# Find a Gene Project (BGGN 213)

Akshara Balachandra, A11933675

May 15, 2019

## Question 1: Name of the protein

**Name**: Sonic Hedgehog protein isoform 1 preprotein (SHH)

**Accession**: NP_000184.1

**Species**: Homo sapiens

## Question 2: BLAST search

**Method**: TBLASTN search against Xenopus ESTs

**Database**: Expressed sequence tags (est)

**Organism**: Xenopus (Taxid: 8353)

### Alignment details

**Chosen match:** Accession BX704643.1, a 874bp clone from Xenopus *tropicalis*.

```
>BX704643.1 BX704643 XGC-tadpole Xenopus tropicalis cDNA
clone TTpA011g13 5', mRNA sequence

cDNA clone TTpA011g13, mRNA sequence.
Length: 874

Score           Expect    Method                      Identities
387 bits(994)   3e-132    Compositional matrix adjust.  197/285(69%)


Positives       Gaps        Frame
220/285(77%)    23/285(8%)  +2


Query  68   RNSERFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEG  127
            RNS+RFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEG
Sbjct  41   RNSDRFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEG  220

Query  128  WDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYYESKAHIHCSVKA  187
            WDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVY+ESKAHIHCSVKA
Sbjct  221  WDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYFESKAHIHCSVKA  400

Query  188  ENSVAAKSGGCFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLTFLDRDDG  247
            ENSVAAKSGGCFPGSA V +E GGTK V++L PGDRVL++D QG L+YSDFL F+D++
Sbjct  401  ENSVAAKSGGCFPGSARVMVEPGGTKAVRELRPGDRVLSSDPQGNLIYSDFLLFIDKEHD  580
```

```
Query  248  AKKVFYVIETREPRERLLLTAAHLLFVAPHNDSATXXXXXXXXXXXXXXXXXXXXRALFAS  307
            KK++YVI+T + R R  +TAAHLLFVA  N + +                      +++FAS
Sbjct  581  VKKLYYVIQTSQTRIR--MTAAHLLFVAQSNGTGS------------------FKSVFAS  700


Query  308  RVRPGQRVYVVAERDGDRRLLPAAVHSVTLSEEAAGAYAPLTAQG  352
            VRPG  +Y   R  D  L   A V  V L EE  G + P     G
Sbjct  701  NVRPGDVIYSADRR--DMTLREAMVEKVDL-EEDIGGFCPRNCPG  826
```



Figure 1: BLAST parameters



Figure 2: BLAST search results

**BX704643 XGC-tadpole Xenopus tropicalis cDNA clone TTpA011g13 5', mRNA sequence**

Sequence ID: BX704643.1  Length: 874  Number of Matches: 2

Range 1: 41 to 826 GenBank  Graphics

| Score | Expect | Method | Identities | Positives | Gaps | Frame |
|---|---|---|---|---|---|---|
| 387 bits(994) | 3e-132 | Compositional matrix adjust. | 197/285(69%) | 220/285(77%) | 23/285(8%) | +2 |

```
Query  68   RNSERFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEG  127
            RNS+RFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEG
Sbjct  41   RNSDRFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEG  220

Query  128  WDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYYESKAHIHCSVKA  187
            WDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVY+ESKAHIHCSVKA
Sbjct  221  WDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYFESKAHIHCSVKA  400

Query  188  ENSVAAKSGGCFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLTFLDRDDG  247
            ENSVAAKSGGCFPGSA V +E GGTK V++L PGDRVL++D QG L+YSDFL F+D++
Sbjct  401  ENSVAAKSGGCFPGSARVMVEPGGTKAVRELRPGDRVLSSDPQGNLIYSDFLLFIDKEHD  580

Query  248  AKKVFYVIETREPRERLLLTAAHLLFVAPHNDSATXXXXXXXXXXXXXXXXXXXXRALFAS  307
            KK++YVI+T + R R  +TAAHLLFVA  N + +                    +++FAS
Sbjct  581  VKKLYYVIQTSQTRIR--MTAAHLLFVAQSNGTGS-----------------FKSVFAS  700

Query  308  RVRPGQRVYVVAERDGDRRLLPAAVHSVTLSEEAAGAYAPLTAQG  352
            VRPG +Y   R D L  A V  VL EE G + P   G
Sbjct  701  NVRPGDVIYSADRR--DMTLREAMVEKVDL-EEDIGGFCPRNCPG  826
```

Figure 3: Alignment details

# Question 3: Novel Protein Information

**Chosen Sequence:**

```
>Xenopus tropicalis protein (from BLAST result)
RNSDRFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEG
WDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYFESKAHIHCSVKA
ENSVAAKSGGCFPGSARVMVEPGGTKAVRELRPGDRVLSSDPQGNLIYSDFLLFIDKEHD
VKKLYYVIQTSQTRIRMTAAHLLFVAQSNGTGSFKSVFASNVRPGDVIYSADRRDMTLRE
AMVEKVDLEEDIGGFCPRNCPG
```

**Name:** *Xenopus* sonic hedgehog

**Species:** Xenopus (Silurana) tropicalis (Common name: western clawed frog, Taxid: 8364)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Pipoidea; Pipidae; Xenopodinae; Xenopus; Silurana

# Question 4: Proof that gene/protein are novel

A BLASTP search against the NR database yielded a top hit to a protein from *Xenopus tropicalis* with 100% coverage and 97.71% identity. The alignment details are given in the screenshots below.

Figure 4: BLASTP parameters



Figure 5: BLASTP results page

**PREDICTED: sonic hedgehog protein [Xenopus tropicalis]**

Sequence ID: XP_017950257.1  Length: 416  Number of Matches: 1

Range 1: 68 to 329 GenPept  Graphics                          ▼ Next Match  ▲ Previous Match

| Score | Expect | Method | Identities | Positives | Gaps |
|---|---|---|---|---|---|
| 533 bits(1373) | 0.0 | Compositional matrix adjust. | 256/262(98%) | 256/262(97%) | 0/262(0%) |

```
Query  1    RNSDRFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEG   60
            RNSDRFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEG
Sbjct  68   RNSDRFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEG   127

Query  61   WDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYFESKAHIHCSVKA   120
            WDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYFESKAHIHCSVKA
Sbjct  128  WDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYFESKAHIHCSVKA   187

Query  121  ENSVAAKSGGCFPGSARVMVEPGGTKAVRELRPGDRVLSSDPQGNLIYSDFLLFIDKEHD   180
            ENSVAAKSGGCFPGSARVMVEPGGTKAVRELRPGDRVLSSDPQGNLIYSDFLLFIDKEHD
Sbjct  188  ENSVAAKSGGCFPGSARVMVEPGGTKAVRELRPGDRVLSSDPQGNLIYSDFLLFIDKEHD   247

Query  181  VKKLYYVIQTSQTRIRMTAAHLLFVAQSNGTGSFKSVFASNVRPGDVIYSADRRDMTLRE   240
            VKKLYYVIQTSQTRIRMTAAHLLFVAQSNGTGSFKSVFASNVRPGDVIYSADRRDMTLRE
Sbjct  248  VKKLYYVIQTSQTRIRMTAAHLLFVAQSNGTGSFKSVFASNVRPGDVIYSADRRDMTLRE   307

Query  241  AMVEKVDLEEDIGGFCPRNCPG   262
            AMVEKVDLEEDIG F P     G
Sbjct  308  AMVEKVDLEEDIGAFAPVTAQG   329
```

Figure 6: Alignment for top result

# Question 5: Multiple sequence alignment

## Relabeled sequences for alignment

```
>Xenopus_sonic_hedgehog (from BLAST result)
RNSDRFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEG
WDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYFESKAHIHCSVKA
ENSVAAKSGGCFPGSARVMVEPGGTKAVRELRPGDRVLSSDPQGNLIYSDFLLFIDKEHD
VKKLYYVIQTSQTRIRMTAAHLLFVAQSNGTGSFKSVFASNVRPGDVIYSADRRDMTLRE
AMVEKVDLEEDIGGFCPRNCPG

>Human |Sonic hedgehog protein (NP_000184.1) [Homo sapiens]
MLLLARCLLLVLVSSLLVCSGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTLGASGRYEGKISRNS
ERFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEGWDEDGHHSEESLH
YEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYYESKAHIHCSVKAENSVAAKSGGCFPGSATVHLEQG
GTKLVKDLSPGDRVLAADDQGRLLYSDFLTFLDRDDGAKKVFYVIETREPRERLLLTAAHLLFVAPHNDS
ATGEPEASSGSGPPSGGALGPRALFASRVRPGQRVYVVAERDGDRRLLPAAVHSVTLSEEAAGAYAPLTA
QGTILINRVLASCYAVIEEHSWAHRAFAPFRLAHALLAALAPARTDRGGDSGGGDRGGGGGRVALTAPGA
ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMAVKSS

>Mexican_walking_fish |sonic hedgehog [Ambystoma mexicanum] ARX70249.1
MGEMRLLLRRLLLGTLVWALLVPSGLTCGPGRGIGKRRQPKKLTPLAYKQFIPNVAEKTLGASGRYEGKI
TRNSDRFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEGWDEDGHHSE
ESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYFESKAHIHCSVKAENSVAAKSGGCFPASAKVT
LEHGVTRPVKDLRPGDRVLAADGQGRLVYS DFLMFLDKEEAVTKVFYVIETERPRQRLRLTAAHLLFAA
RHPANSSSSTGFQSIFASRVRPGHRVLTVDQEGRGLQEATVTRVYLEEGAGAYAPVTSHGTVVIDKVLAS
CYAVIEEHSWAHWAFAPLRLGYGILSIFSPQDYSPHSPPAPSQKEGVHWYSEILYHIGTWVLHSDTIHPW
GMAAKSS
```

```
>Iberian_newt |sonic hedgehog-related protein [Pleurodeles waltl]
MGEMILLRRDLLGGFISTLLVPLGLGCGPGRGIGQRRRPQKLTPLAYKQFIPNVSEKTLGASGRYEVKITRNSERFQELT
PNYNPDIIFKVEENTGADRLMTQRCKDKLNSLAISVMNQWPGVKLRVTEGWDEDGHHSEESLHYEGRAVDITTSDRDRSK
YGMLARLAVEAGFDWVYFESKAHIHCSVKAENSVAVKSGGCFPGSATVTLEQGVRIPVKGLKPGNRVLAVDVEGRLIYSD
FLLFMDEEETARKVFYVIETSLPRERLRLTAAHLLFVAQEHPGNASAGNFRSKFGHRRFRSMFASSVRPGHRVLTEDREG
RGLREATVDRVYLEEATGAYAPVTAHGTVVIDRVLASCYAVIEEHSWAHWAFAPLRVGFGILSFFSPQDYSSHSPPAPSQ
AEGVHWYSEILYRIGTWVLQADTIHPLGMAAKSS

>African_clawed_frog |sonic hedgehog protein-like|XP_018124514.1 [Xenopus laevis]
MSVATGILLLGFTCSLLIPPGLSCGPGRGIGKRRHPKKLTPLVYKQFIPNVAEKTLGASGRYEGKIASNSDRFKELTPNY
NPDIVFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEGWDEDGHHLEESLHYEGRAVDITTSDRDRSKYGM
LARLAVEAGFDWVYYESKAHIHCSVKAENSVAAKSGGCFPAGAEVMVELGGTKAVKDLRPGDRVLSSDAQGNLLYSDFLL
FIDQQHDVKKLFYVIETSHPRRRIRLTAAHLLFVAQVNGSRSFKSVFASSIQPGDIIYSADPRAMALNEVKVEKVGLEED
TGAYAPLTAHGTLVIDQVLSSCYALIEEHTWAHLAFAPLRFGLSLSPYLFPKDSSPSPQSHHQLDSSPSPQSHHQPEGIH
WYSQLLYQLGTWLLDSNSLHPLGMATKSS

>Burmese_python |sonic hedgehog protein|XP_007433256.1 [Python bivittatus]
MLLRRRSGLLPLCLGALFLSAGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTLGASGRYEGKISRNSERFKELTPN
YNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEGWDEDGHHSEESLHYEGRAVDITTSDRDRSKYG
MLARLAVEAGFDWVYYESKAHIHCSVKAENSVAAKSGGCFPGSAWVNLEEGGTKLVKDLRPGDRVLAADVHGQLFYSEFL
AFLDREDPPVHKLFYVIETQRPQTRLLLTAAHLLFVAPPQNHSHSQPQPIFASRVQPGQQVYVLGQGGQTLLQAAVHSVS
LQEEASGAYAPLTAQGTILINQVLASCYAVIEEHSWAHWAFAPFRMAHAVLAMLDPEGLSSPLLFPAAAPEESSSLAGVH
WYSRLLYHIGRWILDSETMHPLGMAS

>Tiger_snake |sonic hedgehog protein |XP_026536709.1 [Notechis scutatus]
MLLLRRTGLLPLCLGALFLSAGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTLGASGRYEGKITRNSERFKELTPN
YNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEGWDEDGHHSEESLHYEGRAVDITTSDRDRSKYG
MLARLAVEAGFDWVYYESKAHIHCSVKAENSVAAKSGGCFPGSAWVNLEGGGTKLVKDLHPGDRVLAADVHGQLFYSEFL
AFLDREEPPVHKLFYVIETQTPQARLLLTAAHLLFVAPPHNHSHSQPQPIFASRVQPGQHVYVLAQGGQTLLPAAVHSVS
LKEEASGAYAPLTAQGTILINQVLASCYAVIEEHSWAHWAFAPFRIAHAALAVLNLEGLSSPLLFPAAVPEEGASLAGVH
WYSRLLYYIGRWILGSEMMHPLGMAS
```

## Alignment

Using local MUSCLE (v3.8) program with bio3d `seqaln()`. The fasta output was converted to CLUSTALW format using `seaview`.

```
CLUSTAL multiple sequence alignment by MUSCLE (3.8)


Iberian_newt            -MGEMILLRRDLLGGF-ISTLLVPLGLGCGPGRGIGQRRRPQKLTPLAYKQFIPNVSEKT
Mexican_walking_fish    MGEMRLLLRRLLLGTL-VWALLVPSGLTCGPGRGIGKRRQPKKLTPLAYKQFIPNVAEKT
African_clawed_frog     ----MSVATGILLLGF-TCSLLIPPGLSCGPGRGIGKRRHPKKLTPLVYKQFIPNVAEKT
Xenopus_sonic_hedgehog  ------------------------------------------------------------
Burmese_python          ----MLLRRRSGLLPLCLGALFLSAGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKT
Tiger_snake             ----MLLLRRTGLLPLCLGALFLSAGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKT
Human                   ----MLLLARCLLLVL-VSSLLVCSGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKT


Iberian_newt            LGASGRYEVKITRNSERFQELTPNYNPDIIFKVEENTGADRLMTQRCKDKLNSLAISVMN
Mexican_walking_fish    LGASGRYEGKITRNSDRFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMN
African_clawed_frog     LGASGRYEGKIASNSDRFKELTPNYNPDIVFKDEENTGADRLMTQRCKDKLNALAISVMN
Xenopus_sonic_hedgehog  -----------RNSDRFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMN
Burmese_python          LGASGRYEGKISRNSERFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMN
Tiger_snake             LGASGRYEGKITRNSERFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMN
```

```
Human                  LGASGRYEGKISRNSERFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMN


Iberian_newt           QWPGVKLRVTEGWDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYF
Mexican_walking_fish   QWPGVKLRVTEGWDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYF
African_clawed_frog    QWPGVKLRVTEGWDEDGHHLEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYY
Xenopus_sonic_hedgehog QWPGVKLRVTEGWDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYF
Burmese_python         QWPGVKLRVTEGWDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYY
Tiger_snake            QWPGVKLRVTEGWDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYY
Human                  QWPGVKLRVTEGWDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYY


Iberian_newt           ESKAHIHCSVKAENSVAVKSGGCFPGSATVTLEQGVRIPVKGLKPGNRVLAVDVEGRLIY
Mexican_walking_fish   ESKAHIHCSVKAENSVAAKSGGCFPASAKVTLEHGVTRPVKDLRPGDRVLAADGQGRLVY
African_clawed_frog    ESKAHIHCSVKAENSVAAKSGGCFPAGAEVMVELGGTKAVKDLRPGDRVLSSDAQGNLLY
Xenopus_sonic_hedgehog ESKAHIHCSVKAENSVAAKSGGCFPGSARVMVEPGGTKAVRELRPGDRVLSSDPQGNLIY
Burmese_python         ESKAHIHCSVKAENSVAAKSGGCFPGSAWVNLEEGGTKLVKDLRPGDRVLAADVHGQLFY
Tiger_snake            ESKAHIHCSVKAENSVAAKSGGCFPGSAWVNLEGGGTKLVKDLHPGDRVLAADVHGQLFY
Human                  ESKAHIHCSVKAENSVAAKSGGCFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLY


Iberian_newt           SDFLLFMDEEE-TARKVFYVIETSLPRERLRLTAAHLLFVAQEHPGNASAGNFRSKFGHR
Mexican_walking_fish   SDFLMFLDKEE-AVTKVFYVIETERPRQRLRLTAAHLLFAAR-HPANSSSST--------
African_clawed_frog    SDFLLFIDQQH-DVKKLFYVIETSHPRRRIRLTAAHLLFVAQ----VNGSRS--------
Xenopus_sonic_hedgehog SDFLLFIDKEH-DVKKLYYVIQTS--QTRIRMTAAHLLFVAQ----SNGTGS--------
Burmese_python         SEFLAFLDREDPPVHKLFYVIETQRPQTRLLLTAAHLLFVAP--PQNHSHSQ--------
Tiger_snake            SEFLAFLDREEPPVHKLFYVIETQTPQARLLLTAAHLLFVAP--PHNHSHSQ--------
Human                  SDFLTFLDRDD-GAKKVFYVIETREPRERLLLTAAHLLFVAP--HNDSATGEPEASSGSG


Iberian_newt           -------RFRSMFASSVRPGHRVLTEDR--EGRGLREATVDRVYL-EEATGAYAPVTAHG
Mexican_walking_fish   -------GFQSIFASRVRPGHRVLTVDQ--EGRGLQEATVTRVYL-EEGAGAYAPVTSHG
African_clawed_frog    --------FKSVFASSIQPGDIIYSADP--RAMALNEVKVEKVGL-EEDTGAYAPLTAHG
Xenopus_sonic_hedgehog --------FKSVFASNVRPGDVIYSADR--RDMTLREAMVEKVDL-EEDIGGFCPRNCPG
Burmese_python         --------PQPIFASRVQPGQQVYVLGQ--GGQTLLQAAVHSVSLQEEASGAYAPLTAQG
Tiger_snake            --------PQPIFASRVQPGQHVYVLAQ--GGQTLLPAAVHSVSLKEEASGAYAPLTAQG
Human                  PPSGGALGPRALFASRVRPGQRVYVVAERDGDRRLLPAAVHSVTLSEEAAGAYAPLTAQG


Iberian_newt           TVVIDRVLASCYAVIEEHSWAHWAFAPLRVGFGILSFFSPQ-------------------
Mexican_walking_fish   TVVIDKVLASCYAVIEEHSWAHWAFAPLRLGYGILSIFSPQ------------------
African_clawed_frog    TLVIDQVLSSCYALIEEHTWAHLAFAPLRFGLSLSPYLFPK-----------DSSPSPQS
Xenopus_sonic_hedgehog ------------------------------------------------------------
Burmese_python         TILINQVLASCYAVIEEHSWAHWAFAPFRMAHAVLAMLDPE----------------GLS
Tiger_snake            TILINQVLASCYAVIEEHSWAHWAFAPFRIAHAALAVLNLE---------------GLS
Human                  TILINRVLASCYAVIEEHSWAHRAFAPFRLAHALLAALAPARTDRGGDSGGGDRGGGGGR


Iberian_newt           ---DYSSHSPPAPS-QAEGVHWYSEILYRIGTWVLQADTIHPLGMAAKSS
Mexican_walking_fish   ---DYSPHSPPAPS-QKEGVHWYSEILYHIGTWVLHSDTIHPWGMAAKSS
African_clawed_frog    HHQLDSSPSPQSHH-QPEGIHWYSQLLYQLGTWLLDSNSLHPLGMATKSS
Xenopus_sonic_hedgehog --------------------------------------------------
Burmese_python         SPLLFPAAAPEESS-SLAGVHWYSRLLYHIGRWILDSETMHPLGMAS---
Tiger_snake            SPLLFPAAVPEEGA-SLAGVHWYSRLLYYIGRWILGSEMMHPLGMAS---
```

7

```
Human                              VALTAPGAADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMAVKSS
```

# Question 6: Phylogenetic tree

I used the simple phylogeny tool from EBI with the following parameters:

- Tree format: `nj`
- Distance correction: `False`
- Excludes Gaps: `False`
- Clustering Methods: `Neighbor-joining`
- Percent identity matrix: `False`

The resulting tree is this.



Figure 7: Phylogenetic tree derived from the sequence alignments in the previous section
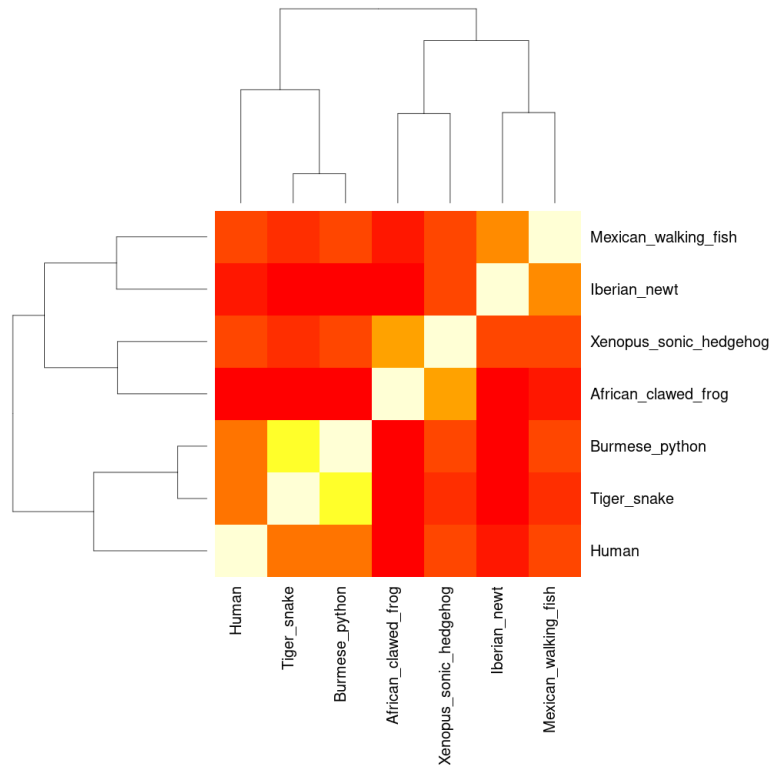
# Question 7: Heatmap of aligned sequences



Figure 8: Heatmap derived from the sequence alignments.

# Question 8: Protein Database Search