DNAMAN
Sequence Analysis Software

Sequence Editing

BIOINFORMATICS PLATFORM

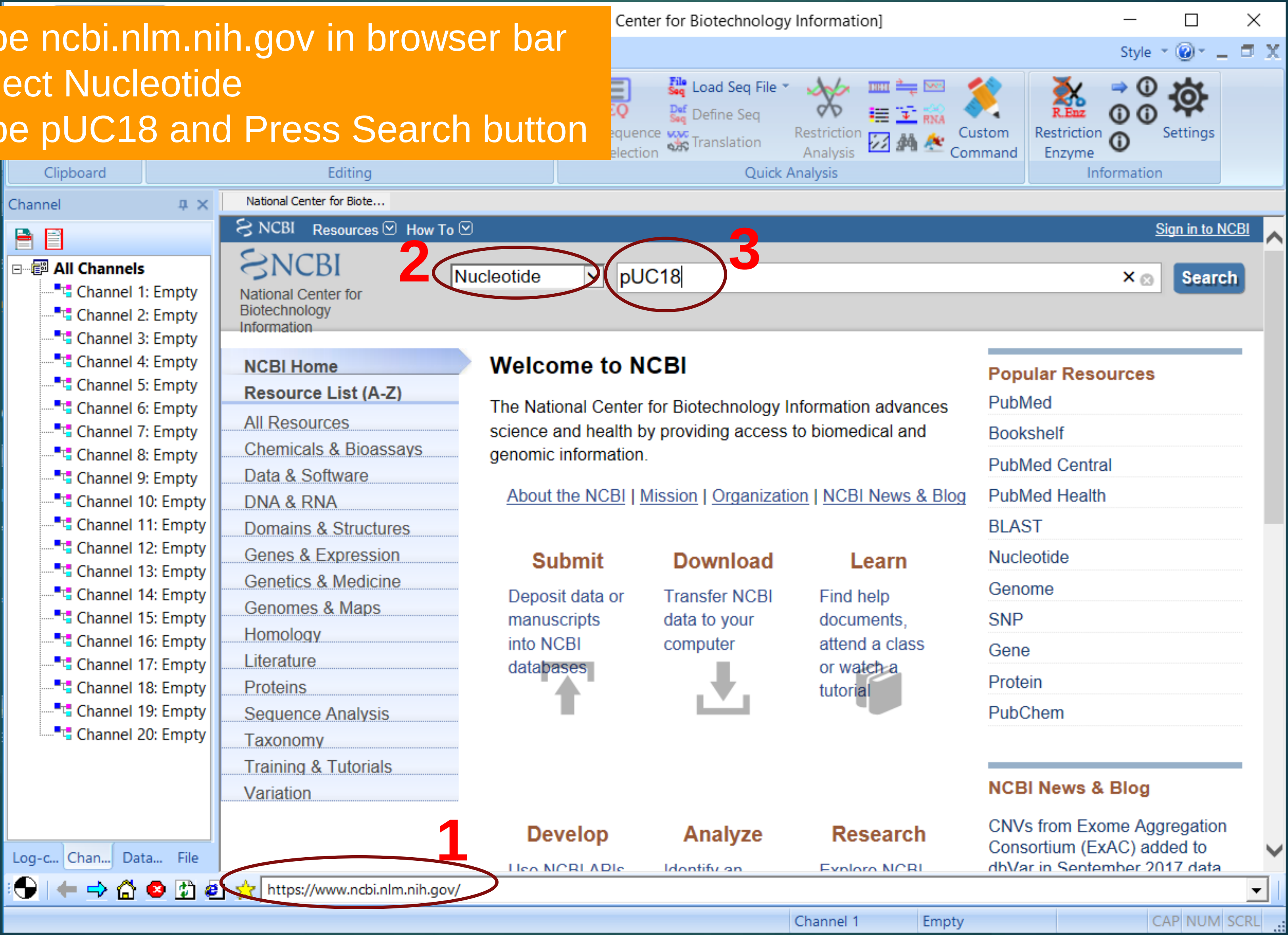
Sequence Editing Features

- Retrieve Sequence From the Internet
- DNAMAN Sequence Format
- Enter Custom Sequence
- Load Existing Sequence
- Analysis Definition
- Work with other Sequence Formats
- Show/Convert Sequence
- Generate Random Sequences

Retrieve Sequence

From the Internet

1. Type ncbi.nlm.nih.gov in browser bar
2. Select Nucleotide
3. Type pUC18 and Press Search button



Center for Biotechnology Information]

National Center for Biote...

Nucleotide

pUC18

https://www.ncbi.nlm.nih.gov/

Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

About the NCBI | Mission | Organization | NCBI News & Blog

Submit

Deposit data or manuscripts into NCBI databases

Download

Transfer NCBI data to your computer

Learn

Find help documents, attend a class or watch a tutorial

Develop

Use NCBI APIs

Analyze

Identify an

Research

Explore NCBI

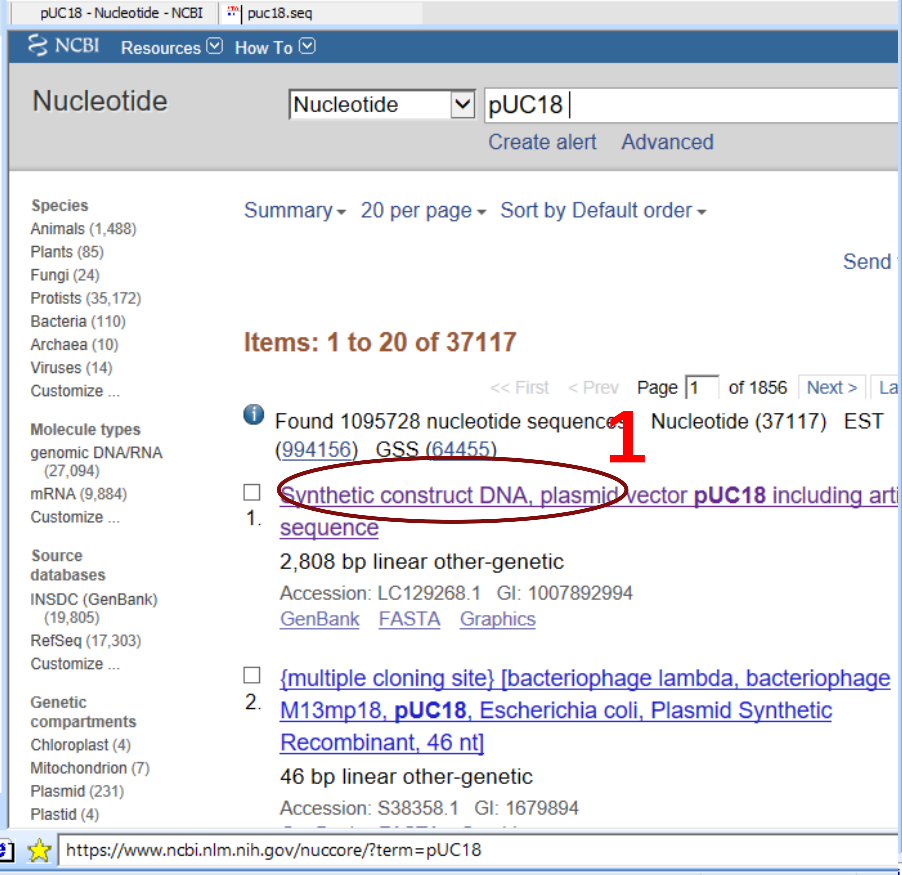
NCBI News & Blog

CNVs from Exome Aggregation Consortium (ExAC) added to dbVar in September 2017 data

Retrieve Sequence

From the Internet

1. Click pUC18 link to open sequence
2. Click Send to
3. Choose Destination to File
4. Choose GenBank format and click Create File button



Search: pUC18

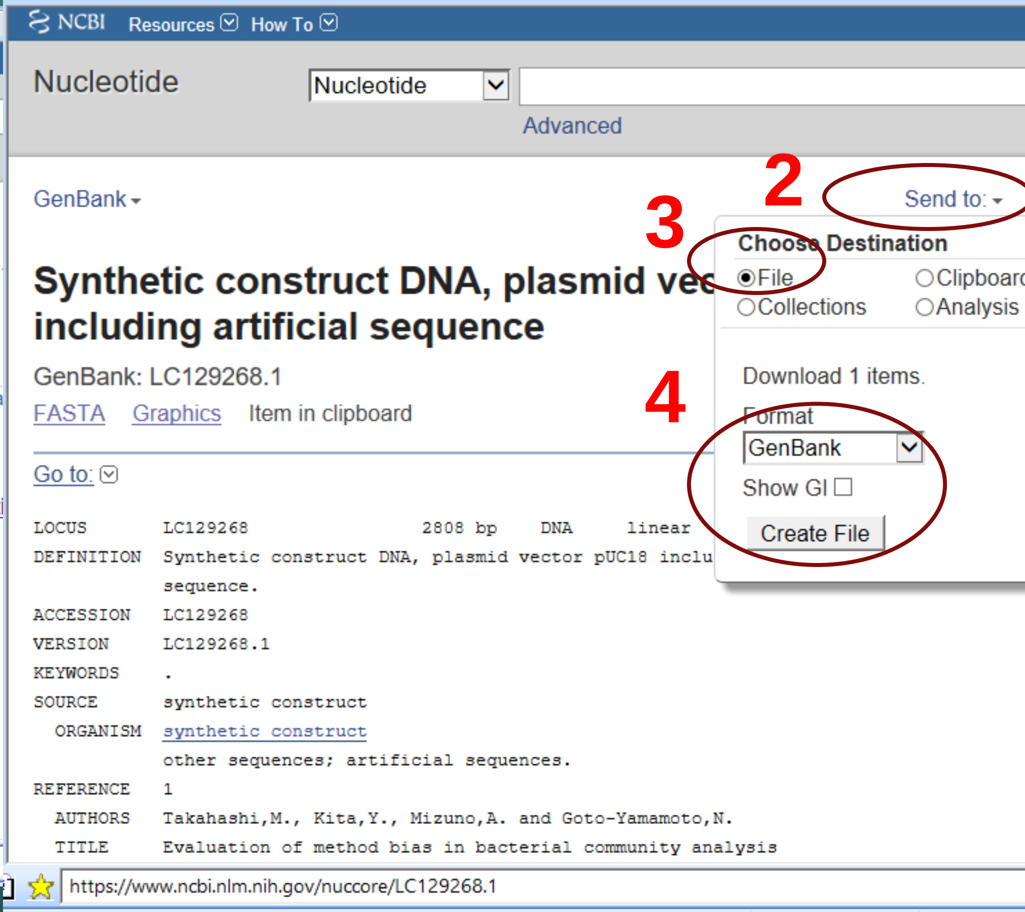
Items: 1 to 20 of 37117

Found 1095728 nucleotide sequences. Nucleotide (37117) EST (994156) GSS (64455)

☐ 1. Synthetic construct DNA, plasmid vector **pUC18** including arti sequence
2,808 bp linear other-genetic
Accession: LC129268.1 GI: 1007892994
GenBank FASTA Graphics

☐ 2. {multiple cloning site} [bacteriophage lambda, bacteriophage M13mp18, **pUC18**, Escherichia coli, Plasmid Synthetic Recombinant, 46 nt]
46 bp linear other-genetic
Accession: S38358.1 GI: 1679894

https://www.ncbi.nlm.nih.gov/nuccore/?term=pUC18



Synthetic construct DNA, plasmid vec including artificial sequence

GenBank: LC129268.1

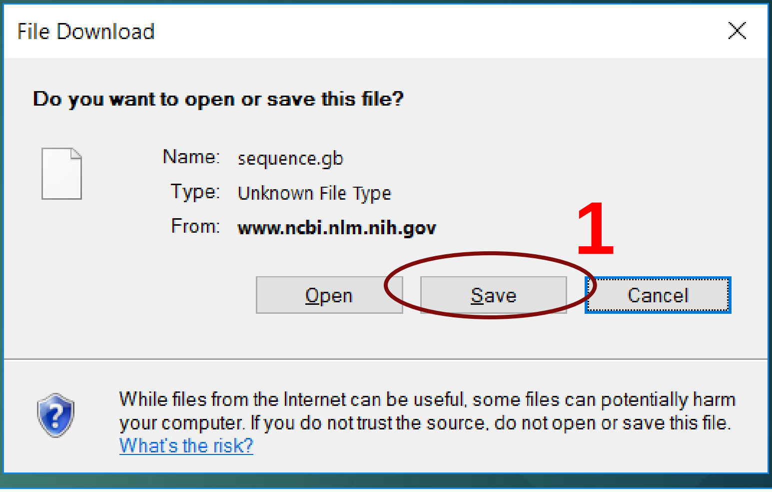
```
LOCUS       LC129268                2808 bp    DNA     linear
DEFINITION  Synthetic construct DNA, plasmid vector pUC18 inclu
            sequence.
ACCESSION   LC129268
VERSION     LC129268.1
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Takahashi,M., Kita,Y., Mizuno,A. and Goto-Yamamoto,N.
  TITLE     Evaluation of method bias in bacterial community analysis
```

https://www.ncbi.nlm.nih.gov/nuccore/LC129268.1

Retrieve Sequence

From the Internet

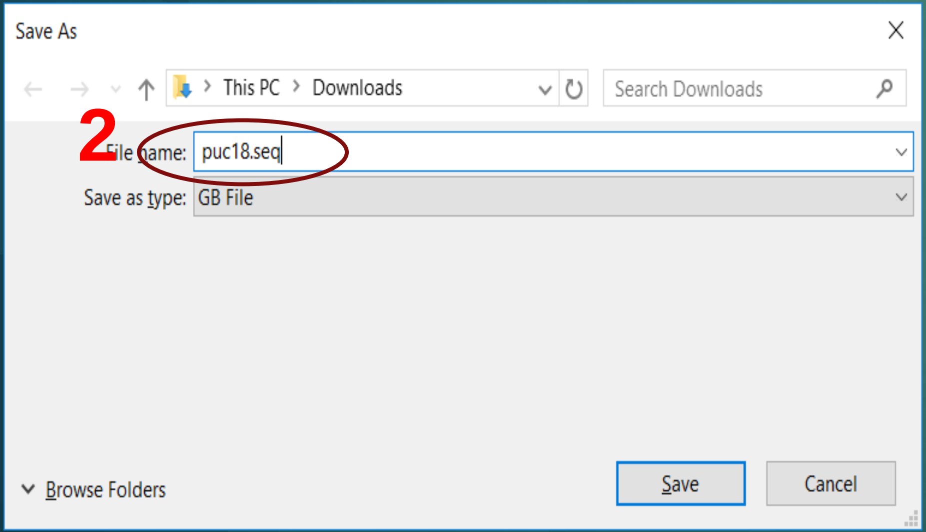
1. Click Save button to download file
2. Enter file name to save
3. Click Open button after file saved



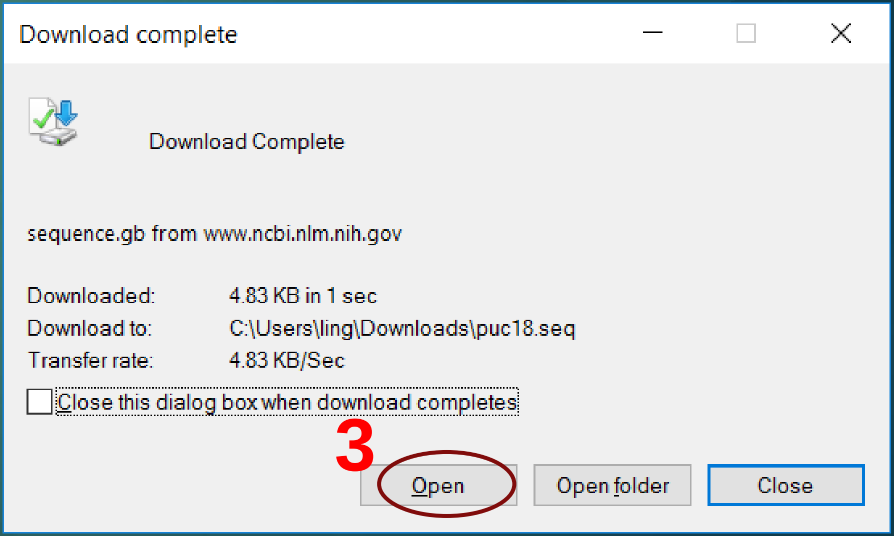
Do you want to open or save this file?

Name: sequence.gb
Type: Unknown File Type
From: **www.ncbi.nlm.nih.gov**

While files from the Internet can be useful, some files can potentially harm your computer. If you do not trust the source, do not open or save this file. What's the risk?



File name: puc18.seq
Save as type: GB File



Download Complete

sequence.gb from www.ncbi.nlm.nih.gov

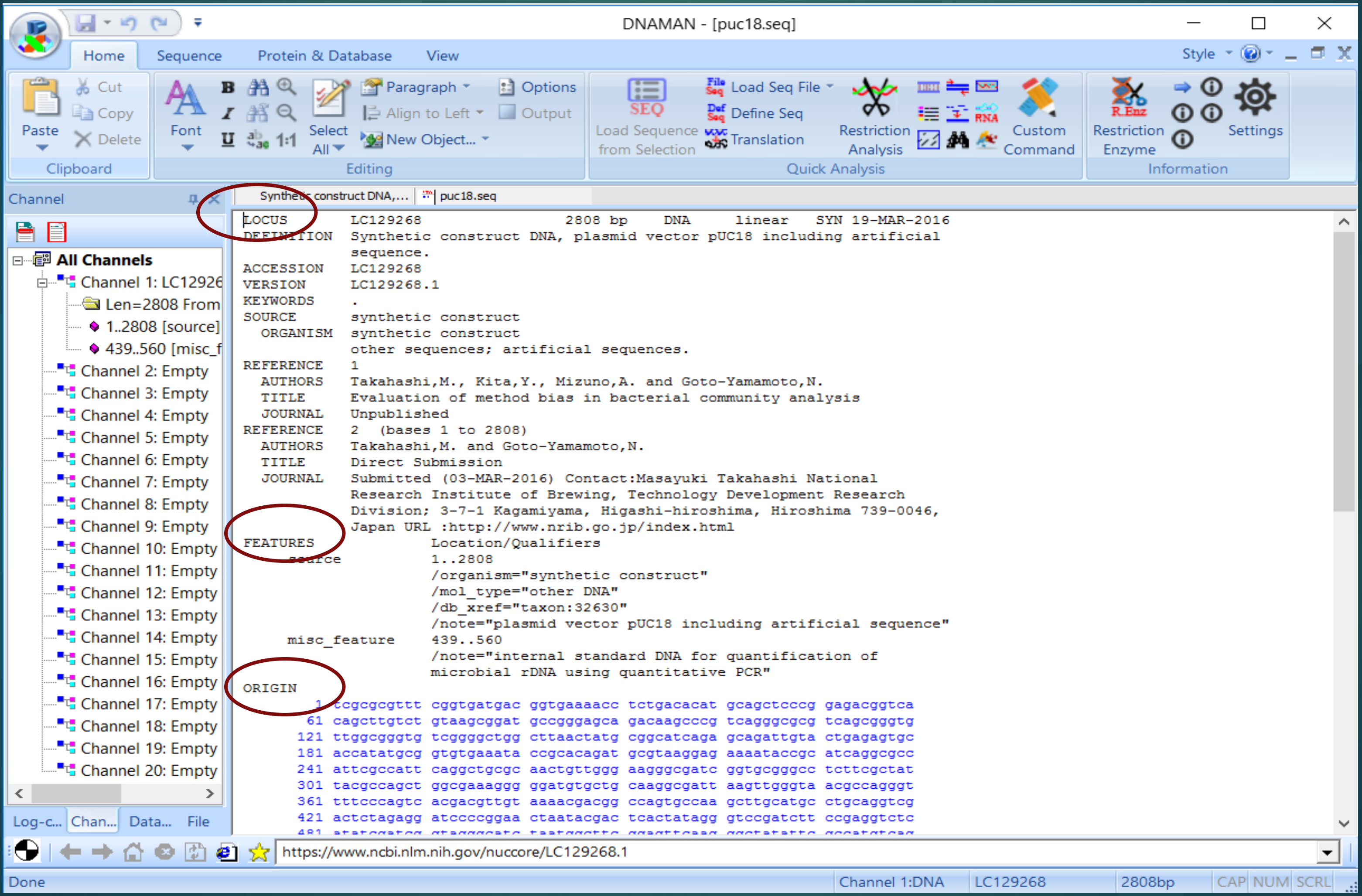
Downloaded: 4.83 KB in 1 sec
Download to: C:\Users\ling\Downloads\puc18.seq
Transfer rate: 4.83 KB/Sec

☐ Close this dialog box when download completes

Retrieve Sequence

From the Internet

Sequence is loaded into DNAMAN

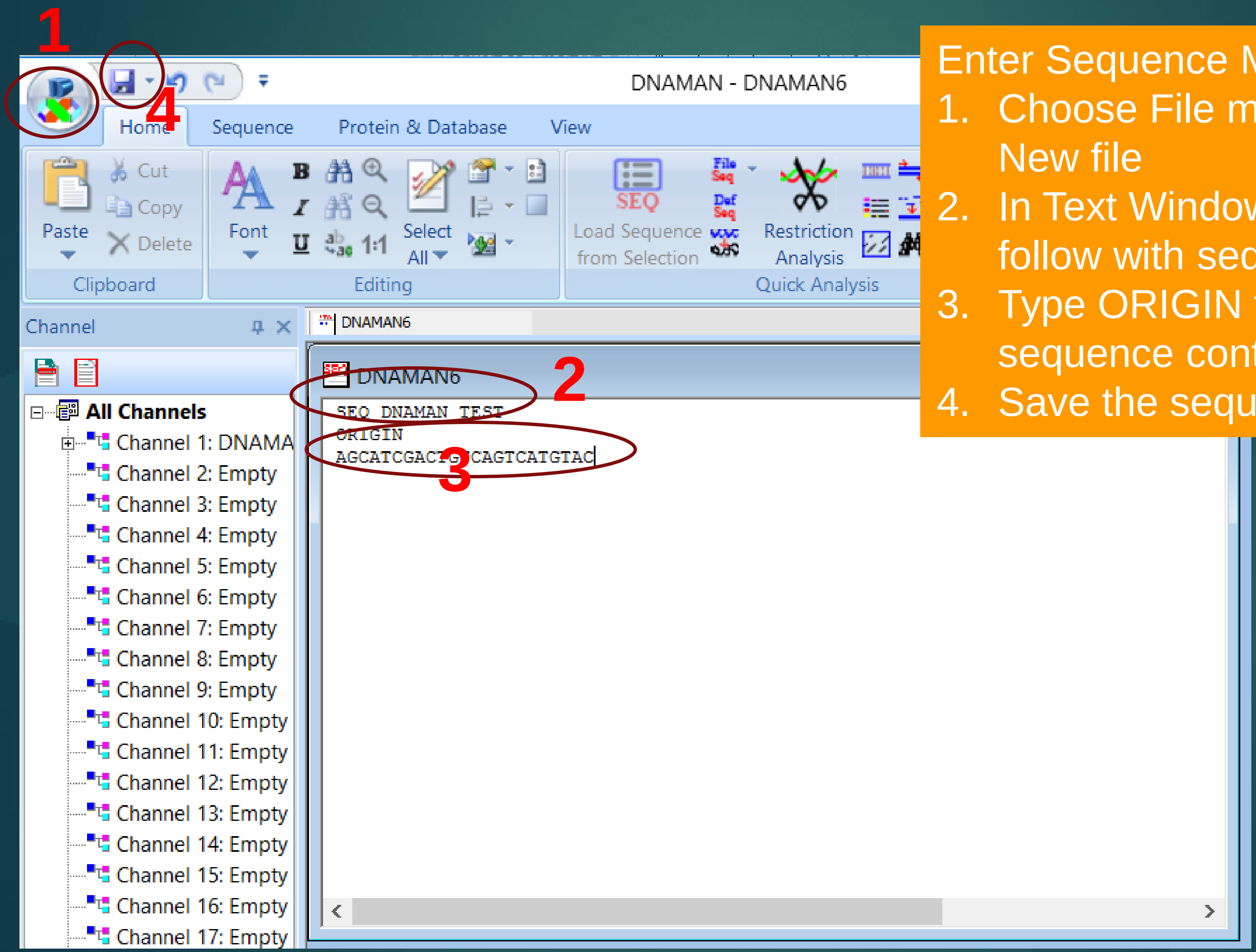


```
LOCUS       LC129268                2808 bp    DNA     linear   SYN 19-MAR-2016
DEFINITION  Synthetic construct DNA, plasmid vector pUC18 including artificial
            sequence.
ACCESSION   LC129268
VERSION     LC129268.1
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Takahashi,M., Kita,Y., Mizuno,A. and Goto-Yamamoto,N.
  TITLE     Evaluation of method bias in bacterial community analysis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2808)
  AUTHORS   Takahashi,M. and Goto-Yamamoto,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAR-2016) Contact:Masayuki Takahashi National
            Research Institute of Brewing, Technology Development Research
            Division; 3-7-1 Kagamiyama, Higashi-hiroshima, Hiroshima 739-0046,
            Japan URL :http://www.nrib.go.jp/index.html
FEATURES             Location/Qualifiers
     source          1..2808
                     /organism="synthetic construct"
                     /mol_type="other DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid vector pUC18 including artificial sequence"
     misc_feature    439..560
                     /note="internal standard DNA for quantification of
                     microbial rDNA using quantitative PCR"
ORIGIN
        1 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca
       61 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg
      121 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc
      181 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc
      241 attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat
      301 tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt
      361 tttcccagtc acgacgttgt aaaacgacgg ccagtgccaa gcttgcatgc ctgcaggtcg
      421 actctagagg atccccggaa ctaatacgac tcactatagg gtccgatctt ccgaggtctc
```

DNAMAN Sequence Format

- DNAMAN uses simplified GenBank Format
- GenBank file is automatically recognized in DNAMAN
- “FEATURES” in GenBank file are parsed out as annotations in DNAMAN
- Sequence contents begin after Keyword “ORIGIN”
- Sequence name is labeled after keyword “Locus” or “Seq”
- Keywords “REFERENCES” and “SOURCES” can be stored in sequence databases, bit not used in Sequence Analysis

Enter Custom Sequence



```
SEQ DNAMAN TEST
ORIGIN
AGCATCGACTG[illegible]CAGTCATGTAC
```

Enter Sequence Manually

1. Choose File menu and click New file
2. In Text Window, type Seq and follow with sequence name
3. Type ORIGIN then enter sequence contents
4. Save the sequence file

Load Existing Sequence

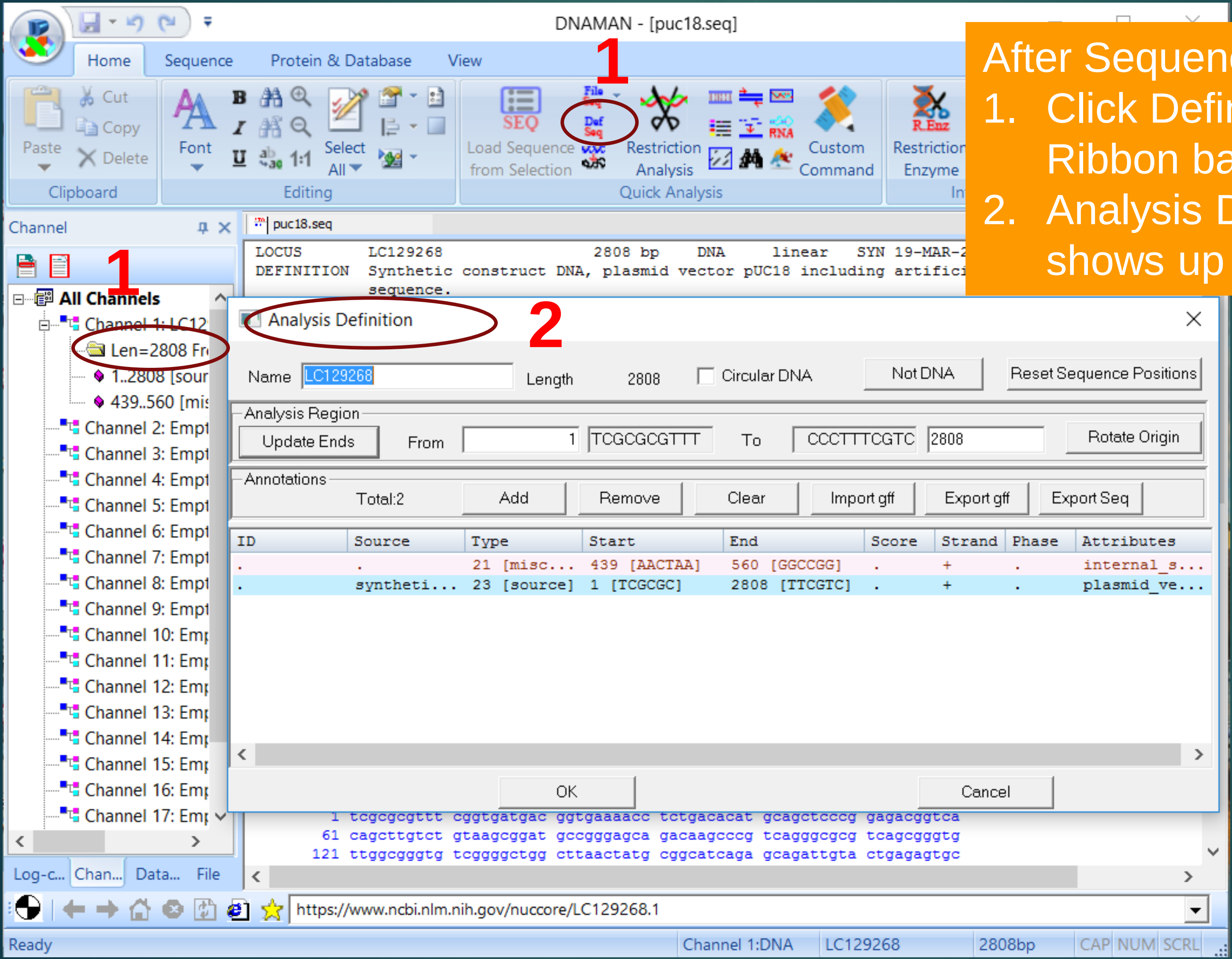
From ***File*** | ***Open*** menu

- Click File | Open menu
- Browse the computer to find sequence files
- Click the file to open
- DNAMAN loads automatically the sequence to channel for analysis

From ***File Control***

- Click the File tab in Sequence Controls on the left of workspace
- Browse the computer directly from the File control
- Click the file to open
- DNAMAN loads automatically the sequence to channel for analysis

Analysis Definition



```
LOCUS       LC129268            2808 bp    DNA     linear   SYN 19-MAR-2
DEFINITION  Synthetic construct DNA, plasmid vector pUC18 including artifici
            sequence.
```

Analysis Definition

Name LC129268 Length 2808 ☐ Circular DNA

Analysis Region

From 1 TCGCGCGTTT To CCCTTTCGTC 2808

Annotations

Total:2

ID	Source	Type	Start	End	Score	Strand	Phase	Attributes
.	.	21 [misc...	439 [AACTAA]	560 [GGCCGG]	.	+	.	internal_s...
.	syntheti...	23 [source]	1 [TCGCGC]	2808 [TTCGTC]	.	+	.	plasmid_ve...

```
  1 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca
 61 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg
121 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc
```

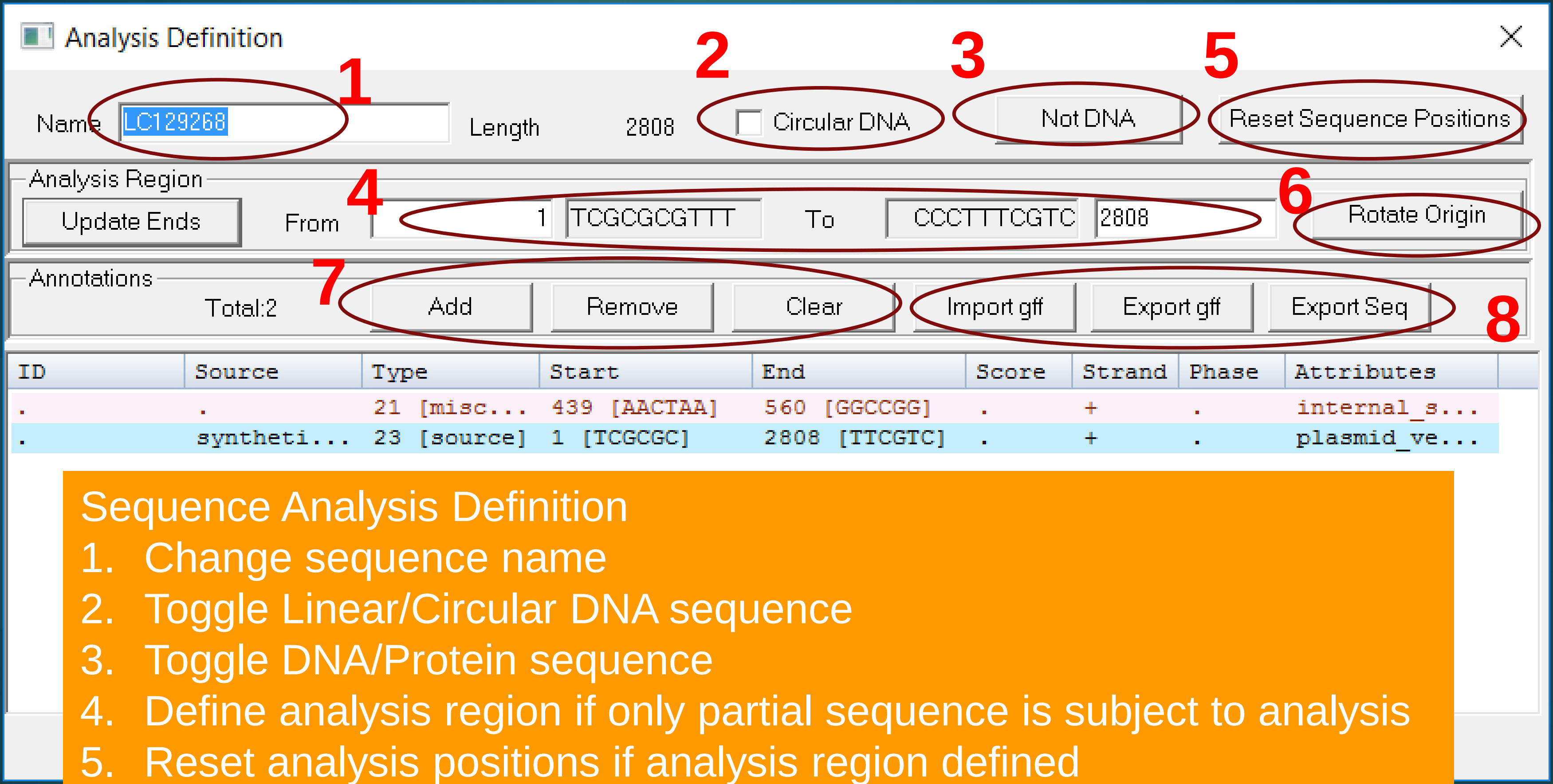
https://www.ncbi.nlm.nih.gov/nuccore/LC129268.1

Channel 1:DNA LC129268 2808bp

After Sequence Loaded in Channel

1. Click Define Sequence on Ribbon bar or Channel control
2. Analysis Definition Dialog box shows up

Analysis Definition



Name LC129268 Length 2808 ☐ Circular DNA

Analysis Region

From 1 TCGCGCGTTT To CCCTTTCGTC 2808

Annotations

Total:2

ID	Source	Type	Start	End	Score	Strand	Phase	Attributes
.	.	21 [misc...	439 [AACTAA]	560 [GGCCGG]	.	+	.	internal_s...
.	syntheti...	23 [source]	1 [TCGCGC]	2808 [TTCGTC]	.	+	.	plasmid_ve...

Sequence Analysis Definition

1. Change sequence name
2. Toggle Linear/Circular DNA sequence
3. Toggle DNA/Protein sequence
4. Define analysis region if only partial sequence is subject to analysis
5. Reset analysis positions if analysis region defined
6. Rotate sequence origin for circular DNA
7. Edit annotations
8. Export/Import annotations or sequence content

Work with other Sequence Formats

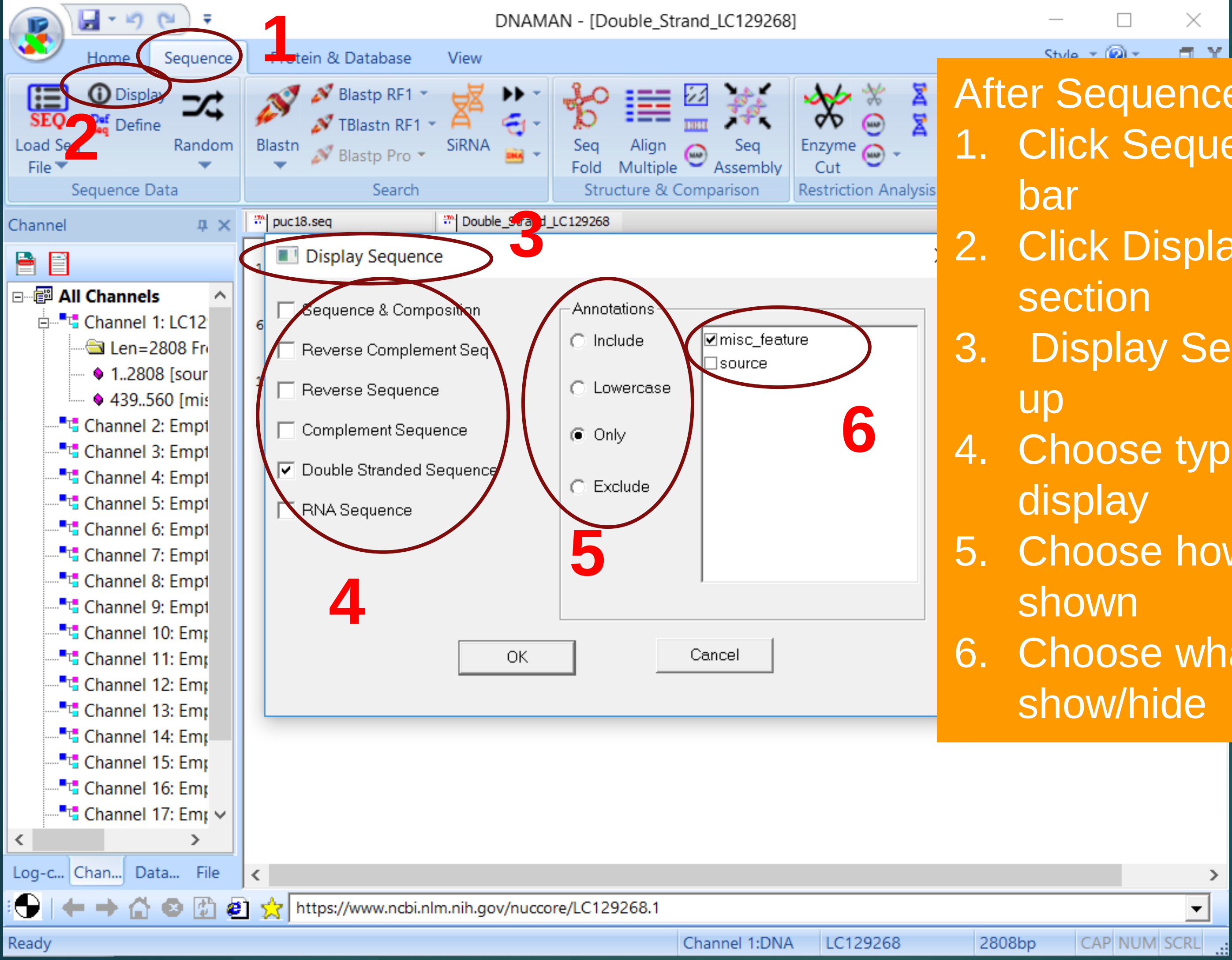
DNAMAN Reads Files

- GenBank
- FASTA (Single and multiple alignment)
- Trace files: ABI and SCF
- GFF3
- GCG multiple alignment (aln)
- GDE multiple alignment
- CLUSTAL multiple alignment
- NBRF multiple alignment
- PHYLIP multiple alignment
- MEGA multiple alignment

DNAMAN Exports Files

- FASTA (Single and multiple alignment)
- GFF3
- GCG multiple alignment
- GDE multiple alignment
- CLUSTAL multiple alignment
- NBRF multiple alignment
- PHYLIP multiple alignment

Show/Convert Sequence



Display Sequence

☐ Sequence & Composition
☐ Reverse Complement Seq
☐ Reverse Sequence
☐ Complement Sequence
☑ Double Stranded Sequence
☐ RNA Sequence

Annotations
○ Include
○ Lowercase
● Only
○ Exclude

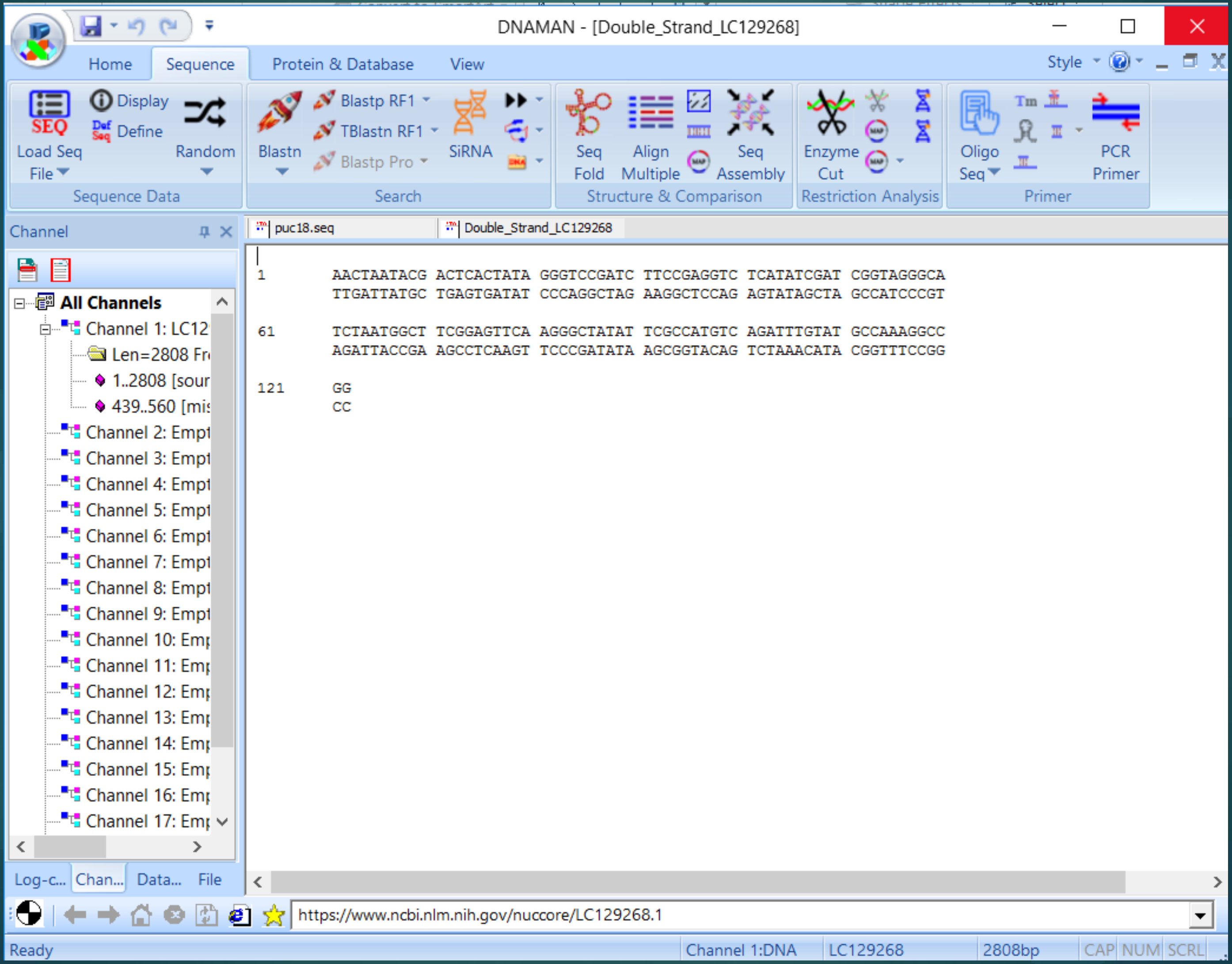
☑ misc_feature
☐ source

After Sequence Loaded in Channel

1. Click Sequence tab in Ribbon bar
2. Click Display in Sequence Data section
3. Display Sequence box shows up
4. Choose type of sequence to display
5. Choose how annotations to be shown
6. Choose what annotations to show/hide

Show/Convert Sequence

Display Results

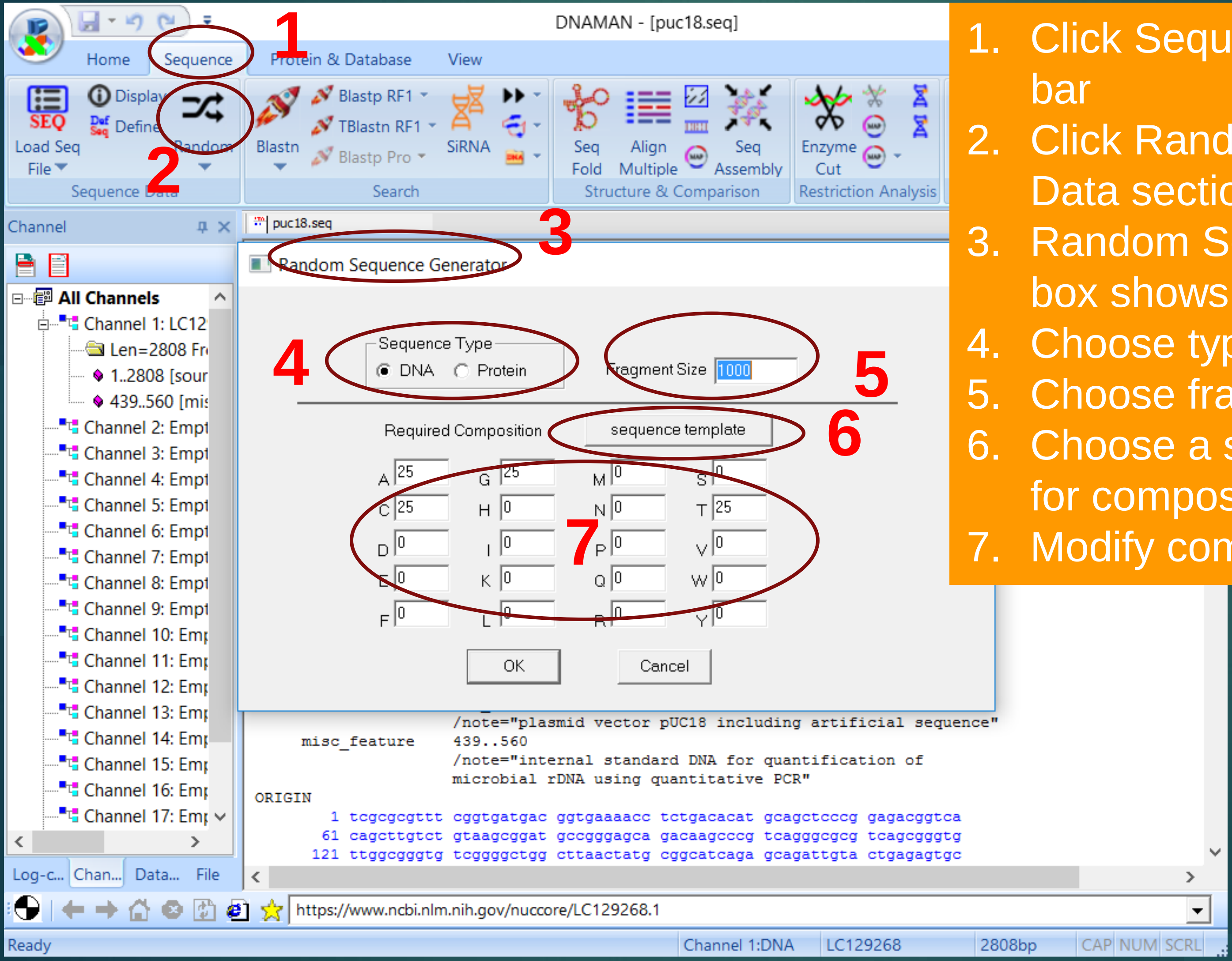


```
1       AACTAATACG ACTCACTATA GGGTCCGATC TTCCGAGGTC TCATATCGAT CGGTAGGGCA
        TTGATTATGC TGAGTGATAT CCCAGGCTAG AAGGCTCCAG AGTATAGCTA GCCATCCCGT

61      TCTAATGGCT TCGGAGTTCA AGGGCTATAT TCGCCATGTC AGATTTGTAT GCCAAAGGCC
        AGATTACCGA AGCCTCAAGT TCCCGATATA AGCGGTACAG TCTAAACATA CGGTTTCCGG

121     GG
        CC
```

Random Sequence



Sequence Type: ● DNA ○ Protein

Fragment Size: 1000

Required Composition

A	25	G	25	M	0	S	0
C	25	H	0	N	0	T	25
D	0	I	0	P	0	V	0
E	0	K	0	Q	0	W	0
F	0	L	0	R	0	Y	0

```
                /note="plasmid vector pUC18 including artificial sequence"
     misc_feature    439..560
                /note="internal standard DNA for quantification of
                microbial rDNA using quantitative PCR"
ORIGIN
        1 tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca
       61 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg
      121 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc
```

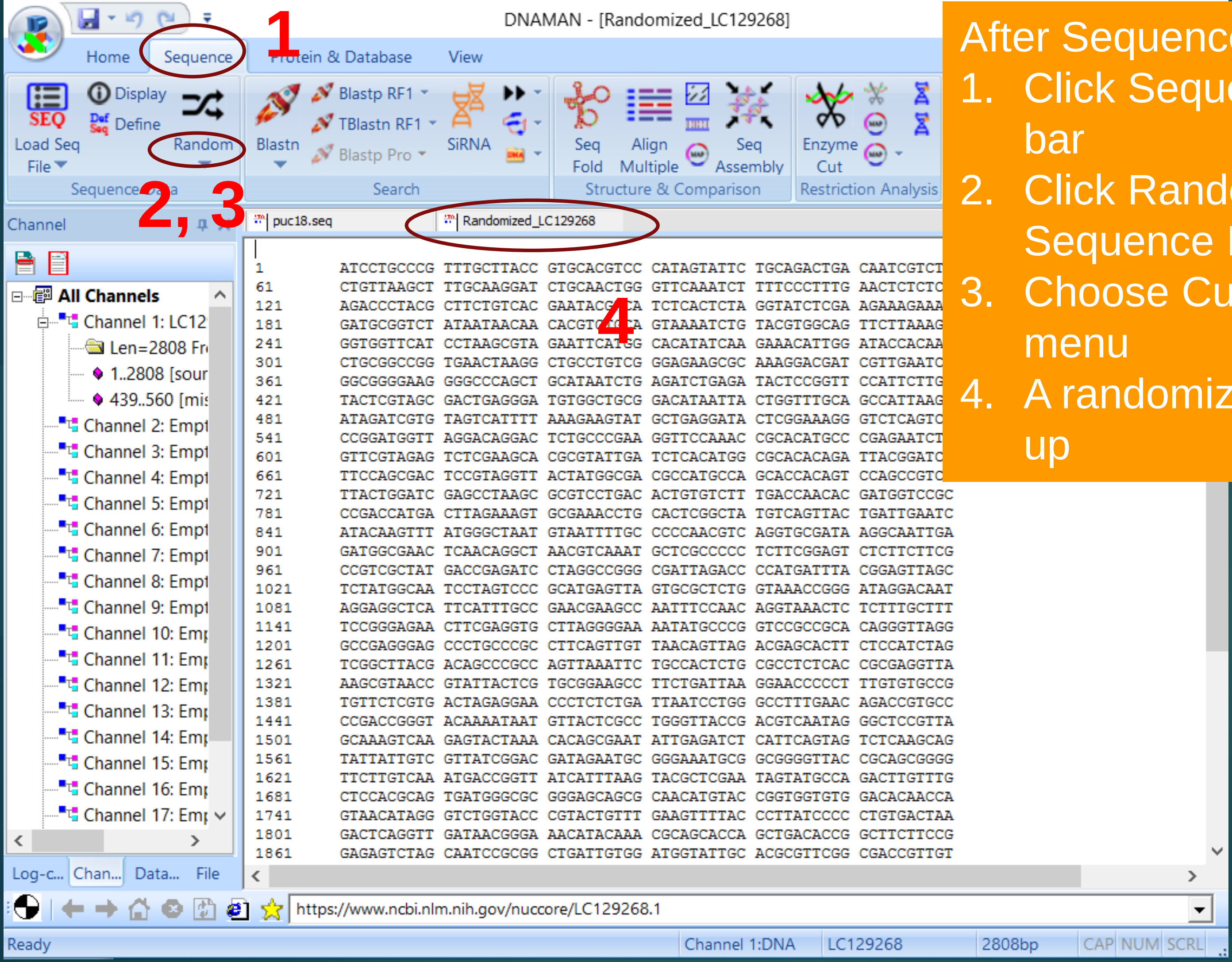
https://www.ncbi.nlm.nih.gov/nuccore/LC129268.1

Ready | Channel 1:DNA | LC129268 | 2808bp | CAP NUM SCRL

1. Click Sequence tab in Ribbon bar
2. Click Random in Sequence Data section
3. Random Sequence Generator box shows up
4. Choose type of sequence
5. Choose fragment size
6. Choose a sequence template for composition if needed
7. Modify composition if needed

Random Sequence

Randomize Current Sequence



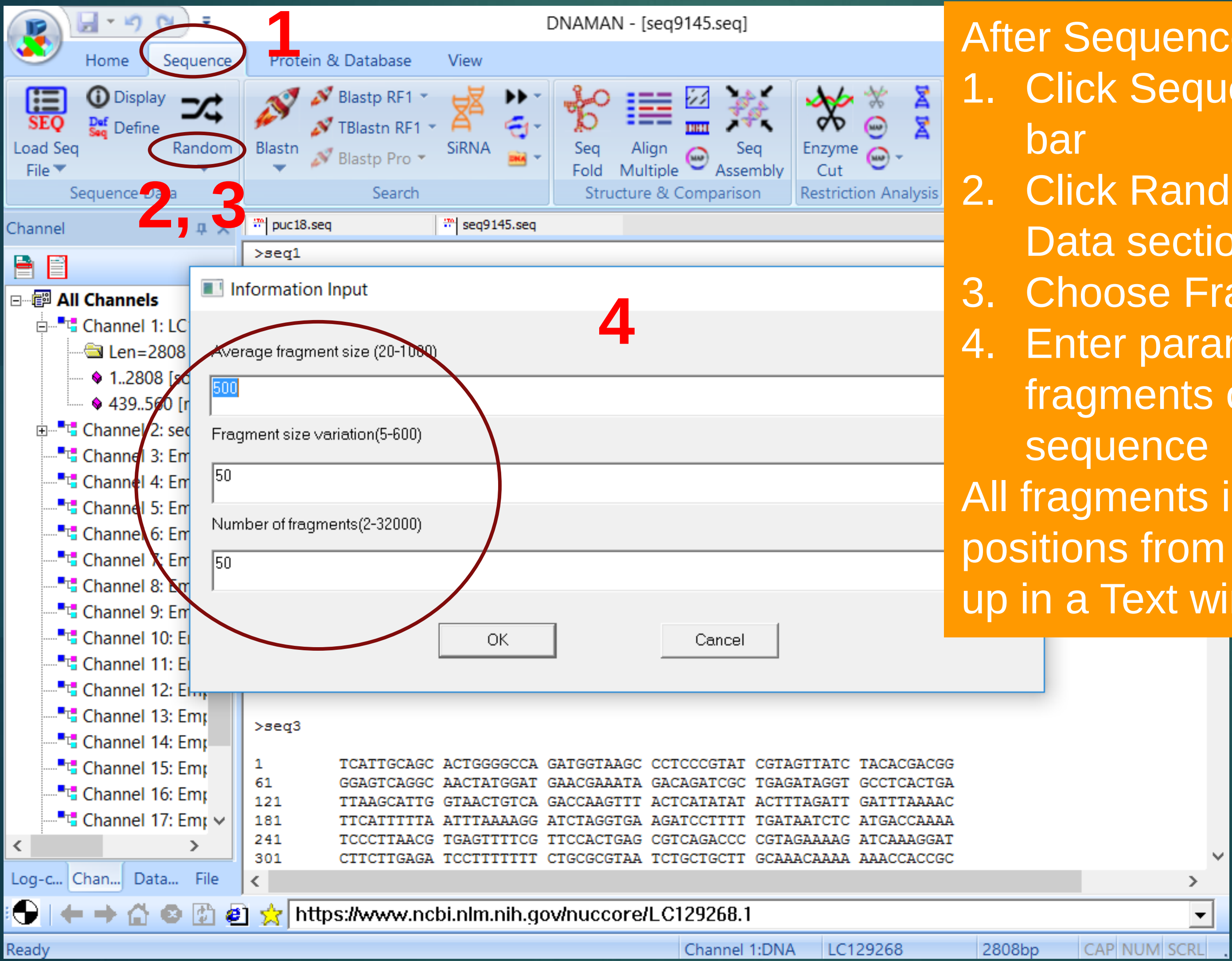
```
1       ATCCTGCCCG TTTGCTTACC GTGCACGTCC CATAGTATTC TGCAGACTGA CAATCGTCT
61      CTGTTAAGCT TTGCAAGGAT CTGCAACTGG GTTCAAATCT TTTCCCTTTG AACTCTCTC
121     AGACCCTACG CTTCTGTCAC GAATACG[illegible]CA TCTCACTCTA GGTATCTCGA AGAAAGAAA
181     GATGCGGTCT ATAATAACAA CACGT[illegible]A GTAAAATCTG TACGTGGCAG TTCTTAAAG
241     GGTGGTTCAT CCTAAGCGTA GAATTCATGG CACATATCAA GAAACATTGG ATACCACAA
301     CTGCGGCCGG TGAACTAAGG CTGCCTGTCG GGAGAAGCGC AAAGGACGAT CGTTGAATC
361     GGCGGGGAAG GGGCCCAGCT GCATAATCTG AGATCTGAGA TACTCCGGTT CCATTCTTG
421     TACTCGTAGC GACTGAGGGA TGTGGCTGCG GACATAATTA CTGGTTTGCA GCCATTAAG
481     ATAGATCGTG TAGTCATTTT AAAGAAGTAT GCTGAGGATA CTCGGAAAGG GTCTCAGTC
541     CCGGATGGTT AGGACAGGAC TCTGCCCGAA GGTTCCAAAC CGCACATGCC CGAGAATCT
601     GTTCGTAGAG TCTCGAAGCA CGCGTATTGA TCTCACATGG CGCACACAGA TTACGGATC
661     TTCCAGCGAC TCCGTAGGTT ACTATGGCGA CGCCATGCCA GCACCACAGT CCAGCCGTC
721     TTACTGGATC GAGCCTAAGC GCGTCCTGAC ACTGTGTCTT TGACCAACAC GATGGTCCGC
781     CCGACCATGA CTTAGAAAGT GCGAAACCTG CACTCGGCTA TGTCAGTTAC TGATTGAATC
841     ATACAAGTTT ATGGGCTAAT GTAATTTTGC CCCCAACGTC AGGTGCGATA AGGCAATTGA
901     GATGGCGAAC TCAACAGGCT AACGTCAAAT GCTCGCCCCC TCTTCGGAGT CTCTTCTTCG
961     CCGTCGCTAT GACCGAGATC CTAGGCCGGG CGATTAGACC CCATGATTTA CGGAGTTAGC
1021    TCTATGGCAA TCCTAGTCCC GCATGAGTTA GTGCGCTCTG GTAAACCGGG ATAGGACAAT
1081    AGGAGGCTCA TTCATTTGCC GAACGAAGCC AATTTCCAAC AGGTAAACTC TCTTTGCTTT
1141    TCCGGGAGAA CTTCGAGGTG CTTAGGGGAA AATATGCCCG GTCCGCCGCA CAGGGTTAGG
1201    GCCGAGGGAG CCCTGCCCGC CTTCAGTTGT TAACAGTTAG ACGAGCACTT CTCCATCTAG
1261    TCGGCTTACG ACAGCCCGCC AGTTAAATTC TGCCACTCTG CGCCTCTCAC CGCGAGGTTA
1321    AAGCGTAACC GTATTACTCG TGCGGAAGCC TTCTGATTAA GGAACCCCCT TTGTGTGCCG
1381    TGTTCTCGTG ACTAGAGGAA CCCTCTCTGA TTAATCCTGG GCCTTTGAAC AGACCGTGCC
1441    CCGACCGGGT ACAAAATAAT GTTACTCGCC TGGGTTACCG ACGTCAATAG GGCTCCGTTA
1501    GCAAAGTCAA GAGTACTAAA CACAGCGAAT ATTGAGATCT CATTCAGTAG TCTCAAGCAG
1561    TATTATTGTC GTTATCGGAC GATAGAATGC GGGAAATGCG GCGGGGTTAC CGCAGCGGGG
1621    TTCTTGTCAA ATGACCGGTT ATCATTTAAG TACGCTCGAA TAGTATGCCA GACTTGTTTG
1681    CTCCACGCAG TGATGGGCGC GGGAGCAGCG CAACATGTAC CGGTGGTGTG GACACAACCA
1741    GTAACATAGG GTCTGGTACC CGTACTGTTT GAAGTTTTAC CCTTATCCCC CTGTGACTAA
1801    GACTCAGGTT GATAACGGGA AACATACAAA CGCAGCACCA GCTGACACCG GCTTCTTCCG
1861    GAGAGTCTAG CAATCCGCGG CTGATTGTGG ATGGTATTGC ACGCGTTCGG CGACCGTTGT
```

After Sequence loaded in channel

1. Click Sequence tab in Ribbon bar
2. Click Random Arrow in Sequence Data section
3. Choose Current Sequence menu
4. A randomized sequence shows up

Random Sequence

Fragment Current Sequence



>seq1

Information Input

Average fragment size (20-1000): 500

Fragment size variation(5-600): 50

Number of fragments(2-32000): 50

```
>seq3

1         TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG
61        GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA
121       TTAAGCATTG GTAACTGTCA GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAAC
181       TTCATTTTTA ATTTAAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAA
241       TCCCTTAACG TGAGTTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT
301       CTTCTTGAGA TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT GCAAACAAAA AAACCACCGC
```

https://www.ncbi.nlm.nih.gov/nuccore/LC129268.1

After Sequence loaded in channel

1. Click Sequence tab in Ribbon bar
2. Click Random Tool in Sequence Data section
3. Choose Fragmentation menu
4. Enter parameters to generate fragments of the current sequence

All fragments in random sizes and positions from the sequence show up in a Text window.