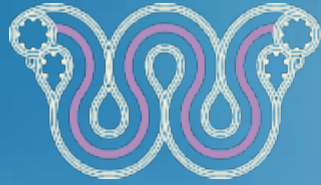




140  
YEARS



PŪTAHI MANAWA

HEALTHY HEARTS FOR  
AOTEAROA NEW ZEALAND

# Teaching 'Generation genome'

Dr Polona Le Quesne Stabej

Waipapa Taumata Rau | The University of Auckland



## Developed by:

Dr. Thierry Lints  
The University of Auckland  
and AKOMICS

Dr. Wilda Laux  
Genomics Aotearoa

Dr. Polona Le Quesne Stabej  
The University of Auckland

Prof. Cris Print, Dr. Tamsin Robb  
The University of Auckland

## With support from:



Genomics Into  
Medicine

This resource is to be used for educational purposes. For any other use, please consult the producers for permission

# DNA Variation & Bioinformatics



## *Analysing DNA to identify who killed baby Tara iti*

A resource developed by Dr. Thierry Lints, Dr. Polona Le Quesne Stabej, Dr. Wilda Laux, Dr. Tamsin Robb & Prof. Cris Print

With funding support from: Pūtahi Manawa- Healthy Hearts for Aotearoa New Zealand, The Genomics into Medicine Initiative and Genomics Aotearoa



A New Zealand Fairy Tern/Tara iti at Waipu River estuary  
[Tergiversation, CC BY-SA 4.0, via Wikimedia Commons](#)

### 1. Introduction

Tara iti (*Sternula nereis davisae*) is probably New Zealand's most endangered indigenous breeding bird. It is estimated that there are fewer than 40 individuals left at present including approximately 9 breeding pairs. We can find Tara iti in the lower half of the Northland Peninsula. Breeding is limited to four regular sites: Waipū, Mangawhai, Pākiri and Papakānui on the South Kaipara Head. Te Ārai and Poutawa rivermouths are also intermittently used.

The most likely threats to Tara iti are:

- **Habitat depletion** – degradation of sand dunes caused by residential development, farming...
- **Predation** – predators such as rats, cats, hedgehogs and mustelids prey upon eggs and chicks.
- **Environmental events** – High tides, floods, and storms can destroy and wash away nests.
- **Death of embryos** – Nesting birds are eaten or chased away by predators, and the embryos die from exposure.
- **Recreational activities** – Beach activities such as drone use, dog walking, bonfires ... disturb the birds and scare them away from their nests.





# Genetic variants and genetic diseases



# DNA detectives

## Who killed tara iti chick



- Read the Department of Fauna report
  - Problem?
  - Which biological samples do we have?
  - What do you think 'SEQUENCING' is?

A. Read the case report from the Department of Fauna.



Department of Fauna  
*Te Tari te Ao Kararehe*

**Case File Number:** DOF-6060842  
**Officer assigned:** Anthony McLeod  
**Date Filed:** 13/1/2020  
**Incident Location:** Waipū River Mouth Wildlife Refuge  
**Case Classification:** Native species predation  
**Case status:** Active investigation

**Event Report:**

At 9:54 AM on Saturday, January 11th 2020, DOF volunteer Susie Marsden discovered a dead fairy tern chick - Tara iti (*Sternula nereis davisae*) while conducting a nest survey at the Waipū River Mouth Wildlife Refuge. The chick was found approximately 260 meters south from the river mouth and 5 meters inland from the high-water mark. Marsden was unable to locate any nest site in the immediate vicinity (although shallow Tara iti nest scrapes in the sand are difficult to identify).

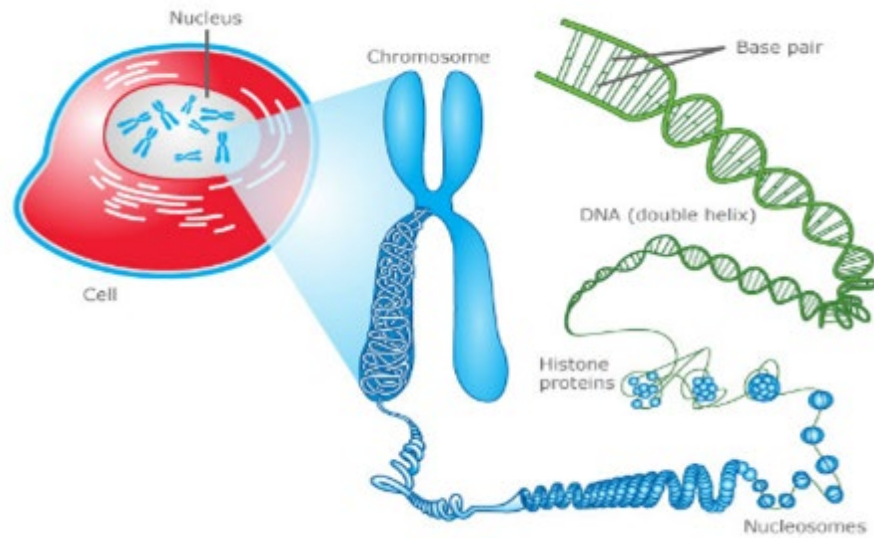
Autopsy of the chick revealed several small puncture wounds on the neck and back. The wounds, although not deep, are the likely cause of death and are consistent with either attack by a conspecific (same species) or predation by a small- to medium-sized predator. The puncture wounds suggest the latter scenario.

The stiffness and dehydrated condition of the body suggested death likely occurred within 48-72 hours prior. Despite a thorough search by Officer McLeod and a volunteer team, no informative animal tracks were found nearby. Strong easterly winds over the previous days could have filled over any predatory mammal tracks with sand.

Skin swabs and wound biopsies have been collected for forensic DNA sequence analysis. The DOF laboratory in Whangarei will file a MinION DNA sequence report within the next week. If sequence data indicate the involvement of an introduced pest, a predator trapping strategy will be rapidly implemented to protect the remaining Tara iti.



# Sequencing: determining the order of bases in DNA



Your **GENOME** – that is, your complete DNA sequence:

- about **three billion base pairs** long
- divided into 23 pairs of large chunks called chromosomes

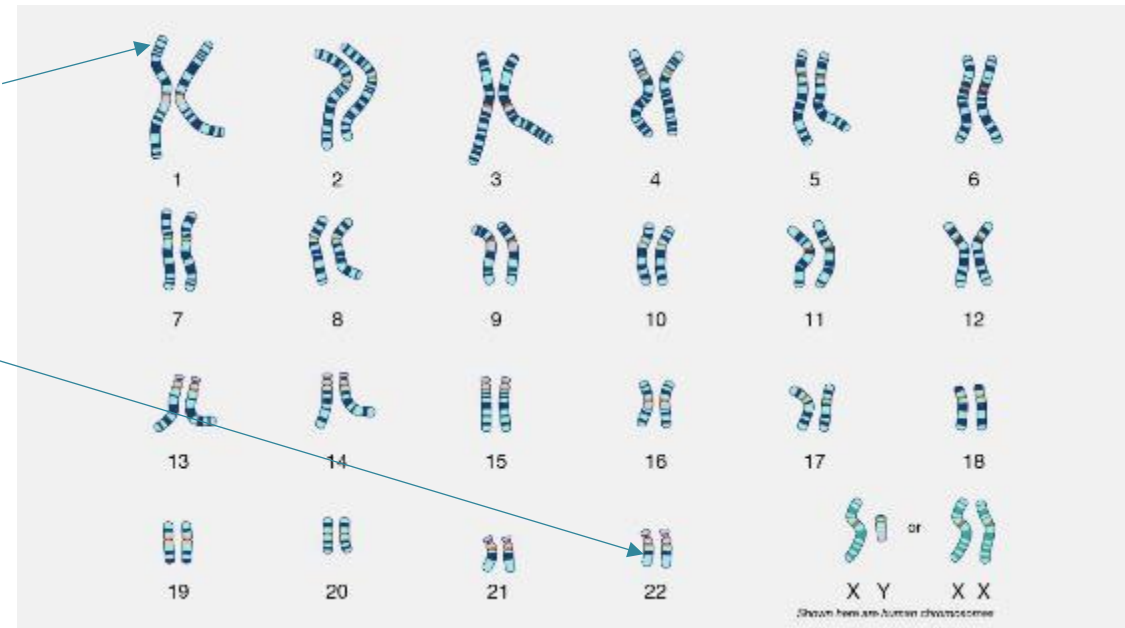
The Human Genome Project

2001: draft human genome sequence

2022: complete human genome sequence

cca 249 million base pairs

cca 47 million base pairs



<https://www.genome.gov/about-genomics/fact-sheets/Chromosomes-Fact-Sheet>

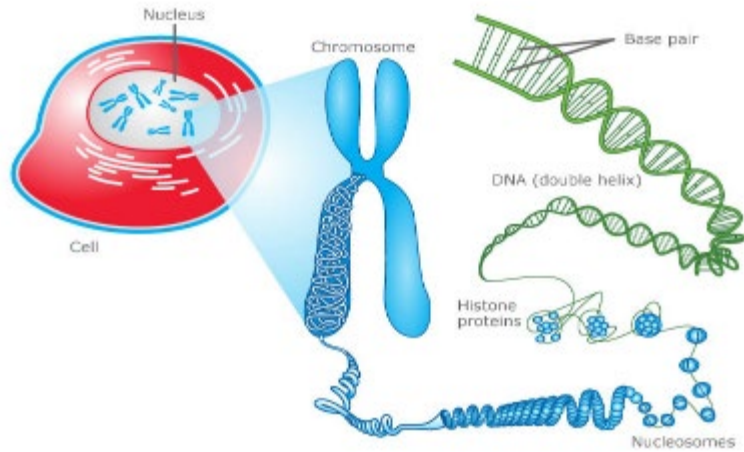
<https://www.genome.gov/genetics-glossary/Karyotype>

For more info on The Human Genome Project see:

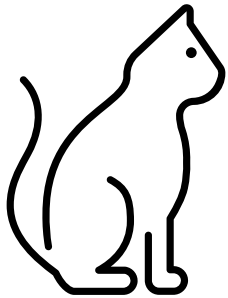
<https://www.genome.gov/about-genomics/educational-resources/fact-sheets/human-genome-project>



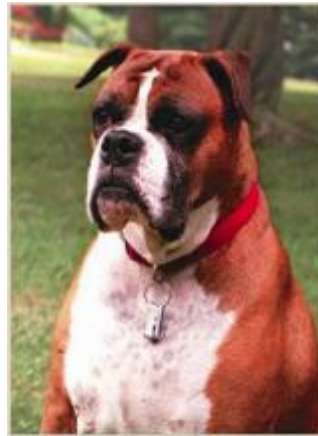
# Sequencing: determining the order of bases in DNA



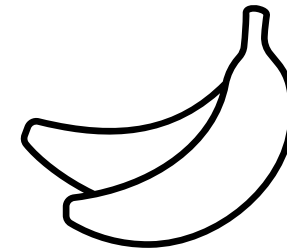
Your **GENOME** – that is, your complete DNA sequence, is about three billion base pairs long and divided into 23 chromosomes



Cat genome (2007)  
cca 2.5 billion base pairs  
18 chromosome pairs +X/Y  
<https://www.nbcnews.com/id/wbna21566038>



Dog genome (2003)  
cca 3 billion base pairs  
38 chromosome pairs  
<https://www.nature.com/articles/ng0703-249>



Banana genome (2012)  
523 million  
<https://www.nature.com/articles/nature11241>



# How sequencing works

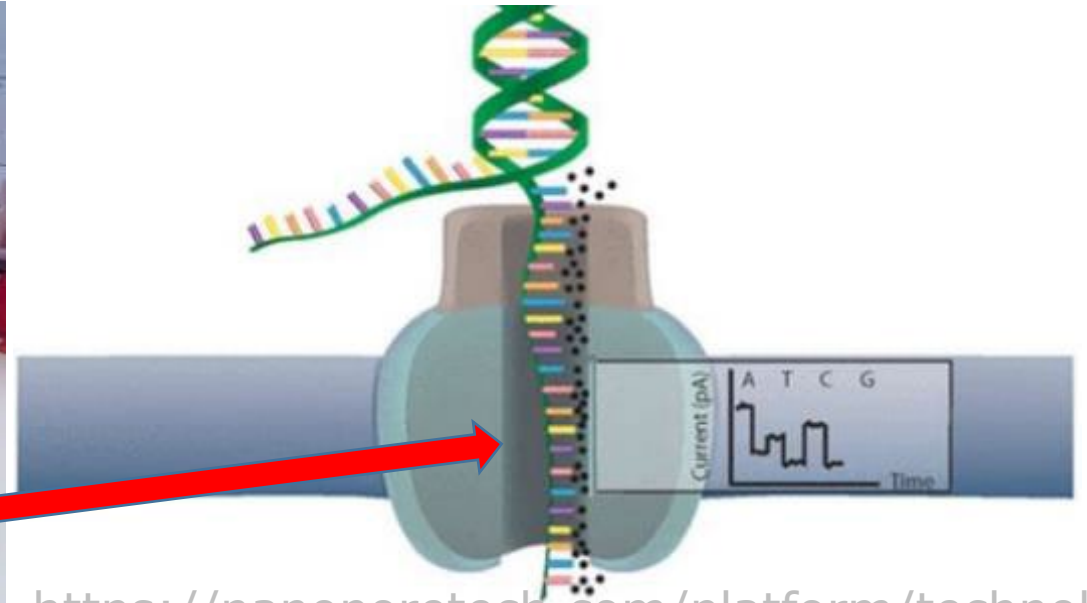
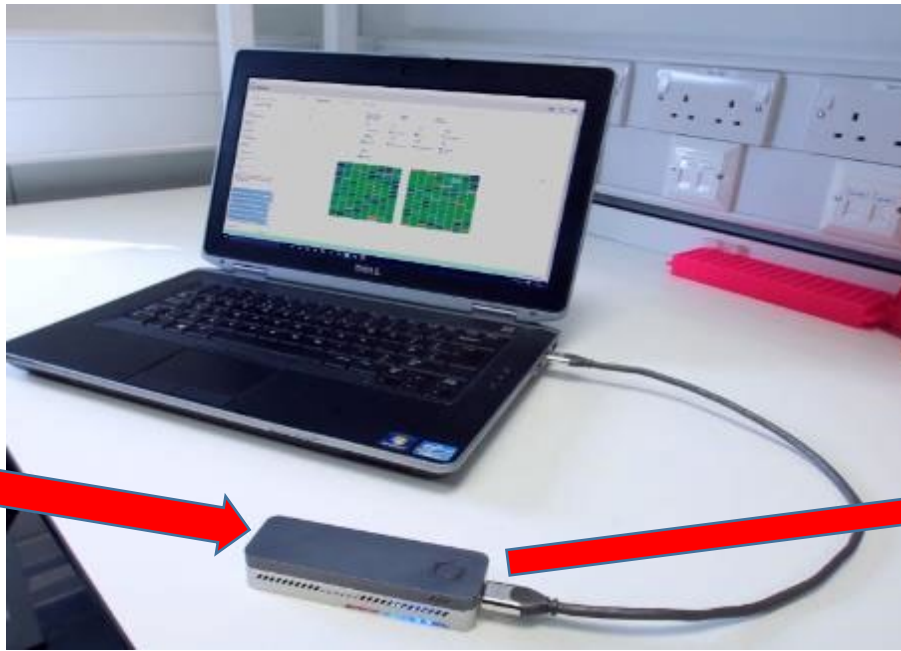
DNA sequencing machine determines the order of letters in the DNA



Illumina (NovaSeq6000)

<https://www.illumina.com/>

Oxford Nanopore Technology (ONT) MinION



<https://nanoporetech.com/platform/technology>



# Genetic variants (DNA changes)

## Traits

Eye  
colour



Hair  
colour



Nose shape



## Genetic diseases



Hearing  
loss



Brain disorders



Heart diseases



Blindness



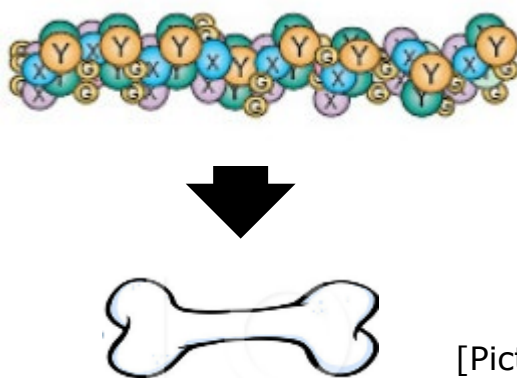
# Genetic variants can cause genetic diseases

Normal collagen gene (recipe for collagen protein)

```
GGGAAGGAGTGGAGGGGAGGCCCAAGGGGGTGTGGAGAAAGGAGCAGAAAGGCGAGCATTGGGGTTTCATAAGCCCAAGGGGAGAAAGGAGCTTACC
CCCGCATGGGTCTTCAAGCAAGTGGACCAAGCTTCTTTTTTAAAAAGTTATTTATTTATTTCTTTTTTTTTTTTTTTTTTTTGGTAAGGTGAATGCACCTT
TTGGTTTTTGGTCATGTTCCGTTGGTCAAAGATAAAAACTAAGTTTGAGAGATGAATGCAAGGAAAAAATATTTCCAAAGTCCATGTGAATGTCT
CCCATTTTTTGGCTTTTGAAGGGGGTTCAGTTTGGGTGCTTGTCTGTTTCCGGGTTGGGGGAAAGTTGGTTGGTGGGAGGAGCCAGTTGGGATGGA
GGGAGTTTACAGGAAGCAGACAGGGCCCAAGCTGGAAGCCGAATTCCTGGTCTGGGGCACCACGTTCCAAAGGGGCCACATCGATGATGGCGAGGCGGGAG
GTCTTGGTGGTTTTGTTTCAATCACTGCTTTGCCCAAGCTCCGGTGTGACTTGGGGTGGGGCGAGACAAAGGAGGGGGCATTACCACTGGGAGT
GATGGAAGAGGGCAGTATGGCTTGGCCAAAAGCCCAAGGGCGGAGAGGTTGGGCGCTGCTGGGGAATCTGGGCACTCACTTCGACCCATCGACAGT
GACGCTGTAGTGAAGCGGCTGTTGCCCTCGGGCGGAGTCTCGATCTCGTTGGAGCCCTGGAGAGCAGGGCTCTTGAAGTTGGCAGCTCTGCTGGTCC
ATGTAGGCCACGCTGTTCTTCAGTGGTAGGTGATGTTCTGGAGGGCTCGGTGGACATCAGGCGCAGGAAGTCACTGGATGGCCACATCGGCAGGGT
CGGAGCCCTGGCGCCCATACTCGAATGCAAGGGAGGGGAGAGAGGAGAGTGAAGCCGCTATCGGGGAACCTCTAGTCTGCTGGCCTCCCTGTCCAG
GGTCTCAGAGAGCTGCCCAATGCACCGTTATATCGAGAGGAGGCACACCTGCCATCGGCAGCTGTGTCTGAACCACTATCAGGAGCTGAGACCCC
TGCTTCCAGGCCAGCTCTGCTCATAC
TCTGAGTCCAGCTCAGCAGCTGGAAT
TGGGCCACACTGGGCTGAGTGGGTAC
TCCAGTACTCTCTGTGTTAGGCGAGG
ATGGCAGGAGTAGGAGGGAGGAGAGG
TTGCGGCTGCGCTCTGGGCTCGGATG
CATCAGCCCGTAGTAGCGCCACCATK
AGGAGGTCAGAGGGCCCGGGGAGCC
GGGATGAGGGGCTACATACAAAGCAGCAGCATCACCAGTGCAGCCGCGAGGACCAGGGGGCCCAATGGGGCCAGGAGACCTTTGATTCATCTTTGCC
AGGAGCACAGCAGAGCCAGGGGAGCCCTGGAGTGGGGGAAATGGTTTGAAGAGGCTGCCAAGAGCCGCAACCCAGCTCTGGAGAGAGAGCCGCC
CAGCTCCCTGCTGGCTCTGGGGCCAGGCTCATAGAGCCAGGCTCAGCCTCTTTCCATAGGACATGCCATAGCTCCTCCAGCCGCTGAGCCAGGGGAG
GACGGATTGGGAGCAGAGAGGCCAAAGCTAGATCAGATTGTTTGTTCAGGATTCTCTCTCTTTGCCACTCCCTGCTCCCTGAACCTCTCTCCAGAGA
GGCAAGGGTGCTGGGTCCCTGGCAAGGGTCCCGAGGTGAGCCTGGGCTGGGGCTCAGGAAGAGGAGAGAGAGGCACTGACTTACTCGGGGACCAGC
AGGACCAGAGGCTCCAGAGGACCTTGTTCACAGGAGAGCCCTGAAGGACAGATAAAAGGCAAGTTCAGGCCCATGAGCGTCAATGTAGCCTGAGG
GCCTGGCTGGAGAGACAAGGCGAGTGTGGGCAACCCATGCCCTTCAATTATCTGGAATCTAGCCGAGGGTGTCCATAGGACAGAGACCTCTCCATAA
TCCTCTCTGTGTACCCCTCAGCTCTCAGAGTACTTGGACATGCCCAAAATCTTCAGACCCAGTCCCACTAGGAGGGGAAAGAAATGACTATCC
AGAGGGGAAACTGAGGCGAAGCTCCCTCCTATCCCAAGCAGCATGGGACTGGGGAGGGGCTGAGCATACTTACAGAGAGGGCCAGGGGGACCT
GGAGGCCAGAGAAAGCCAGGTTGACCTTTATGCTCTGTGCGCTGTTGCGCTGTCTCAGCTTGTGACACAGGGGGCTTGGGTCCTTAGAAGAGAGA
AAGGGACAATGTGAGCGCGAAGTTCATATGCAATGCAAGTGGGCACTGTGATCTGTAGATTCTTAAGGCTATGAGGAGGAGGAGGAGGAGGAGGAGG
GCGGGGCCAGGGGCCAAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG
GCTCGGGCTCAGAGATGAGGCTCCCTCTGCTGGATCTCTCTCTCTCAGTCTGCTCCACTGTGGCCATCTCTCCCACTCCAGGGAACCTCCCT
ACTGCAATCTTCAAGGAGCTGGGGCCAACTCATGGAGAGGCGGCTGCTGCTGGAGAGGGGACTTGGGGCTGAGCTTTAACTCAGTTTTTGGATTAA
GGCCCTGACATCTTGCAGGATCTCTTTCTCTCCCTGCTGGTGAAGTCCGACACCCATCCCAAGGCTCTAAGGAGGCTCAAGAGTCCCTGGCCTG
```

CGGGGCT

Normal collagen:

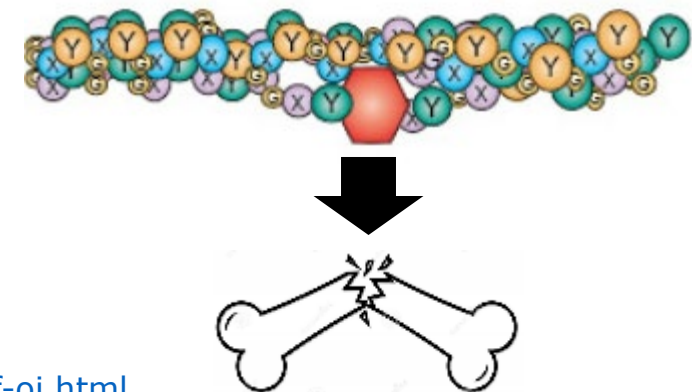


Collagen gene with a disease-causing variant

```
GGGAAGGAGTGGAGGGGAGGCCCAAGGGGGTGTGGAGAAAGGAGCAGAAAGGCGAGCATTGGGGTTTCATAAGCCCAAGGGGAGAAAGGAGCTTACC
CCCGCATGGGTCTTCAAGCAAGTGGACCAAGCTTCTTTTTTAAAAAGTTATTTATTTATTTCTTTTTTTTTTTTTTTTTTTTGGTAAGGTGAATGCACCTT
TTGGTTTTTGGTCATGTTCCGTTGGTCAAAGATAAAAACTAAGTTTGAGAGATGAATGCAAGGAAAAAATATTTCCAAAGTCCATGTGAATGTCT
CCCATTTTTTGGCTTTTGAAGGGGGTTCAGTTTGGGTGCTTGTCTGTTTCCGGGTTGGGGGAAAGTTGGTTGGTGGGAGGAGCCAGTTGGGATGGA
GGGAGTTTACAGGAAGCAGACAGGGCCCAAGCTGGAAGCCGAATTCCTGGTCTGGGGCACCACGTTCCAAAGGGGCCACATCGATGATGGCGAGGCGGGAG
GTCTTGGTGGTTTTGTTTCAATCACTGCTTTGCCCAAGCTCCGGTGTGACTCTGGGGTGGGGCGGAGACACGGGAGGGGGCATTACCACTGGGAGT
GATGGAGAGAGGGCAGTATGGCCTGGCCAAAAGCCCAAGGCCGAGGAGGTGGGGCCCTGCTGGGATTCTGGGCACTCAGCGTGCAGCCATCGACAGT
GACGCTGTAGTGAAGCGGCTGTTGCCCTCGGGCGGAGTCTCGATCTCGTTGGAGCCCTGGAGGAGCAGGGCCTCTTGAAGTTGCCAGTCTGCTGGTCC
ATGTAGGCCACGCTGTTCTTCAGTGGTAGGTGATGTTCTGGAGGGCTCGGTGGACATCAGGCGCAGGAAGTCACTGGATGGCCACATCGGCAGGGT
CGGAGCCCTGGCGCCCATACTCGAATGCAAGGGAGGGGAGAGAGGAGAGTGAAGCCGCTATCGGGGAACCTCTAGTCTGCTGGCCTCCCTGTCCAG
GGTCTCAGAGAGCTGCCCAATGCACCGTTATATCGAGAGGAGGCACACCTGCCATCGGCAGCCTGTGTCTGAACCACTATCAGGAGCTGAGACCCC
TGCTTCCAGGCCAGCTCTGTCTCA
TCTGAGCTGGGCTAGCCCATCTCTAACACTGGC
TTGTCTTGGGGTCTTGTCTGATGATGACGATTTCTT
TGATGGATCCAGGTTGACGCTTGGTTGGGGTCAA
TATGGTAGGGGCACATATGGGCATGGGACCTGGC
ACATCTTGAAGTCACGGCAGGTGCGGGCGGGTTC
CCAGCTCGAGGTCACGGTCAGCAACCAATTTGCA
GAAGTCGAACCAAGCGCTGGGAGGACAGGGGACC
CAGCTCTTGGGGCTGAGGGCCATGAGCAGAG
GGATGAGGGGCTACATACAAAGCAGCAGCATCACCAGTGCAGCCGCGAGGACCAGGGGGCCCAATGGGGCCAGGAGACCTTGAAGTCACTTTGCC
AGGAGCACAGCAGAGCCAGGGGAGCCCTGGAGTGGGGGAAATGGTTTGAAGAGGCTGCCAAGAGCCGCAACCCAGCTCTGGAGAGAGGCCCC
CAGCTCCCTGCTGGCTCTGGGGCCAGGCTCATAGAGCCAGGCTCAGCCTCTTTCCATAGGACATGCCATAGCTCCTCCAGCCGCTGAGCCAGGGGAG
GACGGATTGGGAGCAGAGAGGCCAAAGCTATCAGATTGTTTGTTCAGGATTCTCTCTCTTTGCCACTCCCTGCTCCCTGAACCTCTCTCCAGAGA
GGCAAGGGTGCTGGGTCCCTGGCAAGGGTCCCGAGGTGAGCCTGGGCTTGGGGCTCAGGAAGAGGAGAGAGAGGCACTGACTTACTCGGGGACCAGC
AGGACCAGAGGCTCCAGAGGAGCCTTGTTCACAGGAGAGCCCTGAAGGACAGATAAAAGGCAAGTTCAGGCCCATGAGCGTCAATGTAGCCTGAGG
GCCTGGCTGGAGAGACAAGGCGAGTGTGGGCAACCCATGCCCTTCAATTATCTGGAATCTAGCCGAGGGTGTCCATAGGACAGAGACCTCTCCATAA
TCCTCTCTGTGTACCCCTCAGCTCTCAGAGTACTTGGACATGCCCAAAATCTTCAGACCCAGTCCCACTAGGAGGGGAAAGAAATGACTATCC
AGAGGGGAAACTGAGGCGAAGCTCCCTCCTATCCCAAGCAGCATGGGACTGGGGAGGGGCTGAGCATACTTACAGAGAGGGCCAGGGGGACCT
GGAGGCCAGAGAAAGCCAGGTTGACCTTTATGCTCTGTGCGCTGTTGCGCTGTCTCAGCTTGTGACACAGGGGGCTTGGGTCCTTAGAAGAGAGA
AAGGGACAATGTGAGCGCGAAGTTCATATGCAATGCAAGTGGGCACTGTGATCTGTAGATTCTTAAGGCTATGAGGAGGAGGAGGAGGAGGAGGAGG
GCGGGGCCAGGGGCCAAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG
GCTCGGGCTCAGAGATGAGGCTCCCTCTGCTGGATCTCTCTCTCTCAGTCTGCTCCACTGTGGCCATCTCTCCCACTCCAGGGAACCTCCCT
ACTGCAATCTTCAAGGAGCTGGGGCCAACTCATGGAGAGGCGGCTGCTGCTGGAGAGGGGACTTGGGGCTGAGCTTTAACTCAGTTTTTGGATTAA
GGCCCTGACATCTTGCAGGATCTCTTTCTCTCCCTGCTGGTGAAGTCCGACACCCATCCCAAGGCTCTAAGGAGGCTCAAGAGTCCCTGGCCTG
```

CGGA

Wrong amino acid built in collagen:



[Picture from: <https://oimperfecta.weebly.com/genetics-of-oi.html>]



# DNA detectives: Who killed tara iti chick



## 2. Task

Page 3

### B. Track down the culprit

The Department of Fauna laboratory collected, processed and prepared DNA samples from skin swabs and biopsy tissues of the dead chick, then sequenced the prepared genomic DNA samples. The DNA sequences obtained were further screened through a computer program to remove sequence files closely matching the Tara iti's genome – this makes it possible to focus on foreign DNA sequences only for potential predator DNA sequences.

Results showed that one particular DNA sequence was the 'best guess' for a potential culprit. It is given as follows:

```
ATCAGGCTACATCCTGGAGCGCAAGAAGAAGAAGAGCTTCCGGTGGATGTGGC
TGAACCTTTGACCTGCTGCAGGAGCTGAGCCACGAGGCACGGCGCATGATTGAG
GGCGTGGTGTATGAGATGCGAGTCTACGCGGTCA
```

You are now going to analyse this sequence to properly identify the culprit.

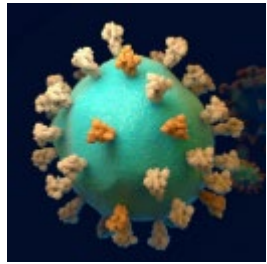
Your aim is to answer two questions:

- Which species does this DNA sequence belong to?
- Does this DNA sequence indicate something special about the culprit?

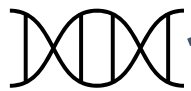
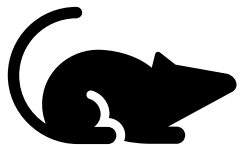
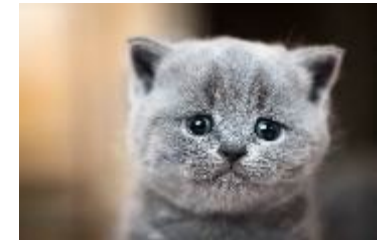
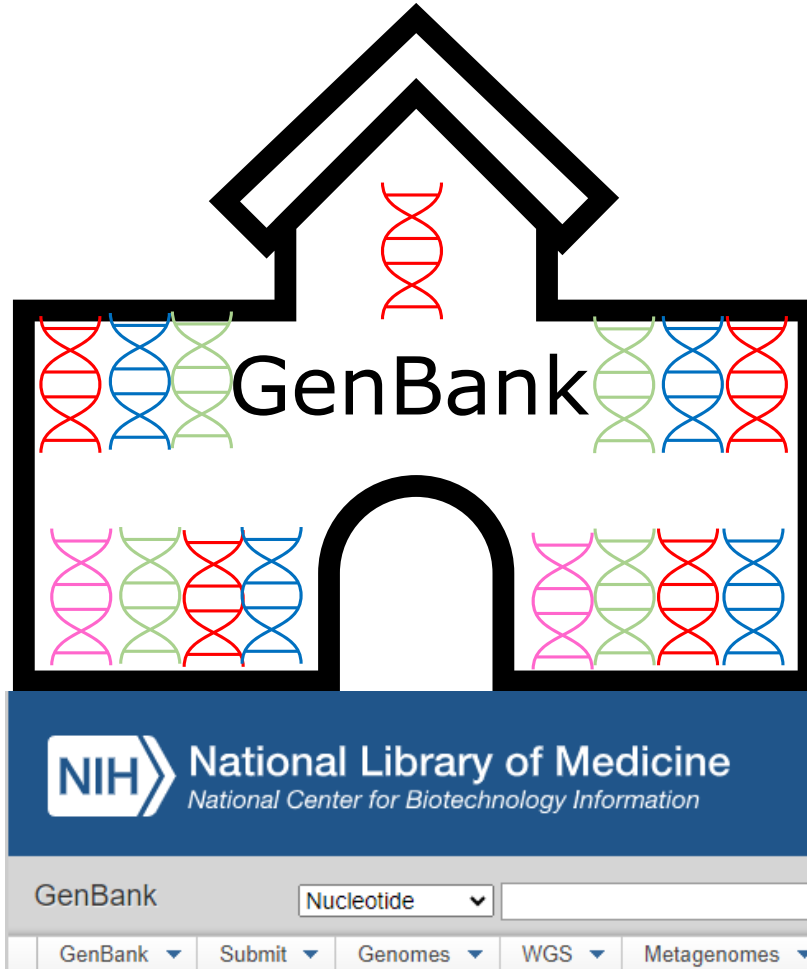
## Clues from the Sequence?



Scientists use the 'GenBank' as a bank where they deposit DNA sequences that have ever been sequenced: from COVID 19 to Human DNA



Credit: NIAID - Biovisualization Program, BCBB, OCICB and Research Technologies Branch, Rocky Mountain Laboratories (RML)- Virion created by Austin Athman © RML labs, remeshed, colored & rendered by Kristen Browne





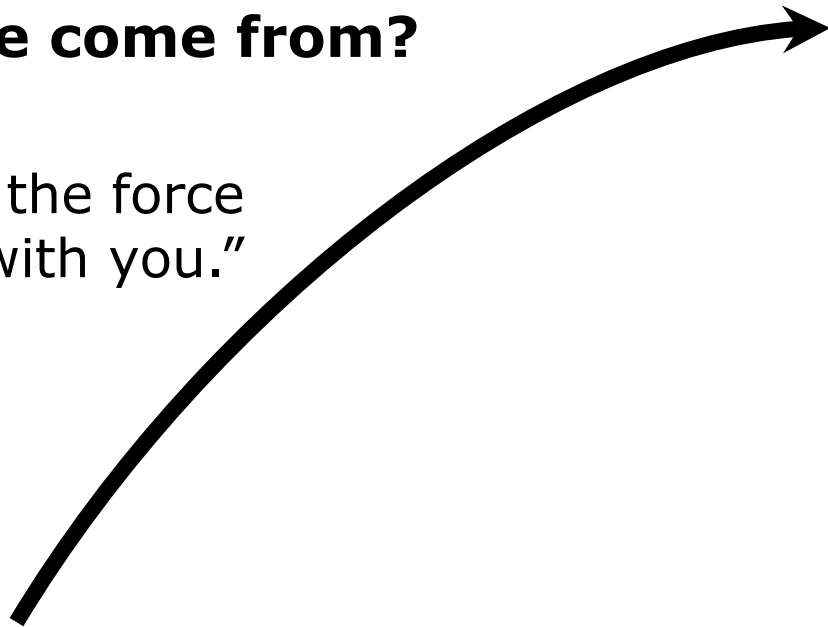
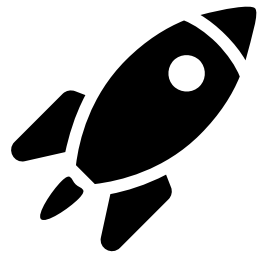


Google

Google

**Which movie does this  
quote come from?**

"May the force  
be with you."



s  
e  
a  
r  
c  
h

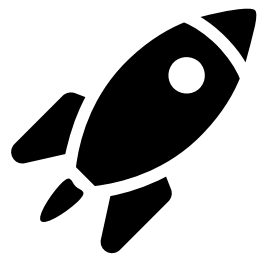




Google

**Which movie does this  
quote come from?**

"May the force  
be with you."



Google



s  
e  
a  
r  
c  
h







The **BLAST** tool can accurately search all sequences in the GenBank database and identify where your **Wound biopsy DNA** sequence comes from:

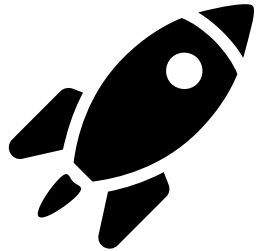
**Wound biopsy seq.**

```
>Wound Biopsy Sequence
ATCAGGCTACATCCTGGAG
CGCAAGAAGAAGAAGAGCT
TCCGGTGGATGTGGC
TGAACCTTGACCTGCTGCA
GGAGCTGAGCCACGAGGC
ACGGCGCATGATTGAG
GGCGTGGTGTATGAGATGC
GAGTCTACGCGGTCA
```

**NCBI  
GenBank**

**Which species and  
gene does this  
portion of a  
sequence belong  
to?**

**B  
L  
A  
S  
T**







# Laptops

**Page 3**

1.) Copy the following DNA sequence:

>Wound Biopsy Sequence

```
ATCAGGCTACATCCTGGAGCGCAAGAAGAAGAAGAGCTTCCGGTGGATGTGGC  
TGAACCTTTGACCTGCTGCAGGAGCTGAGCCACGAGGCACGGCGCATGATTGAG  
GGCGTGGTGTATGAGATGCGAGTCTACGCGGTCA
```

2.) Open up this [link to NCBI BLAST](#) (BLAST performs a sequence similarity search across the hundreds of millions of DNA sequences present in the GenBank database).





# Clue #1: Which species does the Wound Biopsy Sequence belong to?

NIH National Library of Medicine  
National Center for Biotechnology Information

BLAST® » blastn suite

**Important update**  
The core nucleotide database (core\_nt) is now the default nucleotide BLAST database. Learn more about core\_nt.

blastn blastp blastx tblastn tblastx

**Enter Query Sequence**

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#)

Wound Biopsy Sequence  
ATCAGGCTACATCTCTGGAGCGCAAGAGAGAGCTTCGGTGGATGTG  
GC  
TGAACCTTTGACCTGCTGACGAGGAGTGAAGCAGGAGGACGCGCATGATT  
GAG  
GGCGTGGTGTATGAGATGGAGTCTACGGGGTCA

Query subrange  
From   
To

Or, upload file  No file chosen

Job Title  
Wound Biopsy Sequence  
Enter a descriptive title for your BLAST search

☐ Align two or more sequences

**Choose Search Set**

Database  
☒ Standard databases (nr etc.) ☐ rRNA/ITS databases ☐ Genomic + transcript databases ☐ Betacoronavirus ☐ Experimental databases

Core nucleotide database (core\_nt)

Organism options  
Enter organism name or id—completions will be suggested    
Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown

Exclude options  
☐ Models (XMAP) ☐ Uncultured/environmental sample sequences

Limit to options  
☐ Sequences from type material

Entrez Query options  
Enter an Entrez query to limit search

**Program Selection**

Optimize for  
☒ Highly similar sequences (megablast)  
☐ More dissimilar sequences (discontiguous megablast)  
☐ Somewhat similar sequences (blastn)  
Choose a BLAST algorithm

**BLAST** Search database core\_nt using Megablast (Optimize for highly similar sequences)  
☐ Show results in a new window





# Clue #1: Which species does the Wound Biopsy Sequence belong to?

[< Edit Search](#) [Save Search](#) [Search Summary](#) [How to read this report?](#) [BLAST Help Videos](#) [Back to Traditional Results Page](#)

Job Title: Wound Biopsy Sequence  
RID: [BNP3TK7A016](#) Search expires on 08-14 11:52 am [Download All](#) [v](#)  
Program: BLASTN [?](#) [Citation](#) [v](#)  
Database: core\_nt [See details](#) [v](#)  
Query ID: lcl|Query\_6477905  
Description: Wound Biopsy Sequence  
Molecule type: dna  
Query Length: 140  
Other reports: [Distance tree of results](#) [MSA viewer](#) [?](#)

Filter Results

Organism only top 20 will appear ☐ exclude  
Type common name, binomial, taxid or group name  
[+ Add organism](#)

Percent Identity  to  E value  to  Query Coverage  to   
[Filter](#) [Reset](#)

[Descriptions](#) [Graphic Summary](#) [Alignments](#) [Taxonomy](#)

Sequences producing significant alignments [Download](#) [v](#) [Select columns](#) [v](#) Show [100](#) [?](#)

☒ select all 100 sequences selected [GenBank](#) [Graphics](#) [Distance tree of results](#) [MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Panthera onca myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera onca</a>	246	246	97%	4e-61	99.26%	4171	<a href="#">XM_060604458.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Panthera pardus myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera pardus</a>	246	246	97%	4e-61	99.26%	4246	<a href="#">XM_053899761.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Panthera uncia myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera uncia</a>	246	246	97%	4e-61	99.26%	4988	<a href="#">XM_049847407.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Felis catus myosin binding protein C3 (MYBPC3), transcript variant X3, mRNA</a>	<a href="#">Felis catus</a>	246	246	97%	4e-61	99.26%	3912	<a href="#">XM_045039263.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Felis catus myosin binding protein C3 (MYBPC3), transcript variant X2, mRNA</a>	<a href="#">Felis catus</a>	246	246	97%	4e-61	99.26%	4282	<a href="#">XM_019812397.2</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Felis catus myosin binding protein C3 (MYBPC3), transcript variant X1, mRNA</a>	<a href="#">Felis catus</a>	246	246	97%	4e-61	99.26%	4285	<a href="#">XM_019812396.2</a>

- 7.) Write down the names of the top 5 "hits" the DNA sequence matches with
- 8.) Reviewing the 5 top 'hits', can you infer who would be the most likely culprit?



## Answers

7.) Write down the names of the top 5 “hits” the DNA sequence matches with

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

Download

Select columns

Show

100

?

☒ select all 100 sequences selected

[GenBank](#)

[Graphics](#)

[Distance tree of results](#)

[MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Panthera tigris myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera tigris</a>	246	246	97%	4e-61	99.26%	5005	<a href="#">XM_042958215.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Panthera uncia myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera uncia</a>	246	246	97%	4e-61	99.26%	4988	<a href="#">XM_049647407.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Panthera onca myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera onca</a>	246	246	97%	4e-61	99.26%	4171	<a href="#">XM_060604458.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Panthera pardus myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera pardus</a>	246	246	97%	4e-61	99.26%	4246	<a href="#">XM_053899761.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Felis catus myosin binding protein C3 (MYBPC3), transcript variant X2, mRNA</a>	<a href="#">Felis catus</a>	246	246	97%	4e-61	99.26%	4282	<a href="#">XM_019812397.2</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Panthera leo myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera leo</a>	246	246	97%	4e-61	99.26%	5001	<a href="#">XM_042904117.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Felis catus myosin binding protein C3 (MYBPC3), transcript variant X3, mRNA</a>	<a href="#">Felis catus</a>	246	246	97%	4e-61	99.26%	3912	<a href="#">XM_045039263.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Felis catus myosin binding protein C3 (MYBPC3), transcript variant X1, mRNA</a>	<a href="#">Felis catus</a>	246	246	97%	4e-61	99.26%	4285	<a href="#">XM_019812396.2</a>

Use Google to find out the English names: tiger, snow leopard, jaguar, leopard, cat.

8.) Reviewing the 5 top ‘hits’, can you infer who would be the most likely culprit?





## Clue #2: What gene/mRNA does the 'Wound Biopsy Sequence' belong to?

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

Download

Select columns

Show 100

?

☒ select all 100 sequences selected

[GenBank](#)

[Graphics](#)

[Distance tree of results](#)

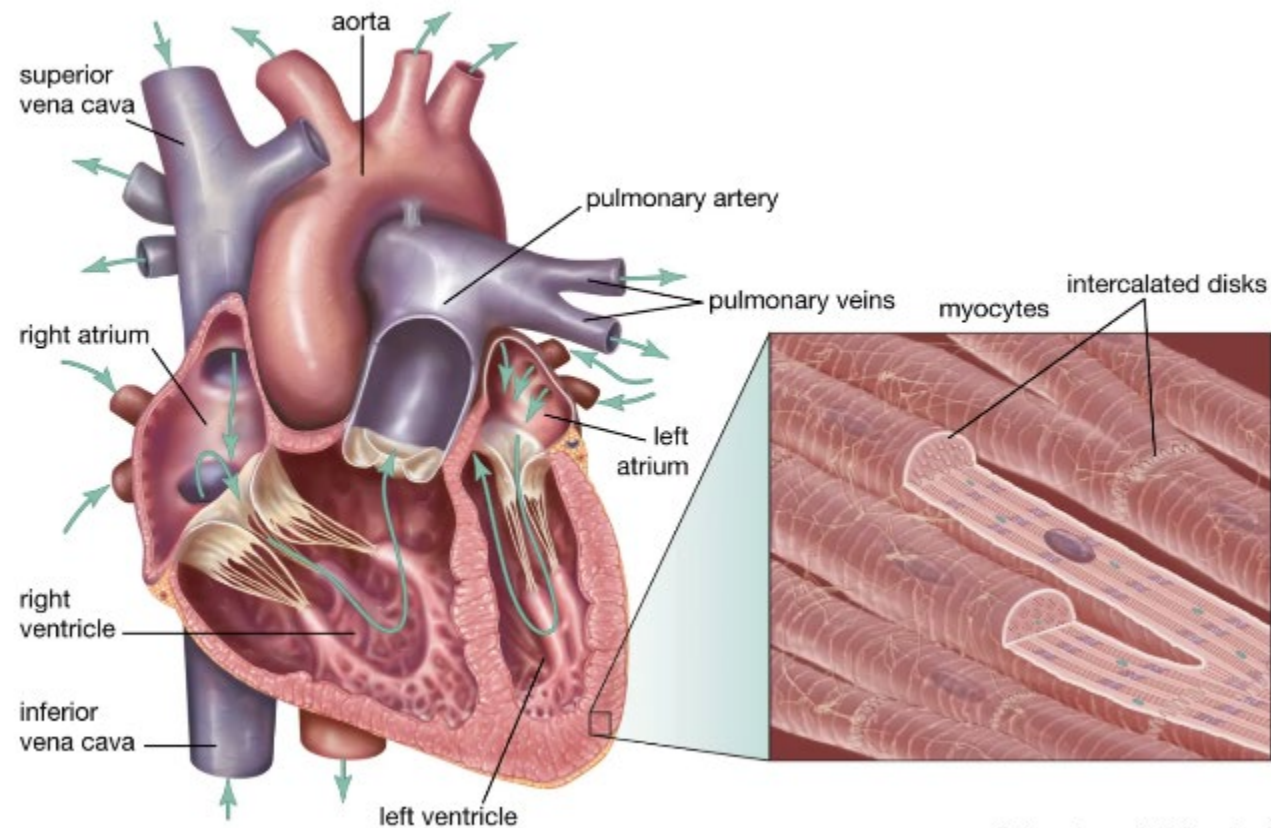
[MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Panthera tigris myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera tigris</a>	246	246	97%	4e-61	99.26%	5005	<a href="#">XM_042958215.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Panthera uncia myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera uncia</a>	246	246	97%	4e-61	99.26%	4988	<a href="#">XM_049647407.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Panthera onca myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera onca</a>	246	246	97%	4e-61	99.26%	4171	<a href="#">XM_060604458.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Panthera pardus myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera pardus</a>	246	246	97%	4e-61	99.26%	4246	<a href="#">XM_053899761.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Felis catus myosin binding protein C3 (MYBPC3), transcript variant X2, mRNA</a>	<a href="#">Felis catus</a>	246	246	97%	4e-61	99.26%	4282	<a href="#">XM_019812397.2</a>

# Myosin Binding Protein C - *MYBPC3*

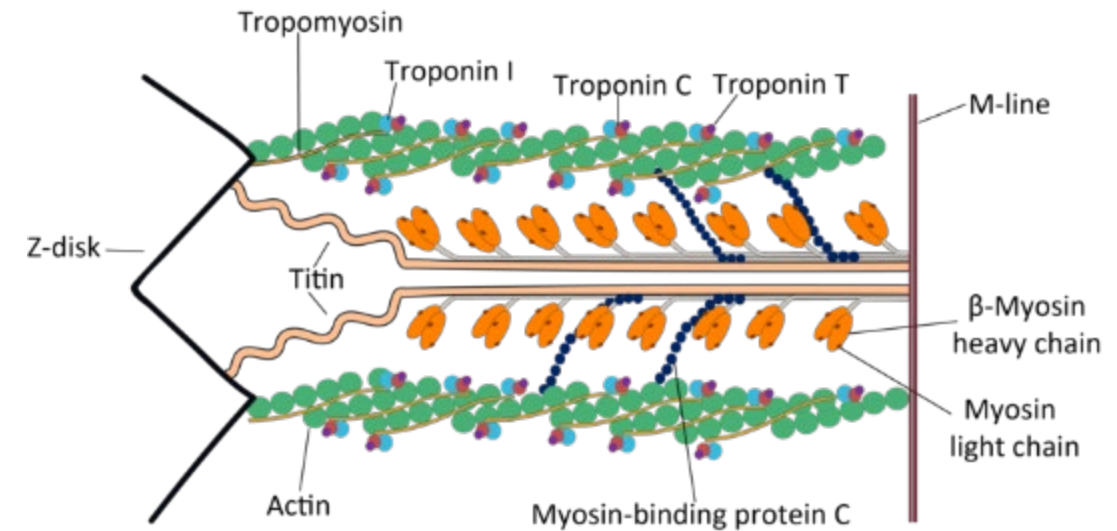


# Myosin Binding Protein C - *MYBPC3*



© Encyclopædia Britannica, Inc.

## Sarcomere – contractile unit



**MYBPC3**

From: <https://www.britannica.com/science/cardiac-muscle>



# Clue #3: Is there anything different in the 'Wound Biopsy Sequence' MYBPC3 compared to GenBank seq?



Page 4

**Job Title** Nucleotide Sequence

**RID** [JN2XDK1K013](#) Search expires on 11-07 04:18 am [Download All](#) ▾

**Program** BLASTN [?](#) [Citation](#) ▾

**Database** core\_nt [See details](#) ▾

**Query ID** lcl|Query\_3554249

**Description** None

**Molecule type** dna

**Query Length** 140

**Other reports** [Distance tree of results](#) [MSA viewer](#) [?](#)

**Filter Results**

**Organism** only top 20 will appear ☐ exclude

Type common name, binomial, taxid or group name

[+ Add organism](#)

**Percent Identity**  to  **E value**  to  **Query Coverage**  to

[Filter](#) [Reset](#)

**Descriptions** Graphic Summary Alignments Taxonomy

**Sequences producing significant alignments** Download ▾ Select columns ▾ Show 100 ▾ [?](#)

☐ select all 1 sequences selected

[GenBank](#) [Graphics](#) [Distance tree of results](#) [MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input type="checkbox"/>	<a href="#">PREDICTED: Panthera tigris myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera tigris</a>	246	246	97%	4e-61	99.26%	5005	<a href="#">XM_042958215.1</a>
<input type="checkbox"/>	<a href="#">PREDICTED: Panthera uncia myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera uncia</a>	246	246	97%	4e-61	99.26%	4988	<a href="#">XM_049647407.1</a>
<input type="checkbox"/>	<a href="#">PREDICTED: Panthera onca myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera onca</a>	246	246	97%	4e-61	99.26%	4171	<a href="#">XM_060604458.1</a>
<input type="checkbox"/>	<a href="#">PREDICTED: Panthera pardus myosin binding protein C3 (MYBPC3), mRNA</a>	<a href="#">Panthera pardus</a>	246	246	97%	4e-61	99.26%	4246	<a href="#">XM_053899761.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Felis catus myosin binding protein C3 (MYBPC3), transcript variant X2, mRNA</a>	<a href="#">Felis catus</a>	246	246	97%	4e-61	99.26%	4282	<a href="#">XM_019812397.2</a>

9.) Deselect the 'select all'

10.) Click on the Felis catus myosin binding protein C3 (MYBPC3), transcript variant X2, mRNA





11) Identify the common and different base(s) between your query and the GenBank database. What is/are the difference(s) if any?

Page 4

Score		Expect		Identities		Gaps		Strand	
246 bits(133)		4e-61		135/136(99%)		0/136(0%)		Plus/Plus	
Wound biopsy seq.	Query	5		GGCTACATCCTGGAGCGCAAGAAGAAGAAGAGCTTCCGGTGGATGTGGCTGAACTTTGAC	64				
GenBank seq (ref)	Sbjct	2528		GGCTACATCCTGGAGCGCAAGAAGAAGAAGAGCTTCCGGTGGATGCGGCTGAACTTTGAC	2587				
	Query	65		CTGCTGCAGGAGCTGAGCCACGAGGCACGGCGCATGATTGAGGGCGTGGTGTATGAGATG	124				
	Sbjct	2588		CTGCTGCAGGAGCTGAGCCACGAGGCACGGCGCATGATTGAGGGCGTGGTGTATGAGATG	2647				
	Query	125		CGAGTCTACGCGGTCA	140				
	Sbjct	2648		CGAGTCTACGCGGTCA	2663				

Do you think this matters?? Just one letter is not the same as in the 'normal' cat?



12.) This single base change is actually known to alter the MYBPC3 protein in cats

13.) Click on 'cds feature' to change the DNA sequence into a protein sequence. Spot any difference(s).

Hands On

Page 5

Descriptions

Graphic Summary

Alignments

Taxonomy

Alignment view

Pairwise

☒ CDS feature

Restore defaults

3 sequences selected

Download

GenBank

Graphics

PREDICTED: Felis catus myosin binding protein C3 (MYBPC3), transcript variant X2, mRNA

Sequence ID: [XM\\_019812397.2](#) Length: 4282 Number of Matches: 1

Range 1: 2528 to 2663 [GenBank](#) [Graphics](#)

▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
246 bits(133)	4e-61	135/136(99%)	0/136(0%)	Plus/Plus

CDS: Putative 1

Query

1

5

G Y I L E R K K K K S F R W M W L N F D

GGCTACATCCTGGAGCGCAAGAAGAAGAAGAGCTTCCGGTGGATGTGGCTGAACCTTTGAC

64

Sbjct

2528

GGCTACATCCTGGAGCGCAAGAAGAAGAAGAGCTTCCGGTGGATGCGGCTGAACCTTTGAC

2587

CDS:myosin-binding p

803

G Y I L E R K K K K S F R W M **P** L N F D

CDS: Putative 1

Query

21

65

L L Q E L S H E A R R M I E G V V Y E M

CTGCTGCAGGAGCTGAGCCACGAGGCACGGCGCATGATTGAGGGCGTGGTGTATGAGATG

124

Sbjct

2588

CTGCTGCAGGAGCTGAGCCACGAGGCACGGCGCATGATTGAGGGCGTGGTGTATGAGATG

2647

CDS:myosin-binding p

823

L L Q E L S H E A R R M I E G V V Y E M

CDS: Putative 1

Query

41

125

R V Y A V

CGAGTCTACGCGGTCA

140

Sbjct

2648

CGAGTCTACGCGGTCA

2663

CDS:myosin-binding p

843

R V Y A V

From: <https://pixabay.com/vectors/dna-amino-acids-biology-code-152135/>



Record what you see. Refer to the table of codons at the link below to identify the names of the amino acids that changed <https://pixabay.com/vectors/dna-amino-acids-biology-code-152135/>

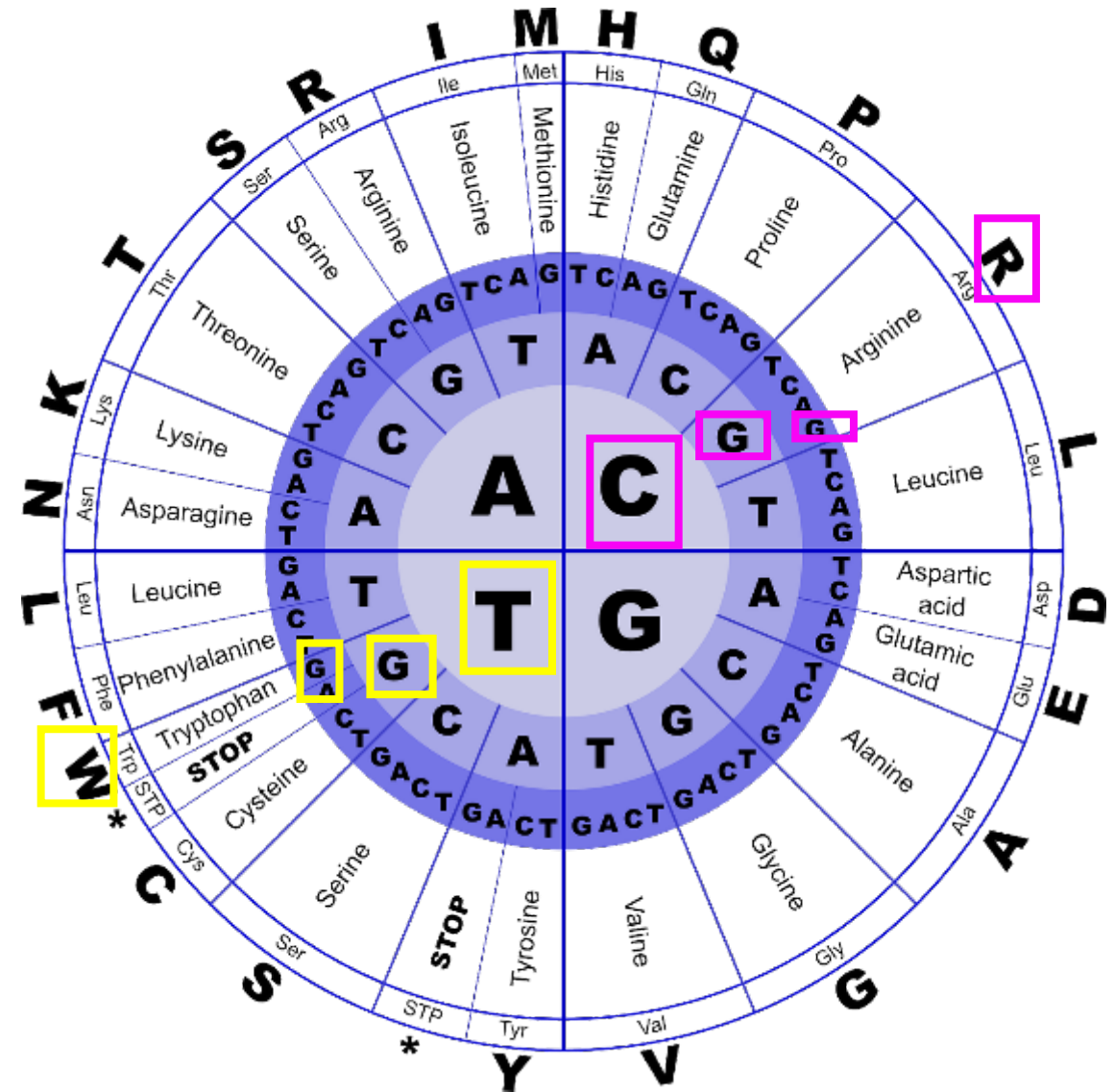
### Wound biopsy sequence

W	M	W	L	N	F	D	
TGG	ATG	TGG	CTG	AACT	TTT	GAC	64
TGG	ATG	CGG	CTG	AACT	TTT	GAC	2587
W	M	R	L	N	F	D	

### GenBank (reference sequence)

Nucleotide C>G

Amino acid R (Arg) > W (Trp)







## 'Normal' MYPBC3 cat protein

1

```
/translation="MPEPGKKPVSAFSKKPRSVEVAASSSAVFEAETERSGVKVRWQR
GGSDISASDKYGLAAEGTRHTLTVRDVGPTDQGPYAVIAGSSKVKFDLKVIEAEKAEP
VPGPAPAPTEAPGGSGEALTSTTEEEGGSPSPKGSSSAAPDGSGASDDPIGLFVMRPQ
DGEVTAGGSITFSARVAGASLLKPPTVKWFKGKWDLSSKVGQHLQLHNSYDRTSKVY
LFELHITDAQTTFAGGYRCEVSTKDKFDSCNFSLTVHEAIGPGDLDLRSAFRRTSLAG
SGRRISDSHEDAGTLDFSSLLKKRDSFRNLRDPRLEAPAEEDVWEILRQASPSEYERI
AFQHGITDLRGMLKRLKGMKRDEKKSTAFQKKLEPAYQVSKGHKIRLTVELADPDAEV
KWLKNGQEIQMSGSKYIFESVGAKRTLTIQCSLADDAAYQCVVGGEKCSTELFVKEP
PVLITRPLEDQLVMVGQRVEFECEVSEEGAQVKWLKDGVELTREETFKYRFKKDGQRH
HLIINEATLEDAGHYALRTSGGQALAEILVQEKKLEVYQSIADLTVGAKDQAVFKCEV
SDEN                                DYSFVPEGFACNLS
AKLH                                VPISGDPAPTVIWQ
KTLT                                TTKDRSIFTVEGAE
KEDE                                SCTVQWEPPAYDGG
QPVLGYILERKKKKSFRRWMRLNFDLLQELSHEARRMIEGVVYEMRVYAVNAIGMSRPS
PASQPFMPIGPPSEPTHLAVEDVSDTTVSLKWRPPERVGAGGLDGYSVEYCREGCSEW
VAALEGLTERTSLLVKDLPTGARLLFRVRAHNMAGPGAPIATKEPVTVQEILQRPRLQ
VPRHLRQTIQRKVGEVPVNLIPFQGKPRPQVTWTKEGQPLAGEEVSIRNSPTDTILFI
RAAHRTHSGTYEVMRLRIENMEDKATLVLRVVDKPSPPQDIRVTEAWGFNVVLEWKPPQ
DDGNTIELGYTVQKADKKTMEWFTVLEHYRRTHCVVSELIIGNGYFRVFSHNTVGPS
DNAATTKEPVFIPRPGITYEPPNYKALDFSEAPSFTRPLVNRSVIAGYNATLCCAVRG
SPKPKVSWFKNGLDLGEDARFRMFSKQGVLTLEIRKTCFPDGGVYVCRATNLQGEAQC
ECRLEVRVPQ"
```

WMRLNFD

p.R818W

## Our 'Wound Biopsy Sequence MYPBPC3 protein

```
/translation="MPEPGKKPVSAFSKKPRSVEVAASSSAVFEAETERSGVKVRWQR
GGSDISASDKYGLAAEGTRHTLTVRDVGPTDQGPYAVIAGSSKVKFDLKVIEAEKAEP
VPGPAPAPTEAPGGSGEALTSTTEEEGGSPSPKGSSSAAPDGSGASDDPIGLFVMRPQ
DGEVTAGGSITFSARVAGASLLKPPTVKWFKGKWDLSSKVGQHLQLHNSYDRTSKVY
LFELHITDAQTTFAGGYRCEVSTKDKFDSCNFSLTVHEAIGPGDLDLRSAFRRTSLAG
SGRRISDSHEDAGTLDFSSLLKKRDSFRNLRDPRLEAPAEEDVWEILRQASPSEYERI
AFQHGITDLRGMLKRLKGMKRDEKKSTAFQKKLEPAYQVSKGHKIRLTVELADPDAEV
KWLKNGQEIQMSGSKYIFESVGAKRTLTIQCSLADDAAYQCVVGGEKCSTELFVKEP
PVLITRPLEDQLVMVGQRVEFECEVSEEGAQVKWLKDGVELTREETFKYRFKKDGQRH
HLIINEATLEDAGHYALRTSGGQALAEILVQEKKLEVYQSIADLTVGAKDQAVFKCEV
SDENV                                VPEGFACNLS
AKLHF                               IGDAPTVIWQ
KTLTQ                               RSIFTVEGAE
KEDEG                              'QWEPPAYDGG
QPVLGYILERKKKKSFRRWMRLNFDLLQELSHEARRMIEGVVYEMRVYAVNAIGMSRPS
PASQPFMPIGPPSEPTHLAVEDVSDTTVSLKWRPPERVGAGGLDGYSVEYCREGCSEW
VAALEGLTERTSLLVKDLPTGARLLFRVRAHNMAGPGAPIATKEPVTVQEILQRPRLQ
VPRHLRQTIQRKVGEVPVNLIPFQGKPRPQVTWTKEGQPLAGEEVSIRNSPTDTILFI
RAAHRTHSGTYEVMRLRIENMEDKATLVLRVVDKPSPPQDIRVTEAWGFNVVLEWKPPQ
DDGNTIELGYTVQKADKKTMEWFTVLEHYRRTHCVVSELIIGNGYFRVFSHNTVGPS
DNAATTKEPVFIPRPGITYEPPNYKALDFSEAPSFTRPLVNRSVIAGYNATLCCAVRG
SPKPKVSWFKNGLDLGEDARFRMFSKQGVLTLEIRKTCFPDGGVYVCRATNLQGEAQC
ECRLEYRVPQ"
```

WMWLNFD

p.R818W



14) Open the National Library of Medicines [PubMed search engine](#)

15) Type 'MYBPC3 in cat' in the search box and learn about MYBPC3 variants by reading the article written by Boeykens et al 2024

The screenshot shows the PubMed search results for the query 'MYBPC3 in cat'. The page header includes the NIH logo and the text 'National Library of Medicine National Center for Biotechnology Information'. The search bar contains the query 'MYBPC3 in cat' and a 'Search' button. Below the search bar are links for 'Advanced', 'Create alert', 'Create RSS', and 'User Guide'. The results section shows 25 results, with the first two results displayed. The first result is titled 'CRISPR/Cas9 gene editing in induced pluripotent stem cells to investigate the feline hypertrophic cardiomyopathy causing MYBPC3/R820W mutation.' and is by Dutton LC, Dudhia J, Guest DJ, Connolly DJ. The second result is titled 'Classification of feline hypertrophic cardiomyopathy-associated gene variants according to the American College of Medical Genetics and Genomics guidelines.' and is by Boeykens F, Abitbol M, Anderson H, Dargatzis T, Ferrari P, Fox PR, Hayward JJ, Haggström J, Davison S, Kittleson MD, van Steenbeek F, Ljungvall I, Lyons LA, Longeri M, Ohlsson Å, Peelmann L, Dufaure de Citres C, Smets P, Turba ME, Broeckx BJG. The page also includes a 'RESULTS BY YEAR' bar chart and a 'PUBLICATION DATE' filter section.

NIH National Library of Medicine  
National Center for Biotechnology Information

Log in

PubMed®

MYBPC3 in cat

Advanced Create alert Create RSS User Guide

Save Email Send to Sort by: Most recent Display options

MY CUSTOM FILTERS

25 results

Page 1 of 3

RESULTS BY YEAR

2005 2024

PUBLICATION DATE

1 year  
5 years  
10 years  
Custom Range

TEXT AVAILABILITY

Abstract  
Free full text

1 CRISPR/Cas9 gene editing in induced pluripotent stem cells to investigate the feline hypertrophic cardiomyopathy causing MYBPC3/R820W mutation.  
Dutton LC, Dudhia J, Guest DJ, Connolly DJ.  
PLoS One. 2024 Oct 10;19(10):e0311761. doi: 10.1371/journal.pone.0311761. eCollection 2024.  
PMID: 39388496 Free PMC article.  
The aim of this study was to generate a cellular model of the feline HCM-causing MYBPC3 mutation R820W. Using CRISPR/Cas9 gene editing we introduced the R820W mutation into a human induced pluripotent stem cell (iPSC) line. ...In summary, we demonstrate successful generati ...

2 Classification of feline hypertrophic cardiomyopathy-associated gene variants according to the American College of Medical Genetics and Genomics guidelines.  
Boeykens F, Abitbol M, Anderson H, Dargatzis T, Ferrari P, Fox PR, Hayward JJ, Haggström J, Davison S, Kittleson MD, van Steenbeek F, Ljungvall I, Lyons LA, Longeri M, Ohlsson Å, Peelmann L, Dufaure de Citres C, Smets P, Turba ME, Broeckx BJG.  
Front Vet Sci. 2024 Feb 2;11:1327081. doi: 10.3389/fvets.2024.1327081. eCollection 2024.  
PMID: 38371598 Free PMC article.  
In silico evaluation followed with joint evidence and data from other publications assisting in the classification of each variant. RESULTS: Two variants, MYBPC3:c.91G > C [A31P] and MYBPC3:c.2453C > T [R818W], were designated as pathogenic. ...DISCUSSION: Rou ...

[OMIA \(Online Mendelian Inheritance in Animals is another good resource:  
https://omia.org/OMIA000515/9685/](https://omia.org/OMIA000515/9685/)



# Classification of feline hypertrophic cardiomyopathy-associated gene variants according to the American College of Medical Genetics and Genomics guidelines

Frédérique Boeykens<sup>1</sup>, Marie Abitbol<sup>2</sup>, Heidi Anderson<sup>3</sup>, Tanushri Dargar<sup>2</sup>, Paolo Ferrari<sup>4, 5</sup>, Philip R Fox<sup>6</sup>, Jessica J Hayward<sup>7</sup>, Jens Häggström<sup>8</sup>, Stephen Davison<sup>9</sup>, Mark D Kittleson<sup>10</sup>, Frank van Steenbeek<sup>11</sup>, Ingrid Jungvall<sup>8</sup>, Leslie A Lyons<sup>12</sup>, Maria Longeri<sup>13</sup>, Åsa Ohlsson<sup>14</sup>, Luc Peelman<sup>1</sup>, Caroline Dufaure de Citres<sup>15</sup>, Pascale Smets<sup>16</sup>, Maria Elena Turba<sup>17</sup>, Bart J G Broeckx<sup>1</sup>

## Abstract

**Introduction:** The correct labeling of a genetic variant as pathogenic is important as breeding decisions based on incorrect DNA tests can lead to the unwarranted exclusion of animals, potentially compromising the long-term health of a population. In human medicine, the American college of Medical Genetics (ACMG) guidelines provide a framework for variant classification. This study aims to apply these guidelines to six genetic variants associated with hypertrophic cardiomyopathy (HCM) in certain cat breeds and to propose a modified criterion for variant classification.

**Methods:** Genetic samples were sourced from five cat breeds: Maine Coon, Sphynx, Ragdoll, Devon Rex, and British Short- and Longhair. Allele frequencies were determined, and in the subset with phenotypes available, odds ratios to determine the association with HCM were calculated. *In silico* evaluation followed with joint evidence and data from other publications assisting in the classification of each variant.

**Results:** Two variants, MYBPC3:c.91G > C [A31P] and MYBPC3:c.2453C > T [R818W], were designated as pathogenic. One variant, MYH7:c.5647G > A [E1883K], was found likely pathogenic, while the remaining three were labeled as variants of unknown significance.

**Discussion:** Routine genetic testing is advised solely for the MYBPC3:c.91G > C [A31P] in the Maine Coon and MYBPC3:c.2453C > T [R818W] in the Ragdoll breed. The human ACMG guidelines serve as a suitable foundational tool to ascertain which variants to include; however, refining them for application in veterinary medicine might be beneficial.



16.) The following table shows different cat breeds with MYBPC3 variants.

Occurrence of MYBPC3 variants in different cat breeds

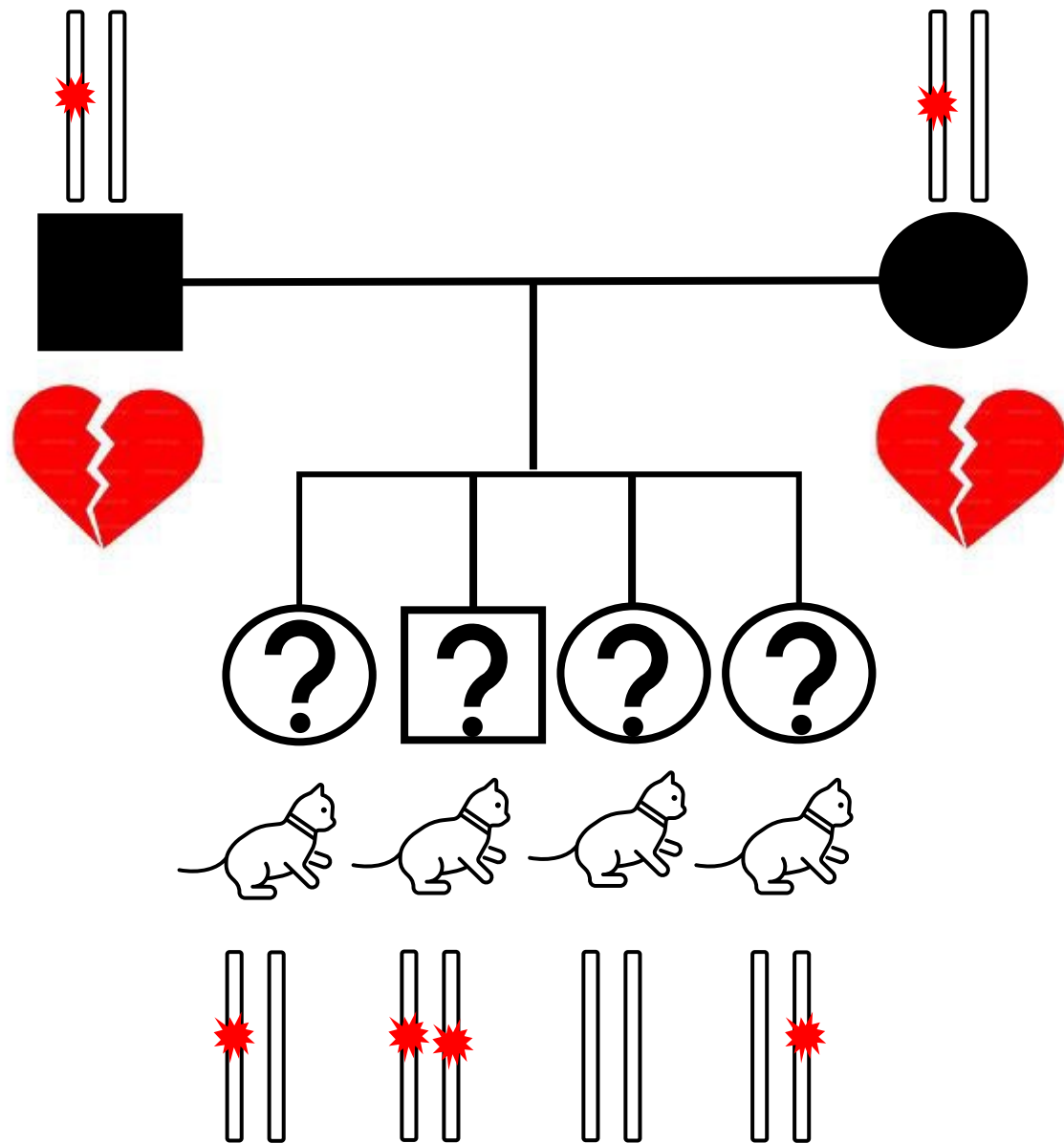
Pathogenic Variant	Maine Coon	Scottish Fold	British Longhair	Munchkin	Ragdoll
91G>C [A31P]	✓	✓	✓	✓	✓
2453C>T [R818W]	●	●	●	●	✓

17.) Note that for technical reasons, the DNA position 2453 in the above table corresponds to position 2573 in *Felis catus* MYBPC3 transcript variant X2 (the first 121 nucleotides of the X2 variant mRNA are untranslated). [https://www.ncbi.nlm.nih.gov/nucleotide/XM\\_019812397.2](https://www.ncbi.nlm.nih.gov/nucleotide/XM_019812397.2)

Based on your sequence data analysis, can you identify which type of cat killed baby Tara iti?







## Hypertrophic cardiomyopathy in cats:

- Heart muscle – thickened (hypertrophied)
- Autosomal dominant
- Homozygous cats: early onset and severe phenotype (disease)
- Hardly any symptoms > congestive heart failure > sudden death

## Hypertrophic cardiomyopathy in human:

- Autosomal dominant
- MYBPC3 mutations – most common HCM cause
- Gene therapy trials (using AAV9 virus)



# Example #2

- 'Predator' sequence:

**TGCACGGAGACCAACTTCATTACATTAATCCTGAGACCCTGGAGACAAT  
TAAGCAGGTTGATCTCTGCAACTACGTCTCTGTCAATGGAGCCAC**

**Go to: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>**



[Descriptions](#)
[Graphic Summary](#)
[Alignments](#)
[Taxonomy](#)

Alignment view

Pairwise



CDS feature


[Restore defaults](#)

3 sequences selected


[Download](#)
[GenBank](#) [Graphics](#)

## Canis lupus familiaris retinoid isomerohydrolase RPE65 (RPE65), mRNA

Sequence ID: [NM\\_001003176.1](#) Length: 2390 Number of Matches: 1

[See 1 more title\(s\)](#) [See all Identical Proteins\(IPG\)](#)

Range 1: 463 to 560 [GenBank](#) [Graphics](#)
[Next Match](#) [Previous Match](#)

Score		Expect	Identities	Gaps	Strand
156 bits(84)		4e-34	94/98(96%)	4/98(4%)	Plus/Plus
Query	1	TGCACGGAGACCAACTTCATTAC----ATTAATCCTGAGACCCTGGAGACAATTAAGCAG			56
Sbjct	463	TGCACGGAGACCAACTTCATTACAAAGATTAATCCTGAGACCCTGGAGACAATTAAGCAG			522
Query	57	GTTGATCTCTGCAACTACGTCTCTGTCAATGGAGCCAC			94
Sbjct	523	GTTGATCTCTGCAACTACGTCTCTGTCAATGGAGCCAC			560



Descriptions

Graphic Summary

Alignments

Taxonomy

Alignment view

Pairwise



CDS feature



Restore defaults

3 sequences selected



Download

[GenBank](#)[Graphics](#)**Canis lupus familiaris retinoid isomerohydrolase RPE65 (RPE65), mRNA**Sequence ID: [NM\\_001003176.1](#) Length: 2390 Number of Matches: 1[See 1 more title\(s\)](#) [See all Identical Proteins\(IPG\)](#)**Range 1: 463 to 560** [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
156 bits(84)	4e-34	94/98(96%)	4/98(4%)	Plus/Plus
CDS: Putative 1	1	C T E T N F I T L I L R P W R Q L S R		
Query	1	TGCACGGAGACCAACTTCATTAC---ATTAATCCTGAGACCCTGGAGACAATTAAGCAG	56	
Sbjct	463	TGCACGGAGACCAACTTCATTACAAAGATTAATCCTGAGACCCTGGAGACAATTAAGCAG	522	
CDS:retinoid isomero	146	C T E T N F I T K I N P E T L E T I K Q		
CDS: Putative 1	20	L I S A T T S L S M E P		
Query	57	GTTGATCTCTGCAACTACGTCTCTGTCAATGGAGCCAC	94	
Sbjct	523	GTTGATCTCTGCAACTACGTCTCTGTCAATGGAGCCAC	560	
CDS:retinoid isomero	166	V D L C N Y V S V N G A T		





The deletion of 4 letters(AAGA) in the RPE65 gene changes the RPE65 protein – the amino acids after the missing 4 DNA letters are all different to the normal RPE65 protein (frameshift mutation).

RPE65 – frameshift → truncating the protein – causing retinal dystrophy / Autosomal recessive

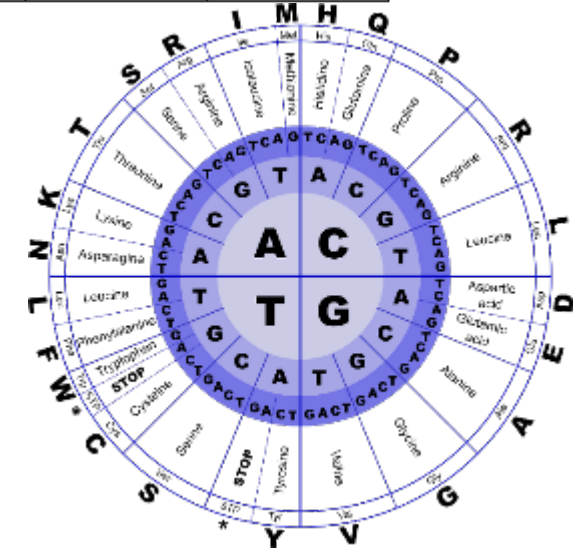


	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Nucleotide GenBank RPE65 sequence	T	T	C																											
Protein GenBank RPE65 sequence	F																													
Nucleotide "Mystery" RPE65 sequence	T	T	C																											
Protein "Mystery" RPE65 sequence	F																													

Range 1: 463 to 560 [GenBank](#) [Graphics](#)

▼ [Next Match](#) ▲ [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
156 bits(84)	4e-34	94/98(96%)	4/98(4%)	Plus/Plus
CDS: Putative 1	1	C T E T N F I T L I L R P W R Q L S R		
Query	1	TGCACGGAGACCAACTTCATTAC----ATTAATCCTGAGACCCTGGAGACAATTAAGCAG	56	
Sbjct	463	TGCACGGAGACCAACTTCATTACAAAGATTAATCCTGAGACCCTGGAGACAATTAAGCAG	522	
CDS:retinoid isomero	146	C T E T N F I T K I N P E T L E T I K Q		
CDS: Putative 1	20	L I S A T T S L S M E P		
Query	57	GTTGATCTCTGCAACTACGTCTCTGTCAATGGAGCCAC	94	
Sbjct	523	GTTGATCTCTGCAACTACGTCTCTGTCAATGGAGCCAC	560	
CDS:retinoid isomero	166	V D L C N Y V S V N G A T		

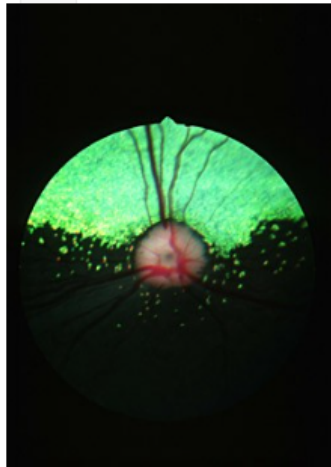




# Gene therapy restores vision to dogs blinded by inherited disease, bringing new hope to childhood sufferers of similar condition

By [H. Roger Segelken](#)

April 27, 2001



G. Aguirre/Cornell University

A retinal fundus photograph of a Briard dog with the inherited eye disease, congenital stationary night blindness. Mutations to the same gene in humans, RPE65, cause Leber congenital amaurosis, beginning in childhood. But gene-therapy experiments in Briards hold hope for curing the disease in humans.

Dogs blinded by an inherited retinal degenerative disease had their vision restored after treatment with genes from healthy dogs, marking the first successful gene therapy for blindness in a large animal. The treatment offers hope for humans with a similar condition.

The achievement, with young dogs suffering from congenital stationary night blindness, which is similar to a childhood disease called Leber congenital amaurosis, is reported in the May 2001 issue of the journal *Nature Genetics* by Gregory M. Acland, Gustavo D. Aguirre, Jharna Ray, Qi Zhang and Susan E. Pearce-Kelling, all at the Cornell University College of Veterinary Medicine; by Tomas S. Aleman, Artur V. Cideciyan, Vibha Anand, Yong Zeng, Albert M. Maguire, Samuel G. Jacobson and Jean Bennett, all at the University of Pennsylvania; and by William W. Hauswirth of the University of Florida, Gainesville.

"We have shown that gene therapy can restore vision in dogs with one of the most clinically severe retinal degenerations," says Acland, a research veterinarian at Cornell's James A. Baker Institute for Animal Health.

2001



News posted 21 February, 2020

## First patient undergoes revolutionary new gene therapy procedure at Manchester Royal Eye Hospital

[Home](#) [News](#) First patient undergoes revolutionary new gene therapy procedure at Manchester Royal Eye Hospital



### Categories

- > News
- > Trust News
- > Research

2020



# Why is genetic diagnosis important?

- Preventative measures (LQTS in NZ)
- Genetic counselling and testing of patients whānau
- Reproductive genetic counselling
- Therapeutic
- Gene therapy



# Genomics careers

- Medical laboratory scientists (diagnostics)
- Bioinformatics
- Computer science
- Genetic variant analysis / curation
- Genetic Counsellors
- Clinical geneticists
- Forensics
- Biologists, functional studies (cell, animal models)
- Pharmacogenomics
- All medical specialities, nurses ...
- Drug development, gene therapies ...



# Discussion

- Direct to consumer tests (Ancerstry, 23 and Me)
- DNA and whakapapa; DNA is tapu
- Storage of DNA samples
- Analysis and storage of genomic data
- Genetic information and insurance:

<https://www.1news.co.nz/2024/05/06/fair-go-should-insurers-discriminate-based-on-genetic-profiles/>

HEALTH

## Remember That DNA You Gave 23andMe?

The company is in trouble, and anyone who has spit into one of its test tubes should be concerned.

By Kristen V. Brown

<https://www.theatlantic.com/health/archive/2024/09/23andme-dna-data-privacy-sale/680057/>





UNIVERSITY OF  
AUCKLAND  
Waipapa Taumata Rau  
NEW ZEALAND

140  
YEARS






- Slides that follow are not part of the presentation, but might be helpful
- Slide 41:
  - this is the mRNA *MYBPC3* sequence (without introns)
  - Start codon (ATG) where the protein synthesis starts is underlined
  - Wound biopsy sequence is underlined
- Slide 42: an example of a laboratory offering the MYBPC3 genetic test
- Slide 43: explanation of how sequencing works



Start of the coding sequence (ATG)  
- At position 122 in

Click here:  
[https://www.ncbi.nlm.nih.gov/nucleotide/XM\\_019812397.2](https://www.ncbi.nlm.nih.gov/nucleotide/XM_019812397.2)

```
1  aggcagatata  gaaatcctgg  ggagcggggg  ggccctctat  atagtgtggg  gggacgggtc
61  ctgctcagat  cctccttggg  tggcctgtga  ctgctcagat  cctgggtgtg  tctctctcaa
121 gatgccttag  ccgggggaga  agccagtctc  agccttcagc  aagaagccaa  ggcctcagga
181  aggcagcc  agcagctctg  ctgtgttcga  ggccgagaca  gagcggctag  gagtaagggt
241  agccttggc  cgggggggca  gtgacatcag  cgccagtgac  aagtatggcc  tagcagccga
301  gggcagcagg  cacactctga  cagtgcggga  cgtgggcccc  accgaccagg  gaccctacgc
361  agtcatcgct  ggctcctcca  aggtcaaggt  tgacctcaag  tgcatagaag  cagagaaggc
421  agagcctgtg  ccggggcctg  ctctgcccc  tactgaggct  cctggaggct  ccggagaagc
481  actgacctcg  accactgagg  aggaaggagg  ctccccaaat  cccaaagggt  caagctcagc
541  agccccgat  ggctctgggt  cctctgatga  ccccatcggc  ctcttgttga  tgcggccaca
601  ggaaggcgag  gtgaccgag  gtggcagcat  caccttctca  gcccgctggg  ccggcgccag
661  cctcctgaaa  ccgcccagag  tcaagtgggt  caagggaag  tgggtggacc  tgagcagcaa
721  ggtggggcaa  cactgcagc  tgcacaacag  ttacgaccgg  accagcaagg  tctacctgtt
781  tgagctgcac  attacagatg  cccagactac  tttcgcagg  ggctaccgct  gtgaggtttc
841  caccgaagg  aaatttgaca  gctgcaact  cagtctcact  gtccatgagg  ccattggccc
901  tggagacctg  gacctcagat  cagctttccg  ccgcacgagc  ctggctggaa  gtggtcggcg
961  gatcagtgac  agccacgagg  acgttggaac  tctgatttc  agctcgtgc  tgaagaagag
1021 agacagtctt  cggaaacctaa  gggaccggag  gctggaggcg  ccagccgagg  aggacgtgtg
1081  ggagatcctg  cggcaggcat  cccctccga  atatgagcgc  atcgcttcc  agcacggcat
1141  caccgacctg  cgcgcatgc  tcaagaggct  caaggcatg  aagcgagacg  agaagaagag
1201  cacagccttt  cagaagaagc  tggagccgc  ctaccagggt  agcaaggccc  acaagatccg
1261  gctgacctg  gagctggccg  accccagcgc  cgaggtcaag  tggctgaaga  atggacagga
1321  gatccagatg  agcggcagca  agtacctctt  cgagtcctgt  ggtgccaaag  gcacgtgcac
1381  catcagccag  tgctcgtgg  ccgatgacgc  ggcttaccag  tgcgtgtggg  cgggggagaa
1441  gtgcagcacc  gagctgttcg  tcaaaagacc  ccccgctgct  atcactcgcc  ccttggaaga
1501  ccagctgggt  atggtggggc  agcgggtgga  gtttgagtgc  gaagtgtctg  aggggggggc
1561  ccaggtcaag  tggctgaagg  atgggggtga  gctgaccagg  gagagacct  tcaaatccg
1621  gttcaagaag  gacgggcaga  gacaccact  catcatcaac  gagcgagcgc  tggaggacgc
1681  ggggcactac  gcgctcgca  ccagcggggg  ccaggcgctg  gccgagctca  tcgttcagga
1741  gaaaaagctg  gaggtgtacc  agagctcgc  agacctgacc  gtgggtgcaa  aggaaccagg
1801  ggtgttcaag  tgtgaagtgt  ccgatgagaa  cgtgcggggc  gtgtggctga  agaaccggaa
1861  ggagctgggt  cccgacagcc  gcgtgaaggt  ggtccacatt  gggcgagtc  acaaacgtac
1921  cattgacgac  gtcacgcctg  ccgacgaggc  ggactacagc  tttgtgctgt  agggcttcgc
1981  ctgcaacctg  tcagcgaaac  tccatttcag  ggaggtcaag  attgacttgc  tgcccaggca
2041  agaacctccc  aagatccacc  tggactgccc  gggcgtgaag  ccgataacca  ttgtggttgt
2101  ggctgggaac  aagctacgcc  tggatgtccc  tatctctggg  gacctgtctc  ccactgtgat
2161  ttggcagaag  accctcacac  agaacaaggt  cccagcaggg  ccattccctg  atgccccagg
2221  ggaggatgca  ggtgacagt  acgagtgggt  gtttgacaag  aagctgtctg  gtgagaccga
2281  gggccgggtg  cgggtggaga  ccaccaagga  ccgacagatc  ttactgtgc  agggggcaga
2341  gaagggaagc  gaggtgtct  acgttgtcac  ggtgaagaac  cccgtggcgc  aggaaccagg
2401  caacctcaca  gtcaaggcca  tcgatgtgcc  agatcccccc  gcggccccc  agatcagcaa
2461  tgtgggtgag  gactcctgca  ccgtgcagt  ggagccgct  gcctacgagc  ggggacagcc
2521  agtcttgggc  tacatcctgg  agcgcaagaa  gaagaagagc  ttccggtgga  tgcggctgaa
2581  ctttgacctg  ctgcaggagc  tggccacga  ggcatggcgc  atgattgagg  gctgtgtgta
2641  tgagatgcga  gtctacgagg  tcaatgccat  cggcatgtcc  agggccagcc  ctgctctcca
2701  gcccttcacg  cctattggcc  cccccagtg  gcccaccac  ctggcgtgg  aggatgtctc
2761  agacaccacc  gtctccctca  agtgccggcc  cccggagcgg  gtgggagcgc  gaggtctgga
2821  tggctacagc  gtggagtatt  gccgggagg  ctgctcggag  tgggtagctg  ccttgagggg
2881  gctgacagaa  cgcacctcgt  tgcgtgtgaa  ggacctgccc  accggggccc  ggctgctctt
2941  ccgtgtgcgg  gcgcacaata  tggcggggcc  cggggccccc  atcgccacca  aggaagccgt
3001  gacgggtcag  gagatactgc  agcggccacg  gctccagggt  ccccgccacc  tgcgccagac
3061  catccagagg  aaggctgggg  agccgtgaa  cctctcatt  cctttccagg  gcaagccccg
3661  cttcactcgc  cctctgtgta  accgctcagt  cattgctggc  tacaatgcta  ccctctgtctg
3721  tgctgtccgg  ggtagcccca  agcccaaagg  ttctgtgttc  aagaatggcc  tggacctggg
3781  agaagatgcc  cgcttcgca  tgttcagcaa  gcagggtgtg  ttgaccttgg  agatcagaaa
3841  gacctgcccc  tttgatgggg  gcctctatgt  ctgcaggccc  accaaactac  aaggcgaggc
3901  acagtgtgag  tggcgcctag  aagtacgagt  gctcagtgta  tctggtggc  tcccgagag
3961  ggccagggtg  cacaggatgc  cagccatgt  actgggagcc  tggaaaggaa  ctggacaagc
4021  cgcctctcgc  tgcgtgatat  gtgtgtgcaa  gtgtgtatct  tagggttggg  ctagtggaa
4081  catcagtagg  gcagtgaat  taaggtagc  gggcctgggg  ggctcgccac  acggactgaa
4141  ctgagctccc  actggaagct  ctgaggtctt  tgggacaaat  ggggtgtagg  ctgacaaaat
4201  ggtgactggg  gacaggcctt  tctggcctgc  catgtctct  caatgtataa  gggcaataaa
4261  agatcagtcg  ggtcacagg  a
```

 **National Library of Medicine**  
National Center for Biotechnology Information

Nucleotide

GenBank

**PREDICTED: Fells catus myosin binding protein C3 (MYBPC3), transcript variant X2, mRNA**  
NCBI Reference Sequence: XM\_019812397.2  
[FASTA](#) [Graphics](#)

**MYBPC3** gene – an instruction  
for a building block of the **heart**

>Wound Biopsy Sequence  
ATCAGGCTACATCCTGGAGCG  
CAAGAAGAAGAAGAGCTTCCG  
GTGGATGTGGCTGAACTTTGA  
CCTGCTGCAGGAGCTGAGCC  
ACGAGGCACGGCGCATGATTG  
AGGGCGTGGTGTATGAGATGC  
GAGTCTACGCGGTCA





## 17.) Can you identify which type of cat killed baby Tara iti?



HOME DOG CAT COW HORSE BLOG MORE CONTINUING EDUCATION

### Hypertrophic Cardiomyopathy (Maine Coon, Ragdoll types)

**Gene:** MYBPC3

**Transmission:** Autosomal dominant

The animal only has to have one copy of the mutation to be at risk of developing hypertrophic cardiomyopathy. Animals with two copies of the mutation generally have more severe symptoms and an earlier onset of the disease than animals with just one copy of the mutation. Offspring are potentially at risk of developing the disease if at least one parent carries the mutation.

**Mutations:**

Maine Coon mutation: Substitution, MYBPC3 gene; c.91 G>C, p.(A31P), exon2

Ragdoll mutation: Substitution, MYBPC3 gene; c.2453 C>T, p.(R818W)

**Medical system:** Cardiac

**Breeds:** American Bobtail, Domestic Cat, Highland Lynx, Maine Coon, Munchkin, Pixie-bob, Ragamuffin, Ragdoll, Scottish Fold, Siberian

**Age of onset of symptoms:** By 4 years of age.

Hypertrophic cardiomyopathy is the most common heart disease in cats. In this disease the heart muscle is enlarged and left ventricular walls are thickened and are not dilated. Hypertrophic cardiomyopathy affects young cats as a genetic disease and older cats as a secondary disease to hyperthyroidism. At present, the causative mutations are known for the Maine Coon and Ragdoll breeds of cat. Homozygous animals, with two copies of the mutation in question, can develop a more severe hypertrophic cardiomyopathy.

Note that hypertrophic cardiomyopathy in the cat has complex genetics with a number of known gene mutations involved as well as unknown mutations yet to be identified. Characterized mutations include:

Maine Coon cat



[https://commons.wikimedia.org/wiki/File:Maine\\_Coon\\_male\\_NO\\_Sigdalskauen\\_Balder.jpg](https://commons.wikimedia.org/wiki/File:Maine_Coon_male_NO_Sigdalskauen_Balder.jpg)  
Tbjornstad (talk | contribs) / Creative Commons Attribution-Share Alike 4.0 International license.

Ragdoll cat



Look at: <https://labgenvet.ca/en/disease/hypertrophic-cardiomyopathy-maine-coon-ragdoll/>