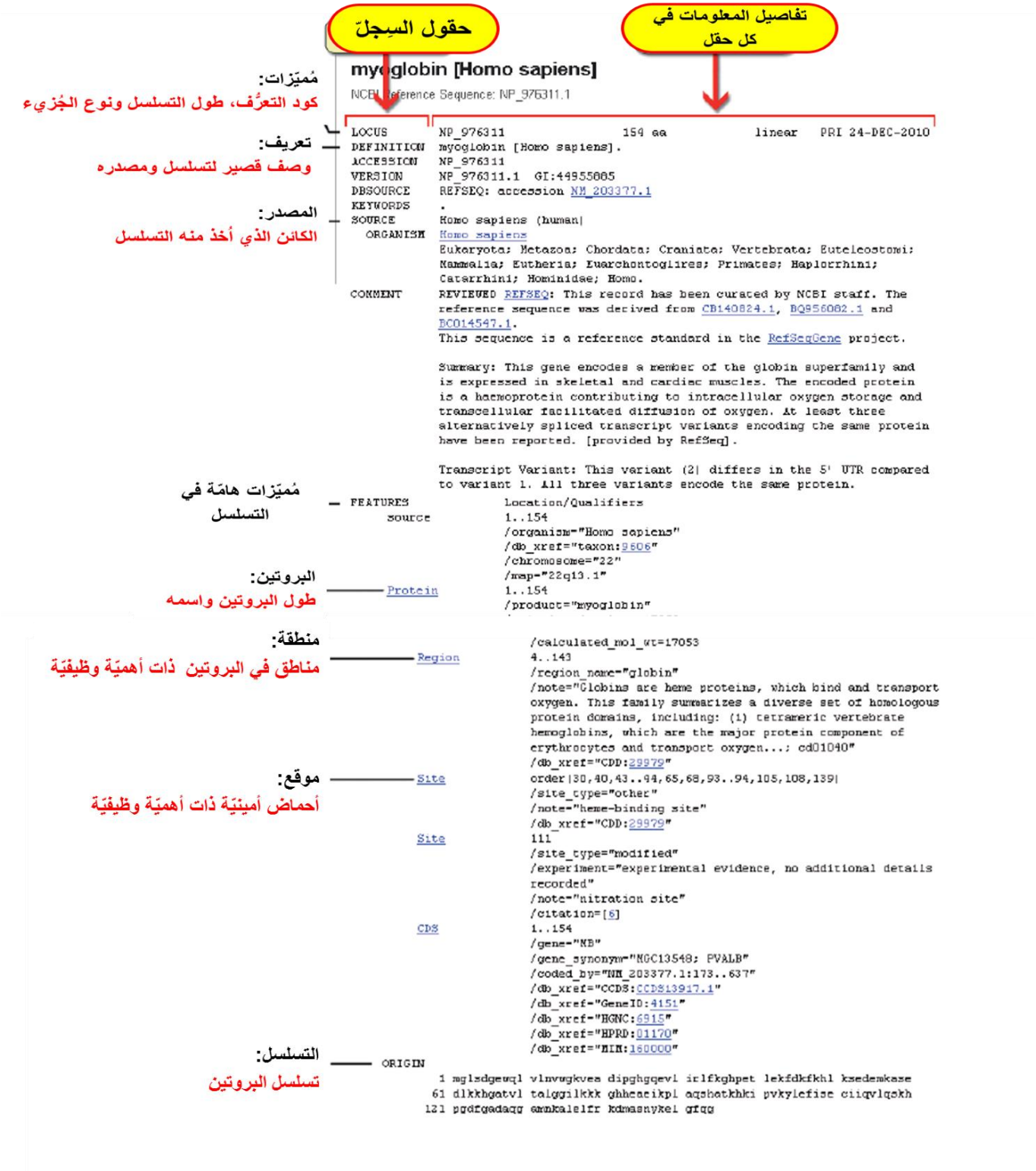
فيما يلي عدد من الحقول الهامة في قاعدة بيانات تسلسلات النوكلوئتيدات:

اسم الحقل	تفصيل المعلومات
LOCUS	يُشير إلى كود التعرُّف، طول التسلسل، ونوع الجُزيء
DEFINITION	يصف باختصار التسلسل ومصدره
ORGANISM	يُشير إلى الكائن الذي أخذ منه التسلسل
FEATURES • الحقل الفرعي GENE • الحقل الفرعي EXON • الحقل الفرعي CoDing Sequence (CDS)	يصف مُميّزات هامة في التسلسل. • يُشير إلى حُدود الجين بما في ذلك حُدود الاكسونات والإنترونات • يُشير إلى حُدود كل اكسون • يُشير إلى حدود التسلسل المُشفِّر فقط
ORIGIN	يعرض تسلسل نوكلوئتيدات ال RNA رسول.

مبنى السِجلّ والربط بين قواعد البيانات

قواعد البيانات المُختلفة مُترابطة بشكل كبير فيما بينها، مثلًا من السِجلّ الذي يصف جُزيء RNA رسول للميوغلوبين، يُمكن إيجاد سِجلّات مُناسبة من قواعد بيانات أخرى مثلًا تسلسل البروتين ميوغلوبين من قاعدة بيانات تسلسلات البروتينات. هُنا أيضًا يتمّ ترتيب المعلومات بحقول ثابتة تحتوي على معلومات وروابط مُناسبة.

مثلًا في الحقل FEATURES يتواجد الحقل الفرعي Protein الذي يصف طول البروتين واسمه؛ في الحقل الفرعي Region تتواجد مناطق في البروتين ذات أهميّة وظيفيّة له، المقطع المُمتّد من الحامض الأميني 4 وحتى الحامض الأميني 143 ذا أهمية لوظيفة البروتين، يُسمّى هذا المقطع غلوبين. الحقل الفرعي Site يُشير إلى أحماض أمينيّة ذات أهميّة وظيفيّة للبروتين، مثل ربط مجموعة الهيم (קבוצת ההם). في الحقل ORIGIN يُمكن أنْ نرى تسلسل الأحماض الأمينيّة للبروتين.



تلخيص

تعرّفنا في هذه الجولة الإرشاديّة على البحث في قواعد البيانات بمُساعدة مُحرِّك البحث Entrez. تعلَّمنا عن مبنى السِجلّات، طريقة تحليل السِجلات واستخلاص المعلومات البيولوجيّة العمليّة.

أنتم مدعوون لاختبار الأداة والتمرُّن على استعمالها، نتمنى أنْ تستمتعوا بوقتكم. بالنجاح!