

## Bioinformatics Appendix for Human Physiology

### 1. Amino Acid Sequences For Species

>Unknown

QTAYAGDAANGMRVFNTYCSDCSITEGKNKIGPSLWNVVGPKPASPISDFNYSDAMRKNNDIIWTE  
DRISTYITNPQGLLPQVVKMAFPGLKDPQKCADVIQFLSQQH

>Homo\_sapiens\_1\_healthy

MGDVEKGKKIFIMKCSQCHTVEKGGKHKHTGPNLHGLFGRKTGQAPGYSYTAANKNKGIIWGEDT  
LMEYLENPKKYIPGTMIFVGIKKKEERADLIAYLKATNE

>Homo\_sapiens\_2\_disease

MGDVEKGKKIFIMKCSQCHTVEKGGKHKHTGPNLHGLFGRKTGQAPGYSYTAANKNKGIIWGEDT  
LMEYLENPKKYIPGTMIFVGIKKKEERADLIAYLKATNE

>Pan\_troglodytes

MGDVEKGKKIFIMKCSQCHTVEKGGKHKHTGPNLHGLFGRKTGQAAGYSYTVANKNKGITWGEDT  
LMEYLENPKKHIPETKMIFVGIKKKEERADLIAFLKATNE

>Apis\_mellifera

MGIPAGDPEKGGKIFVQKCAQCHTIESGGKHKVGNLYGVYGRKTGQAPGYSYTDANKGKGITW  
NKETLFEYLENPKKYIPGTMVFAGLKKPQERADLIAIEQASK

>Pseudomonas\_sp\_Ant30\_3

QAYAAGDAEDGGKLFKRICGGCHQVGESARGSFQPQLNGIFGRASGSTTDYVYSDAMKAAGIVW  
TRETLYAYLEDPKAVVPGTRMIFWGMSSDAQKIDDLVLAYLETFQPPE

>Caldithrix\_abyssi

ENSSFGEGMSLAELGAKLYKSKACVTCHSVDGSPLVGPTFKGVFGHTVKNLNDGSSVKADENYLRE  
SILKPQAKVVEGFQPMPTYSILKPREVDALIEYIKSLGE

### 2. Amino Acid Abbreviation Codes

Amino Acid	3-Letter Code	1-Letter Code
Alanine	Ala	A
Cysteine	Cys	C
Aspartic acid or aspartate	Asp	D
Glutamic acid or glutamate	Glu	E
Phenylalanine	Phe	F
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Lysine	Lys	K
Leucine	Leu	L
Methionine	Met	M
Asparagine	Asn	N
Proline	Pro	P
Glutamine	Gln	Q
Arginine	Arg	R
Serine	Ser	S
Threonine	Thr	T
Valine	Val	V
Tryptophan	Trp	W
Tyrosine	Tyr	Y

### 3. Description of Species

**Eukarya:**

1. **Unknown:** Mammal that lives in the ocean.
2. **Homo sapiens 1:** Healthy human individual.
3. **Homo sapiens 1:** Human individual with possible genetic disease in cytochrome C.

4. *Pan troglodytes*: Chimpanzee.
5. *Apis mellifera*: European honey bee.

**Bacteria:** (ocean dwellers)

6. *Pseudomonas species Ant30-3*: Aerobic bacteria found growing in the Antarctic in very, very cold places. It's a relative of *Pseudomonas aeruginosa*, which can be found in the environment and cause infections in humans such as pneumonia.
7. *Caldithrix abyssi*: Anaerobic isolated from a hydrothermal vent where temperature can rise up to 400°C. Uses nitrogen instead of oxygen in metabolism. Name literally means "cauldron (hot) threads of the abyss".

#### 4. Description of Programs and Databases

**BLAST** (Basic Local Alignment Search Tool) is a program that compares DNA and amino acid sequences to a known database of sequences of organisms. It calculates the significance of matches. BLAST can be used to test for genetic diseases. BLAST can also be used to infer functional and evolutionary relationships between sequences. BLAST is maintained by the National Center for Biotechnology Information (NCBI) which is a subdivision of the National Institutes of Health (NIH). Research groups worldwide contribute to the database.

**Seaview** is a graphical user interface that allows visualization of the alignment of multiple sequences. It can also organize phylogenetic trees based on evolutionary relationships. It is maintained by the institution Pôle Rhône-Alpes de Bioinformatique in France. Different research groups contribute to its database.

To download Seaview4

1. go to website: <http://doua.prabi.fr/software/seaview>
2. Go to "Download SeaView" section and select appropriate format
3. Click "Save file"
4. Open and extract Seaview
5. Go to location Seaview file is saved. Open application "seaview.exe"

#### **TOOLKIT CREDITS:**

Developed by Angela Gee, (Los Angeles Trade Tech Community College, CA) and Marissa Pantastico-Caldas, Los Angeles Trade Tech Community College (CA) with data and guidance by John Kirkpatrick (University of Rhode Island, RI) with and support by the rest of the C-DEBI Collaborative Toolkit Team.

#### **WEBSITE:**

[http://www.coexploration.org/C-DEBI/toolkits\\_biology.html](http://www.coexploration.org/C-DEBI/toolkits_biology.html)