



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

BIOINFORMATICS LAB

Episode III – Phylogenetic Analysis

Federico M. Giorgi, PhD

Department of Pharmacy and Biotechnology
First Cycle Degree in Genomics

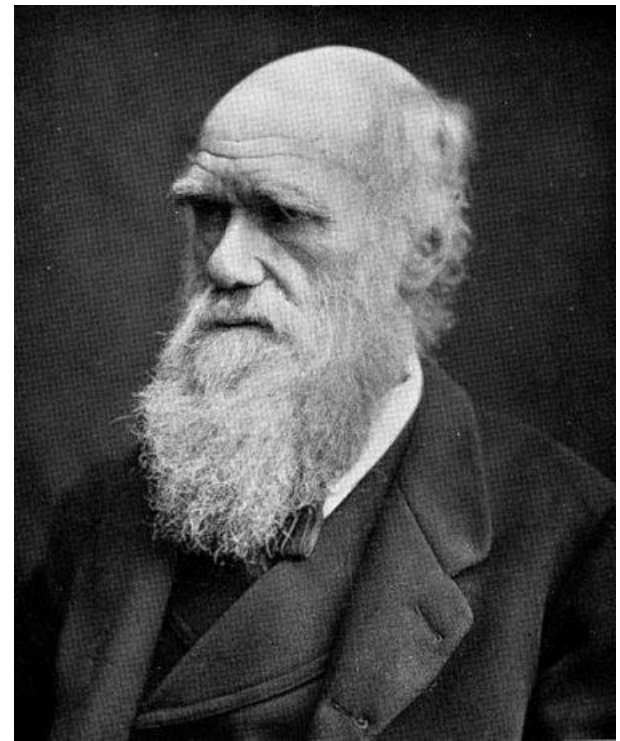
Phylogeny

The history of the evolution of a species or group



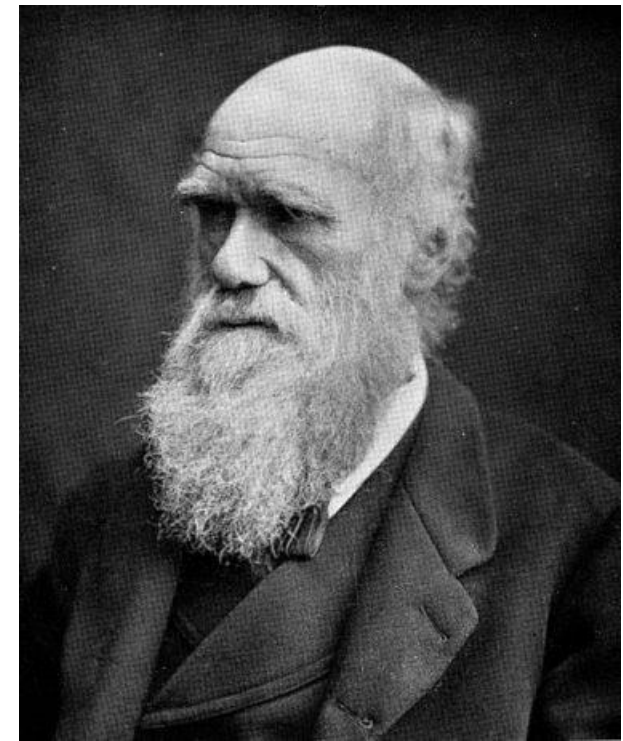
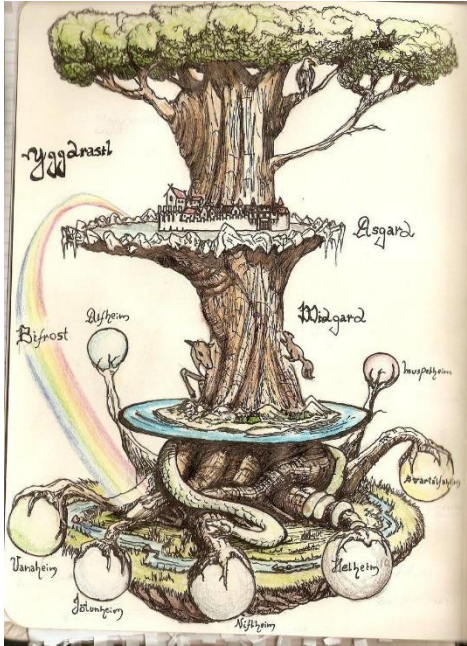
Phylogeny

The history of the evolution of a species or group



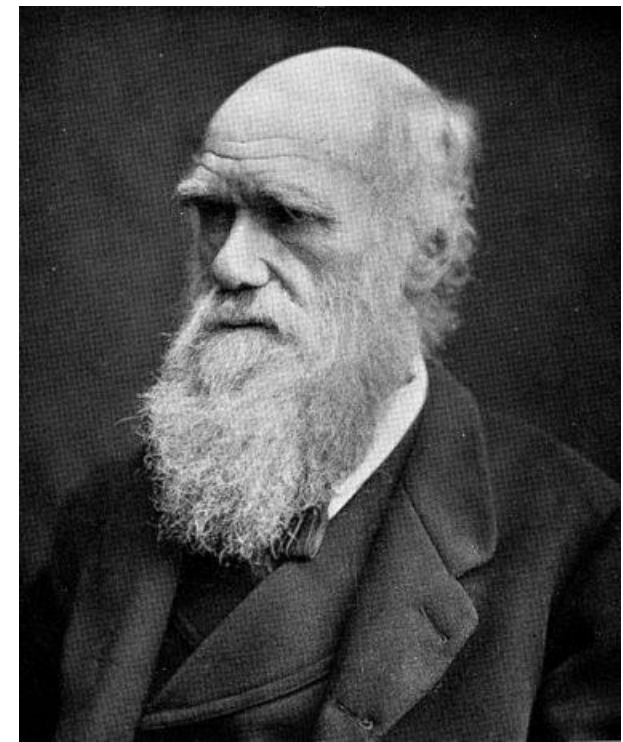
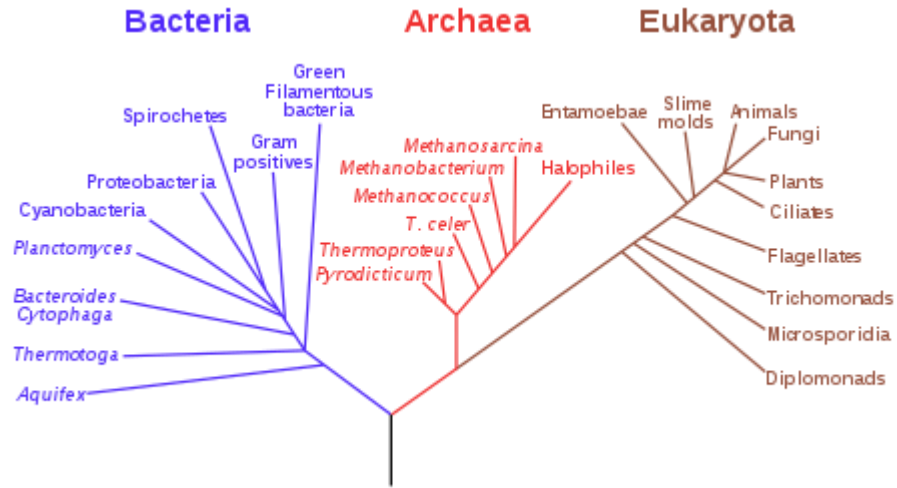
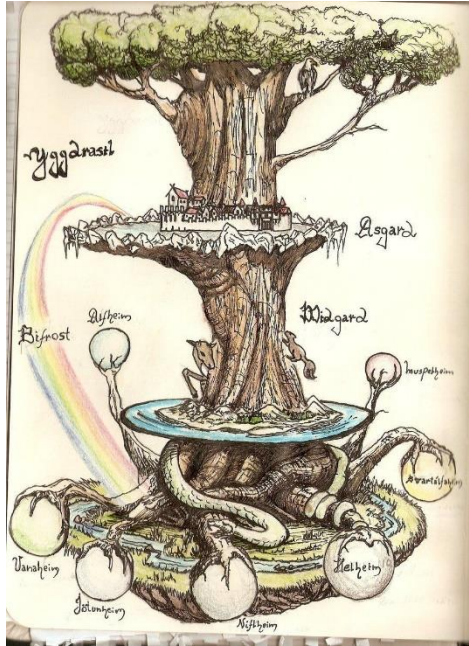
Phylogeny

The history of the evolution of a species or group



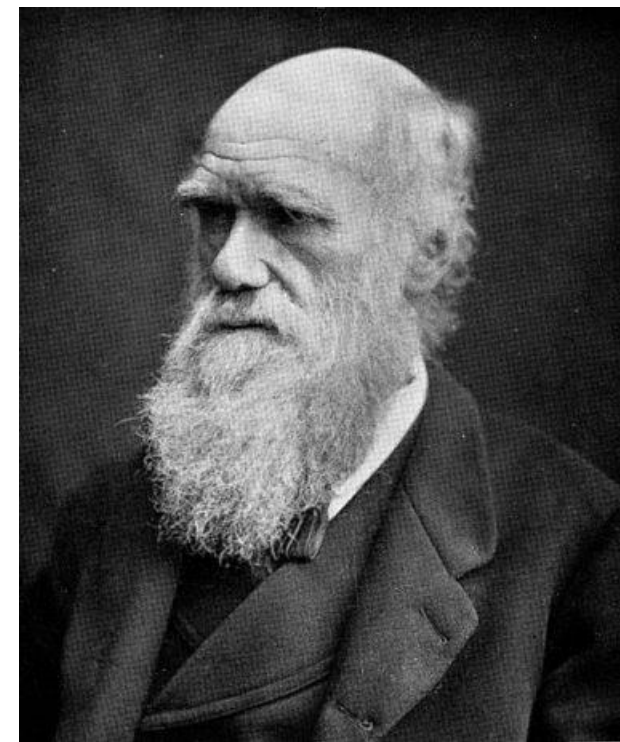
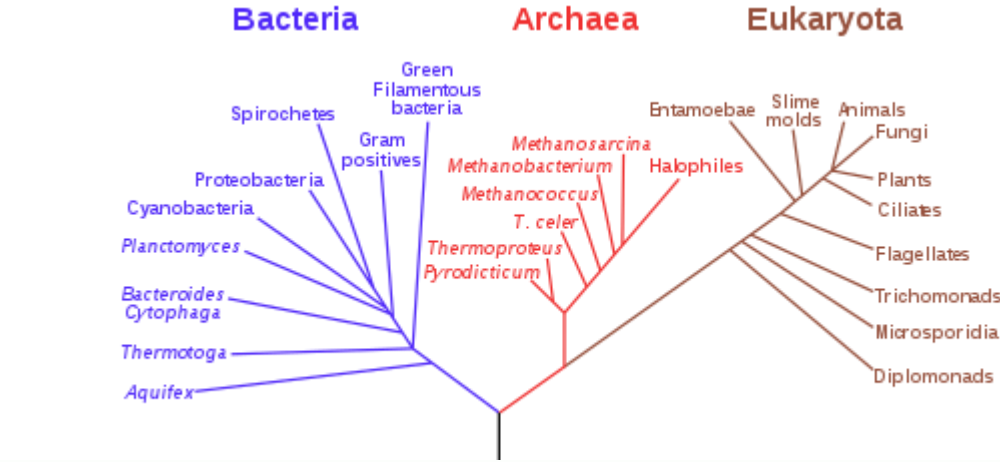
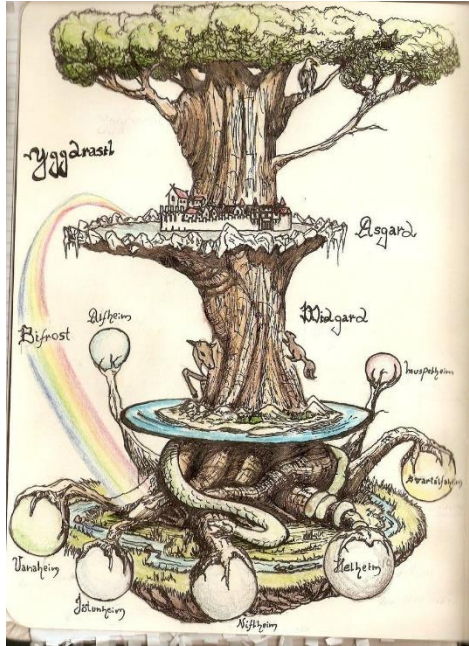
Phylogeny

The history of the evolution of a species or group



Phylogeny

The history of the evolution of a species or group



Q5E940	BOVIN	-----MPREDRATWKS	NYFLKII	QLLDD	YPKCFIVGADNVGS	KMQQIRMSLRGK	AVVLMGKNTMMRKAIRGHLENN	PALE	76
RLA0	HUMAN	-----MPREDRATWKS	NYFLKII	QLLDD	YPKCFIVGADNVGS	KMQQIRMSLRGK	AVVLMGKNTMMRKAIRGHLENN	PALE	76
RLA0	MOUSE	-----MPREDRATWKS	NYFLKII	QLLDD	YPKCFIVGADNVGS	KMQQIRMSLRGK	AVVLMGKNTMMRKAIRGHLENN	PALE	76
RLA0	RAT	-----MPREDRATWKS	NYFLKII	QLLDD	YPKCFIVGADNVGS	KMQQIRMSLRGK	AVVLMGKNTMMRKAIRGHLENN	PALE	76
RLA0	CHICK	-----MPREDRATWKS	NYFMKII	QLLDD	YPKCFVVGADNVGS	KMQQIRMSLRGK	AVVLMGKNTMMRKAIRGHLENN	PALE	76
RLA0	RANSY	-----MPREDRATWKS	NYFLKII	QLLDD	YPKCFIVGADNVGS	KMQQIRMSLRGK	AVVLMGKNTMMRKAIRGHLENN	SALE	76
Q7ZUG3	BRARE	-----MPREDRATWKS	NYFLKII	QLLDD	YPKCFIVGADNVGS	KMQMOTIRLSLRGK	AVVLMGKNTMMRKAIRGHLENN	PALE	76
RLA0	ICTPU	-----MPREDRATWKS	NYFLKII	QLLDD	YPKCFIVGADNVGS	KMQMOTIRLSLRGK	AVVLMGKNTMMRKAIRGHLENN	PALE	76
RLA0	DROME	-----MVRENKA	AWKAQYFIKVV	ELFDEF	PKCFIVGADNVGS	KMQMOTIRLSRGL	AVVLMGKNTMMRKAIRGHLENN	POLE	76
RLA0	DICDI	-----MSGAG	-SKRKKLFTEKAT	KLFTT	YDKMIVAEADFVGS	SLOKIRKSIRGI	CAVLMGKNTMIRKVIIRD	LADSK	75
RLA0	DICDI	-----MSGAG	-SKRKNVFEKAT	KLFTT	YDKMIVAEADFVGS	SLOKIRKSIRGI	CAVLMGKNTMIRKVIIRD	LADSK	75
Q54LPO	DICDI	-----MAKLSKQ	QKQMYTEKLS	SLIQQ	YSKILLVHVDNVGS	NQMASVRKSLRGK	AILLMGKNTIRITALKKNL	CAV	76
RLA0	SULAC	-----MIGLAVTTT	KKIAKWKVDEVAELT	EKLKLT	KTIITLIANIEGFP	PADKLHEIRKKLRGK	ADIKVTKNL	LFNIALKNAG	79
RLA0	SULTO	-----MRIMAVITQ	ERKIAKWKIEEYKELE	EKLRE	YHTIIIANIEGFP	PADKLHDIRKKMRGM	AEIKVTKNL	LFNIALKNAG	80
RLA0	SULSO	-----MKRLALALK	QQRKVASWKL	EYKELTEL	LKNSNTILLGNLE	GFPADKLHEIRKKLRGK	ADIKVTKNL	LFNIALKNAG	80
RLA0	AERPE	-----MSVVSIVG	QMYKREKPIPE	WKTLLMLRELE	ELFSKIRVVLFADLTC	PTFVVRVVKLWKK	YPMVAKKRIIL	RAMKAGLE	86
RLA0	PYRAE	-----MMLAIG	KRRYVRTQYPA	RKVKIVSEAT	ELLQKYPYVFLFDLHGLSS	RILHEVYRRLRY	GVIKIIP	TLFKTAFTKVYGG	85
RLA0	METAC	-----MAEERHHT	EHIPQWKKDE	TENIKELIQS	HKVFQMGVIEGILAT	KMOKIRRDLDKDV	AVLKVS	RNTLTERALNQLG	78
RLA0	METMA	-----MAEERHHT	EHIPQWKKDE	TENIKELIQS	HKVFQMVRIEGILAT	KIIRRDLDKDV	AVLKVS	RNTLTERALNQLG	78
RLA0	ARCFU	-----MAAVRGS	---PPEYKVR	AEEIKRMIS	KPVVAIVSFRNVPAG	QMKIRFRGK	AETIKVKNT	LLERALD	75
RLA0	METKA	-----MAVKAK	QPPSGYE	PKVAEWKRRE	VKELKELMDE	YENVGLVDLE	GIPAPQLOE	IRAKLRE	88
RLA0	METTH	-----MAHVAEW	KKKEVQELHDL	IKGEVVG	IANLADIPARQL	QKMRQLRDS	ALIRMS	KKLLISLAL	88
RLA0	METTL	-----MITAESE	HKIAPWKIEE	VNKLKEL	LKNGQI	VALVDMMEV	PARLOE	IRDKIR	82
RLA0	METVA	-----MIDAKSE	HKIAPWKIEE	VNALKEL	LKNSVIAL	IDMMEV	PARLOE	IRDKIR	82
RLA0	METJA	-----METKVK	AHVADWKIEE	VYKTLK	LKLIKSPVVAIV	DDMDV	PARLOE	IRDKIR	81
RLA0	PYRAB	-----MAHVAEW	KKKEVEELAN	LKIKSPV	VIALVDVSSMP	AYPLSQM	RRLIRENG	GLLRVSRNTL	81
RLA0	PYRHO	-----MAHVAEW	KKKEVEELAN	LKIKSPV	VIALVDVSSMP	AYPLSQM	RRLIRENG	GLLRVSRNTL	81
RLA0	PYRFU	-----MAHVAEW	KKKEVEELAN	LKIKSPV	VIALVDVSSMP	AYPLSQM	RRLIRENG	GLLRVSRNTL	81
RLA0	PYRKO	-----MAHVAEW	KKKEVEELAN	LKIKSPV	VIALVDVAGV	PAYPLSK	MRDKLR	GKALLR	81
RLA0	HALMA	-----MSAESER	KTETIP	WKKQEEVD	AIVEMIES	YESVGV	VNIAGIP	SRLODM	81
RLA0	HALVO	-----MSESEVR	QTEVIP	QWKKRE	VDLVDV	IESYES	SGVGVG	VAGIP	81
RLA0	HALSA	-----MSAEQRT	TEEVPE	WKRQEV	AELVDL	LETDS	SGVGVN	VNTGIP	81
RLA0	THEAC	-----MKEVS	QOKKELVNETT	RIKASRS	VAVLVD	TAGIRT	ROI	ODIRG	81
RLA0	THEVO	-----MRKIN	PKKKEIV	SELAG	DITKSK	AVAV	VDIKGV	RT	81
RLA0	PICTO	-----MTEPA	QWKIDFV	KNLE	INSRK	VAAIV	SIKGL	RNN	81
RLA0	ruler	1.....10.....20.....30.....40.....50.....60.....70.....80.....90							

The Five Practical Stages of Phylogenetic Analysis

1. Sequence Acquisition

2. Multiple Sequence Alignment

3. Specify Model for Nucleotide/Amino acid evolution

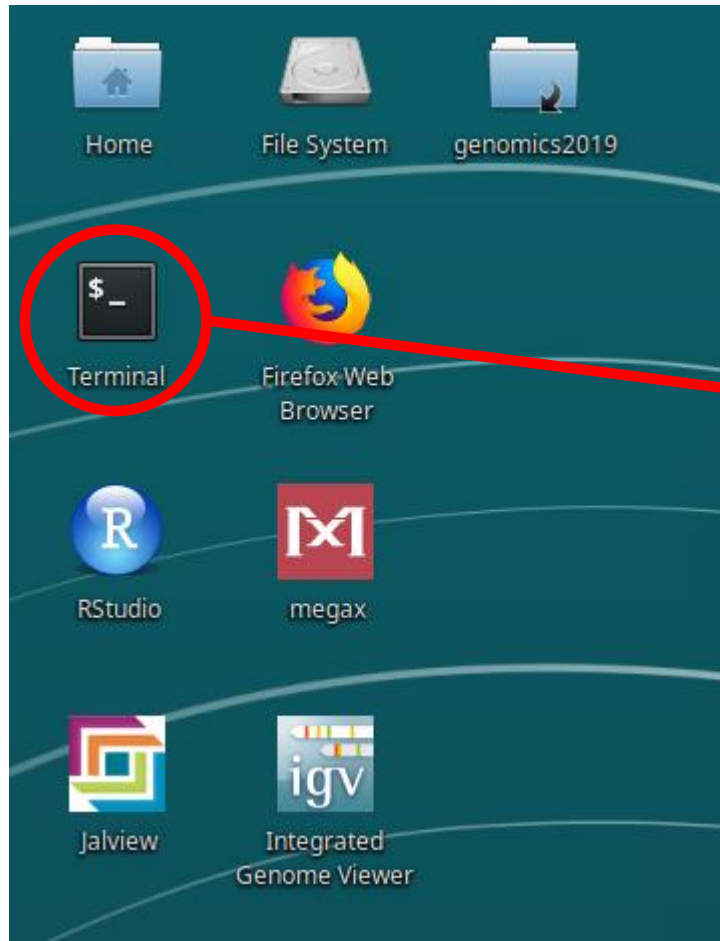
statistical models to estimate the number of changes that occurred between multiple sequences
(simplest is Hamming distance = nr. differences)

4. Tree Building

5. Tree Evaluation



Let's open The Terminal!



Terminal

Reminders:

- userid **student**
- password **4genomics4**



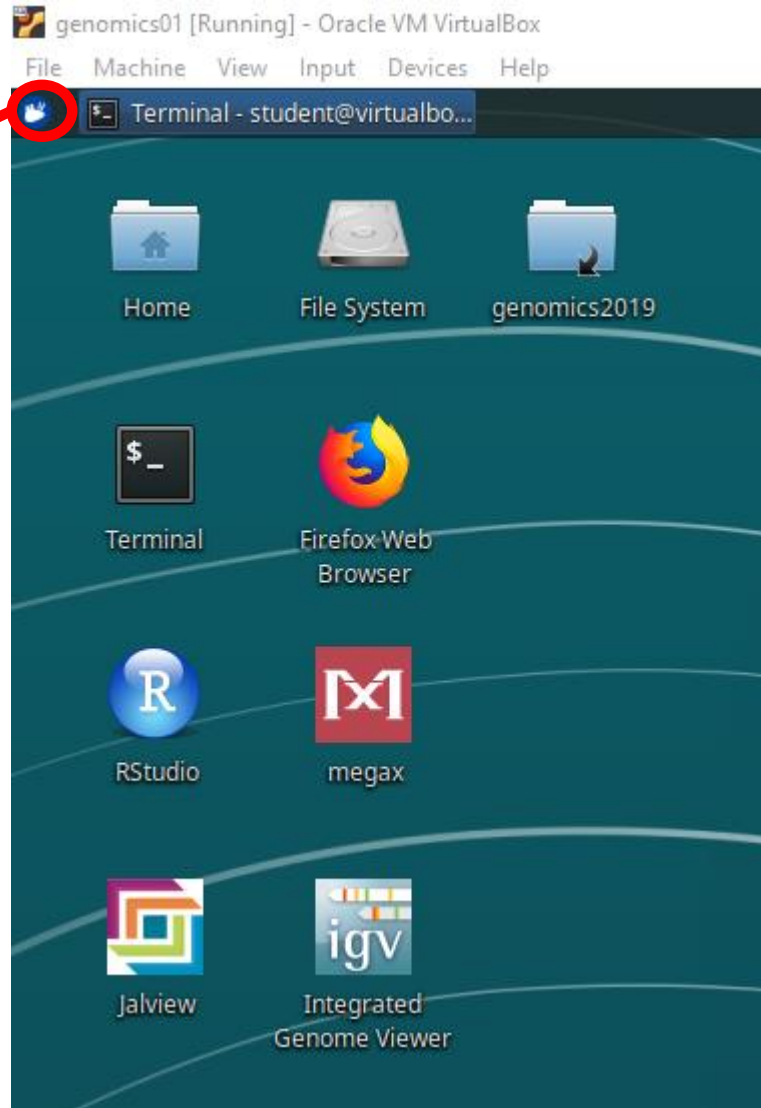
Sequences Exercises

(Open exercises_03_phylogeny.pdf)



Turn Unix off CORRECTLY

- Please turn it off nicely

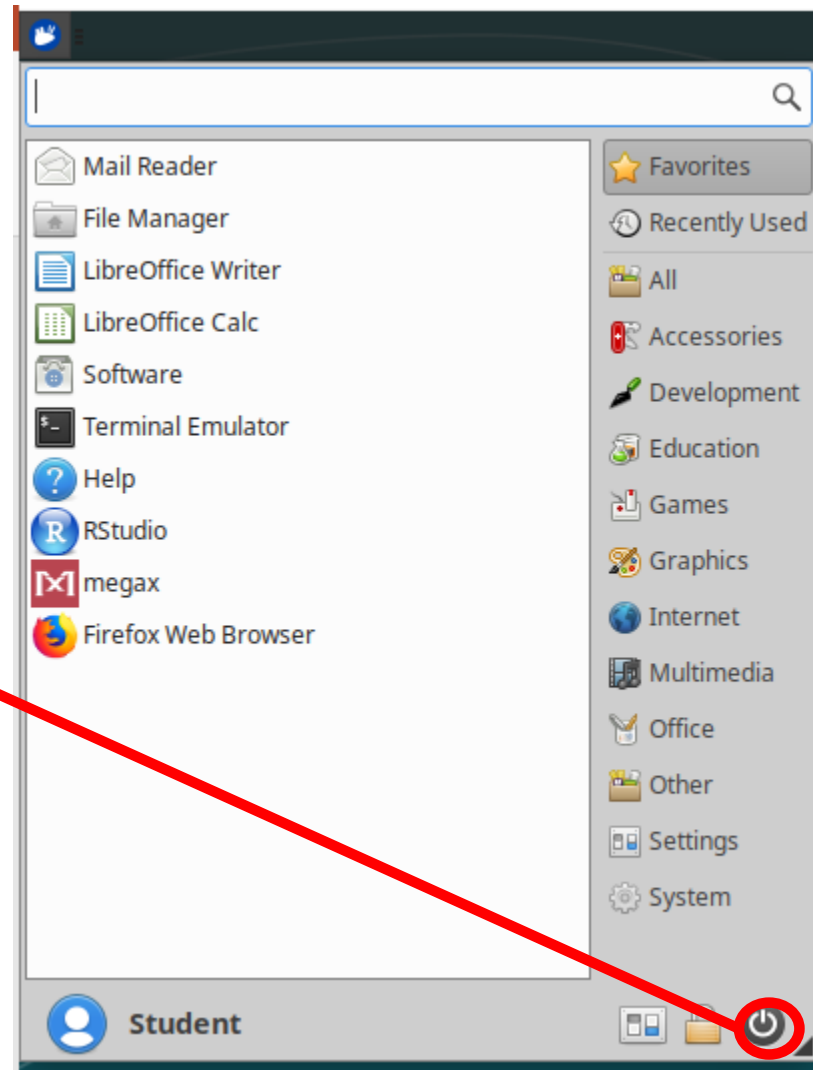


Click on the mouse

Turn Unix off CORRECTLY

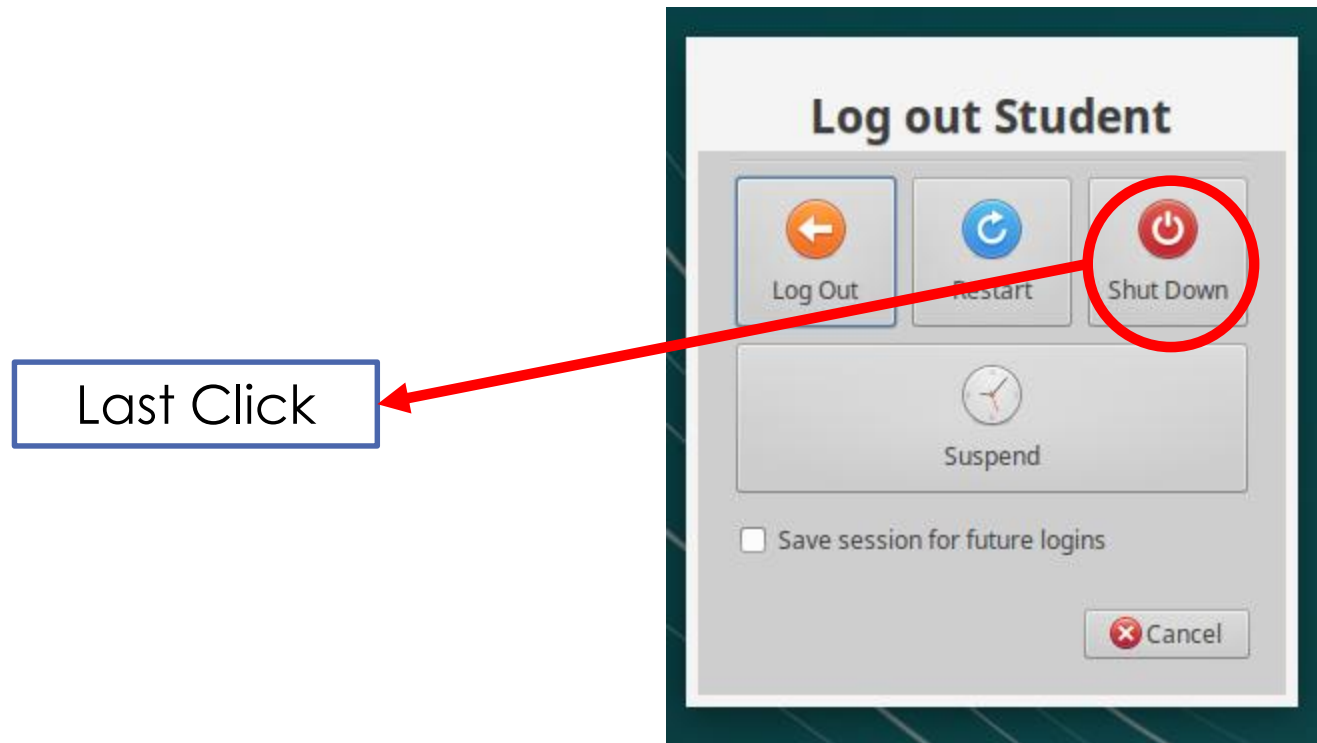
- Please turn it off nicely

Click Again



Turn Unix off CORRECTLY

- Please turn it off nicely





ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Federico M. Giorgi, PhD

Department of Pharmacy and Biotechnology

federico.giorgi@unibo.it

www.giorgilab.org