

# Our Animal DNA: Comparing genes across the Tree of Life

## Practical 1: find a protein sequence and run BLAST

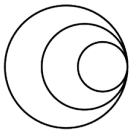
### Step 1: Find the gene

We're going to find the protein sequence of the human gene *KRT71*. To find it, we're going to go to [ensembl.org](https://ensembl.org).

The screenshot shows the Ensembl genome browser homepage. At the top, there is a navigation bar with links for BLAST/BLAT, VEP, Tools, BioMart, Downloads, Help & Docs, and Blog. A search bar is located in the top right corner, with the text "Search all species..." and a magnifying glass icon. Below the navigation bar, there are several sections: "Tools" with links to BioMart, BLAST/BLAT, and Variant Effect Predictor; a "Search" section with a dropdown menu for "All species" and a search box; "All genomes" with a dropdown menu for "-- Select a species --"; and "Favourite genomes" with links to Human, Mouse, and Zebrafish. A red arrow points to the search box, and the word "Search" is written in red text next to it. On the right side of the page, there is a "News" section with several items, including "Ensembl Release 101 (August 2020)" and "Other news from our blog".

Type the gene name *KRT71* into the search box and hit [Go](#).

Human *KRT71* should be your first search result.



wellcome  
connecting  
science

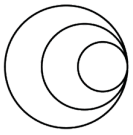
EMBL-EBI



The screenshot shows the Ensembl search interface. At the top, there are navigation links: BLAST/BLAT, VEP, Tools, BioMart, Downloads, Help & Docs, and Blog. Below this is a search bar with 'krt71' entered and a search button. The results show 460 matches. A sidebar on the left allows filtering by category (Gene: 117, Transcript: 143, Phenotype: 1, GeneTree: 94, ProbeFeature: 14, Protein Family: 91) and species (Human: 4, Mouse: 34, African green monkey: 4, African savanna elephant: 4, Algerian mouse: 4, Alpine marmot: 4, and 86 more species). The main content area lists the top results: KRT71 (Human Gene) with ENSG00000139648, KRT71-201 (Human Transcript) with ENST00000267119, and Krt71 (Mouse Gene, Strain: reference (CL57BL6)) with ENSMUSG00000051879. Each result includes a brief description and links to various data tables.

Click on the top result to go the gene page, which looks like this:

The screenshot shows the Ensembl gene page for KRT71 (ENSG00000139648) in Human (GRCh38.p13). The page is divided into a left sidebar with navigation options (Summary, Splice variants, Transcript comparison, Gene alleles, Sequence, Secondary Structure, Comparative Genomics, Genomic alignments, Gene tree, Gene gain/loss tree, Orthologues, Paralogues, Ensembl protein families, Ontologies, Genetic Variation, Variant table, Variant image, Structural variants, Gene expression, Pathway, Regulation, External references, Supporting evidence, ID History, Gene history) and a main content area. The main content area includes a 'Summary' section with fields for Name, CCDS, UniProtKB, RefSeq, and Annotation method. Below this is a genomic track visualization showing contigs, genes, regulatory build, and enhancers. Red arrows and text annotations highlight key features: 'More things to see about the gene' points to the 'Genetic Variation' section in the sidebar; 'Export, configure, share and bookmark' points to the 'Configure this page', 'Custom tracks', 'Export data', 'Share this page', and 'Bookmark this page' buttons at the bottom of the sidebar.



You can find out lots of information about the gene by clicking on some of the links on the left of the page, including its sequence, its orthologues, genetic variation between individuals, where the protein is found and processes the protein is involved in.

## Step 2: Export the sequence

We want to export the protein sequence. Click on the blue button [Export data](#) (see above image).

The protein sequence is also known as the **peptide sequence**. Select [peptide sequence](#) on the page. You will also need to deselect any other selected options, including the Genomic sequence, where you need to select “None” from the drop-down. Then click [Next>](#).

Export data

### Export Configuration - Feature List

**Tip**  
For sequence export, please go to the relevant sequence page (see lefthand menu) and use the new "Download sequence" button

Gene to export: ENSG00000139648.7 (KRT71)

Output: FASTA sequence \*

Strand: Feature strand \*

5' Flanking sequence (upstream): 0 \* (Maximum of 1000000)

3' Flanking sequence (downstream): 0 \* (Maximum of 1000000)

[Next >](#)

Fields marked \* are required

#### Options for FASTA sequence

Genomic: None \*

Select/deselect all:

cDNA:

Coding sequence:

Peptide sequence:

5' UTR:

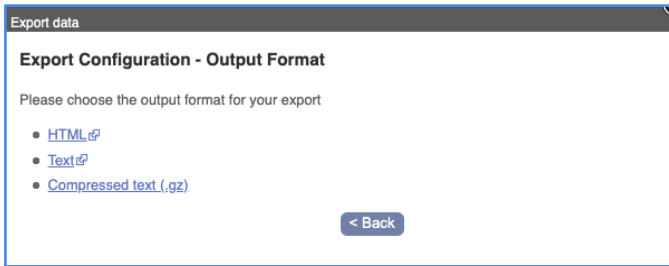
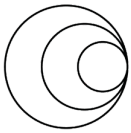
3' UTR:

Exons:

Introns:

Fields marked \* are required

You can export in different formats, we'll go for [Text](#).



This sequence is in a format called **FASTA**. FASTA is the standard way of writing sequences for both proteins and DNA. It consists of a header line at the top, which is indicated by a > sign at the beginning of the line and contains information about what the sequence is, and lines of sequence of fixed length (usually 60 bases or amino acids) underneath. This is what the sequence should look like:

```

>ENSG00000139648:ENST00000267119.6 peptide: ENSP00000267119 pep:protein_coding
MSRQFTCKSGAAAKGGFSGCSAVLSGGSSSSFRAGSKGLSGGFGSRSLSLGGVRSLNVA
SGSGKSGGYGFGRGRASGFAGSMFGSVALGPVCPTVCPGGIHQVTVNESLLAPLNVELD
PEIQVRAQEREQIKALNNKFASFIDKVRFLQQNQVLETKWELLQQLDLNCKNNLEPI
LEGYISNLRKQLETLSGDRVRLDSELNVRDVVEDYKKRYEEEINKRTAAENEFVLLKKD
VDAAYANKVELQAKVESMDQEIKFFRCLFEAEITQIQSHISDMSVILSMDNRRNLDLDSI
IDEVRTQYEEIALKSKAEAEALYQTKFQELQLAAGRHGDDLKNTKNEISELTRLIQRIRS
EIENVKQASNLETAIADAEQRGNALKDARAKLDELEGALHQAKEELARMLREYQELMS
LKLALDMEIATYRKLLESEECRMSGEFPPVSIISSTSGGSVYGFRRPSMVSGGYVANS
SNCISGVCSVRGGEGRSRGSANDYKDTLKGSSLSAPSCKTSR

```

**Header** (points to the line starting with >)

**Amino acid 1-letter codes** (points to the sequence lines)

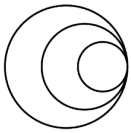
If you have something different, such as extra sequence beyond these 10 lines, go back to the export menu and make sure you've deselected some of the extra options, especially the **Genomic** option.

We can use this sequence in BLAST.

### Step 3: Open BLAST

We're going to use this human gene again for Step 9, so keep this tab open.

Open a new tab in your browser and go to [rapid.ensembl.org](http://rapid.ensembl.org). We're going to use this sequence to find the *KRT71* gene in Darwin Tree of Life species.



Tools

**BLAST >**

[All tools](#)

Search our genomes for your DNA or protein sequence

**BLAST**

Search

Camarhynchus parvulus (Small) for

Go

e.g. [Camarhynchus parvulus 2:361680-384534](#) or [Clytia hemisphaerica IPR001650](#)

Ensembl Rapid Release is a new site designed to make our data available more quickly. Release of data occurs on a two-week cycle, meaning we can make our gene sets available with minimal delay once the annotation is complete. For each species we provide a gene set along with additional features such as protein feature annotation and BLAST functionality.

It is important to note that Ensembl Rapid Release is by nature not as fully featured as a typical data release on [www.ensembl.org](http://www.ensembl.org). Currently we do not provide homology data, gene symbol assignment (which uses homology data), data archiving or programmatic access. We are however working on adding more functionality over the coming months to further improve usability.

[More details](#) about Ensembl Rapid Release and the current and planned features.

All genomes

-- Select a species --

[View and download available data for all species](#)

Favourite genomes



**Camarhynchus parvulus (Small tree finch) -**  
GCA\_902806625.1  
Camarhynchus\_parvulus\_V1.1



**Clytia hemisphaerica (Z4C2) -**  
GCA\_902728285.1  
GCA902728285v1

Latest Genomes

We have 10 new species this release:

- [Carassius auratus red var.](#) (Red crucian carp) - GCA\_009069565.1
- [Corvus moneduloides](#) (New Caledonian crow) - GCA\_009650955.1
- [Daubentonia madagascariensis](#) (Aye-aye) - GCA\_004027145.1
- [Denticeps clupeioides](#) (Denticle herring) - GCA\_900700375.2
- [Ficedula albicollis](#) (Collared flycatcher) - GCA\_000247815.2
- [Hipposideros galeritus](#) (Cantor's roundleaf bat) - GCA\_004027415.1
- [Meleagris gallopavo](#) (Turkey) - GCA\_000146605.4
- [Ovis aries](#) (Sheep) - GCA\_000298735.2
- [Verasper variegatus](#) (Spotted halibut) - GCA\_013332515.1
- [Zalophus californianus](#) (California sea lion) - GCA\_009762305.2

[View all species and download data](#)

Click on **BLAST** to open it. This is what the interface looks like:

**BLAST search**

New job Clear form

**Sequence data:**

Maximum of 30 sequences (type in plain text, FASTA or sequence ID)

Paste in the sequence

Or upload sequence file  No file chosen

DNA  
 Protein

**Search against:**

X

[Add/remove species](#)

DNA database   
 Protein database

Choose the species to search

Choose what sequences to match to

**Search tool:**

**Search Sensitivity:**

**Description (optional):**

**Additional configurations:**

#### Step 4: add the sequence

Copy and paste the sequence we found of human *KRT71* into the box. This should include the header line (starts with >) and all the lines of protein one-letter codes.

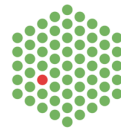
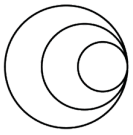
**Sequence data:**

```
>ENSG00000139648:ENST00000267119.6 peptide: ENSP00000267119 pep:prc
MSRQFTCKSGAAAKGGFSGCSAVLSGGSSSSFRAGSKGLSGGFGSRSLYSLGGVRSLNVA
SSGKSGGYGFRGRASGFAGSMFGSVALGPVCPVCPGPIHQVTVNESLLAPLNVELD
PEIQKVRAQEREQIKALNNKFASFIDKVRFLFLEQQNQVLETWELLQQLDNNCKNNLEPI
LEGYISNLRKQLETLSGDRVRLDSELNRNVDVVEDYKKRYEEEINKRTAAENEFVLLKGD
VDAAYANKVELQAKVESMDQEIKFFRCLFEAEITQIQSHISDMSVILSMDNNRNLDDLSI
IDEVTRTQYEEIALKSKAEAEALYQTKFQELQLAAGRHGDDLKNTKNEISELTRLIQRIRS
EIEENVKQASNLETAIADAEQRGDNALKDARAKLDELEGALHQAKEELARMLREYQELMS
LKLALDMEIATYRKLLESECRMSGEFPPSPVSIISSTSGGSVYGFPRPSMVSGGYVANS
SNCISGVCSVRGGEGRSRGSANDYKDTLKGSSLSAPSKKTSR
```

[Add more sequences](#) (1 sequence added, 29 more sequences allowed)

DNA  
 Protein

The tool has automatically detected that the sequence is a protein sequence.



### Step 5: Choose species to search against

Click on [Add/remove species](#) to choose some species.

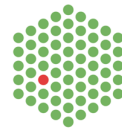
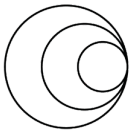
The screenshot shows the 'Add/remove species' window. On the left, the 'Species Selector' has a search box with the text 'Start typing the name of a species...'. Below it are tabs for 'All Divisions' and 'Other Vertebrates', with sub-tabs for 'Birds and Reptiles', 'Bony fishes', and 'Others'. A list of species is shown under 'Birds and Reptiles', with *Camarhynchus parvulus (Small tree finch)* selected. On the right, the 'Selected species' panel shows two instances of *Camarhynchus parvulus (Small tree finch)*. Red arrows and text annotations highlight the search box ('Search for a species'), the selected species in the list ('Choose species by type'), and the species in the 'Selected species' panel ('Species you've selected').




There is a finch species selected by default. You will need to remove this by clicking on the cross next to the species on the right.

Find species by typing in their name. We're going to add Sea otter first. Its latin name is *Enhydra lutris kenyoni*, so type that into the search box.

The screenshot shows the 'Add/remove species' window with the search box containing 'Enhydra lutris kenyoni'. A dropdown menu is open, showing the search results: 'Enhydra lutris kenyoni (Enhydra\_lutris\_kenyoni)'. Below the search box, the 'Birds and Reptiles' category is expanded, showing 'Accipiter gentilis'.




Click on the species to add it to your list of species to search. Now find Black swan (*Cygnus atratus*) and Atlantic cod (*Gadus morhua*) in the same way. We now have a mammal, a bird and a fish to compare to.



Selected species <span>3</span>	
	Gadus morhua (Atlantic cod) - GCA... X
	Cygnus atratus (Black swan) - GCA... X
	Enhydra lutris kenyon (Sea otter) - ... X

Click on [Apply](#) to select these species.

These species now appear in the BLAST interface.

Search against:	 Gadus morhua (Atl... X	 Cygnus atratus (Bla... X	 Enhydra lutris keny... X
<a href="#">Add/remove species</a>			

### Step 6: choose the database to search

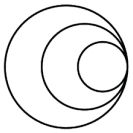
Select [Protein database](#) to compare the human protein sequence to proteins from these species.

<input type="radio"/> DNA database	Genomic sequence ▼
<input checked="" type="radio"/> Protein database	Proteins (Ensembl) ▼

### Step 7: Run BLAST

This is what your input form should look like now:





BLAST search

New job Clear form

Sequence data:

```
>human_KRT7.1
MSRQFTCKSGAAAKGGFSGCSAVLSGGSSSSFRAGSKLSGGFGSRSLYSLGGVRSLNVA
SGSGKSGGYGFRGRASGFAGSMFGSVALGPVCPVCPGGIHQVTNESLLAPLNVELD
PEIQKVRQEREQIKALNNKFAFIDKVRFLQQQVLETKWELLQQLDNLNCKNNLEPI
LEGYISNLRKQLETLSGDRVRLDSELNRVDDVEDYKRYEIEINKRTAAENEFVLLKLD
VDAAYANKVELQAKVESMDQEIKFRCLEAEITQIQSHISDMSVILSDMNNRNLDDSI
IDEVRTQYEEIALKSKAEAEALYQTKFQELQAAGRGGDLKNTKNEISELTRLIQRIRS
EINVKKQASNLETAIADAEQRGNALKDARAKLDELEGALHQAKEELARMLREYQELMS
LKLALDMEIATYRKLESEECRMSGEPSPVSIISSTSGGSVYGFPSMVSGGYVANS
SNCISGVCSVRGGGRSRGSANDYKDTLGRGSSLSAPSCKTSR
```

*Add more sequences (1 sequence added, 29 more sequences allowed)*

DNA  
 Protein

Search against:

Gadus morhua (Atl... x)  Cygnus atratus (Bla... x)  Enhydra lutris keny... x

*Add/remove species*

DNA database  Genomic sequence  
 Protein database  Proteins (Ensembl)

Search tool: BLASTP

Search Sensitivity: Normal

Description (optional):

Additional configurations:

[General options](#) [Scoring options](#) [Filters and masking options](#)

**Run**

Human sequence added

Three species chosen

Protein selected

Click on the green Run button at the bottom to run BLAST.



As the job is running, you can see it in a table.

Recent jobs

7 Refresh

Show/hide columns (1 hidden) Filter

Analysis	Jobs	Submitted at
BLASTP	Job 1: ENSG00000139648:ENST00000267119.6 peptide: ENSP00000267119 pep:protein_coding <b>Running</b>	14/10/2020, 11:51 (BST)
	Job 2: ENSG00000139648:ENST00000267119.6 peptide: ENSP00000267119 pep:protein_coding <b>Done: 92 hits found</b> <a href="#">View results</a>	
	Job 3: ENSG00000139648:ENST00000267119.6 peptide: ENSP00000267119 pep:protein_coding <b>Done: 100 hits found</b> <a href="#">View results</a>	

Some jobs are still running

Click to see the results of completed jobs

Step 8: view results

Choose the otter job first, indicated by the pawprint (mammal) icon. The bird and fish icons represent the black swan and Atlantic cod jobs respectively. Click on [View results](#).

This will show you the results for this query. You will see a table listing all the otter proteins that match to the human protein. There's a picture of this on the next page.

There are lots of lines to this table. Each line represents one protein in sea otter which has some sequence similarity to the human protein sequence we used as input. We call each one a **hit**.

The table tells us what that sea otter protein is and where it is found in the sea otter genome. Anything in blue is a link. Feel free to click on the links to see the gene, protein, genomic region and sequence of the hits before you go back to explore the page more. If you click back into BLAST from any page, you'll get to your job table again.

The table tells us where in the human protein the similarity starts and ends (subject start and end) and where in the otter protein (query start and end).

The most important information is in the E-value and %ID (percentage identity) columns. The %ID is how much of the otter protein is the same as the human protein. You'll see that the top hit in the table is [ENSELKPP00000012757](#). It is 92.93% identical to the human protein.

The e-value is a probability that the match is due to chance. This is calculated from the BLAST score, which adds value for matches and introduces penalties for gaps and mismatches. This is then combined with the full length and identity to determine how likely it is that this level of similarity could happen by chance. The smaller the number for the e-value, the better the hit. For the first six hits, the value is zero, but for the seventh it is  $2e-177$ , which means the probability that this is due to chance is two times E (a mathematical constant around 2.718) to the power of -177, an infinitesimally small number. In practice, you need to consider this number that E is to the power of: the higher the number after the minus, the smaller the e-value and the better the hit.

The top hit is most likely to be the orthologue of the human gene. Other hits may be similar genes in the otter genome, perhaps paralogues. Paralogues are when a gene is copied in a genome, then the two versions of the gene change over time in the same species, so they can have similar but distinct functions.

Navigate to the other jobs

**BLAST**

Web Tools

- Web Tools
- BLAST
- Ticket
  - Job 1: ENSG00000139648
  - Job 2: ENSG00000139648
  - Job 3: ENSG00000139648

Configure this page

Custom tracks

Export data

Share this page

Download results file

New job

### Results for Job 3: ENSG00000139648:ENST00000267119.6 peptide: ENSP00000267119 pep:protein\_coding

**Job details**

Job name: ENSG00000139648:ENST00000267119.6 peptide: ENSP00000267119 pep:protein\_coding

Species: Enhydra lutris kenyoni (Sea otter) - GCA\_002288905.2

Assembly: ASM228890v2

Search type: BLASTP (NCBI Blast)

**Results table**

Show All entries

Show/hide columns (2 hidden)

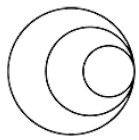
Filter

Subject name	Gene hit	Subject start	Subject end	Subject ori	Genomic Location	Orientation	Query start	Query end	Length	Score	E-val	%ID
<a href="#">ENSELKP00000012757</a>	<a href="#">ENSELKG00000009690</a>	45	525	Forward	<a href="#">KZ291803.1:10712066-10720480</a>	Reverse	45	523	481	756	0.0	92.93
<a href="#">ENSELKP00000013176</a>	<a href="#">ENSELKG00000009999</a>	46	540	Forward	<a href="#">KZ291803.1:10767343-10777116</a>	Reverse	46	523	495	686	0.0	78.38
<a href="#">ENSELKP00000013095</a>	<a href="#">ENSELKG00000009892</a>	44	511	Forward	<a href="#">KZ291803.1:10747552-10758829</a>	Reverse	43	523	494	628	0.0	72.67
<a href="#">ENSELKP00000012948</a>	<a href="#">ENSELKG00000009752</a>	44	535	Forward	<a href="#">KZ291803.1:10731961-10738939</a>	Reverse	42	523	493	608	0.0	72.21
<a href="#">ENSELKP00000012936</a>	<a href="#">ENSELKG00000009752</a>	44	543	Forward	<a href="#">KZ291803.1:10731961-10738939</a>	Reverse	42	523	501	600	0.0	71.06
<a href="#">ENSELKP00000013114</a>	<a href="#">ENSELKG00000009892</a>	44	469	Forward	<a href="#">KZ291803.1:10747552-10758829</a>	Reverse	43	523	494	536	0.0	65.18

How much of the two proteins are the same?

Which otter proteins match human KRT71?





### Step 9: Learn about the human protein

Let's go back to the human protein to find its function. Go back to the human gene page you left open in another tab. If you lost this, go back to page one for instructions on how to find it again.

The description at the top of the page gives the full name of the gene: [keratin 71](#). Have you heard of keratin before? Do you know where in the body you would find it?

<b>Gene: KRT71</b> ENSG00000139648	
<b>Description</b>	keratin 71 [Source:HGNC Symbol;Acc: <a href="#">HGNC:28927</a> ]

The menu on the left-hand side of the page lists different things we can look at for the gene. There are lots of things listed, but we're going to explore GO terms and Phenotypes.

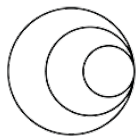
GO  
describes  
the  
function  
of the  
gene

Gene-based displays

- [-] Summary
  - Splice variants
  - Transcript comparison
  - Gene alleles
- [-] Sequence
  - Secondary Structure
- [-] Comparative Genomics
  - Genomic alignments
  - Gene tree
  - Gene gain/loss tree
  - Orthologues
  - Paralogues
  - Ensembl protein families
- [-] Ontologies
  - GO: Biological process
  - GO: Molecular function
  - GO: Cellular component
- [-] Phenotypes
- [-] Genetic Variation
  - Variant table
  - Variant image
  - Structural variants
- Gene expression
- Pathway
- Regulation
- External references
- Supporting evidence
- [-] ID History
  - Gene history

Phenotypes  
describe the  
effect of  
altering this  
gene

Click on [Phenotypes](#) to see the effect of altering this gene. Phenotypes are observable differences between individuals. This is what the page will look like:



**Phenotypes**

Phenotypes, diseases and traits associated with this gene ENSG00000139648

Phenotype, disease and trait	Source
<a href="#">Hypotrichosis 13</a>	<a href="#">MIM morbid</a>
<a href="#">Woolly hair</a>	<a href="#">Orphanet</a>

Phenotype, disease and trait annotations associated with variants in this gene

Phenotype, disease and trait	Source(s)	Number of variants	Show/hide details
ALL variants with a phenotype annotation	-	4	Show
Annotated by HGMD	<a href="#">HGMD-PUBLIC</a>	1	Show
<a href="#">Echocardiography</a>	<a href="#">dbGaP</a>	1	Show
<a href="#">Hypotrichosis 13</a>	<a href="#">ClinVar</a>	1	Show
<a href="#">INSULIN</a>	<a href="#">dbGaP</a>	1	Show

Phenotype, disease and trait annotations associated orthologues of this gene in other species

Phenotype, disease and trait	Source	Species	Gene
<a href="#">waved hair</a>	<a href="#">MGI</a>	Mouse ( <i>Mus musculus</i> )	<a href="#">ENSMUSG000000051879</a> Krt71
<a href="#">whorled hair</a>	<a href="#">MGI</a>	Mouse ( <i>Mus musculus</i> )	<a href="#">ENSMUSG000000051879</a> Krt71
<a href="#">abnormal hair follicle inner root sheath morphology</a>	<a href="#">MGI</a>	Mouse ( <i>Mus musculus</i> )	<a href="#">ENSMUSG000000051879</a> Krt71
<a href="#">abnormal vibrissa morphology</a>	<a href="#">MGI</a>	Mouse ( <i>Mus musculus</i> )	<a href="#">ENSMUSG000000051879</a> Krt71
<a href="#">abnormal hair follicle morphology</a>	<a href="#">MGI</a>	Mouse ( <i>Mus musculus</i> )	<a href="#">ENSMUSG000000051879</a> Krt71
<a href="#">alopecia</a>	<a href="#">MGI</a>	Mouse ( <i>Mus musculus</i> )	<a href="#">ENSMUSG000000051879</a> Krt71

Effects seen in human

Click on the data source to get descriptions of the effects and how they were discovered

Effects seen in other species

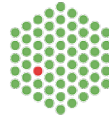
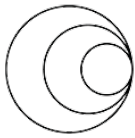
There are two phenotypes linked to the gene itself: Woolly hair and Hypotrichosis 13. Can you click around to find out what Hypotrichosis 13 means? What do these phenotypes suggest about what the gene does? Note that some websites may ask for a donation, you can just dismiss this banner.

Go to [GO: Biological process](#) in the menu on the left. This will open a table listing the functions the protein has, like the one below:

**GO: Biological process**

Accession	Term	Evidence	Annotation source	Transcript IDs
<a href="#">GO:0031069</a>	hair follicle morphogenesis	IEA	Ensembl	<a href="#">ENST00000267119</a>
<a href="#">GO:0031424</a>	keratinization	TAS	Reactome	<a href="#">ENST00000267119</a>
<a href="#">GO:0045109</a>	intermediate filament organization	IMP	UniProt	<a href="#">ENST00000267119</a>
<a href="#">GO:0070268</a>	cornification	TAS	Reactome	<a href="#">ENST00000267119</a>

We can see that the gene is linked to hair development.



**Questions:**

Based on what you've learned about the human protein:

3. Knowing *KRT71*'s involvement in hair follicles, now why do you think the fish and bird genes are so different?
  
  
  
  
  
  
  
  
  
  
  
  
  
  
4. Orthologues are genes that started out the same and diverged over time between species. Would you class these genes as orthologues?