


# SERIAL CLONER 2.1

## May 2010

### Handling Features/Annotations

Since version 2.0, Serial Cloner handles annotations and features. Here are some descriptions and hints about these new possibilities.

The symbol  indicate that this function has been added or updated in V2.1

#### I. Accessing features in the Sequence window

A new tab has been added (*Features*) that gives access to features characteristics and management. A list of Features is shown on the left; double clicking on one will select the corresponding nucleotide sequence. To edit (*add/delete/modify*) Features, one has to be in the *'Edit'* mode. The Edit button is used to toggle between the *'protected'* mode where no unwanted modification can be done and the *'Edit'* mode.

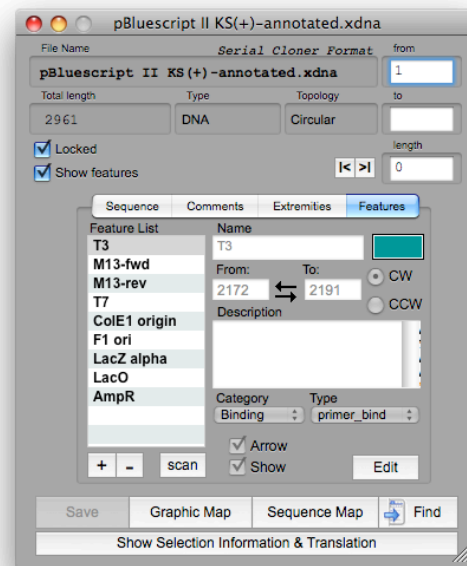
One can select whether a feature will actually be displayed on the map (using the *'Show'* checkbox). Shown features are listed in Bold characters. To decide to Show or Hide all features at once, press the [ALT] key before clicking on the checkbox.

A feature can have or not a pointing arrow (*'Arrow'* checkbox).

To change the Feature color, click on the *colored box* and select another color.

The order in the list dictate the order by which Features will be drawn, and this can be important in case they overlap. To alter the order simply drag the Feature at another position in the list.

The + and – button are used to add and delete a Feature respectively. A pop-up message will ask for confirmation before deleting a Feature (there is no undo). This pop-message can be avoided if you press the [ALT] key before clicking on the – button. The button then gets red to indicate that no confirmation will be asked for. This is particularly useful when deleting multiple Features.



 2.1

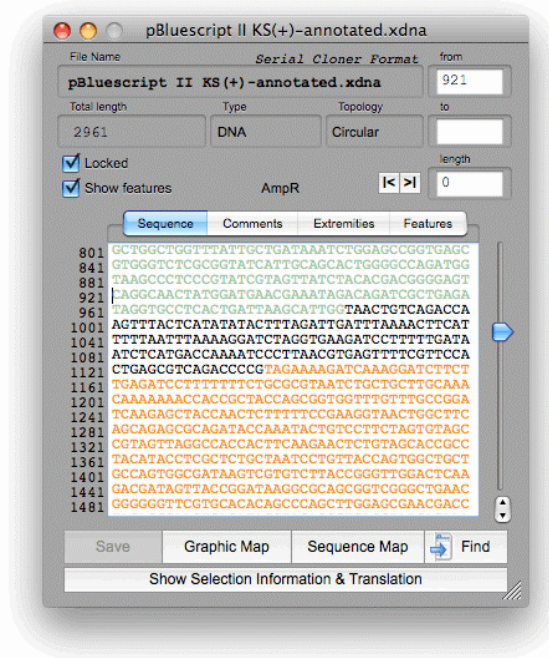
#### II. Display of Features in the Sequence Window

Features are visible in color in the sequence window. It is possible to show/hide them using the *'Show Features'* checkbox.

When the insertion point is located within a Feature, the *corresponding Feature's name* is indicated above the Sequence Field.

2.1

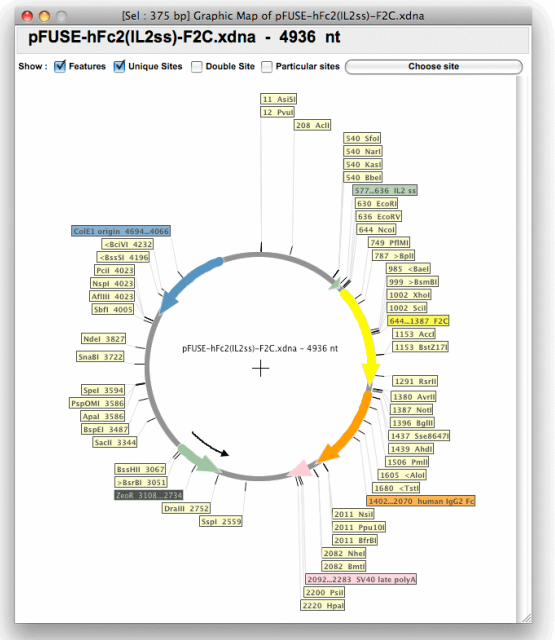
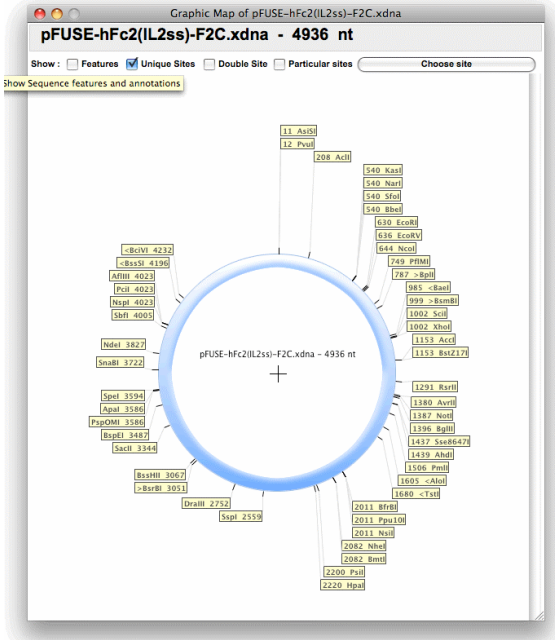
NB: note that since version 2.1 it is now possible to change the Sequence window size (as well as the character size under Windows AND MacOSX). The scrolling of the sequence when selecting has also been improved.



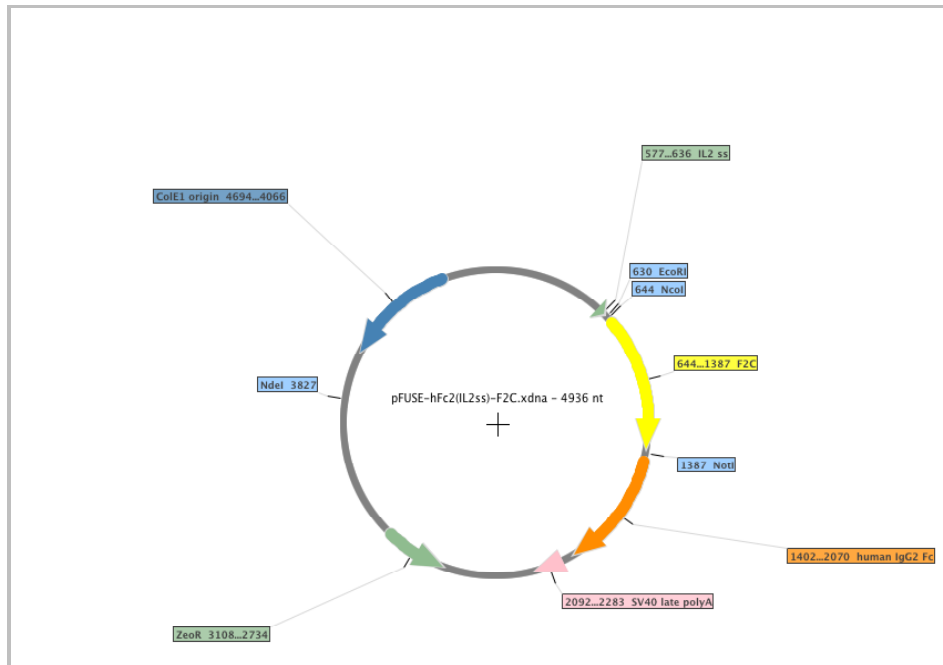
### III. Graphical display of Feature.

Features can be displayed in the Graphic Map. They are represented by colored boxes with or without a pointing arrow. It is possible to switch between a Graphic Map with and without Feature using the checkbox present in the top left of the window (*Features*).

A *simple click* on a Feature name selects the corresponding portion in the map. A *double Click* selects the corresponding sequence in the parent Sequence Window.



Note that it is now possible (since V2.0) to *move* the Restriction site and Features name boxes to highlight some of them before printing or copying the map. It is also possible to display only certain sites together with the Features. Finally, it is now possible (since v2.0) to resize the Graphic map in Height AND Width.



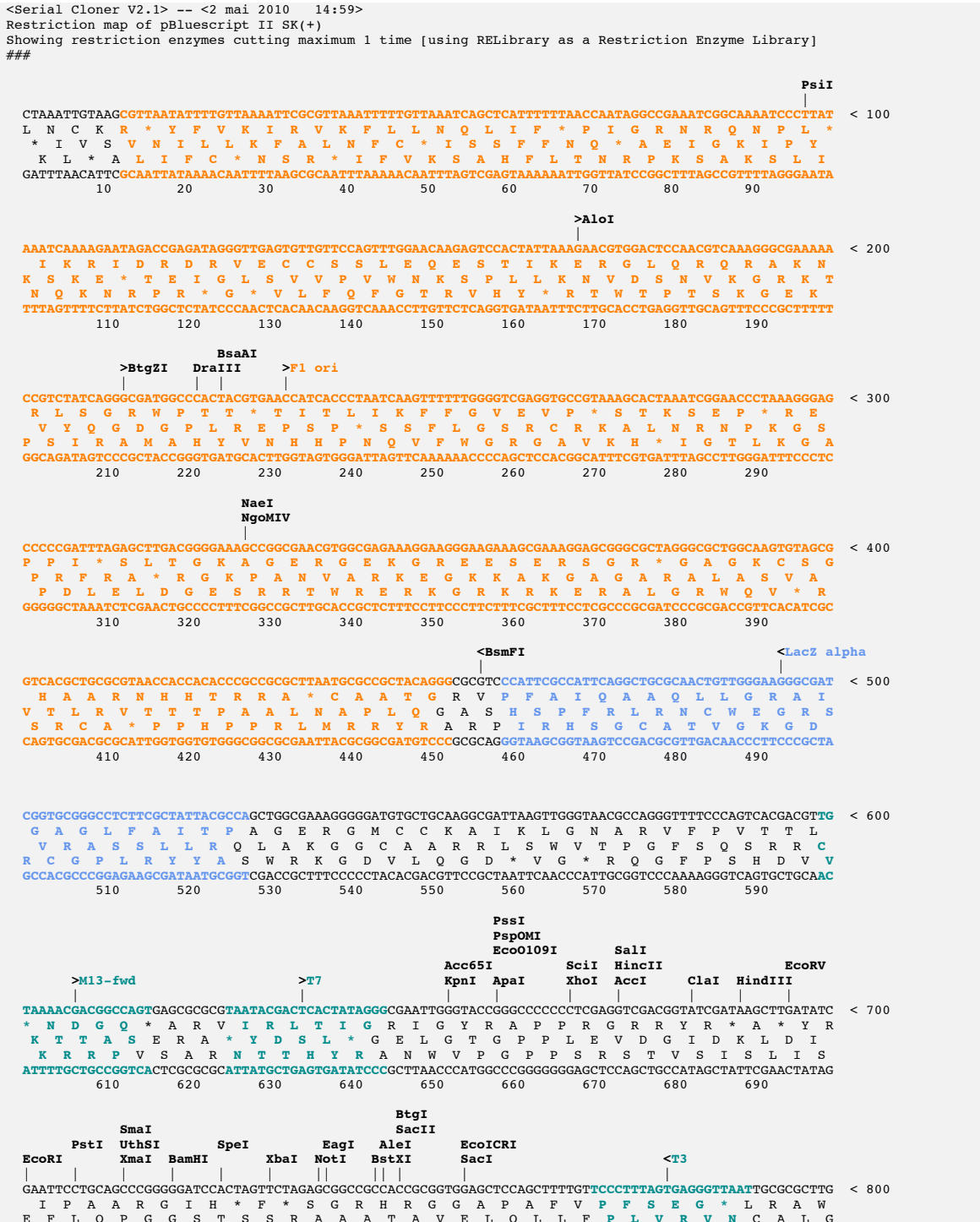
2.1

### IV. Text visualization of the Features

With Serial Cloner 2.1, it is now possible to see the Features with the Sequence Map. A checkbox present at the top left of the Sequence Map window allows to toggle on and off Feature display.

The ">" and "<" is used to indicated the sense of the Feature.

A legend is pasted at the end of the map.



N S C S P G D P L V L E R P P P R W S S S F C S L \* \* G L I A R L  
CTTAAGGACGTGGGCCCCCTAGGTGATCAAGATTCGCCGCGGTGGCCACCCTCGAGGTCGAAAACAAGGAAATCACTCCCAATTAACGCGCGAAC  
710 720 730 740 750 760 770 780 790

<M13-rev  
CGGTAATCATGGTCATAGCTGTTCCTGTGTGAATTTGTTATCCGCTCACAATCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGGT < 900  
R N H G H S C F L C E I V I R S Q F H T T Y E P E A \* S V K P G V  
V I M V I A V S C V K L L S A H N S T Q H T S R K H K V \* S L G C  
A \* S W S \* L F P V \* N C Y P L T I P H N I R A G S I K C K A W G A  
CGCATTAGTACCAGTATCGACAAGGACACACTTTAACAAATAGGCGAGTGTAAAGGTGTGTGTATGCTCGGCTTCGTATTACATTTCCGACCCAC  
810 820 830 840 850 860 870 880 890

<LacO  
CCTAATGAGTGAGTAACACATTAATTCGGTTGGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCTGCGCAGCTGATTAATGAATCGGCCAACG < 1000  
P N E \* A N S H \* L R C A H C P L S S R E T C R A S C I N E S A N A  
L M S E L T H I N C V A L T A R F P V G K P V V P A A L M N R P T  
\* \* V S \* L T L I A L R S L P A F Q S G N L S C Q L H \* \* I G Q R  
GGATTACTACFCGATTGCAATGTAATTAACGCAACGCGAGTGACGGCGAAAGGTGACGCCCTTTGGACAGCAGCGTACGTAATTAATTAATCGCGTTC  
910 920 930 940 950 960 970 980 990

>SapI  
CGCGGGGAGAGCGGTTTTCGATTTGGGCGCTTCCGCTTCTCGCTCACTGACTCGTTCGCTCGGTCGTTCGGCTGCGGAGCGGATCAGCTCAC < 1100  
R G E A V C V L G A L P L P R S L T R C A R S F G C G E R Y Q L T  
R G E R R F A Y W A L F R F L A H \* L A A L G R S A A A S G I S S L  
A G R G G L R I G R S S A S S L T D S L R S V V R L R R A V S A H  
GCGCCCTTCCGCAACGCATAACCCGAGAAAGCGAGTGACTGAGCAGCGAGCCAGCAGCCGAGCCGCTCGCCATAGTCGAGT  
1010 1020 1030 1040 1050 1060 1070 1080 1090

PciI  
NspI  
AflIII  
TCAAAGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAGCCGAGAACCGTAAAAAGGCCG < 1200  
Q R R \* Y G Y P Q N Q G I T Q E R T C E Q K A S K R P G T V K R P  
K G R N T V I H R I R G \* R R K E H V S K R P A K G Q E P \* K R P  
S K A V I R L S T E S G D N A G K N M \* A K G Q Q K A R N R K K A A  
AGTTTCGCGCATTATGCCAATAGGTGTCTAGTCCCTATTCGCTTCTTGACACTCGTTCGCGGTCGTTTTCGGTCCCTGGCATTTCCTCCGCG  
1110 1120 1130 1140 1150 1160 1170 1180 1190

CGTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATA < 1300  
R C W R F S I G S A P L T S I T K I D A Q V R G G E T R Q D Y K D T  
V A G V F P \* A P P P \* R A S Q K S T L K S E V A K P D R T I K I  
L L A F F H R L R P P D E H H K N R R S S Q R W R N P T G L \* R Y  
GCAACGACGCAAAGGATCCAGGCGGGGACTGCTCGTAGTGTTCGCTGCGAGTTCAGTCCACCGCTTGGGCTGTCTGATATTTCTAT  
1210 1220 1230 1240 1250 1260 1270 1280 1290

CCAGGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCGACCCCTGCCGTTACCGGATACCTGTCGCGCTTTCCTCCCTCGGGAAGCGTGGCG < 1400  
R R F P L E A P S C A L L F R P C R L P D T C P P F S L R E A W R  
P G V S P W K L P R A L S C S D P A A Y R I P V R L S P F G K R G A  
Q A F P P G S S L V R S P V P T L P L T G Y L S A F L P S G S V A  
GGTCCGCAAAGGGACCTTCGAGGAGCACGCGAGAGGACAGGCTGGGACGGCGAATGGCCTATGGACAGCGGAAAGAGGGAAAGCCCTTCGACCCG  
1310 1320 1330 1340 1350 1360 1370 1380 1390

<BseYI  
CTTTCATAGTACGCTGCTGAGTATCTCAGTTCGGTGTAGGTCGTCGCTCCAAGCTGGGCTGTGTGCACGAAACCCCGTTTCAGCCCGACCGCTGCG < 1500  
F L I A H A V G I S V R C R S F A P S W A V C T N P P F S P T A A  
F S \* L T L \* V S Q F G V G R S L Q A G L C A R T P R S A R P L R  
L S H S S R C R Y L S S V \* V V R S K L G C V H E P P V Q P D R C A  
GAAAGATATCGAGTCCCATAGACTCAAGCCACATCCAGAACGAGGTTTCGACCCGACACAGTCTGGGGGGCAAGTCGGGCTGGCGCAGC  
1410 1420 1430 1440 1450 1460 1470 1480 1490

<ColE1 origin  
CCTTATCCGGTAACTATCGCTTGAAGTCCAAACCCGGTAAAGACGACTTATCGCCACTGGCAGCAGCCTGGTAAACAGGATTAGCAGCGGAGGTATGT < 1600  
P Y P V T I V L S P T R \* D T T Y R H W Q Q P L V T G L A E R G M \*  
L I R \* L S S \* V Q P G K T R L I A T T G S S H W \* Q D \* Q S E V C  
L S G N Y R L E S N P V R H D L S P L A A A T G N R I S R A R Y V  
GGAATAGGCCATTGATAGCAGAATCAAGTTCGGCCATTCTGTCTGAATAGCGGTGACCGTTCGTCGGTACCATTTGCTCAATCGTCTCGCTCCATACA  
1510 1520 1530 1540 1550 1560 1570 1580 1590

AlwNI  
AGCGGCTACAGAGTTCGAGTGGGCTAAGTACGCTACACTAGAGGACAGTATTGGTATCTGCGCTCGCTGAGCCAGTACCTTCGGA < 1700  
A V L Q S S \* S G V L T T A T L E G Q Y L V S A L C \* S Q L P S E  
R R C Y R V L E V A \* L R L A H \* K D S I W Y L R S A E A S Y L R K  
G G A T E F L K W W P N Y G Y T R R T V F G I C A L L K P V T F G  
TCCGCCAGATGCTCAAGAACCTCACCCGATTGATCCGATGTGATCTTCTGTCATAAACCATAGACGCGAGACGACTTCGGTCAATGGAAAGCT  
1610 1620 1630 1640 1650 1660 1670 1680 1690

AAAAGAGTTGGTGTCTTGTATCCGGCAAACAACCCGCTGGTAGCGTGGTTTTTTTGGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTC < 1800  
K E L V A L D P A N K P L V A V V F L F A S S R L R A E K K D L  
K S W \* L L I R Q T N H R W \* R V F F C L Q A A D Y A Q K K R I S  
K R V G S S \* S G K Q T T A G S C G G F F V C K Q Q I T R R K K G S Q  
TTTTCTCAACCATGAGAACATAGCCGTTTGTGGCGCCATCGCCCAAAAAAACAACGTTCTGCTCTAATGCGCGTCTTTTTTTCTAGAG  
1710 1720 1730 1740 1750 1760 1770 1780 1790

AAGAAGATCCCTTGTATCTTTCTACCGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCAC < 1900  
K K I L \* S F L R G L T L S G T K T H V K G F W S \* D Y Q K G S S P  
R R S F D L F Y G V \* R S V E R K L T L R D F G H E I I K K D L H  
E D P L I F S T G S D A Q W N E N S R \* G I L V M R L S K R I F T  
TTCTTCTAGGAACTAGAAAAGATGCCCCAGACTGCGAGTCACTTGTCTTTGAGTGAATCCCTAAAACAGTACTCTAATAGTTTTTCTAGAAAGTG  
1810 1820 1830 1840 1850 1860 1870 1880 1890

CTAGATCCTTTTAAATTAATAAGTAAATTTAAATCAATCAAAGTATATAGTAACTTGGTCTGACAGTTCCAATGCTTAATCAGTGGAGCACCT < 2000  
R S F \* I K N E V L N Q S K V Y M S K L G L T V T N A \* S V R H L  
L D P F K L K M K F \* I N L K Y I \* V N L V \* Q L P M L N Q \* G T Y  
\* I L L N \* K \* S F K S I \* S I Y E \* T W S D S Y Q C L I S E A P  
GATCTAGGAAATTTAATTTTACTTCAAATTTAGTTAGATTTTATATATACTCATTTGAACCCAGACTGTCAATGGTTACGAATTAGTCACTCCGTGGA  
1910 1920 1930 1940 1950 1960 1970 1980 1990

>HaeIV

AhdI

ATCTCAGCAGTCTGTCTATTTCGTTTCATCCATAGTTCGCTGACTCCCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTG < 2100  
S Q R S V Y F V H P \* L P D S P S C R \* L R Y G R A Y H L A P V L  
L S D L S I S F I H S C L T P R R V D N Y D T G G L T I W P Q C C  
I S A I C L F R S S I V A \* L P V V \* I T T I R E G L P S G P S A A  
TAGAGTCTGATAGCAGATAAAGCAAGTAGGTATCAACGGACTGAGGGGACGACATCTATTGATGCTATGCCCTCCCGAATGGTAGACCGGGTACAGC  
2010 2020 2030 2040 2050 2060 2070 2080 2090

<BsaI

CAATGATACCCGGAGACCCAGCTCACCAGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTGCACTTTATC < 2200  
Q \* Y R E T H A H R L Q I Y Q \* T S Q P E G P S A E V V L Q L Y P  
N D T A R P T L T G S R F I S N K P A S R K G R A Q K W S C N F I  
M I P R D P R S P A P R L S A I N Q P A G R A E R R S G P A T L S  
GTTACTATGGCCCTCTGGTGGAGTGGCCAGGCTAAATAGTCTTATTGGTGGTGGCCCTCCCGGCTCGCGTCTTACCAGGACGTTGAAATAG  
2110 2120 2130 2140 2150 2160 2170 2180 2190

CGCTCCATCCAGTCTATTAATTTGGTCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTCCTTGTACAGGCATCGTG < 2300  
P P S S L L I V A G K L E \* V V R Q L I V C A T L L P L L Q A S W  
R L H P V Y \* L L P G S \* S K \* F A S \* \* F A Q R C C H C Y R H R G  
A S I Q S I N C R E A R V S S S P V N S L R N V V A I A T G I V  
GCGGAGTAGGTCAGATAAATTAACACGGCCCTTCGATCTCATTATCAAGCGGTCAATTATCAACCGCTGCAACACGGTAACGATGTCCTGAGCAC  
2210 2220 2230 2240 2250 2260 2270 2280 2290

<AmpR

GTGTCAAGTCTCGTCTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTAGCT < 2400  
C H A R R L V W L H S A P V P N D Q G E L H D P P C C A K K R L A  
V T L V V W Y G F I Q L R F P T I K A S Y M I P H V V Q K S G \* L  
V S R S S F G M A S F S S G S Q R S R R V T \* S P M L C K K A V S S  
CACAGTCCGAGCAGCAAAACATACCAAGTAAGTTCGAGCCCAAGGTTGCTAGTTCGCTCAATGTAAGTACTAGGGGTACAACACGTTTTTTCGCCAATCGA  
2310 2320 2330 2340 2350 2360 2370 2380 2390

CCTTCGGTCTCCGATCGTTGTGTCAGAAAGTAAAGTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCCTTACTGTCCATCCGCT < 2500  
P S V L R S L S E V S W P Q C Y H S W L W Q H C I I L L L S C H P \*  
L R S S D R C Q K \* V G R S V I T H G Y G S T A \* F S Y C H A I R  
F G P P I V R S K L A A V L S L M V M A A L H N S L T V M P S V  
GGAAGCCAGGAGGCTAGCAACAGTCTTCAATCAACCGGCTCACAATAGTACCAATACCGTCTGACGTTAAGAGAAATGACAGTACGGTAGGCA  
2410 2420 2430 2440 2450 2460 2470 2480 2490

TatI

ScaI

>BcgI

>BcgI

BsaHI

AAGATGCTTTTCTGTACTGGTGGTACTCAACCAAGTCACTTCTGAGAATAGTGTATGCGGGCAGCAGGTTGCTCTTGCCCGGCTCAATACGGGATAAT < 2600  
D A F L \* L V S T Q P S H S E N S V C G D R V A L A R R Q Y G I I  
K M L F C D Q \* V L N Q V I L R I V Y A A T E L L L P G V N T G \* Y  
R C F S V T G E Y S T K S F \* E \* C M R R P S C S C P A S I R D N  
TTCTACGAAAGCAGTACCACTCATGAGTGGTTCAGTAAGACTTTATCACATACGCGGCTCAACGAGAACGGCCGAGTTATGCGCCCTATTA  
2510 2520 2530 2540 2550 2560 2570 2580 2590

XmnI

ACCCGCGCACATAGCAGAACTTTAAAGTGTCTCATTTGGAAAAAGTCTTCGCGGGCGAAAACTCTCAAGGATCTTACCCTGTTGAGATCCAGTTCGA < 2700  
P R H I A E L \* K C S S L E N V L R G E N S Q G S Y R C \* D P V R  
R A T \* Q N F K S A H H W K T F F G A K T L K D L T A V E I Q F D  
T A P H S R T L K V L I I G K R S S G R K L S R I L P L L R S S S M  
TGGCCGGTGTATGCTTTGAAATTTACAGAGTGAACCTTTTGAAGAAGCCCGCTTTTGAAGTTTCTAGAATGGCGACAACCTTAGGTCAAGCT  
2610 2620 2630 2640 2650 2660 2670 2680 2690

TGTAACCCACTCGTGCACCAACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGG < 2800  
C N P L V H P T D L Q H L L L S P A F L G E Q K Q E G K M P Q K R E  
V T H S C T Q L I F S I F Y F H Q R F W V S K N R K A K C R K K G  
\* P T R A P N \* S S A S F T F T S V S G \* A K T G R Q N A A K K G  
ACATTTGGGTGAGCAGCTGGGTTGACTAGAACTAGAAATGAAAAGTGGTTCGCAAGACCCAGTCTTTTGGTCTTCCGTTTTCGGGCTTTTTCGGGCTTTTTC  
2710 2720 2730 2740 2750 2760 2770 2780 2790

AATAAGGGCGACACGGAATGTTGAATACTCATACTCTCTCTTTTCAATATTATTGAAGCATTATTCAGGGTTATTGTCTCATGAGCGGATACATATTT < 2900  
\* G R H G N V E Y S S Y S S F F N I I E A F I R V I V S \* A D T Y L  
N K G D T E M L N T H T L P F S I L L K H L S G L L S H E R I H I \*  
I R A T R K C \* I L I L F L F Q Y Y \* S I Y Q G Y C L M S G Y I F

```

TTATCCCGCTGTGCCCTTACAACCTATGAGTATGAGAAGGAAAAAGTTATAATAACTTCGTAAATAGTCCCAATAACAGAGTACTCGCCTATGTATAAA
      2810      2820      2830      2840      2850      2860      2870      2880      2890

```

```

GAATGTATTTAGAAAAATAACAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCAC < 2961
N V F R K I N K * G F R A H F P E K C H
M Y L E K * T N R G S A H I S P K S A X
E C I * K N K Q I G V P R T F P R K V P X
CTTACATAAATCTTTTATTGTTTATCCCAAGGCGGTGTAAGGGGCTTTTCACGGTG
      2910      2920      2930      2940      2950      2960

```

```

Features :
T3          : [790 : 771 - CCW]
M13-fwd    : [599 : 616 - CW]
M13-rev    : [828 : 808 - CCW]
T7         : [626 : 645 - CW]
ColE1 origin : [1824 : 1196 - CCW]
F1 ori     : [13 : 453 - CW]
LacZ alpha : [528 : 460 - CCW]
LacO      : [856 : 834 - CCW]
AmpR     : [2635 : 1976 - CCW]

```

## V. File Format

Serial Cloner uses an extended DNA Strider file format. It is still readable by DNA Strider and by any software that respects the recommendations of the DNA Strider format. Strider will just ignore additional information, and in particular features and annotations. Of course, if one then saves the sequence using DNA Strider for example, Feature data will be lost.

2.1

Serial Cloner 2.1 can now export file sequence in a GenBank format that will contain the Annotations/Features. Special Serial\_Cloner fields that will, for example, contain the color associated with the Feature, will be present. These fields will of course be properly parsed by Serial Cloner.

## VI. Import of feature-enriched sequences

Serial Cloner 2.x imports features-containing sequences saved by *ApE* or *Vector NTI*. Feature colors should be correctly imported.

2.1

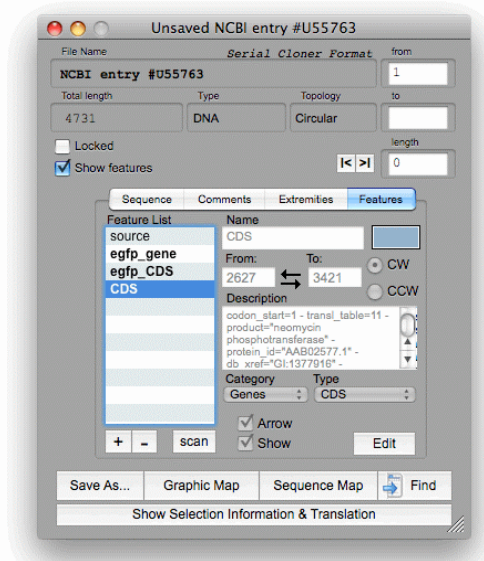
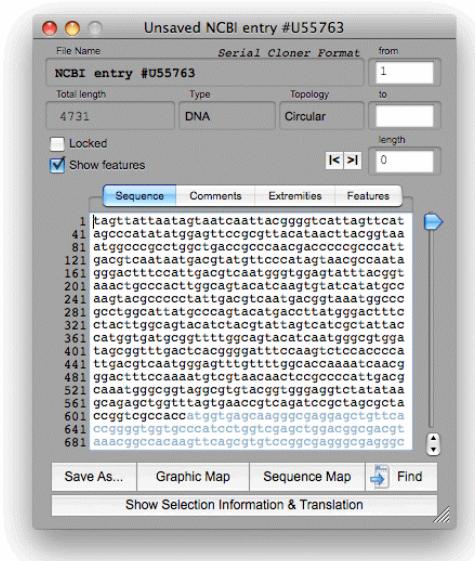
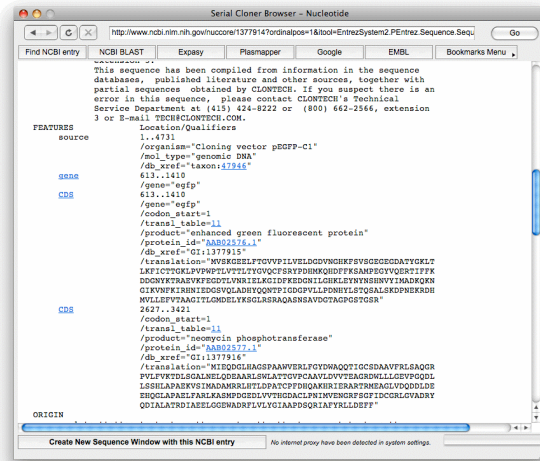
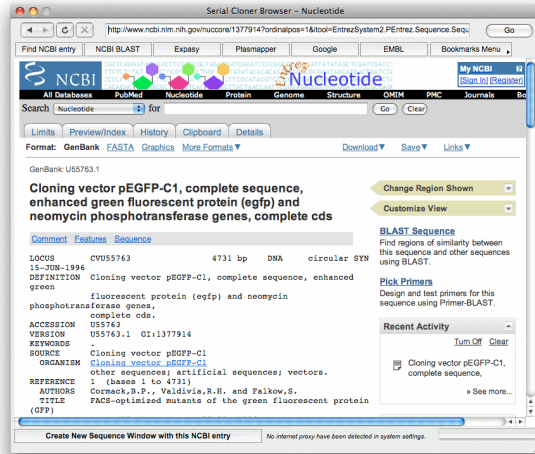
With Version 2.1 of Serial Cloner, I also provide with the MacOS version a companion Software created by Davide Cittaro that will transform a MacVector file to a GenBank formatted file (see the short description in the Companion Software\_Read-Me.pdf file). This Genbank format will be read by Serial Cloner.

Import of *Genbank* and *EMBL* have been improved and features should now be recognized and parsed. This is true for sequences saved in files or imported from the Genbank/EMBL web interface using Serial Cloner mini-Web Browser (using its bottom-left button). See the example bellow. [Note that other formats devoid of Features are still recognized by Serial Cloner : *Fasta*, *pDRAW32*, *raw text*].

Some Features are still ignored (like the join Feature) as no splicing tools are yet available in Serial Cloner.

Note that when Serial Cloner opens or paste a Sequence that was in an external format, a message will indicate that a particularly formatted sequence has been recognized. It is possible to prevent such a message from showing, which may be useful if you often open such external files (for example if you systematically open Vector NTI files).





The pEGFP-C1 has been obtained from the NCBI web interface using Serial Cloner mini-Web Browser. Because a GenBank formatted sequence has been detected by Serial Cloner, the button "Create a New Sequence Window..." at the bottom left was enabled. Clicking on it imported the Genbank entry in a new Sequence Window, correctly parsing the Features.

## VII. Copy/Paste

### A. Copying

There are several way to copy a sequence in Serial Cloner; some will keep Feature information.

- **Copy** (shortcut : **[C]-C**) : this is a basic Copy function. The selected portion of the sequence will be copied in the Clipboard and no Feature information will be kept.



Ex: copying between 12 and 1840 of pBluescript II SK (+):

```
CGCTTAATATTTTGTAAAATTTCGCGTTAAATTTTGTAAAATCAGCTCATTTTTAAACCAATAGGCCGAAATCGGCAAA
ATCCCTTATAAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAGAA
CGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACCCCTAATCAAGTT
TTTTGGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCTAAAGGGAGCCCCGATTAGAGCTTGACGGGAAAGCCG
GCGAACGTGGCGAGAAAGGAAGGAAAGCAAGGAGGAGCGGGCGTAGGGCGCTGGCAAGTGTAGCGGTACAGCTGCG
CGTAACCACCACACCCGCGCGCTTAATGCGCGCTACAGGGCGCTCCCATTCGCCATTACAGGCTGCGCAACTGTTGGG
AAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAAGTTGGGTA
ACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGCGCCAGTGAGCGCGCTAATACGACTCACTATAGGGCAATT
GGGTACCGGGCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGCAGCCCGGGGATCCACTAGTT
CTAGAGCGGCGCCACCGCGGTGGAGTCCAGCTTTTGTTCCTTTAGTGAGGGTTAATGCGCGCTTGGCGTAATCATG
GTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAG
CCGTTGGGTGCCATAAGTGAAGTAACTCACATTAATGCGTTCGCTCACTGCCCGCTTCCAGTCCGGAAACCTGTCCG
TGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGCGGTTTGCATTTGGGCGCTCTTCCGCTTCCGCTCAC
TGAATCGCTGCGCTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAA
CAGGGGATAACGAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCG
TTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT
ATAAAGATAACAGGCGTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCGACCCCTGCCGCTTACCGGATACCTGT
CCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTGCG
TCCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAA
CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTA
CAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT
ACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTTGTTTTGCAAGCA
CGAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCA
```

- *Copy as Fasta* (shortcut : C-K) : The selected portion of the sequence will be copied in the Clipboard in a Fasta format

Ex: copying between 12 and 1840 of pBluescript II SK (+):

```
> [pBluescript II SK(+) - 2961 bp] gi|58063|emb|X52328|ARBL2SKP pBluescript II SK(+)
vector DNA, phagemid excised from lambda ZAPII - Selection From [12] To [1840]
CGCTTAATATTTTGTAAAATTTCGCGTTAAATTTTGTAAAATCAGCTCATTTTTAAACCAATAGGCCGAAATCGGCAAA
ATCCCTTATAAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAGAA
CGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACCCCTAATCAAGTT
TTTTGGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCTAAAGGGAGCCCCGATTAGAGCTTGACGGGAAAGCCG
GCGAACGTGGCGAGAAAGGAAAGGAAAGCAAGGAGGAGCGGGCGTAGGGCGCTGGCAAGTGTAGCGGTACAGCTGCG
CGTAACCACCACACCCGCGCGCTTAATGCGCGCTACAGGGCGCTCCCATTCGCCATTACAGGCTGCGCAACTGTTGGG
AAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAAGTTGGGTA
ACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGCGCCAGTGAGCGCGCTAATACGACTCACTATAGGGCAATT
GGGTACCGGGCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGCAGCCCGGGGATCCACTAGTT
CTAGAGCGGCGCCACCGCGGTGGAGTCCAGCTTTTGTTCCTTTAGTGAGGGTTAATGCGCGCTTGGCGTAATCATG
GTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAG
CCGTTGGGTGCCATAAGTGAAGTAACTCACATTAATGCGTTCGCTCACTGCCCGCTTCCAGTCCGGAAACCTGTCCG
TGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGCGGTTTGCATTTGGGCGCTCTTCCGCTTCCGCTCAC
TGAATCGCTGCGCTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAA
CAGGGGATAACGAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCG
TTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT
ATAAAGATAACAGGCGTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCGACCCCTGCCGCTTACCGGATACCTGT
CCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTGCG
TCCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAA
CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTA
CAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT
ACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTTGTTTTGCAAGCA
CGAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCA
```

Note that the "*Copy all sequence as Fasta*" will copy all opened sequence in Fasta and join them in the Clipboard. This format is useful when multi-aligning sequences (form exemple using CLUSTAL).

2.1

- Copy Formatted (shortcut : [C]-[SHIFT]-K)

The full sequence will be copied as a Fasta-formatted sequence. Features will be shown in their colors and a legend will be added at the end of the sequence.

Ex: copying the sequence of pBluescript II SK (+):

```
> [pBluescript II SK(+) - 2961 bp] gi|58063|emb|X52328|ARBL2SKP pBluescript II SK(+)
vector DNA, phagemid excised from lambda ZAPII
CTAAATTGTAAGCGTTAATATTTTGTAAAAATTCGCGTTAAATTTTTGTAAAAATCAGCTCATTTTTTAACCAATAGGCCG
AAATCGGCAAAAATCCCTTATAAATCAAAAAGAAATAGACCGAGATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCA
CTATTAAGAAGCTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACC
CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAGCACTAAATCGGAACCCCTAAAGGGAGCCCCGATTAGAGCTTGAC
GGGAAAAGCCGGCGAACGTGGCGAGAAAAGGAAGGAAGAAAGCGAAAAGGAGCGGGCGTAGGGCGCTGGCAAGTGTAGCG
GTCACGCTGCGCGTAACCAACCACACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTCCATTCGCCATTCAGGCTGCG
CAACTGTTGGGAAGGGCGATCGGTGCGGGCCCTTCGCTATTACGCCAGCTGGCGAAAAGGGGATGTGCTGCAAGCCGAT
TAAGTTGGGTAACGCGAGGTTTCCAGTCACGACGTTGTAAAAACGACGGCCAGTGAGCGCGTAATACGACTCAATA
TAGGGCGAATTGGGTACCGGGCCCCCTCGAGGTGACGCGTATCGATAAGCTTGATATCGAATTCCTGCAGCCCGGGG
ATCCACTAGTTCTAGAGCGGCCACCAGCGGTGGAGTCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGCGCTTG
GCATAATCATGGTCATACTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTCACACAACATACGAGCCGGAAGCAT
AAAGTGTAAAGCCTGGGTAGCCATAATGAGTGAAGTAACTACATTAATGCGTTGCGCTCACGTCGCCCTTCCAGTCCG
GAAAACCTGTCGTGCCAGTGCATTAATGAATCGGCCAACGCGGGGAGAGGCGGTTTGCATTTGGGGCGCTTCCGCT
TCCTCGCTCACTGACTCGCTGCGCTCGTCTGCTGCGCTGCGCGGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGT
ATCCACAGAATCAGGGGATAACGAGGAAAGAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCG
CGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC
CCGACAGGACTATAAGATAACAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCGCTTAC
CGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGT
AGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACATATCGT
CTTGAGTCCAACCGGTAAGACACGACTTATCCGCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT
AGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACACGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGC
TGAAGCCAGTTACCTTCGGAAGAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTTTTT
GTTTCAAGCAGCAGATTCGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCAGACGCTCA
GTGGAAACGAAAACCTCACGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATAAA
AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCT
ATCTCAGCGATCTGCTATTTTCGTTTACCTATAGTTGCTGACTCCCGCTGTTAGATAACTACGATACGGGAGGGCTT
ACCATCTGGCCCCGAGTGTGCAATGATACCCGAGACCCAGCTCACCGGCTCCAGATTTATCAGCAATAAACCGCCAG
CCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATGTTGCGGGAAAGCTAGA
GTAAGTAGTTCGCCAGTTAATAGTTTGCAGCAAGTGGTGGCCAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTC
CTTCCGTCCTCCGATCGTTGTCAGAAAGTAAAGTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTC
CTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCATTTGAGAATAGTGTATGCG
GCACCGAGTTGCTTTCGCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTAAAAGTGTCTCATCATG
AAAAACGTTCTTCGGGGCGAAAACCTCAAGGATCTTACCCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCC
AACTGATCTTACGATCTTTTACTTTTACCAGCGTTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGG
AATAAGGGCGACACGAAATGTTGAATACTCATACTCTTCTTTTCAATATATTGAAGCATTTATCAGGGTTATTGTC
TCATGAGCGGATACATATTTGAATGATTTAGAAAAATAACCAATAGGGGTTCCGCGCACATTTCCCGGAAAAGTGCCA
C
```

```
Features :
T3 : [790 : 771 - CCW]
M13-fwd : [599 : 616 - CW]
M13-rev : [828 : 808 - CCW]
T7 : [626 : 645 - CW]
ColE1 origin : [1824 : 1196 - CCW]
Fl ori : [13 : 453 - CW]
LacZ alpha : [528 : 460 - CCW]
LacO : [856 : 834 - CCW]
AmpR : [2635 : 1976 - CCW]
```

- Copy as GenBank

2.1

The full sequence will be copied as a GenBank formatted sequence.

One interesting use of such a format is to paste it in a database (FileMaker, Excel, etc...) so that anybody in the lab can import the sequence with its features by copying the text back into Serial Cloner.

The format is exactly the same than the one saved using the *Export as Genbank* function.

Ex: copying the sequence of pBluescript II SK (+):

```

LOCUS      pBluescript_II_SK(+)      2961 bp      DNA      circular      02/05/10
DEFINITION
ACCESSION
VERSION
SOURCE
ORGANISM
COMMENT    Serial Cloner Genbank Format
COMMENT    SerialCloner_Type=DNA
COMMENT    SerialCloner_Comments=gi|58063|emb|X52328|ARBL2SKP pBluescript II SK(+)
vector DNA
COMMENT    , phagemid excised from lambda ZAPII
COMMENT    SerialCloner_Ends=0,0,,0,
FEATURES   Location/Qualifiers
    primer_bind   complement(771..790)
                    /label=T3
                    /SerialCloner_Color=&h008B8B
                    /SerialCloner_Show=True
                    /SerialCloner_Protect=False
                    /SerialCloner_Arrow=True
    primer_bind   599..616
                    /label=M13-fwd
                    /SerialCloner_Color=&h008B8B
                    /SerialCloner_Show=True
                    /SerialCloner_Protect=False
                    /SerialCloner_Arrow=True
    primer_bind   complement(808..828)
                    /label=M13-rev
                    /SerialCloner_Color=&h008B8B
                    /SerialCloner_Show=True
                    /SerialCloner_Protect=False
                    /SerialCloner_Arrow=True
    primer_bind   626..645
                    /label=T7
                    /SerialCloner_Color=&h008B8B
                    /SerialCloner_Show=True
                    /SerialCloner_Protect=False
                    /SerialCloner_Arrow=True
    rep_origin    complement(1196..1824)
                    /label=ColE1 origin
                    /SerialCloner_Color=&h4682B4
                    /SerialCloner_Show=True
                    /SerialCloner_Protect=False
                    /SerialCloner_Arrow=True
    rep_origin    13..453
                    /label=F1 ori
                    /SerialCloner_Color=&hFF8000
                    /SerialCloner_Show=True
                    /SerialCloner_Protect=True
                    /SerialCloner_Arrow=True
    CDS            complement(460..528)
                    /label=LacZ alpha
                    /SerialCloner_Color=&h6495ED
                    /SerialCloner_Show=True
                    /SerialCloner_Protect=False
                    /SerialCloner_Arrow=True
    misc_binding  complement(834..856)
                    /label=LacO
                    /SerialCloner_Color=&h6495ED
                    /SerialCloner_Show=True
                    /SerialCloner_Protect=False
                    /SerialCloner_Arrow=True
    CDS            complement(1976..2635)
                    /label=AmpR
                    /SerialCloner_Color=&h8FBC8F
                    /SerialCloner_Show=True
                    /SerialCloner_Protect=False
                    /SerialCloner_Arrow=True
ORIGIN
1 CTAATTTGTA AGCGTTAATA TTTTGTAA ATTCGCGTTA AATTTTGT AAATCAGCTC

```

```

61 ATTTTAAAC CAATAGGCCG AAATCGGCAA AATCCCTTAT AAATCAAAAG AATAGACCGA
121 GATAGGGTTG AGTGTGTGTC CAGTTTGGAA CAAGAGTCCA CTATTAAGA ACCTGACTC
181 CAACGTCAA GGGCGAAAAA CCGTCTATCA GGGCGATGGC CCACTACGTG AACCATCACC
241 CTAATCAAGT TTTTGGGGT CGAGGTGCCG TAAAGCACTA AATCGGAACC CTAAGGGGAG
301 CCCCCGATTT AGAGCTTGAC GGGGAAAGCC GGCAGACGTG GCGAGAAAGG AAGGGAAGAA
361 AGCGAAAGGA GCGGGCGCTA GGGCGCTGGC AAGTGTAGCG GTCACGCTGC GCGTAACCAC
421 CACACCCGCC GCGCTTAATG CGCCGCTACA GGGCGCGTCC CATTGCGCAT TCAGGCTCGC
481 CAACTGTTGG GAAGGGCGAT CCGTGCAGGC CTCTTCGCTA TTACGCCAGC TGGCGAAAGG
541 GGGATGTGCT GCAAGGCGAT TAAGTTGGGT AACGCCAGGG TTTTCCCAGT CACGACGTTG
601 TAAAACGACG GCCAGTGAGC GCGCGTAATA CGACTACTA TAGGGCGAAT TGGGTACCGG
661 GCCCCCCTC GAGGTCGACG GTATCGATAA GCTTGATATC GAATTCCTGC AGCCCGGGGG
721 ATCCACTAGT TCTAGAGCGG CCGCCACCGC GGTGGAGCTC CAGCTTTTGT TCCCTTTAGT
781 GAGGGTTAAT TGCGCGCTTG GCGTAATCAT GGTCTATAGT GTTTCCTGTG TGAATTTGTT
841 ATCCGCTCAC AATTCACAC AACATACGAG CCGGAAGCAT AAAGTGTAAG GCCTGGGGTG
901 CCTAATGAGT GAGCTAATC ACATTAATTTG CGTTGCGCTC ACTGCCCGCT TTCCAGTCGG
961 GAAACCTGTC GTGCCAGCTG CATTAAATGAA TCGGCCAAGC CGCGGGGAGA GCGGGTTTGC
1021 GTATTGGGCG CTCTTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC
1081 GGCAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT ATCCACAGAA TCAGGGGATA
1141 ACGCAGGAAA GAACATGTGA GCAAAAAGCC AGCAAAAAGC CAGGAACCGT AAAAAGGCCG
1201 CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC CCCCAGACGA GCATCACAAA AATCGACGCT
1261 CAAGTCAGAG GTGGCGAAAC CCGACAGGAC TATAAAGATA CCAGGCGTTT CCCCTGGAA
1321 GCTCCCTCGT GCGCTCTCCT GTTCCGACCC TGCCGCTTAC CGGATACCTG TCCGCTTTC
1381 TCCCTTCGGG AAGCGTGGCG CTTTCTCATA GCTCAGCTG TAGGTATCTC AGTTCGGTGT
1441 AGTTCGTTCG CTCCAAGCTG GGCTGTGTGC ACGAACCCCG CGTTCAGCCC GACCGCTCGC
1501 CCTATCCGGT TAACTATCGT CTTGAGTCCA ACCCGGTAAG ACACGACTTA TCGCCACTGG
1561 CAGCAGCCAG TGGTAACAGG ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTCTT
1621 TGAAGTGGTG GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TCGCTCTGTC
1681 TGAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA CAAACCACCG
1741 CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC GCGCAGAAAA AAAGGATCTC
1801 AAGAAGATCC TTTGATCTTT TCTACGGGGT CTGACGCTCA GTGGAACGAA AACTCACGTT
1861 AAGGGATTTT GGTCAATGAGA TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAATTTAAA
1921 AATGAAGTPT TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT
1981 GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCGTTCATCC ATAGTTGCCT
2041 GACTCCCGT CGTGTAGATA ACTACGATAA GGGAGGGCTT ACCATCTGGC CCCAGTGTG
2101 CAATGATACC GCGAGACCCA CGCTCACCGG CTCCAGATTT ATCAGCAATA AACGAGCCAG
2161 CCGGAAGGGC CGAGCGCAGA AGTGGTCCCTG CAACTTTATC CGCCTCCATC CAGTCTATTA
2221 ATTGTTGCGG GGAAGCTAGA GTAAGTAGTT CGCCAGTTAA TAGTTTGGC AACTGTTGTTG
2281 CCATTGCTAC AGGCATCGTG GTGTACGCTC CGTCGTTTGG TATGGCTTCA TTCAGCTCCG
2341 GTTCCCAACG ATCAAGGCGA GTTACATGAT CCCCATGTT GTGCAAAAAA GCGGTTAGCT
2401 CCTTCGGTCC TCCGATCGTT GTCAGAAAGTA AGTTGGCCGC AGTGTATCA CTCATGGTTA
2461 TGTACGCACT GCATAATCTT CTTACTGTCA TGCCATCCGT AAGATGCTTT TCTGTGACTG
2521 GTGAGTACTC AACCAAGTCA TTCTGAGAAT AGTGTATGCG GCGACCGAGT TGCTCTTGCC
2581 CGCGTCAAT ACGGGATAAT ACCGCGCCAC ATAGCAGAAC TTTAAAAGTG CTCATCATTG
2641 GAAAACGTTT TTCGGGGCGA AAACCTCTCA GGATCTTACC GCTGTTGAGA TCCAGTTCGA
2701 TGTAAACCCAC TCGTGACCC AACTGATCTT CAGCATCTTT TACTTTCACC AGCGTTTCTG
2761 GGTGAGCAAA AACAGGAAG CAAAATGCCG CAAAAAAGGG AATAAGGGCG ACACGGAAT
2821 GTTGAATACT CATACTCTTC CTTTTCAAT ATTATTGAAG CATTATCAG GGTATTGTC
2881 TCATGAGCGG ATACATATTT GAATGTATTT AGAAAAATAA ACAAATAGGG GTTCCGCGCA
2941 CATTTCCCG AAAAGTGCCA C

```

//

## B. Pasting



When pasting a sequence in a Sequence Window, Serial Cloner tries to guess the format (Fasta, Generic GenBank, Serial Cloner GenBank) and to parse the data, including the Features present in a GenBank format. Note that it is even possible to copy a full web page that contains a GenBank entry (from Safari or Firefox for example) and to paste it directly in a Sequence Field. It should be properly parsed (Sequence, comments, Features).

As indicated above, copying and pasting sequences in Genbank Serial Cloner Format is also a good way to store and share annotated sequence in a text format that can be imported later in Serial Cloner by any user.

If Serial Cloner cannot detect a particular format, then it will parse it as a raw text filtering out non-ATGC letter (in a DNA window) or non-degenerate DNA letters (in a Degenerate DNA window).

## VIII. Adding Features to a Sequence

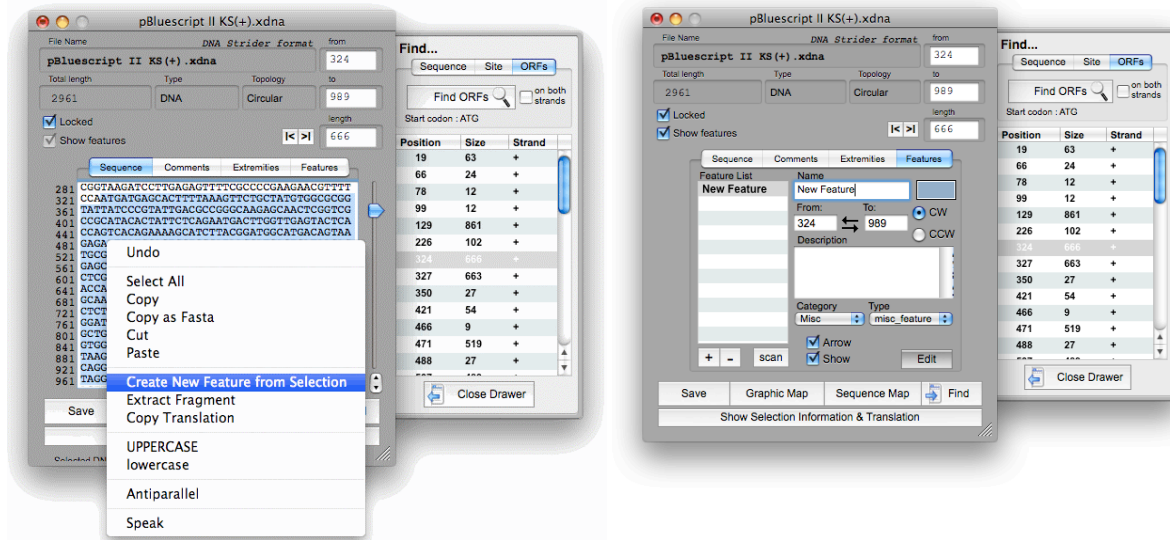
Several ways exist to add new Features to a sequence.

### A. Manual entry

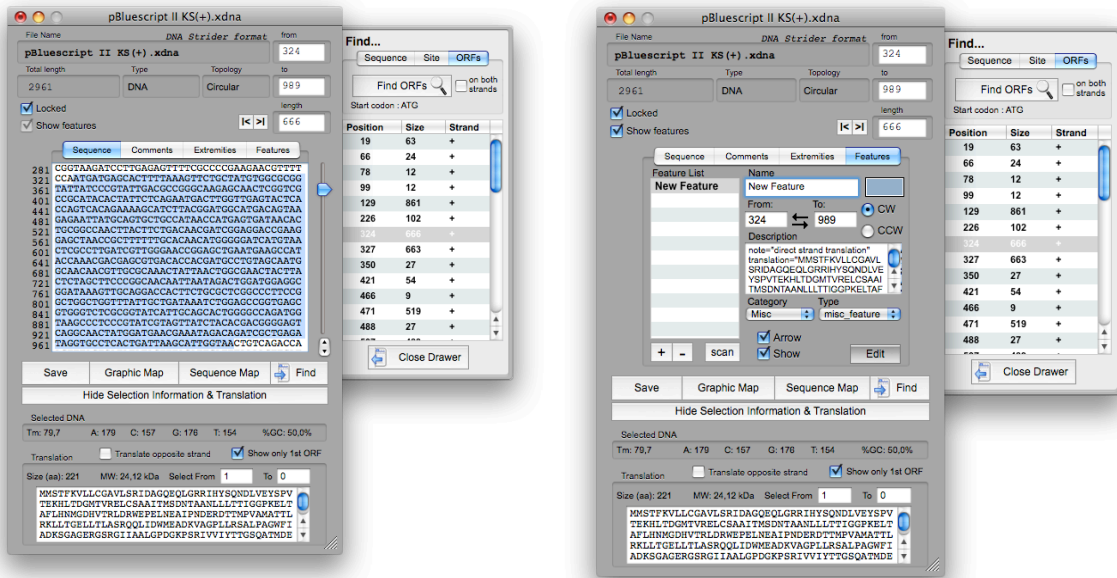
Simply go in the Feature section (*Features* Tab) and press the '+' button. Then, enter numerically the coordinates of the feature and set the different values. If a portion of sequence has been selected in the Sequence Field before pressing the '+' button, then the coordinates of the current selection will be automatically proposed as default values for the features (and in some cases the translation pasted as indicated bellow).

### B. After selecting a nucleotide fragment in the sequence window.

First select a portion of sequence in the sequence window either directly or using the ORF search for example. Then either display a contextual menu (Right-Click or [CTRL]-Click) and select '*Create New Feature From Selection*'. Alternatively, use the '*New Feature From Selection*' (shortcut : [C]-@) from the '*Features*' menu.

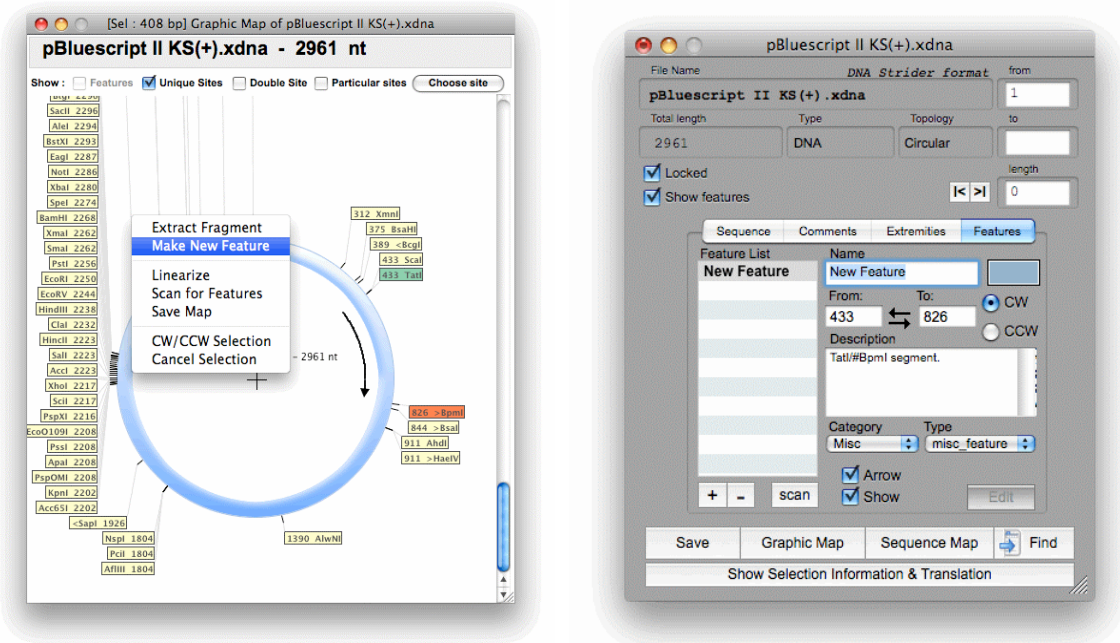


Note that if the '*Selection information and Translation*' was shown before creating the Feature, then the current translation will be added in the features comments (see screenshots bellow).



**C. After selecting a nucleotide fragment in the Graphic Window.**

First, select a fragment by choosing an upstream (simple click) and downstream (double-click or [SHIFT]-Click) restriction site. Clockwise and Counter-Clockwise can be toggled using the spacebar. Then display a contextual menu by right-clicking (or [CTRL]-Clicking) on the map and select 'Make New Feature'. Alternatively, use the 'New Feature From Selection' (shortcut : [C]-@) from the 'Features' menu. The feature will be created between these two sites and the names of the enzymes added in the comment field. The Feature will contain the restriction sites used at each end.



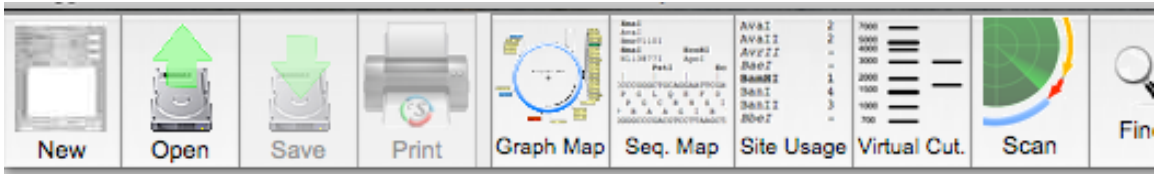


### D. By automatic scanning

Serial Cloner now includes a database that allows to automatically detect the presence of frequent sequences like the  $\beta$ -lactamase, the ColE1 or the f1 sequence. By default, Serial Cloner comes with tens of pre-defined Features (see the list bellow). A new interface is provided in Serial Cloner 2.1 that allows the addition of user-defined sequences.

To use the database and detects the presence of these domains in your sequences, simply use the *Scan* button present in the *Toolbar* or the use the '*Scan Sequence*' command ([ALT]-[C]-S, in the '*Features*' menu). This can be done either while displaying the *Sequence Window*, the *Graphic Map*. In the latter case, the associated *Sequence Window* must not have been closed. It is also allowed to scan in the *Build a Construct...* window. Since V2.1, it is also possible to scan directly in the *Sequence Map* window.

2.1



It is also possible to scan using the '*Scan*' button present in the '*Features*' tab of the *Sequence Window*. By default, all are '*Shown*' but the '*Show*' checkbox can be used to alter this state.

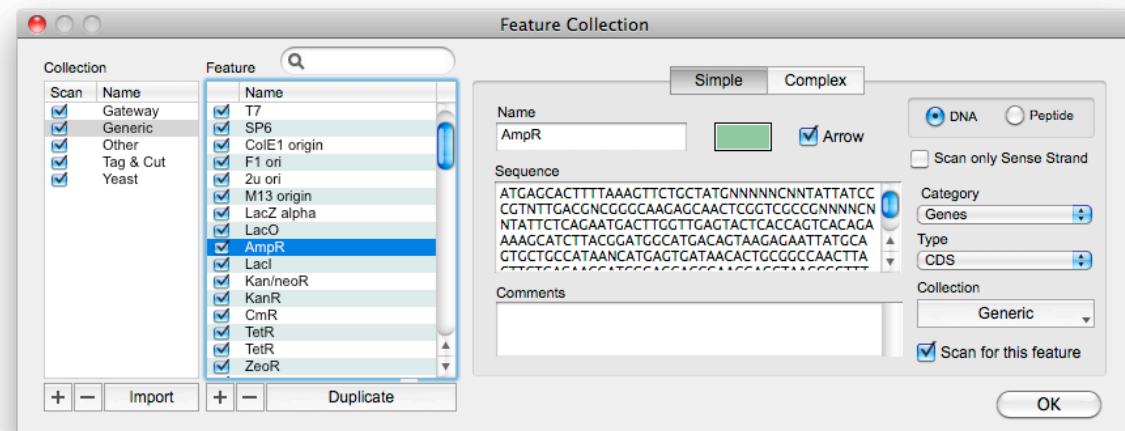
Note that if you scan more than once the same sequence, the same Features will of course be found and the new Features will replace the identical new ones. BUT if you change something on an automatically found annotation (like its name, color, etc.), it will not anymore be automatically erased upon a new scan (to protect your modifications). A new scan will thus result in a duplication of this Feature and the unwanted one will have to be deleted manually.

2.1

## IX. Add/Delete/Modify Features to be scanned for

### A. Manage Feature interface.

The new Feature menu proposes a '*Manage Features...*' command. It will open a new window that will allow to add/delete/modify Features and Feature collections. It can be used for 'Simple' definition or 'Complex' definition of Features.



### a. Feature Collections

On the left are listed the different Collections. The checkbox indicates whether this Collection should be used when scanning or not. As scan time depends on the number of Features to scan for, it is a good idea to unselect a Collection that does not correspond to the work you routinely carry-out if scanning time starts to bother you.

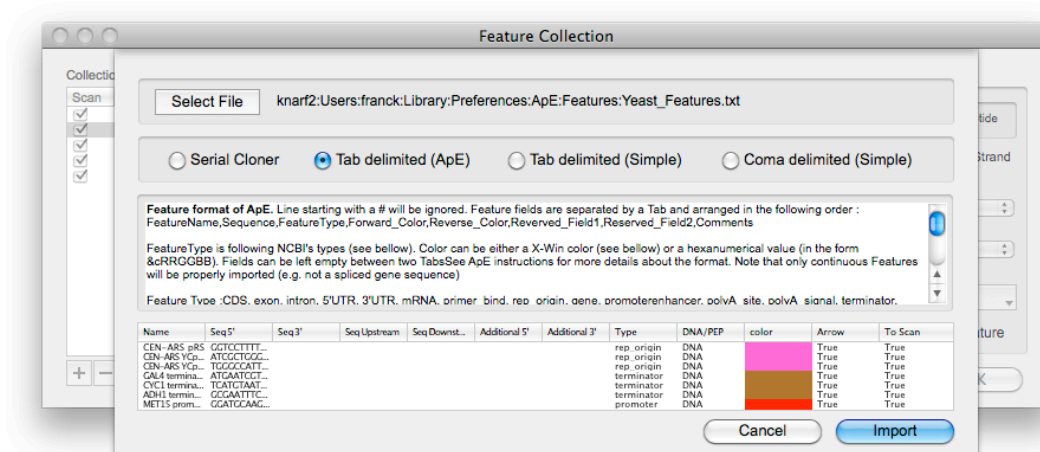
Note that you can also delete the Collection but this will not be undoable so it may be safer to simply unselect it. Also, even in Collection you want to use for scanning, some Features will never be meaningful for your work. What you can do is to create a new Collection (named for example 'Ignored Features') and store the Features you currently never use in this Collection. To change the Collection to which a Feature belong, use the button under 'Collection' (on the right) and select another one.

Following the same track, it is possible to have a Feature present in two different Collections. Simply Duplicate the Feature (button 'Duplicate') and change the Collection associated to the duplicated one. You can thus create Collections of Features that you only scan in certain circumstances (for example if you want, sometimes, to indicate all the Cysteines in a sequence or all the 'ATG', or the NcoI and NotI sites, etc...).

### b. Import a Feature Collection

You can also import Feature Collections either saved by Serial Cloner, by other software, or simply organized as a table. External Collections can be either in an ApE format or generic. The format is described in the Import Window.

If you want to send your Feature Collections to a colleague, you can use the Export command (see bellow). Note also that the Feature Collections are stored in ~/Library/Application Support/SerialCloner/Features/ (MacOSX) or Users/~/.AppData/Roaming/SerialCloner/Features/ (Windows 7) or \Documents and Settings\Admin\Application Data\SerialCloner\Features (Windows XP) ['Admin' may be replaced by a username in certain system config).

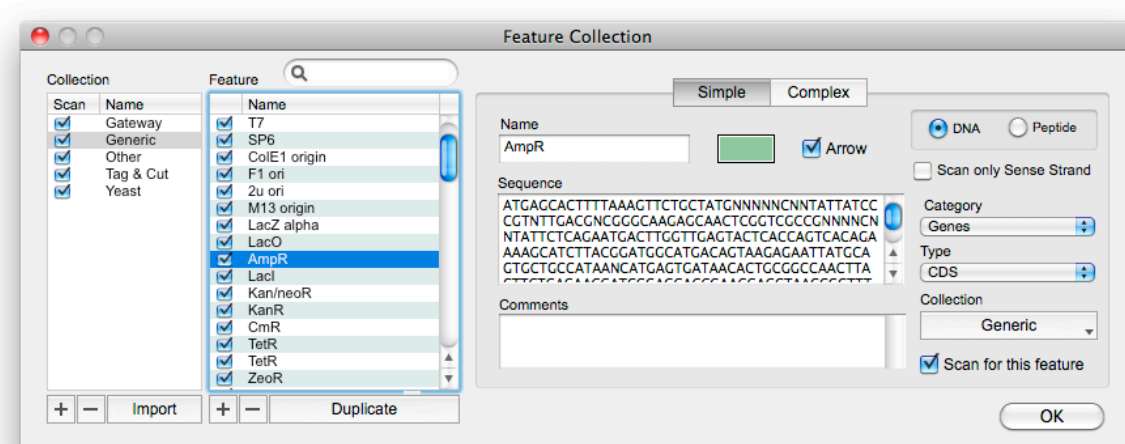


### c. Export a Feature Collection

It is possible to export the currently selected Feature Collection and save it independently. This may be useful to share a special Feature Collection with your colleagues. You can also use this export function to send me a Collection of Feature that you want to share with the community. I will put them for download in the Serial Cloner Web site.

To Export, press the [ALT] key. The 'Import' button will be replaced by an '*Export*' button. The Feature Collection will be saved in the Serial Cloner format.

### d. Feature List



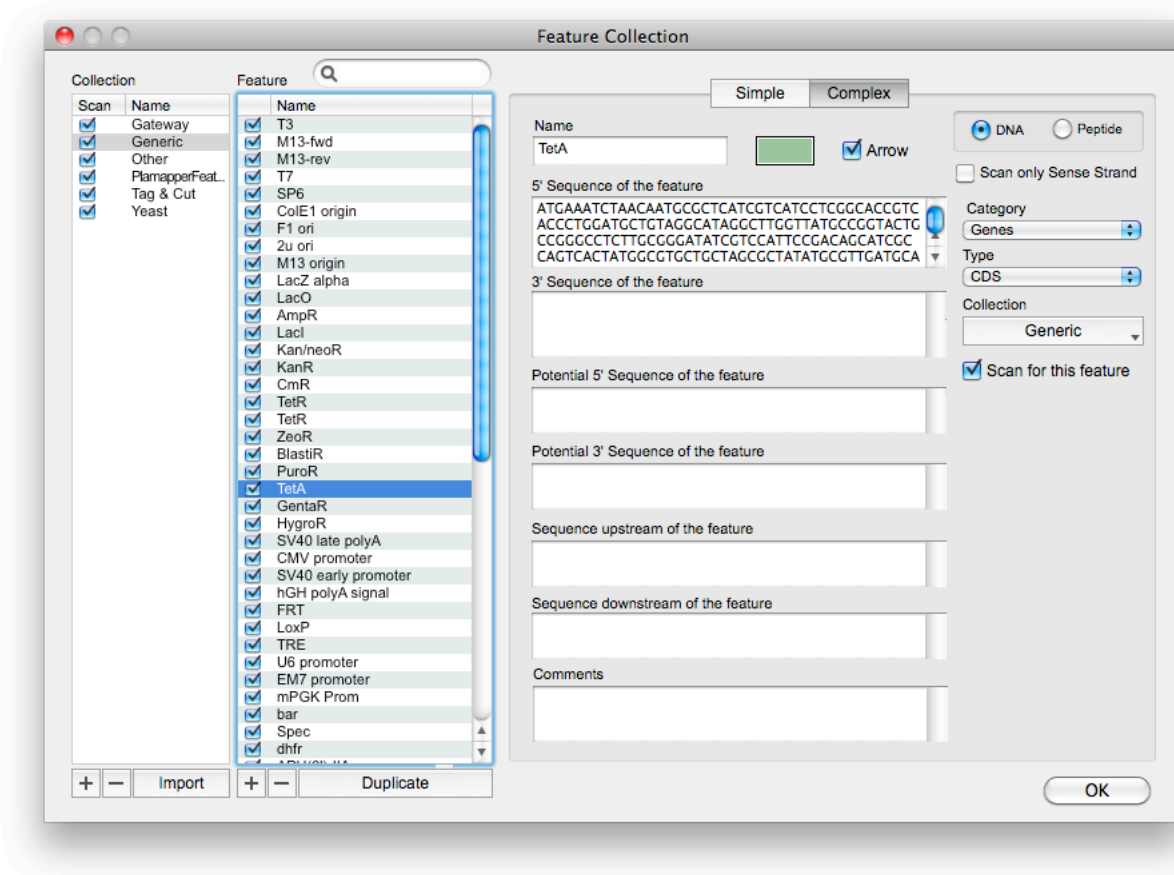
The second list displays the Features contained in the currently selected Collection. It is possible to search for a Feature inside this collection using the search field. As for the Feature Collection before, the Checkbox indicates whether the Feature should be used when scanning. Buttons under the list allows addition or deletion of Features. Be careful, there is no undo. It may be safer just to uncheck the Feature in general if you do not want to scan for it. It is also possible to duplicate a Feature. This can be useful when generating a series of very similar Features or when one Features is to be stored in different collections that will not be scanned at the same time.

### e. Feature definition

The right side of the window allows users to enter Feature details. Name fields are explicit and should be easy to understand in the '*Simple*' definition mode. It is possible to change the collection the Feature is attached to here.

Note also that Features can be entered in *DNA* or *Protein* mode. In the former mode, degenerate nucleotide alphabet can be used. The *Protein* mode is particularly interesting for tags and protease cleavage sites since different codon bias may have been used in different constructs. Look at the 'Tag & Cut' Collection for examples. Note also that it is also to impose scanning only in the sense strand.

The '*Complex*' mode may require more explanations.



- *5' Sequence of the Feature* : if used alone, it is the exact equivalent of the Sequence of the Feature used in the 'Simple' definition mode. When used with the 3' field mode (see bellow), it will determine the consensus of the 5' end of the sequence
- *3' Sequence of the Feature* : Indicates the sequence of the 3' end of the Feature. Any intervening sequence present between the 5' and the 3' sequence of the Feature will be considered as part of the Feature. Serial Cloner will look for :  $[5' \text{ Sequence } ]\text{-NNNNN-}[3' \text{ Sequence } ]$
- *Potential 5' and 3' sequences* : these sequences may or may not be present in the sequence to be scanned. If present, they will be incorporated to the detected Feature upon scanning. Good examples are the START and STOP codons of the EGFP that are not always present (depending of Nter or Cter fusion) but that belong to the Feature when present.
- *Sequence Upstream and Downstream of the Feature* : Indicates sequences that must be present for the Feature to be scored as 'detected' when scanning, but that do not belong to the Feature itself. A good example is a particular amino-acid that should be highlighted if present in a certain context but not always (like the cystein in the CVLL sequence). Following this track, we plan to allow users to enter complex consensus sequences in the future (like  $[S/T]X[RK]$ ).

## **Serial Cloner 2.1 : Internal database of Features**

The following features will be looked for upon scanning. Note that proper detection depends on correct matching to the consensus sequence stored in Serial Cloner. Please, do not hesitate to signal any lack of detection so that I can update the consensus sequence for this Feature. You can also propose to add your preferred Feature in the database (submitting by email a consensus sequence or a Collection). If I received enough of them, I will put them on my Web site for user to download it or incorporate them in a future version of Serial Cloner.

2.1

### **Internal Database :**

#### **A. Generic:**

- T3
- M13-fwd
- M13-rev
- T7
- SP6
- ColE1 origin
- F1 ori
- 2u ori
- M13 origin
- LacZ alpha
- LacO
- AmpR
- LacI
- Kan/neoR
- KanR
- CmR
- TetR
- TetR
- ZeoR
- BlastIR
- PuroR
- TetA
- GentaR
- HygroR
- SV40 late polyA
- CMV promoter
- SV40 early promoter
- hGH polyA signal
- FRT
- LoxP
- TRE
- U6 promoter
- EM7 promoter
- mPGK Prom
- bar

- Spec
- dhfr
- APH(3')-IIA
- aadA
- tsr
- pat
- aph
- oriVR6Kgamma
- oriS
- oriRg
- R6K ori
- 35S CMV
- PBAD
- AML-P
- NOS-P
- TrpC
- CaMV 35S
- mPGK Prom
- EGFP
- GST
- mRFP1
- mCherry
- EYFP
- ECFP
- hRluc1
- Rluc
- AcGFP
- IRES [1]
- IRES [2]
- Amp Promoter
- Lac Promoter
- HIV-1 5LTR
- HIV-1 psi
- HIV RRE
- Delta U3
- SV40 intron
- Synthetic intron
- SV40 ORI
- lacI
- tac promoter
- rrnB
- rrnB T1
- rrnBT2
- bGH PA

## **B. Yeast**



- KanMX
- URA3
- TRP1
- MET15
- LYS2
- LEU2
- HIS3
- ADE2
- 2 micron origin
- CEN-ARS pRS
- CEN-ARS YCp-lac
- CEN-ARS YCp50
- GAL4 terminator
- CYC1 terminator
- ADH1 terminator
- MET15 promoter
- GAL1-10 promoter
- GAL1 promoter
- ADH1 promoter
- Gal4 activation domain
- LexA DNA-binding domain
- Gal4 DNA-binding domain
- HIS1
- PDR5

### **C. Other**

- IL2 ss
- human IgG1 Fc
- human IgG2 Fc
- human IgG3 Fc
- human IgG4 Fc
- mouse IgG2A Fc
- mouse IgG2B Fc
- mouse IgG3 Fc
- rabbit IgG Fc
- rat IgG2B Fc

### **D. Tags & Cut**

- 6His
- Cys
- C-Myc tag
- HA tag
- VSV-G tag
- HSV tag
- Flag tag
- T7 tag

- GluGlu tag
- E- tag
- S-tag
- PC tag
- Strp tag II
- SBP tag
- Avitag
- TEV tag [1]
- TEV tag [2]
- PreScission site
- Factor Xa site
- Thrombin site
- SV40 NLS1
- SV40 NLS2

#### **E. Gateway**

- attB4
- attB3
- AttB2
- AttB1
- attP4
- attP3
- AttP2
- AttP1
- attL4
- attL3
- AttL2
- AttL1
- attR4
- attR3
- AttR2
- AttR1
- ccdB