

# Artemis

## Learning objectives:

- Loading files into Artemis
- Search options in Artemis
- Selecting and extracting sequences and features in Artemis
- Optional exercise: Structural and functional annotation in Artemis

## Introduction

Artemis is a free DNA viewer and annotation tool written by Kim Rutherford (Rutherford *et al.*, 2000). It is routinely used by the Parasite Genomics Group at the Wellcome Sanger Institute for annotation and analysis of both prokaryotic and eukaryotic genomes. The program allows the user to view simple sequence files, EMBL/Genbank entries and the results of sequence analyses in a highly interactive and intuitive graphical format. Artemis is designed to present multiple sets/types of information within a single context. This manifests itself as the ability to zoom in to inspect DNA sequence motifs and zoom out to view local gene architecture, several kilobases of a genome or even an entire genome in one screen. It is also possible to perform some analyses within Artemis with the output stored for later access.

## Aims

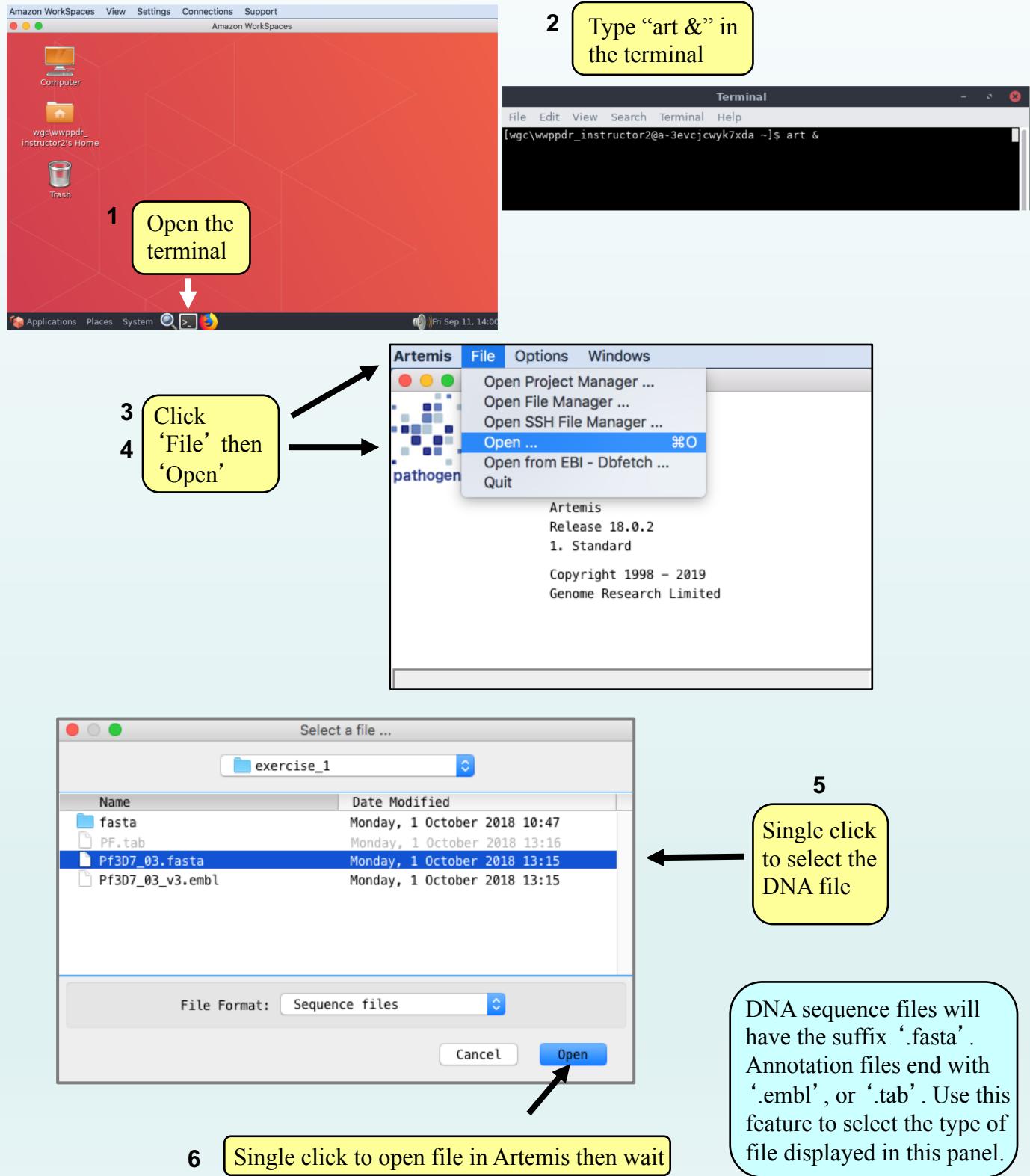
The aim of this Module is for you to become familiar with the basic functions of Artemis using a series of worked examples. These examples are designed to take you through the most immediately useful functions. However, there will be time, and encouragement, for you to explore other menus; nooks and crannies of Artemis that are not featured in the exercises in this manual. Like all the Modules in this workshop, the key is ‘if you don’t understand please ask’.

# Artemis Exercise 1 Part I

## 1. Starting up the Artemis software

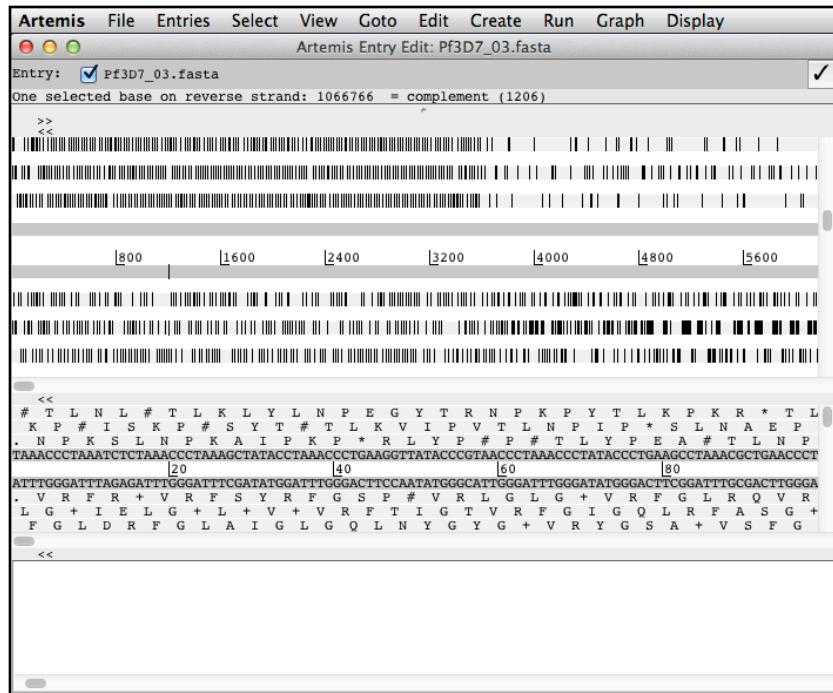
Open the terminal on workspaces and type “art” then hit return.

Navigate to the directory Module\_1\_Artemis, exercise\_1 containing the file Pf3D7\_03.fasta.



## 2. Loading annotation files (entries) into Artemis

Hopefully you will now have an Artemis window like this! If not, ask a demonstrator for assistance.



Now follow the numbers to load up the annotation file for *Plasmodium falciparum* 3D7 chromosome 3.

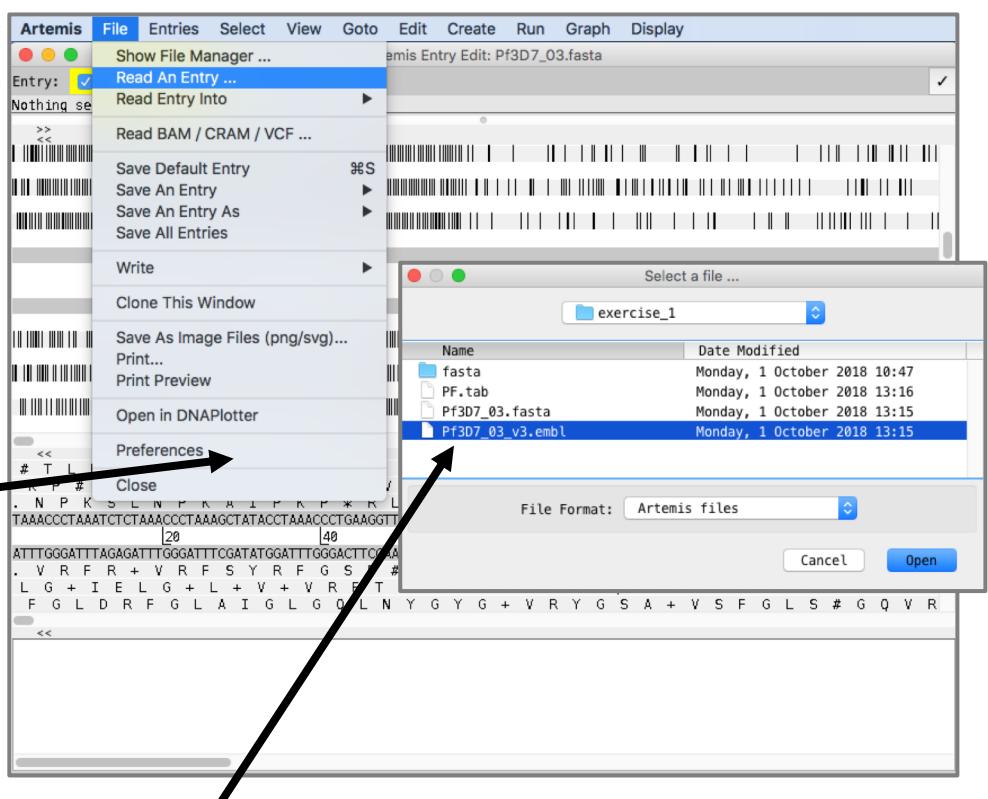
1

Click 'File' then  
'Read an Entry'

Entry = file

2

Single click to  
select  
Pf3D7\_03.embl file

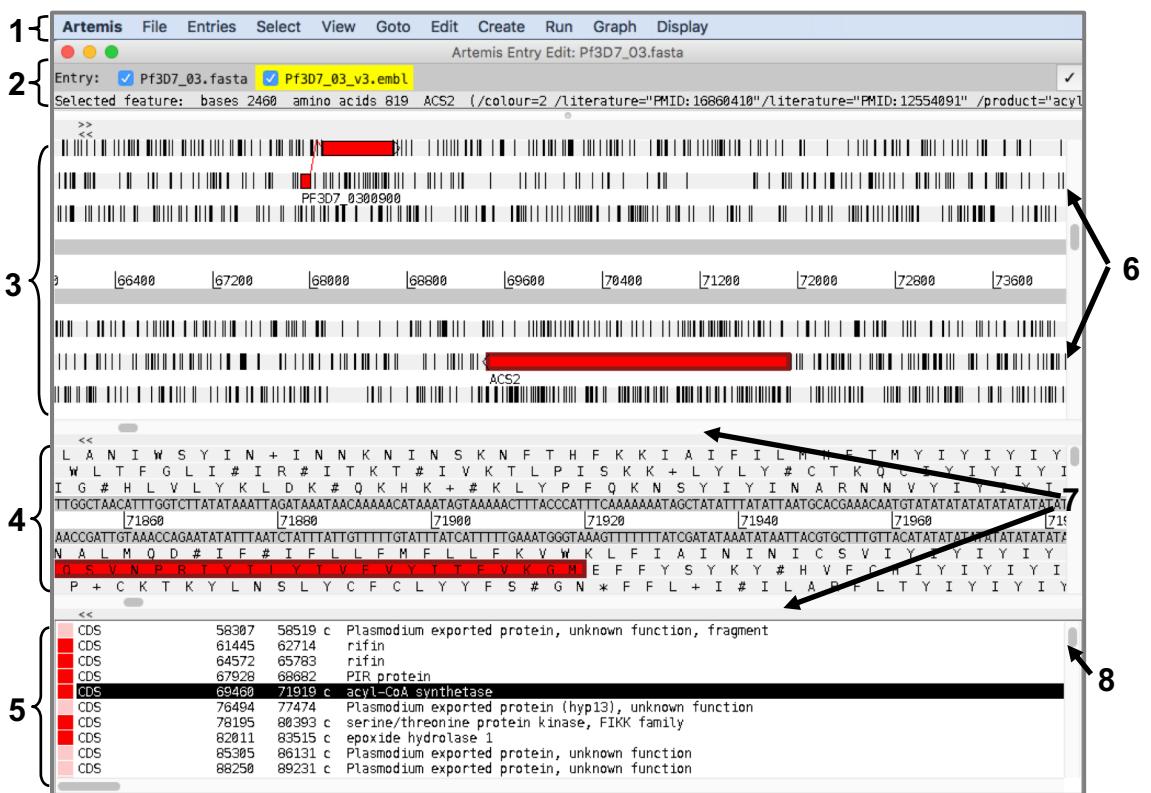


3 Single click to open file in Artemis then wait

What's an "Entry"? It's a file of DNA and/or features which can be overlaid onto the sequence information displayed in the main Artemis view panel.

### 3. The basics of Artemis

Now you have an Artemis window open let's look at what's in there.



1. Drop-down menus. There's lots in there so don't worry about them right now.
2. Shows what entries are currently loaded (bottom line) and gives details regarding the feature selected in the window below; in this case an acyl-CoA synthetase (selected line).
3. This is the main sequence view panel. The central 2 grey lines represent the forward (top) and reverse (bottom) DNA strands. Above and below those are the 3 forward and 3 reverse reading frames. Stop codons are marked as black vertical bars. Genes and other features (eg. Pfam matches) are displayed as coloured boxes. We will refer to genes as coding sequences or CDSs from now on.
4. This panel has a similar layout to the main panel but is zoomed in to show nucleotides and amino acids. Double click on a gene in the main view to see the zoomed view of the start of that gene. Note that both this and the main panel can be scrolled left and right (7, below) zoomed in and out (6, below).
5. This panel lists the various features in the order that they occur on the DNA with the selected gene highlighted. The list can be scrolled (8, below).
6. Sliders for zooming view panels.
7. Sliders for scrolling along the DNA.
8. Slider for scrolling feature list.

## 4. Getting around in Artemis

The 3 main ways of getting to where you want to be in Artemis are the ‘Goto’ drop-down menu, the Navigator and the Feature Selector. The best method depends on what you’re trying to do and knowing which one to use comes with practice.

### 4.1 The ‘Goto’ menu

The functions on this menu (ignore the Navigator for now) are shortcuts for getting to locations within a selected feature or for jumping to the start or end of the DNA sequence. Most are self-explanatory, so feel free to try any of them.



It may seem that ‘Goto’ ‘Start of Selection’ and ‘Goto’ ‘Feature Start’ do the same thing. Well they do if you have a feature selected but ‘Goto’ ‘Start of Selection’ will also work for a region which you have highlighted by click-dragging in the main window. So yes, give it a try! This is a very commonly used feature, so it is worth memorizing the keyboard shortcuts for these, **ctrl<left arrow>** and **ctrl <right arrow>** respectively.

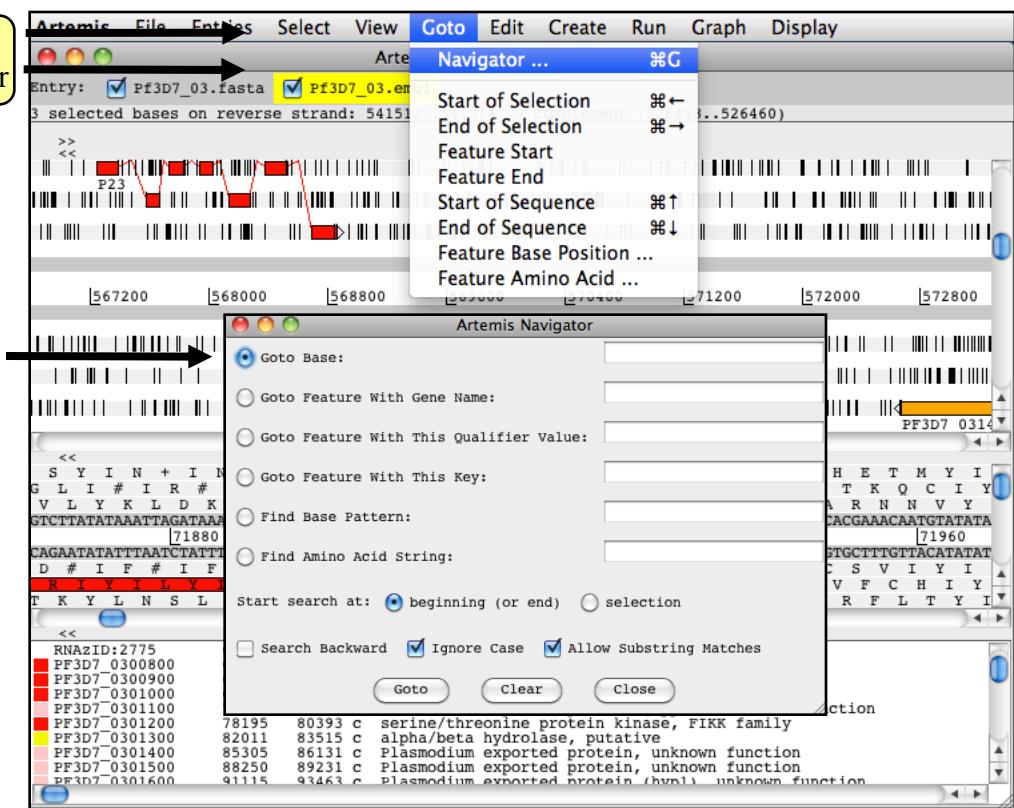
#### Suggested tasks:

1. Zoom out, highlight a large region of sequence by clicking the left hand button and dragging the cursor, then go to the start and end of the highlighted region.
2. Select a gene then go to the start and end.
3. Go to the start and end of the genome sequence.
4. Select a gene. Within it, go to a base (nucleotide) and/or amino acid of your choice.

## 4.2 Navigator

The Navigator panel is fairly intuitive so open it up and give it a try.

Click 'Goto' then Navigator



Suggestions of where to go:

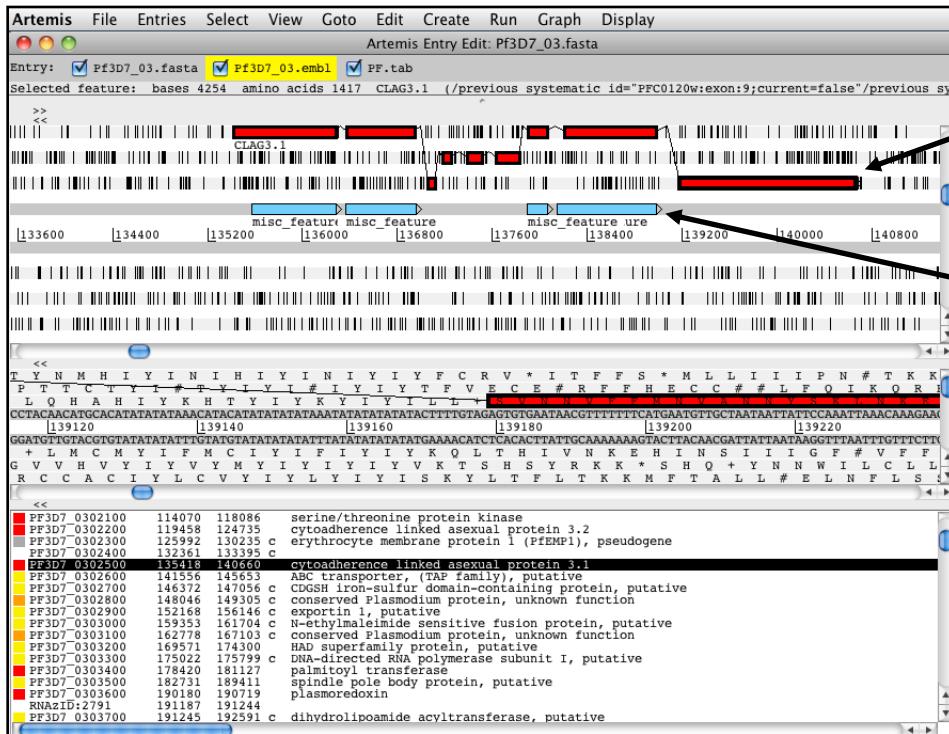
1. Think of a number between 1 and 1067971 and go to that base (notice how the cursors on the horizontal sliders move with you).
2. Your favourite gene name (it may not be there so you could try 'VAR').
3. Use 'Goto Feature With This Qualifier value' to search the contents of all qualifiers for a particular term. For example using the word 'pseudogene' will take you to the next feature with the word 'pseudogene' in any of its qualifiers. Note how repeated clicking of the 'Goto' button takes you through the pseudogenes as they occur on the chromosome.
4. tRNA genes. Type 'tRNA' in the 'Goto Feature With This Key' .
5. Amino acid consensus sequences (real or made up!). You can use 'X' s. Note that it searches all six reading frames regardless of whether the amino acids are encoded or not.

What are Keys and Qualifiers? See **Appendix IV**

Clearly there are many more features in Artemis which we will not have time to explain in detail. Before getting on with this next section it might be worth browsing the menus. Hopefully you will find most of them easy to understand.

## Artemis Exercise 1 Part II

This part of the exercise uses the files and data you already have loaded into Artemis from Part I. By a method of your choice go to the region located between bases 134000 to 141000 on the DNA sequence. This region encodes the *CLAG3.1* gene which codes for cytoadherence linked asexual protein. You can use either the Navigator, Feature Selector or Goto functions discussed previously to get there. The region you arrive at should look similar to that shown below.



Once you have found this region have a look at some of the information that is available to you:

Information to view:

### **Annotation**

If you click on a particular feature you can view the annotation attached to it: select a CDS feature (or any other feature) and click on the ‘Edit’ menu and select ‘Selected Feature in Editor’, or simply push ‘E’. A window will appear containing all the annotation that is associated with that CDS.

### **Viewing amino acid or protein sequence**

Click on the view menu and you will see various options for viewing the bases or amino acids of the feature you have selected, in two formats i.e. EMBL or FASTA. This can be very useful when using other programs that are not integrated into Artemis e.g. those available on the Web that require you to cut and paste sequence into them.

### **Plots/Graphs**

Feature plots can be displayed by selecting a CDS feature then clicking ‘View’ and ‘Feature Plots’. The window which appears shows plots predicting hydrophobicity, hydrophilicity and coiled-coil regions for the protein product of the selected CDS.

### **Load additional files**

The results from the Pfam protein motif searches are not shown, but can be viewed by loading the appropriate file. Click on ‘File’ then ‘Read an Entry’ and select the file PF.tab. Each Pfam match will appear as a coloured blue feature in the main display panel on the grey DNA lines. To see the details click the feature then click ‘View’ then ‘Selection’ or click ‘Edit’ then ‘Selected Features in Editor’. You can also run Pfam by going to the Run menu and selecting ‘Pfam search’. For this you need to select one CDS.

### **Viewing the results of database searches**

Click the ‘View’ menu, then select ‘Search Results’ and then ‘Fasta results’. The results of the database search will appear in a scrollable window.

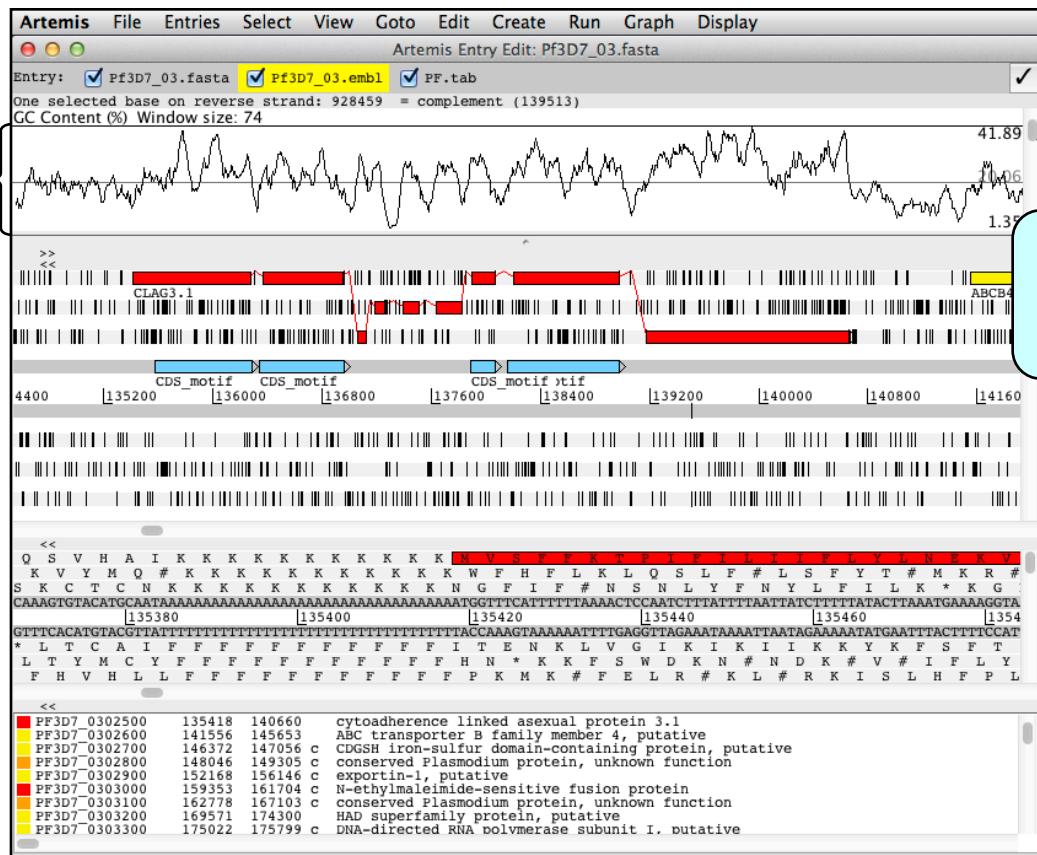
Further information on specific Pfam entries can be found on the web at  
<http://pfam.xfam.org/>

In addition to looking at the fine details of the annotated features it is also possible to look at the characteristics of the DNA covering the region displayed. This can be done by adding to the display various plots showing different characteristics of the DNA.

### To view the graphs:

Click on the ‘Graph’ menu to see all those available. Some of the most useful plots for *P. falciparum* is the ‘GC Content (%)’ as shown below. G+C content is a very good indicator of coding capacity in Malaria. On average, the coding regions are ~23% G+C and the non-coding regions are ~19%. Have a look at the G+C content for this region by selecting the appropriate graph. Left click within the graph window and then select by clicking on the exons to see how this relates to the G+C peaks on the graph.

DNA plot



Sliders for  
adjusting  
the window  
size

## Artemis Exercise 1 Part III

In this part of the Module we will be looking at methods of selecting and extracting features. We are going to extract different genes and regions and perform some more detailed analysis on it. We will aim to write and save new EMBL format files which will include just the annotation and DNA for this region.

In Artemis you can select genes fitting different search criteria. One possibility is to look for a specific product, for example *rifin*, as shown below.

1

Click 'Select' then 'Feature Selector'

3

Type search term

4

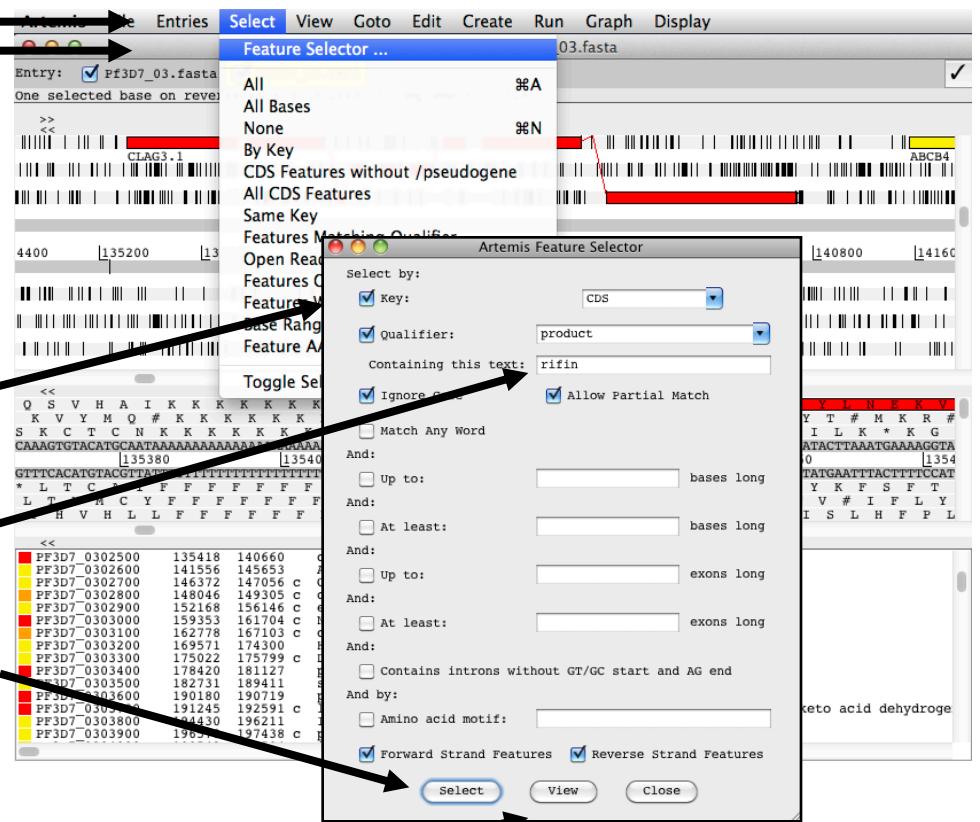
Click to select features containing search term

5

Click to view selected features

6

Double click to bring features into main view window.



All features with key "CDS" with qualifier "product" containing text "rifin"					
CDS	46369	47579	c	A-type rifin	(PMID:18197962)
CDS	55390	56584	c	B-type rifin	(PMID:18197962)
CDS	61445	62714		A-type rifin	(PMID:18197962)
CDS	64572	65783		A-type rifin	(PMID:18197962)
CDS	1015795	1016942		A-type rifin	(PMID:18197962)
CDS	1018874	1019973		B-type rifin	(PMID:18197962)
CDS	1027571	1028929		A-type rifin	(PMID:18197962)

The genes listed in 6 (on the previous page) are only those fitting your selection criterion. They can be copied or moved in to a new entry so they can be viewed in isolation from the rest of the information within Pf3D7\_03.embl. To create a new entry go to ‘Create’ and choose ‘New Entry’.

In the next step of the exercise choose one of the selected genes and write out a FASTA-file of the sequence.

Artemis

**File** Entries Select View Goto Edit Create Run Graph Display

Entry:

Selected f

>> <<

PF3D7\_0324000 1005952 1006800 exported protein family 1  
PF3D7\_0324100 1007413 1008202 c Pfmc-2TM Maurer's cleft two transmembrane protein  
PF3D7\_0324200 1010342 1011140 c exported protein family 3  
PF3D7\_0324300 1012355 1013464 exported protein family 4  
PF3D7\_0324400 1015795 1016942 rifin  
PF3D7\_0324500 1018874 1019973 rifin  
PF3D7\_0324600 1021998 1023008 stevor  
PF3D7\_0324700 1024030 1025411 erythrocyte membrane protein 1 (PfEMP1), exon 2, pseudogene  
PF3D7\_0324800 1027571 1028929 rifin  
Pfalciparum REP 1028959 1029560  
PF3D7\_0324900 1030822 1038254 c erythrocyte membrane protein 1, PfEMP1  
Pfalciparum REP 1040961 1050620  
Pfalciparum REP 1050649 1053812  
Pfalciparum REP 1055849 1057919  
Pfalciparum REP 1063240 1067971

Artemis Entry Edit: Pf3D7\_03.fasta

Con no. 2 RIF (/colour=2 /db xref="UniProtKB:097309"/db xref="MPMP:PF3D7\_0324000")

Amino Acids Of Selected Features

Amino Acids Of Selected Features To Qualifier

PIR Database Of Selected Features

**Bases Of Selection**

Upstream Bases Of Selected Features

Downstream Bases Of Selected Features

Upstream+Feature+Downstream Bases ...

All Bases

Codon Usage Of Selected Features

Raw Format

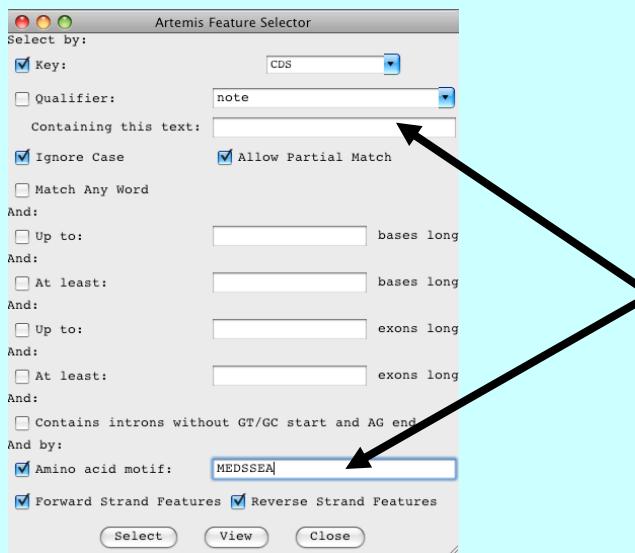
**FASTA Format**

EMBL Format

Genbank Format

Click ‘File’ then  
‘Write ‘Bases of  
Selection’ ‘FASTA  
Format’

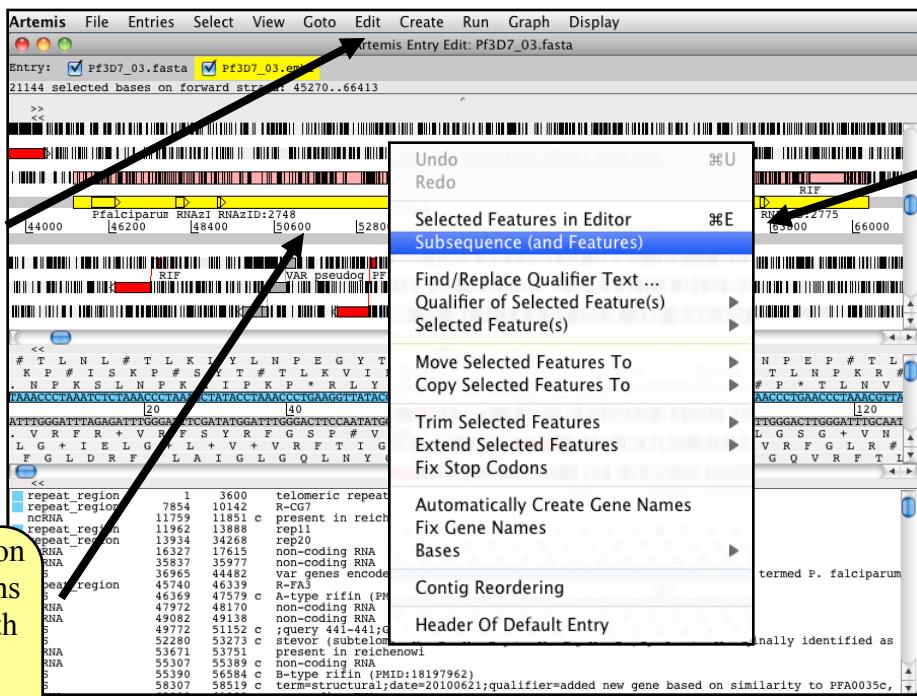
**Additional methods of selecting/extracting features using the Feature Selector**  
 It is worth noting that the Feature Selector can be used in many other ways to select and extract subsets of features from the genome such as text or amino acid searches.



Space for a search term or amino acid motif

In the next part of the exercise we will be looking at the region containing the *rifin* genes in more detail. They are located at the end of the chromosomes, in the subtelomeric region. We are going to extract this region from the whole chromosome sequence. Then we will aim to write and save new EMBL format files which will include just the annotation and DNA for this region.

3



Click  
‘Subsequence  
(and Features)’

## **2** | Click ‘Edit’

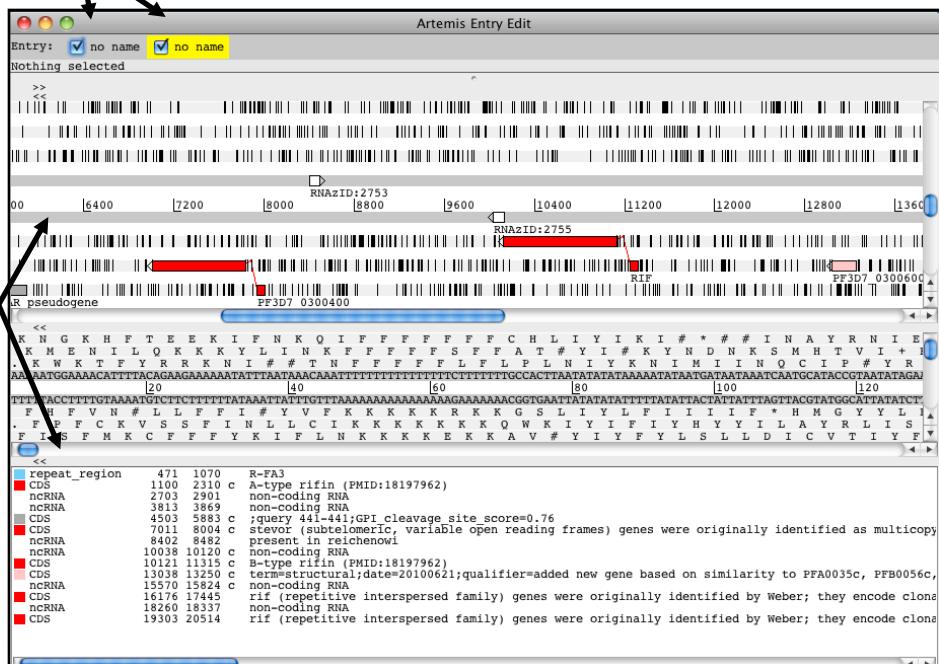
Select the region containing rifins by clicking with the left mouse button and dragging.

Note the entry names have changed

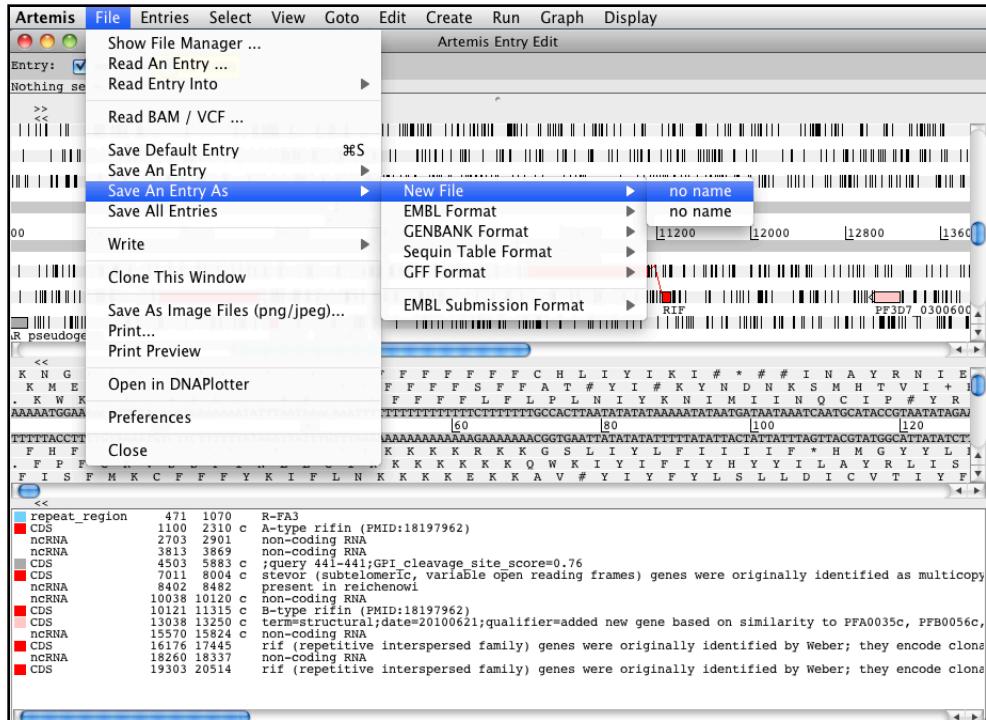
4

A new Artemis window will appear displaying only the region that you have highlighted.

Note the bases have been renumbered from the first base you selected.



Note that the two entries on the grey Entry line are now denoted ‘no name’, they represent the same information in the same order as the original Artemis window but simply have no assigned name. So click on the File menu then ‘Save an entry as’ and then ‘New file’. Another menu will ask you to choose one of the entries listed. At this point they will both be called ‘no name’. Left click on the top entry in the list. A window will appear asking you to give this file a name. The new files can be saved in different formats.



Once you have finished this exercise remember to close this Artemis session down completely before starting the next exercise.

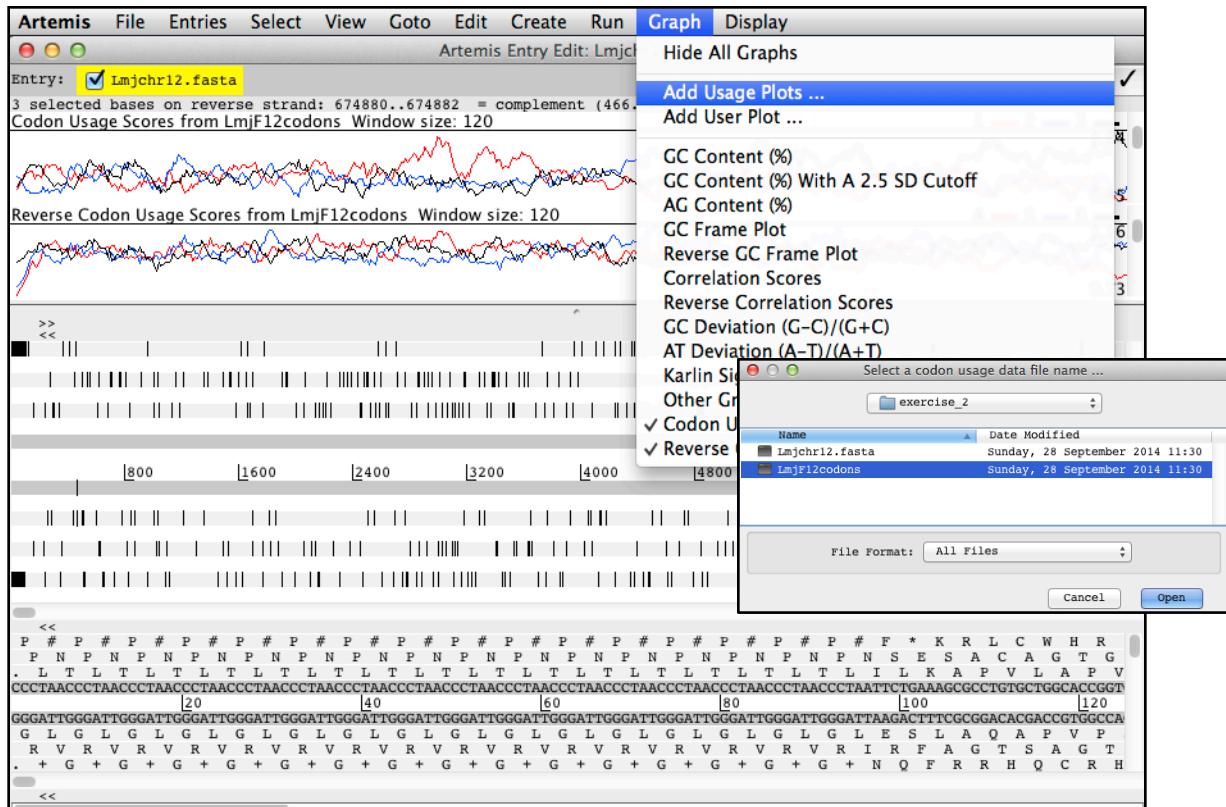
## Optional exercise

We are now switching to a different organism. The following exercise demonstrates how to use Artemis as a tool for structural annotation. Given a length of chromosome with no existing annotation Artemis can mark up ORFs above a given size. This also shows how codon usage plots can be exploited in gene model prediction.

If you haven't already closed the previous session of Artemis, do so now. Double click the ARTEMIS Icon on your Desktop and navigate to the directory Module\_1\_Artemis, optional\_exercise and open the sequence file Lmjchr12.fasta.

Next, open the codon usage table file LmjF12codons by selecting 'Add Usage Plots' from the Graph menu. Codon usage is a very good indicator of coding capacity in *Leishmania* genomes where there is a much more prominent codon bias for some amino acids.

Note, we will cover the use of RNAseq data in gene prediction later on during the course.



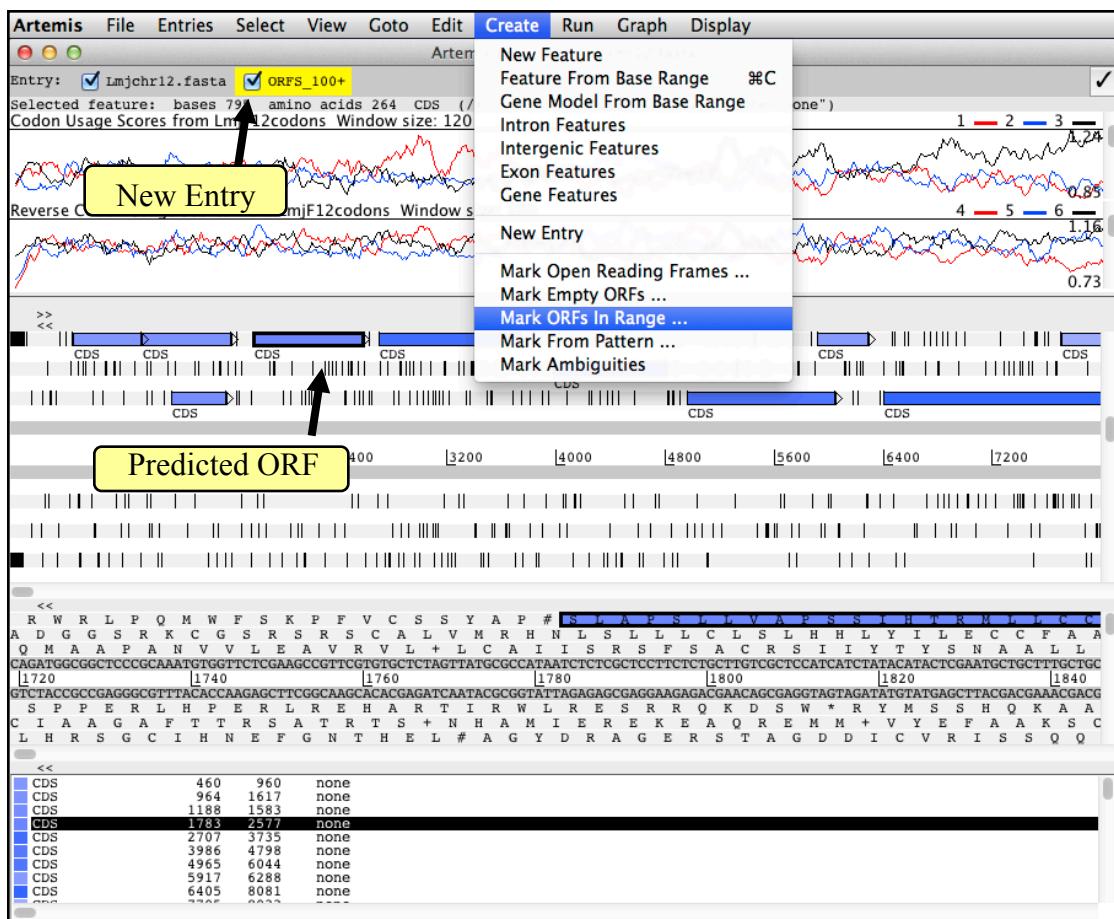
Select the first 100 kbs of sequence on the positive strand either by highlighting the sequence in the sequence window (use shift and click to select the final base) or choose the ‘Base Range’ option in the select menu and enter ‘1..100000’.

With this region selected, select ‘Mark ORFs in Range’ from the Create menu. When prompted for minimum ORF size enter 100. Note that this results in the creation of a new entry called ‘ORFS\_100+’. You can experiment with a range of ORF sizes by de-selecting this entry and repeating the first steps in this process.

Note that the marked up ORFs vary in colour from pale to navy blue. This colouring reflects the codon usage support for this model with darker blue being highly supported by codon usage.

Try selecting some of the newly created features in the gene window. Double clicking on one of these will bring up the predicted peptide sequence in the bottom window. You can rapidly move to the N- or C-terminus of the predicted peptide by holding down ctrl, and then left or right arrow respectively.

Note that we have chosen only to generate ORFs for the positive strand for this example. In a genome not organized into transcription units we would normally do likewise for the reverse strand as well.



Although some of these predictions are likely to be correct, there is considerable overlap between predicted ORFs, and many are small and unsupported by codon usage. To validate/negate our predicted models we need to do further sequence comparison. This can be done with a tool such as ACT (to be discussed later in the Comparative Genomics Module), or with one of the integrated Blast options in Artemis. Select the ORF at position 12745, click on it, then select RUN>NCBI Searches>blastx. This will open a browser window with NCBI results.

The screenshot shows the Artemis software interface with several panels:

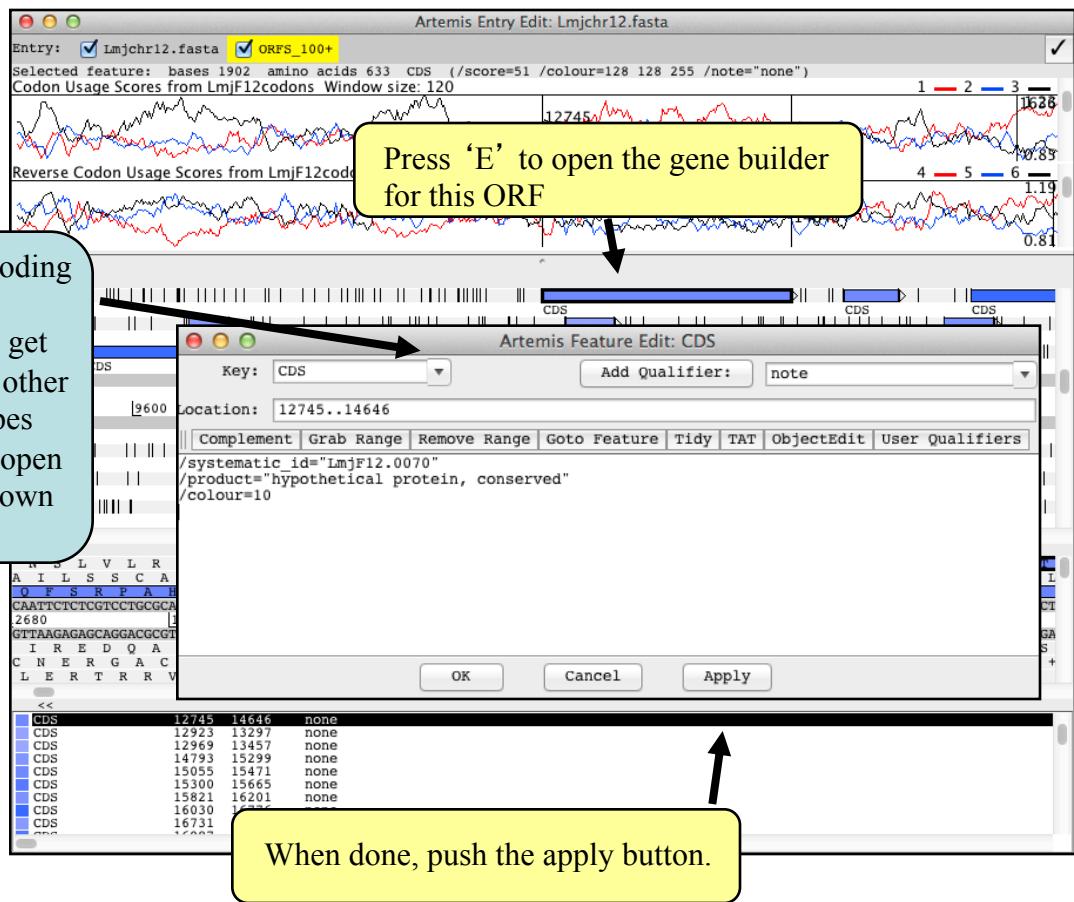
- Artemis Entry Panel:** Shows an entry for "Lmjchr12.fasta" and "ORFS\_100+". It displays "Selected feature: bases 1902 amino acids 633 CDS (/score=64)" and "Codon Usage Scores from LmjF12codons Window size: 120". Below this are two line graphs: "Reverse Codon Usage Scores from LmjF12codons Window size: 120" (red and blue lines) and a genomic track with CDS regions.
- NCBI Searches Panel:** A dropdown menu listing various search options:
  - blastp
  - tblastn
  - blastn
  - blastx**
  - tblastxOther listed options include: Pfam Search, Rfam Search, Run fasta on selected features against, Run sigcleave (0) on selected features, Run pepstats on selected features, Run blastp on selected features against, Run tblastn on selected features against, Run hth on selected features, Run smart on selected features, Run clustalx (PROTEIN) on selected features, Run jalview (PROTEIN) on selected features, Run tblastx on selected features against, Run blastx on selected features against, Run fastx (%uniprot) on selected features, Run clustalx (DNA) on selected features, and Run jalview on selected features.
- Genomic Track Panel:** Shows a detailed view of the genomic sequence with CDS annotations and amino acid translations (e.g., N S L V L R T A L A S S C T P M D R T R \* A I L S S C A Q L L R L R V L Q W T E H V). It also shows codon usage scores for positions 8800, 9600, 10400, 11200, and 12000.
- Sequence Panel:** Displays the DNA sequence with its corresponding protein translation and amino acid positions (e.g., 2680, 12700, 12720, 12740).
- Table Panel:** A table showing CDS details for each position:

CDS	Start	End	Type
CDS	12745	14646	none
CDS	12923	13297	none
CDS	12969	13457	none
CDS	14793	15299	none
CDS	15055	15471	none
CDS	15300	15665	none
CDS	15821	16201	none
CDS	16030	16776	none
CDS	16731	17075	none

conserved hypothetical protein [Leishmania major strain Friedlin]								
Sequence ID: ref XP_001681612.1		Length: 620		Number of Matches: 1				
> See more title(s)								
Range 1: 1 to 620					GenPept	Graphics		
Score	Expect	Method	Identities		Positives	Gaps		
1238 bits(3203)	0.0	Compositional matrix adjust.	620	620(100%)	620	620(0%)		
Query 40	0.0	MHTITPTFTSFSPPSPVPASTPSTAIDHLAHCAGLRLPQVRMSGEFSLLHLQGTVT	620	620(100%)	219			
Sbjct 1	0.0	HHTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT	620	620(100%)	676			
Query 220	0.0	DKGGSSADLTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	399	399(100%)	399			
Sbjct 61	0.0	DKCXSAAADLTTPSAAAFAVWPQTQPPSKCCTAEGTLCVLEKLRKGARHSDAATPGE	399	399(100%)	120			
Query 400	0.0	VLLRKRNVEQYVIVYNSK1PSVYCNQGLAKMHEERERENSPLFKYPLAVECEAQEE	579	579(100%)	579			
Sbjct 121	0.0	VLLRKRNVEQYVIVYNSK1PSVYCNQGLAKMHEERERENSPLFKYPLAVECEAQEE	579	579(100%)	180			
Query 580	0.0	ARRVVLQLERNCQDNEQARHRKECLEREARLRAAEVVAQYKACTAREADAREXRRK	759	759(100%)	759			
Sbjct 181	0.0	ARRVVLQLERNCQDNEQARHRKECLEREARLRAAEVVAQYKACTAREADAREXRRK	759	759(100%)	300			
Query 760	0.0	GEAVSEATAAKTLTIRRERADANLDVARVASCGNCNEEDDRRLAERLAAERTLAEENFR	939	939(100%)	939			
Sbjct 241	0.0	GEAVSEATAAKTLTIRRERADANLDVARVASCGNCNEEDDRRLAERLAAERTLAEENFR	939	939(100%)	360			
Query 940	0.0	AQSRQAERFPAQGQARFPAQGQARFPAQGQARFPAQGQARFPAQGQARFPAQGQAR	1119	1119(100%)	1119			
Sbjct 301	0.0	AQSRQAERFPAQGQARFPAQGQARFPAQGQARFPAQGQARFPAQGQARFPAQGQAR	1119	1119(100%)	360			
Query 1120	0.0	RANSADDVILQOMQNSWFLDAVERDREBAAKQRKMDTAVNVLRAQKQADAAQERDR	1299	1299(100%)	1299			
Sbjct 361	0.0	RANSADDVILQOMQNSWFLDAVERDREBAAKQRKMDTAVNVLRAQKQADAAQERDR	1299	1299(100%)	420			
Query 1300	0.0	DRQVAAEYAAKELENPQRFVEHAAQRQRQERQELDAAEATAYVQKQADAAHQARRQS	1479	1479(100%)	1479			
Sbjct 421	0.0	DRQVAAEYAAKELENPQRFVEHAAQRQRQERQELDAAEATAYVQKQADAAHQARRQS	1479	1479(100%)	480			
Query 1480	0.0	FPLFWPAQSGPAAEKA1DANRREFDLRLRQAKQERDERRAEQEAARARALVEYDTL	1659	1659(100%)	1659			
Sbjct 481	0.0	FPLFWPAQSGPAAEKA1DANRREFDLRLRQAKQERDERRAEQEAARARALVEYDTL	1659	1659(100%)	540			
Query 1660	0.0	AREVEREKENRKEERKEAHLERLTIEJAEAKERKGVADGRICRACAAQAVVHPATEARMLLYRC	1839	1839(100%)	1839			
Sbjct 541	0.0	AREVEREKENRKEERKEAHLERLTIEJAEAKERKGVADGRICRACAAQAVVHPATEARMLLYRC	1839	1839(100%)	600			
Query 1840	0.0	PTTCOLLSADASASYDFQVQCRR 1899	1899	1899(100%)	1899			
Sbjct 601	0.0	PTTCOLLSADASASYDFQVQCRR 620	620	620(100%)	620			

Not surprisingly, the top hit is to a gene on chromosome 12 in *L. major*, a hypothetical protein

Now that we know that this is a real gene we can make a few adjustments. First, open the gene builder window by selecting the ORF and pressing E. This will open a text window where we can add annotations on the gene. Start by deleting the current ‘automatic’ annotations in this window. Try entering the text in the gene builder shown below to record gene ID, predicted product and a colour code that will distinguish this gene from the automatically generated ORFs.



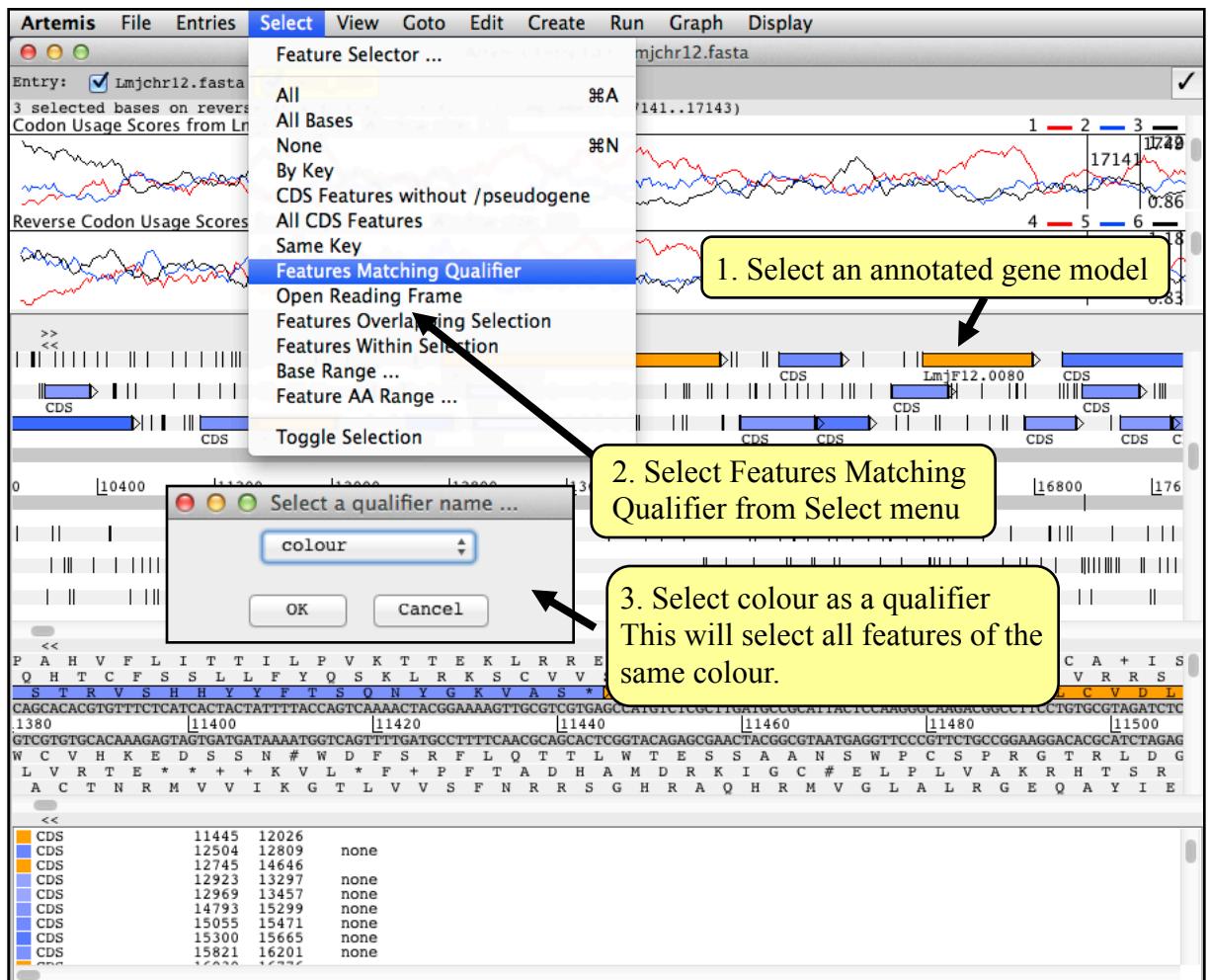
Based on the NCBI blast results we can adjust the N-terminus of this model to the correct start codon. To automatically position the sequence window at the N-terminus of the gene model push **ctrl-<left arrow>**.

Go to **Edit>Trim Selected Feature>To Next Met** (or **ctrl-T**), then reposition the sequence window at the new start as described above. Continue until the start resembles the NCBI blast results. If trimmed passed the desired start codon the model can be reset through **Edit>Extend Selected Feature>To Previous Stop Codon**, or **ctrl-Q**.



There are more than 20 protein coding genes in the first 100 kbs of chromosome 12. See how many of these you can find by repeating the steps in the past slides.

**IMPORTANT!!** Any changes made to the predicted ORFs will be written to an entry file called ORFS\_100+. When you're done with gene predictions follow the steps below to save these entries to the sequence file instead. Make sure all of the annotated features have a /colour=10 in their gene builder window.



Artemis File Entries Select View Goto Edit Create Run Graph Display

Entry:  Lmjchr12.fasta  ORFS\_100+

3 selected features total bases 3231 total amino acids 11445 Codon Usage Scores from LmjF12codons Window size 11 Reverse Codon Usage Scores from LmjF12codons Window size 11

>> <<

CDS CDS LmjF12.0060 CDS

0 [10] 00 [11200] 2000 [12000]

5. After the features have been copied to Lmjchr12.fasta, de-select ORFS\_100+. Only annotated ORFs should remain.

4. From the Edit menu, select 'copy selected features', then select the sequence file Lmjchr12.fasta

Undo ⌘U  
Redo ⌘R  
Selected Features in Editor ⌘E  
Subsequence (and Features)  
Find/Replace Qualifier Text ...  
Qualifier of Selected Feature(s)  
Selected Feature(s)  
Move Selected Features To  
Copy Selected Features To ▶ Lmjchr12.fasta  
ORFS\_100+  
Trim Selected Features  
Extend Selected Features  
Fix Stop Codons  
Automatically Create Gene Names  
Fix Gene Names  
Bases ▶ 16000 16800 17600  
Contig Reordering  
Header Of Default Entry

CDS	11445	12026	
CDS	12504	12809	none
CDS	12745	14646	
CDS	12923	13297	none
CDS	12969	13457	none
CDS	14793	15299	none
CDS	15002	15471	

T E K L R R E P C L A \* C R I T P R A R R R P S C A + I S R K S C V V S H V S L D A A L L O G O D G L P V R R S Y G K V A S \* S A M S R I M E H Y S K G K T A F I C V D C ACGAAAAAGTTCGCTGTGAGCCATGTCGCTTGATGCCGATTACTCCAAGGGCAAAGGCCCTTCCTGTGCGTAGATCTC [11440] [11460] [11480] [11500] TGCCTTTCAACGCAAGCACTCGGTACAGAGCGAACTACCGCGTAATGAGGTTCCCCTGCCTGGAAAGGACACGCATCTGAG R F L Q T T L W T E S S A A N S W P C S P R G T R L D G P F T T A D H A M D R K I G C # E L P L V A K R H T S R V S F N R R S G H R A Q H R M V G L A L R G E Q A Y I E

Artemis

**File** Entries Select View Goto Edit Create Run Graph Display

Entry:  3 selected Codon Usage

Show File Manager ...  
Read An Entry ...  
Read Entry Into  
Read BAM / VCF ...  
Save Default Entry ⌘S  
Save An Entry  
**Save An Entry As**  
Save All Entries  
Write  
Clone This Window  
Save As Image Files (png/svg)...  
Print...  
Print Preview  
Open in DNAPlotter  
Preferences  
Close

Artemis Entry Edit: Lmjchr12.fasta

6. From the File menu, select save an Entry as > EMBL format > Lmjchr12.fasta.

o acids 1074 (LmjF12.0060 to 172)

New File  
**EMBL Format** ► Lmjchr12.fasta  
GENBANK Format  
Sequin Table Format  
GFF Format  
EMBL Submission Format

LmjF12.0080

4 — 5 — 6 —  
1.18  
0.83

2800 13600 14400 15200 16000 16800 17600

P A H V F L I T T I L P V K T T E K L R R E P C L A \* C R I T P R A R R P S C A + I S  
Q H T C F S S L F Y Q S K L R K S C V S H V S L D A A L Q G Q D G L P V R R D L S  
S T R V P S H Y Y F T S Q N Y G K V A S \* A M S R L M P H Y S K G K T A F L C V D L  
CAGCACACGTGTCTCATCACTACTATTTACCAAGTCAAAACATCGGAAAGTTCGGCTGAGGCCATGTCCTGGCTTGATGCCGATTACTCCAAAGGGCAAGACGCCCTTCCTGTGGCTGATGCTC  
1380 [11400] [11420] [11440] [11460] [11480] [11500]  
GTCTGTGCAACAAAGAGTAGTCATGATAAAATGGTCAGTTTGATGCCCTTCACACCGCAGCTCGGTACAGAGGCAGTACAGGGCTAATGAGGTTCGGCTTGCTGCCGAAAGGACACCCATCTAGAG  
W C V H K E D S S N # W D F S R F L Q T T L W T E S S A A N S W P C S P R G T R L D G  
L V R T E \* \* + + K V L \* F + P F T A D H A M D R K I G C # E L P P L V A K R H T S R  
A C T N R M V V I K G T L V V S F N R R S G H R A Q H R M V G L A L R G E Q A Y I E  
<<  
CDS 11445 12026  
CDS 12745 14646  
CDS 16030 16776