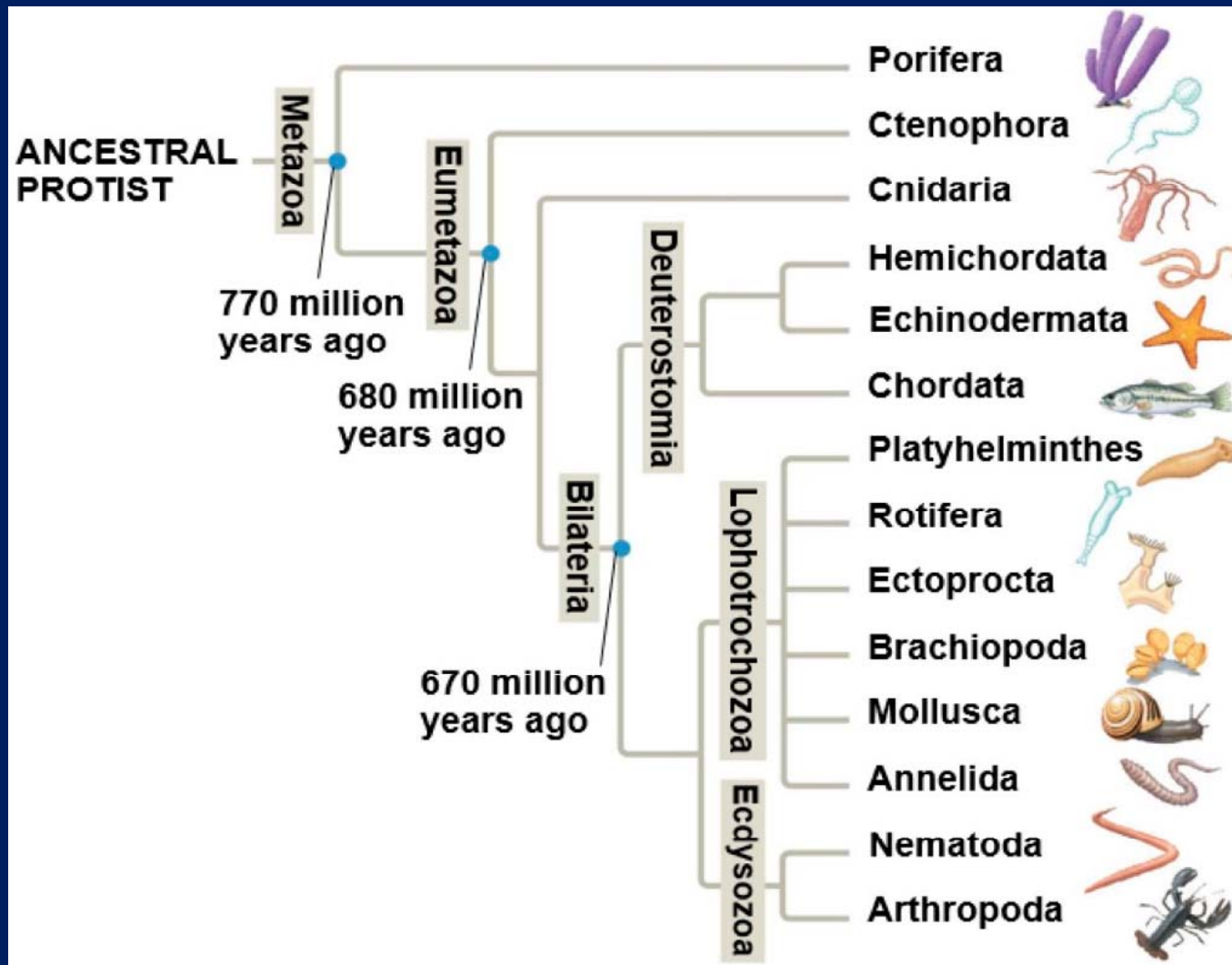


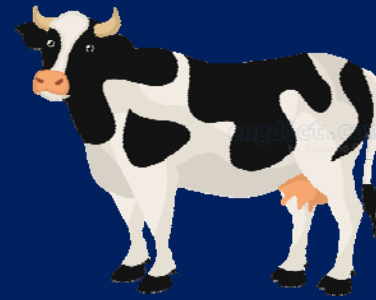
ในวิชาชีววิทยา ที่เกี่ยวกับวิวัฒนาการของสิ่งมีชีวิต นิสิตคงเคยเห็นแผนผังวิวัฒนาการ แบบนี้ เรียกว่า cladogram หรือ phylogenetic tree
วันนี้เราจะมาลองสร้างดูบ้าง โดยใช้ข้อมูลโปรตีนและโปรแกรม Bioinformatics ครับ



ลองเอาสิ่งมีชีวิต 4 ชนิดนี้มาเป็นตัวอย่างครับ คือ ปลา วัว คน ลิงชิมแปนซี โดยทั้ง 4 ชนิดนี้เป็นสัตว์มีกระดูกสันหลังทั้งหมด แต่ปลาต่างจากพวกที่เหลือ เพราะไม่ใช่สัตว์เลี้ยงลูกด้วยนม



ปลาปักเป้าจุดดำ



วัว



คน

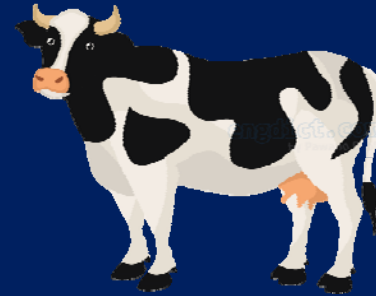


ลิงชิมแปนซี

ในอดีต การศึกษาวิวัฒนาการ อาศัยความรู้ทางกายวิภาค ซากฟอสซิล ฯลฯ ในการจัด
จำแนก ต่อมาเมื่อมีเทคโนโลยีชีวภาพ ความรู้ระดับโมเลกุล DNA โปรตีน ก็นำมาใช้ใน
การศึกษาวิวัฒนาการของสิ่งมีชีวิตได้อีกทางหนึ่งครับ



Tetraodon nigroviridis



Bos taurus

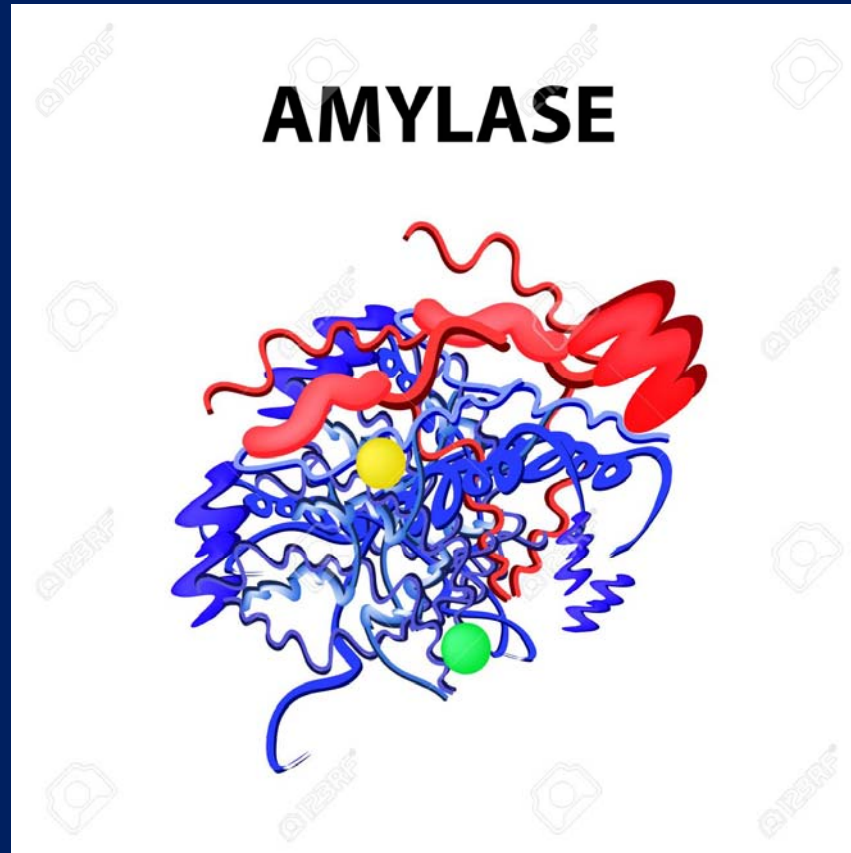


Homo sapiens



Pan troglodytes

สำหรับโปรตีนที่เลือกมาใช้ในการสร้าง cladogram ต้องเป็นโปรตีนที่พบในสิ่งมีชีวิตทั้งหมดที่เราศึกษา และมีข้อมูลลำดับกรดอะมิโนอยู่ในฐานข้อมูล ตัวอย่างวันนี้ ลองใช้เอนไซม์ amylase ซึ่งเป็นเอนไซม์ย่อยแป้งให้เป็นน้ำตาล พบในสิ่งมีชีวิตอย่างกว้างขวาง ตั้งแต่แบคทีเรีย เชื้อรา พืช และสัตว์ครึ่ง



จากนั้นก็นำคำว่า amylase กับชื่อวิทยาศาสตร์ของสิ่งมีชีวิตทั้ง 4 ชนิดไปค้นหาข้อมูลในฐานข้อมูล NCBI Protein แล้ว copy ลำดับกรดอะมิโน ในรูป FASTA format

[amylase \[Tetraodon nigroviridis\]](#)

1. 513 aa protein

Accession: CAD20312.1 GI: 20502194

[Nucleotide](#) [Taxonomy](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[Amylase, alpha 2A \(pancreatic\) \[Bos taurus\]](#)

6. 511 aa protein

Accession: AAI05234.1 GI: 75773673

[Nucleotide](#) [Taxonomy](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[Amylase, alpha 2A \(pancreatic\) \[Homo sapiens\]](#)

1. 511 aa protein

Accession: AAH07060.1 GI: 13937901

[Nucleotide](#) [PubMed](#) [Taxonomy](#) [Order cDNA Clone](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[alpha-amylase 2B isoform X1 \[Pan troglodytes\]](#)

2. 511 aa protein

Accession: XP_016812681.1 GI: 1034066219

[BioProject](#) [Nucleotide](#) [Taxonomy](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

ลำดับกรดอะมิโน ในรูป FASTA format เราคลิกดูได้ตรงลิงก์ FASTA ของแต่ละข้อมูล
ครับ

น่าสังเกตว่า วัว คน และลิงชิมแปนซี มีกรดอะมิโนของเอนไซม์ amylase 511 กรดอะมิ
โนเท่ากัน ส่วนปลาปักเป้าจุดดำ มี 513 กรดอะมิโน

[amylase \[Tetraodon nigroviridis\]](#)

1. [513 aa protein](#)

Accession: CAD20312.1 GI: 20502194

[Nucleotide](#) [Taxonomy](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[Amylase, alpha 2A \(pancreatic\) \[Bos taurus\]](#)

6. [511 aa protein](#)

Accession: AAI05234.1 GI: 75773673

[Nucleotide](#) [Taxonomy](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[Amylase, alpha 2A \(pancreatic\) \[Homo sapiens\]](#)

1. [511 aa protein](#)

Accession: AAH07060.1 GI: 13937901

[Nucleotide](#) [PubMed](#) [Taxonomy](#) [Order cDNA Clone](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[alpha-amylase 2B isoform X1 \[Pan troglodytes\]](#)

2. [511 aa protein](#)

Accession: XP_016812681.1 GI: 1034066219

[BioProject](#) [Nucleotide](#) [Taxonomy](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

เพื่อประหยัดเวลา นิสิตที่ลองทำตามตัวอย่างนี้ สามารถเอาตัวเลขหลัง GI: ไปค้นใน
ฐานข้อมูล NCBI โปรตีนได้โดยตรงครับ

[amylase \[Tetraodon nigroviridis\]](#)

1. 513 aa protein

Accession: CAD20312.1 GI: **20502194**

[Nucleotide](#) [Taxonomy](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[Amylase, alpha 2A \(pancreatic\) \[Bos taurus\]](#)

6. 511 aa protein

Accession: AAI05234.1 GI: **75773673**

[Nucleotide](#) [Taxonomy](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[Amylase, alpha 2A \(pancreatic\) \[Homo sapiens\]](#)

1. 511 aa protein

Accession: AAH07060.1 GI: **13937901**

[Nucleotide](#) [PubMed](#) [Taxonomy](#) [Order](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

[alpha-amylase 2B isoform X1 \[Pan troglodytes\]](#)

2. 511 aa protein

Accession: XP_016812681.1 GI: **1034066219**

[Nucleotide](#) [Taxonomy](#)

[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

NCBI Resources How To

Protein Protein 1034066219 Advanced

COVID-19 Information
[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2](#)

GenPept

alpha-amylase 2B isoform X1 [Pan troglodytes]
NCBI Reference Sequence: XP_016812681.1
[Identical Proteins](#) [FASTA](#) [Graphics](#)

เอาข้อมูลลำดับกรดอะมิโนของ amylase จากสัตว์ทั้ง 4 ชนิด มาแปะไว้ในโปรแกรม NotePad หรือ Word ก็ได้

```
Untitled - Notepad
File Edit Format View Help

>CAD20312.1 amylase [Tetraodon nigroviridis]
MRFFILVTLFGLSLAQHTPHFKHGRTSIVHLFEWRWTDIAAECERFLGPNGFAGVQISPPNEHIILDSPW
RPWWQRYQPIGYNLCSRSGSENELRDMITRCNNVGVNIYVDVINHMCRASSGEGTHSSCGSWFNAGTKD
FPSVSYSLDFNDHKCRTGSGNIESYDADVNVQRDCRLENLLDLALEKDYVRGKVADYMNRLVDMGVAGF
RVDACKHMWPGDLSAVYGRLLNLTNWFPGGSRPFYQEVIDLGSELITAKEYFHLGRVTEFKYGAKLGT
VFRKWNNEKLMYTKNWGEGWGFMPNGNAVVFIDNHDNQRGHGAGGAIVTFWDFRLHKMAVAYMLAHPYG
VTRVMSSFRWNRHIVNGKDQNDWMGPPSHPDGSKSVPINPDETCGDGWVCEHRWRQIKNMVIFRNVVNG
QPHSNWWDNNSNQVAFGRGNRGFIIFNDDWDLVTLNTGLPAGTYCDVISGQKEGGRCTGKQIHVGADG
RAHFRISNRDEDPFVAIHAESKL

>AAI05234.1 Amylase, alpha 2A (pancreatic) [Bos taurus]
MKFFLLLSVIVFCWAQYAPHTKTGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQISPPSENAVITDPS
RPWWERYQPVSYKLCSTRSGNESEFKDMVTRCANNVGVRIYVDAVINHMTGSGVSAGTSSTCGSYFNPGTRD
FPAVPYSGWDFNDEKCNTGNGEIKSYDVAYQVRDCRLVGLLDLALAKDYVRSTVAEYLNRLIDIGVAGFR
IDASKHMWPGDIKAVLDKHLNLTSWFPEGSRPFYQEVIDLGGETITSSDYFGNGRVTEFKYGVKLGTV
LRKWSGEKMAYLKNWGEWGFMPSDRALVFDNHDNQRGHGAGGASILTFWDARLYKMGVGFMLAHPYGF
TRVMSSYHWPRHFEDGKDVNDWVGPPNNGVIKEVTINPDTTCGNGWVCEHRWRQIRNMVIFRNVVDGQP
FTNWWDNNGSNQVAFGRGNKGFVFNDDWALSATLQTLPPGTYCDVISGDKIGDNCTVIEINVSCDGNA
YFSISNSAEDPFIAHTESKL

>AAH07060.1 Amylase, alpha 2A (pancreatic) [Homo sapiens]
MKFFLLLFITGFCWAQYSPNTQQGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQVSPPNENVAIYNPF
RPWWERYQPVSYKLCSTRSGNEDEFNRMVTRCANNVGVRIYVDAVINHMCNGNAVSAGTSSTCGSYFNPGSRD
FPAVPYSGWDFNDGKCKTGSGDIENYNDATQVRDCRLTGLLDLALAKDYVRSKIAEYMNHLIDIGVAGFR
LDASKHMWPGDIKAILDKHLNLSNWFPAGSKPFYQEVIDLGGEPIKSSDYFGNGRVTEFKYGAKLGTGTV
IRKWNGEKMSYLNWGEWGFVPSDRALVFDNHDNQRGHGAGGASILTFWDARLYKMAVGFMLAHPYGF
TRVMSSYRWPRQFQNGNDVNDWVGPPNNGVIKEVTINPDTTCGNDWVCEHRWRQIRNMVIFRNVVDGQP
FTNWWYDNGSNQVAFGRGNRGFVFNDDWSFSLTLQTLPPGTYCDVISGDKINGNCTGIKIYVSDDGKA
HFSISNSAEDPFIAHTESKL

>XP_016812681.1 alpha-amylase 2B isoform X1 [Pan troglodytes]
MKFFLLLFITGFCWAQYSPNTQQGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQVSPPNENVAIHNP
RPWWERYQPVSYKLCSTRSGNEDEFNRMVTRCANNVGVRIYVDAVINHMGSNAVSAGTSSTCGSYFNPGSRD
FPAVPYSGWDFNDGKCKTGSGDIENYNDATQVRDCRLVGLLDLALAKDYVRSKIAKYMNHLIDIGVAGFR
LDASKHMWPGDIKAILDKHLNLSNWFPAGSKPFYQEVIDLGGEPIKSSDYFGNGRVTEFKYGAKLGTGTV
IGKWNGEKMSYLNWGEWGFMPSDRALVFDNHDNQRGHGAGGASILTFWDARLYKMAVGFMLAHPYGF
TRVMSSYRWPRQFQNGNDVNDWVGPPNNGVIKEVTINPDTTCGNDWVCEHRWRQIRNMVIFRNVVDGQP
```


ตอนนี้เราก็มีข้อมูลลำดับกรดอะมิโนครบทั้ง 4 ชนิดแล้ว
สำหรับบรรทัดที่เป็นคำอธิบาย เริ่มด้วยเครื่องหมาย > นั้น เราสามารถเปลี่ยนให้เข้าใจ
ง่ายโดยแก้ไขให้เป็นชื่อสัตว์เฉย ๆ ก็ได้ แต่ต้องเป็นอักษรภาษาอังกฤษ ห้ามใช้อักษรไทย

```
Untitled - Notepad
File Edit Format View Help

>CAD20312.1 amylase [Tetraodon nigroviridis]
MRFFILVTLFGLSLAQHTPHFKHGRTSIVHLFEWRWTDIAAECERFLGPNGFAGVQISPPNEHIILDSPW
RPWWQRYQPIGYNLCRSRSGSENELRDMITRCNNVGVNIYVDVINHMCRASSGEGTHSSCGSWFNAGTKD
FPSVSYSYLDNFNDHKCRTGSGNIESYDADVNQVRDCRLENLLDLALEKDYVRGKVADYMNRLVDMGVAGF
RVDACKHMWPGDLSAVYGRLLNLTNWFPGGSRPFYQEVLDLSELITAKEYFHLGRVTEFKYGAKLGT
VFRKWNNEKLMYTKNWGEGWGFMPNGNAVVFIDNHDNQRGHGAGGAAIVTFWDFRLHKMAVAYMLAHPYG
VTRVMSSFRWNRHIVNGKDQNDWVGPPSHPDGSTKSVPINPDETCGDGWVCEHRWRQIKNMVIFRNVVNG
QPHSNWWDNNSNQVAFGRGNRGFIFNDDWDLVTLNTGLPAGTYCDVISGQKEGGRCTGKQIHVGADG
RAHFRISNRDEDPFVAIHAESKL

>AAI05234.1 Amylase, alpha 2A (pancreatic) [Bos taurus]
MKFFLLLSVIVFCWAQYAPHTKTGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQISPPSENAVITDPS
RPWWERYQPVSYKLCSTRSGNESEFKDMVTRCANNVGVRIYVDAVINHMTGSGVSAGTSSTCGSYFNPGRTRD
FPAVPYSGWDFNDEKCNTGNGEIKSYDVAYQVRDCRLVGLLDLALAKDYVRSTVAEYLNRLIDIGVAGFR
IDASKHMWPGDIKAVLDKHLNLTSWFPEGSRPFIYQEVLDLGGGETITSSDYFGNGRVTEFKYGVKLGTV
LRKWSGEKMYLKNWGEGWGFMPSDRALVFVDNHDNQRGHGAGGASILTFWDARLYKMGVGFMLAHPYGF
TRVMSSYHWPRHFEDGKDVNDWVGPPNNGVVIKEVTINPDTTCGNGWVCEHRWRQIRNMVIFRNVDGQP
FTNWWDNNGSNQVAFGRGNKGFIVFNDDWALSATLQTLPPGTYCDVISGDKIGDNCTVIEINVSCDGN
YFSISNSAEDPFIAIHESKL

>AAH07060.1 Amylase, alpha 2A (pancreatic) [Homo sapiens]
MKFFLLLFTIGFCWAQYSPNTQQGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQVSPPNENVAIYNPF
RPWWERYQPVSYKLCSTRSGNEDEFNRNMVTRCANNVGVRIYVDAVINHMCGNAVSAGTSSTCGSYFNPGSRD
FPAVPYSGWDFNDEKCKTGSGDIENYNDATQVRDCRLTGLLDLALAKDYVRSKIAEYMNHLIDIGVAGFR
LDASKHMWPGDIKAILDKHLNLSNWFPAGSKPFYQEVLDLGGPEIKSSDYFGNGRVTEFKYGAKLGTV
IRKWNGEKMSYLKNWGEGWGFVPSDRALVFVDNHDNQRGHGAGGASILTFWDARLYKMAVGFMLAHPYGF
TRVMSSYRWRPQFQNGNDVNDWVGPPNNGVVIKEVTINPDTTCGNDWVCEHRWRQIRNMVIFRNVDGQP
FTNWDYDNGSNQVAFGRGNRGFIVFNDDWSFSLTLQTLPPGTYCDVISGDKINGNCTGIKIYVSDDGKA
HFSISNSAEDPFIAIHAESKL

>XP_016812681.1 alpha-amylase 2B isoform X1 [Pan troglodytes]
MKFFLLLFTIGFCWAQYSPNTQQGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQVSPPNENVAIHNP
RPWWERYQPVSYKLCSTRSGNEDEFNRNMVTRCANNVGVRIYVDAVINHMMSGNAVSAGTSSTCGSYFNPGSRD
FPAVPYSGWDFNDEKCKTGSGDIENYNDATQVRDCRLTGLLDLALAKDYVRSKIAEYMNHLIDIGVAGFR
LDASKHMWPGDIKAILDKHLNLSNWFPAGSKPFYQEVLDLGGPEIKSSDYFGNGRVTEFKYGAKLGTV
IGKWNGEKMSYLKNWGEGWGFMPSDRALVFVDNHDNQRGHGAGGASILTFWDARLYKMAVGFMLAHPYGF
TRVMSSYRWRPQFQNGNDVNDWVGPPNNGVVIKEVTINPDTTCGNDWVCEHRWRQIRNMVIFRNVDGQP
```

ลองแก้คำอธิบายหลังเครื่องหมาย > ให้เหลือแค่ชื่อสัตว์ภาษาอังกฤษครับ

คือ Puffer Cow Human และ Chimpanzee

ตอนนี้เราได้ข้อมูลพร้อมสำหรับการสร้าง Cladogram แล้ว สัปดาห์หน้ามาดูกันต่อครับ

สำหรับไฟล์ .txt ของข้อมูลนี้ จะใส่ไว้ให้ในเว็บไซต์รายวิชาด้วย

```
CLAPexample - Notepad
File Edit Format View Help

>Puffer
MRFFILVTLFGLSLAQHTPHFKHGRTSIVHLFEWRWTDIAAECERFLGPNFGAGVQISPPNEHIILDSPW
RPWWQRYQPIGYNLCSRSSENELRDMITRCNNVGVNIYVDVINHMCRASSGEGTHSSCGSWFNAGTKD
FPSVSYSLDFNDHKCRTGSGNIESYDADVNQVRDCRLENLLDLALEKDYVRGKVADYMNRLVDMGVAGF
RVDACKHMWPGDLSAVYGRLLNLTNWFPGGSRPFYQEVIDLGSELITAKEYFHLGRVTEFKYGAKLGT
VFRKWNNEKLMYTKNWGEGWGFMPNGNAVVIDNHDNQRGHGAGGAAIVTFWDFRLHKMAVAYMLAHPYG|
VTRVMSSFRWNRHIVNGKDQNDWMGPPSHPDGSTKSPINPDETCDGDGWVCEHRWRQIKNMVIFRNVVNG
QPHSNWWDNNSNQVAFGRGNRGFIFNDDWDLVTLNTGLPAGTYCDVISGQKEGGRCTGKQIHVGADG
RAHFRISNRDEDPFVAIHAESKL

>Cow
MKFFLLLSVIVFCWAQYAPHTKTGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQISPPSENAVITDPS
RPWWERYQPVSYKLCRSGNESEFKDMVTRCNNVGVRIYVDVINHMTGSGVSAAGTSSSTCGSYFNPGRTRD
FPAVPYSGWDFNDEKCNTNGEIKSYDVAYQVRDCRLVGLLDLALAKDYVRSTVAEYLNRLIDIGVAGFR
IDASKHMWPGDIKAVLDKLNHLNLSWFPEGSRPFYQEVIDLGGETITSSDYFGNGRVTEFKYGVKLGTV
LRKWSGEKMAYLKNWGEWGFMPSDRALVFVDNHDNQRGHGAGGASILTFWDARLYKMGVGFMLAHPYGF
TRVMSSYHWPRHFEDGKDVNDWVGPPNNNGVIKEVTINPDTTCGNDWVCEHRWRQIRNMVIFRNVVDDGQP
FTNWWDNNGSNQVAFGRGNKGFIVFNDDWALSATLQTLPPGTCDVISGDKIGDNCTVIEINVSCDGN
YFSISNSAEDPFIAHITESKL

>Human
MKFFLLLTIGFCWAQYSPNTQQGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQVSPPNENVAIYNPF
RPWWERYQPVSYKLCRSGNEDEFNRNMVTRCNNVGVRIYVDVINHMCNAVSAGTSSSTCGSYFNPGRSRD
FPAVPYSGWDFNDGKCKTGSGDIENYNDATQVRDCRLTGLLDLALALEKDYVRSKIAEYMNHLIDIGVAGFR
LDASKHMWPGDIKAILDKLNHLNSNWFPAGSKPFYQEVIDLGGEPKSSDYFGNGRVTEFKYGAKLGTV
IRKWNGEKMSYLNKNWGEWGFVPSDRALVFVDNHDNQRGHGAGGASILTFWDARLYKMAVGFMLAHPYGF
TRVMSSYRWPRQFQNGNDVNDWVGPPNNNGVIKEVTINPDTTCGNDWVCEHRWRQIRNMVIFRNVVDDGQP
FTNWWYDNNGSNQVAFGRGNRGFIVFNDDWSFSLTLQTLPPGTCDVISGDKINGNCTGIKIYVSDDGKA
HFSISNSAEDPFIAHITESKL

>Chimpanzee
MKFFLLLTIGFCWAQYSPNTQQGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQVSPPNENVAIHNP
RPWWERYQPVSYKLCRSGNEDEFNRNMVTRCNNVGVRIYVDVINHMSGNAVSAGTSSSTCGSYFNPGRSRD
FPAVPYSGWDFNDGKCKTGSGDIENYNDATQVRDCRLVGLLDLALALEKDYVRSKIAKYMNHLIDIGVAGFR
LDASKHMWPGDIKAILDKLNHLNSNWFPAGSKPFYQEVIDLGGEPKSSDYFGNGRVTEFKYGAKLGTV
IGKWNGEKMSYLNKNWGEWGFMPSDRALVFVDNHDNQRGHGAGGASILTFWDARLYKMAVGFMLAHPYGF
TRVMSSYRWPRQFQNGNDVNDWVGPPNNNGVIKEVTINPDTTCGNDWVCEHRWRQIRNMVIFRNVVDDGQP
```