

Opening Thought Question – 4/25

Last week we learned that sea otters underwent a genetic bottleneck.

1. What is a bottleneck? (Define bottleneck)
2. Predict whether sunflower sea stars have experienced a bottleneck due to sea star wasting disease.
3. How would you test your prediction?

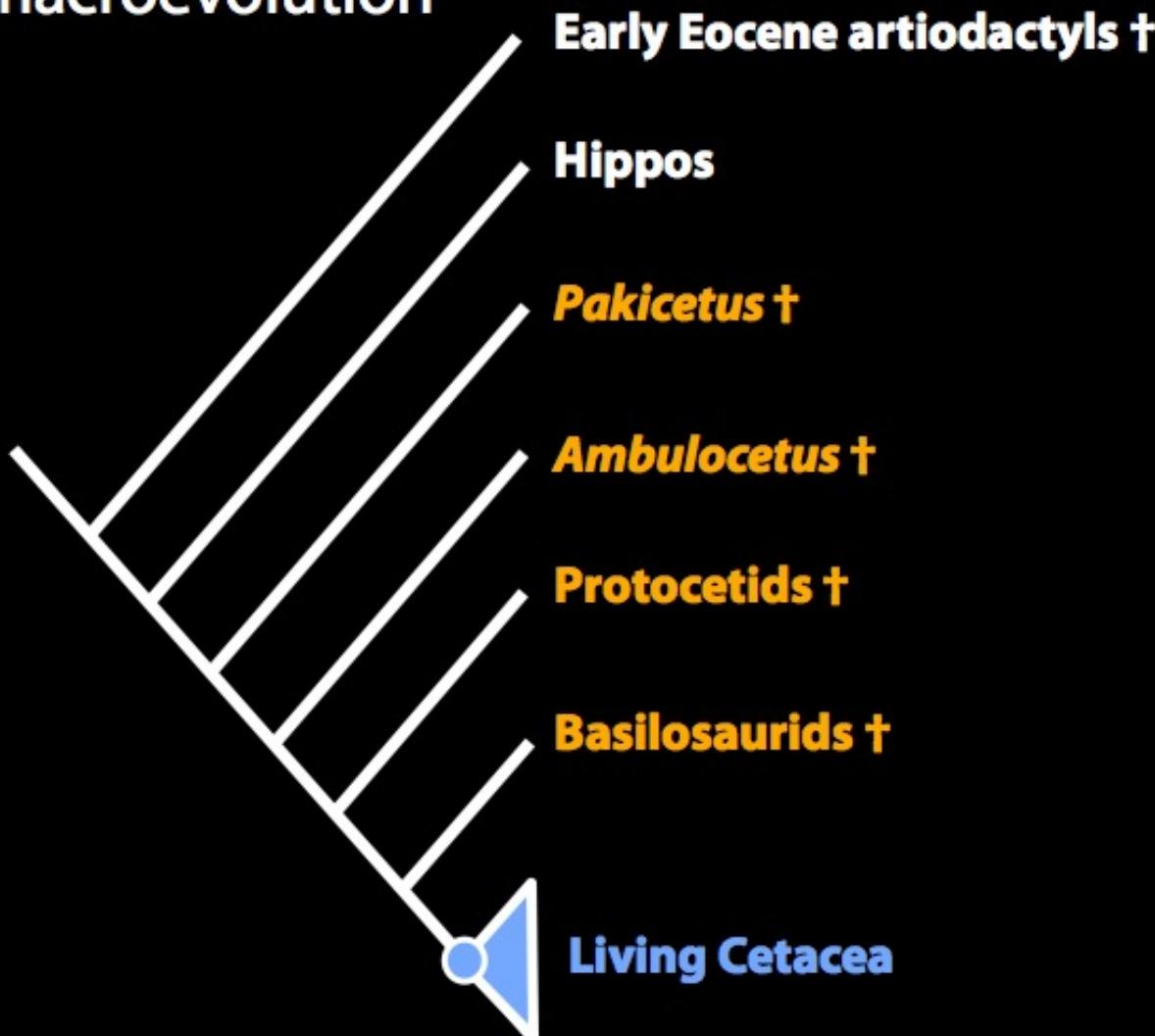
Whales are marine tetrapods

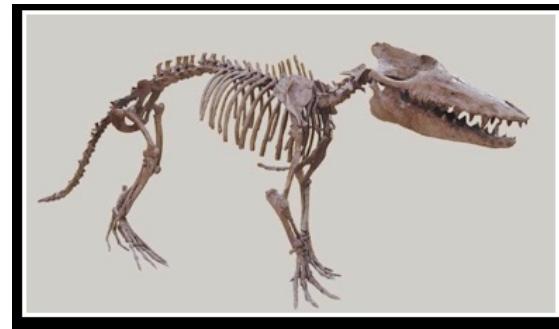


What are Marine Tetrapods?

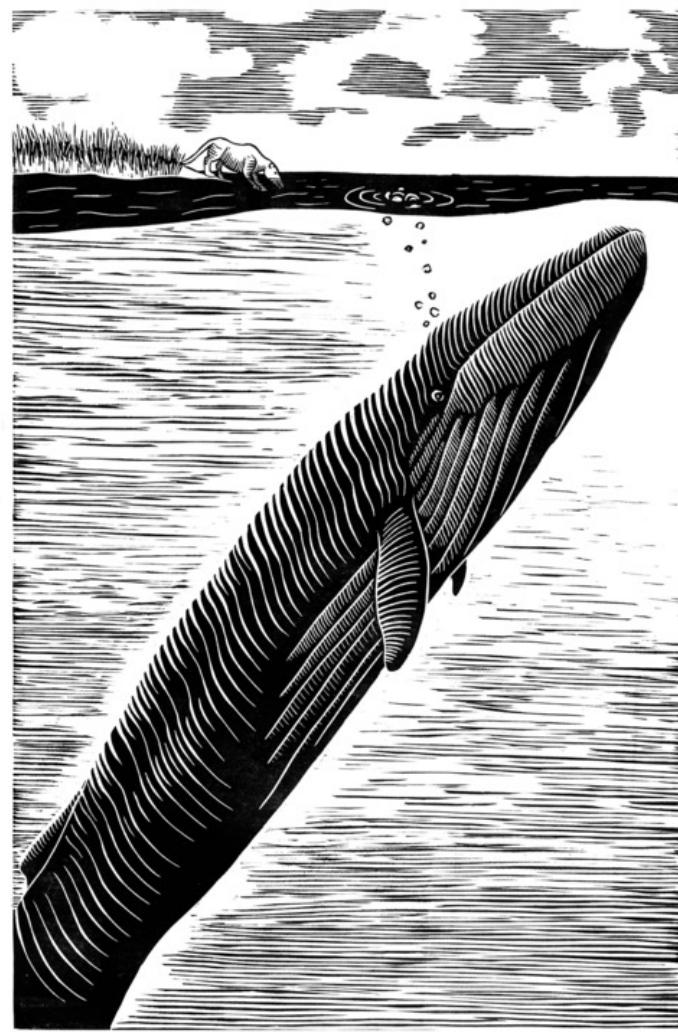
- “Four-footed”
 - Evolved from ancestors with 2 sets of limbs
- Terrestrial ancestry
- Obtain food from the sea
- Seabirds, Marine Reptiles, Marine Mammals

macroevolution



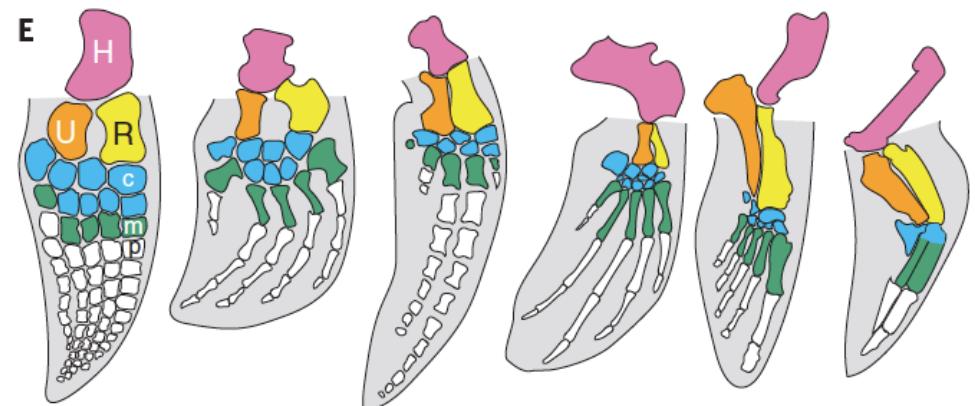


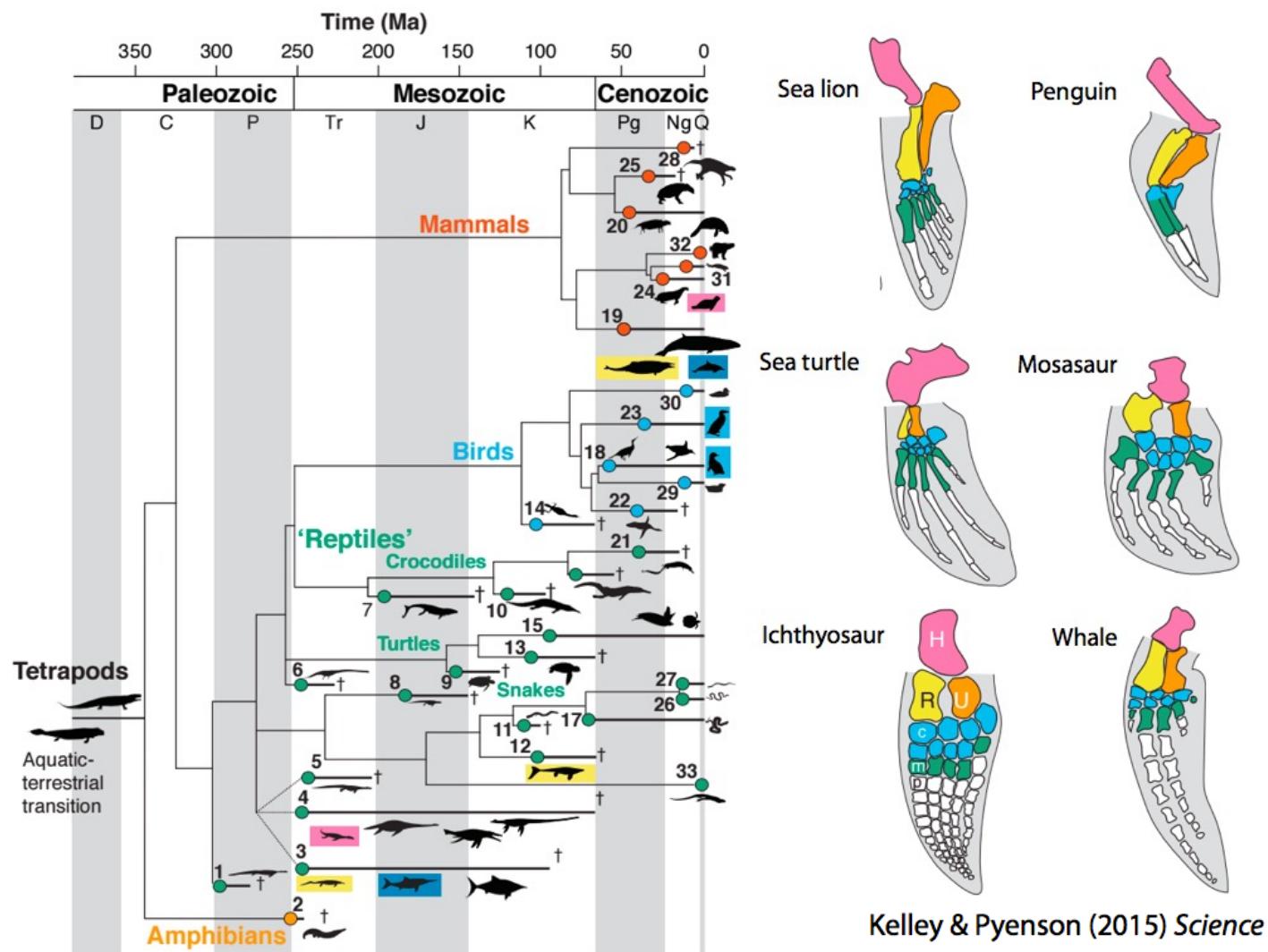
Art by Carl Buell & Alex Boersma



Adaptations to Oceanic Life

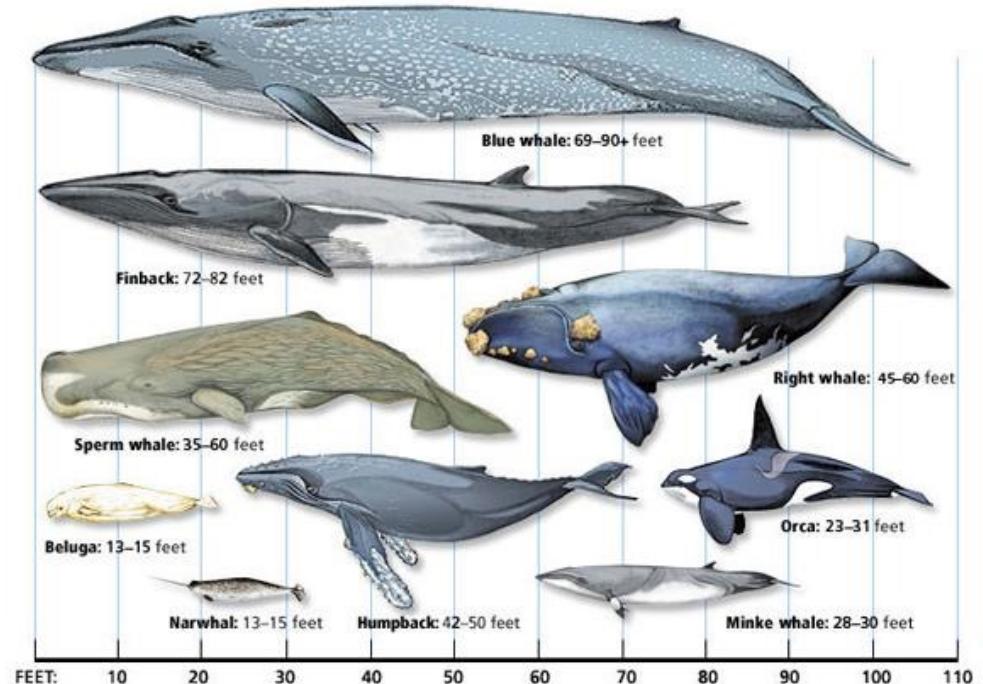
- Locomotion





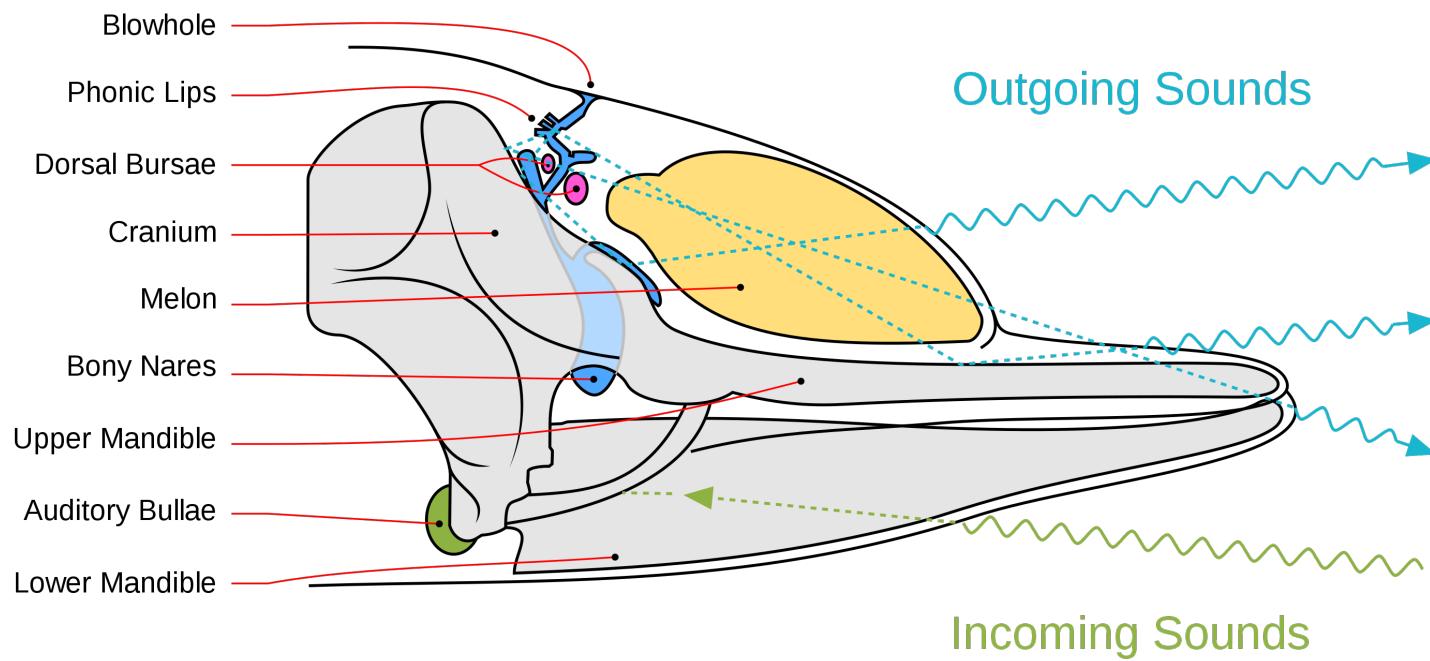
Adaptations to Oceanic Life

- Locomotion
- Gravity effects

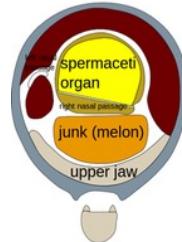


Adaptations to Oceanic Life

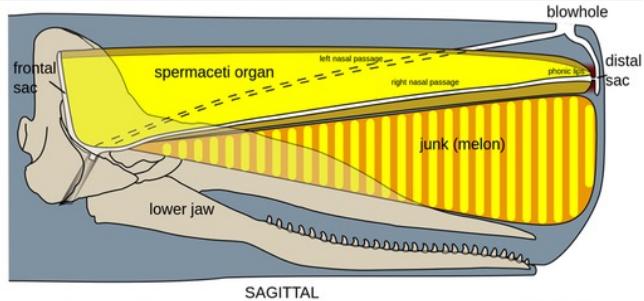
- Locomotion
- Gravity effects
- Alternative sensory systems



Echolocation in sperm whales



TRANSVERSE



Adaptations to Oceanic Life

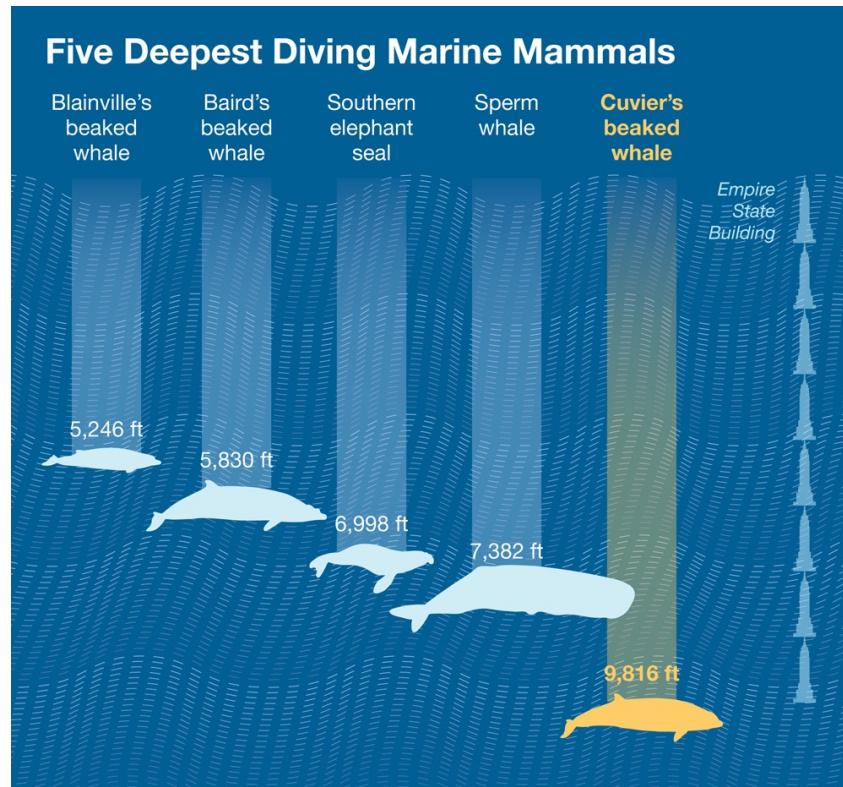
- Locomotion
- Gravity effects
- Alternative sensory systems
- O₂ capture



Photo copyright of Hadoram Shirihai

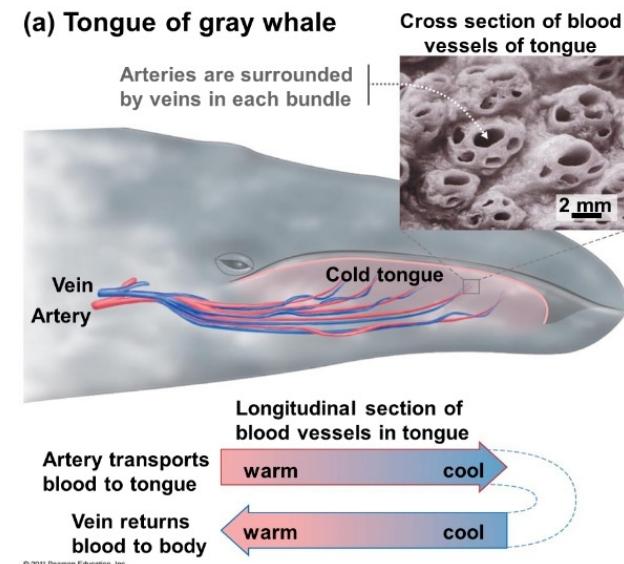
Adaptations to Oceanic Life

- Locomotion
- Gravity effects
- Alternative sensory systems
- O₂ capture
- Diving physiology



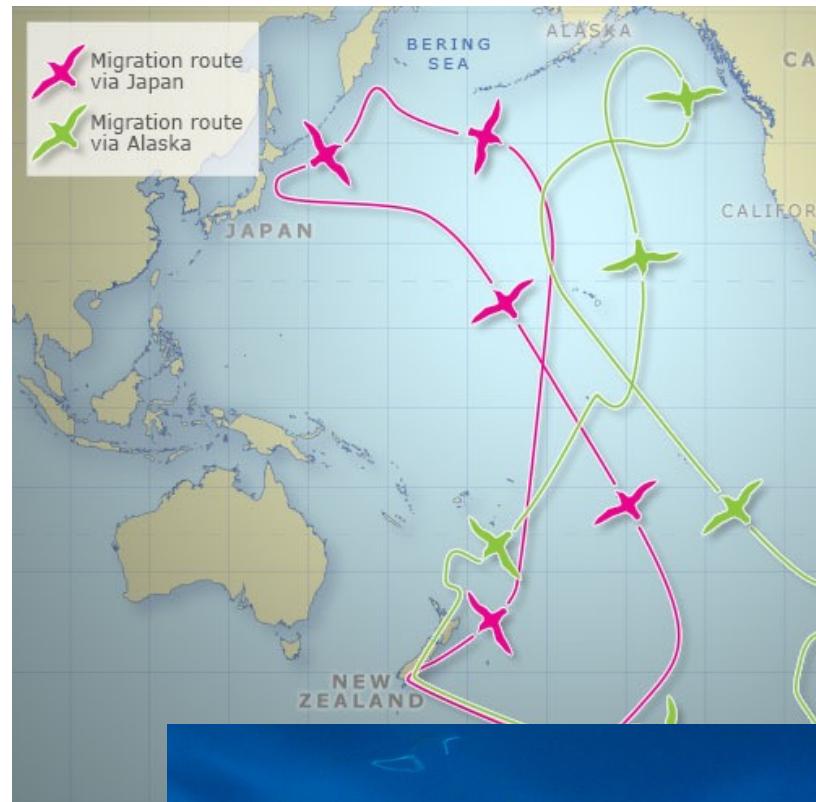
Adaptations to Oceanic Life

- Locomotion
- Gravity effects
- Alternative sensory systems
- O₂ capture
- Diving physiology
- Heat conservation



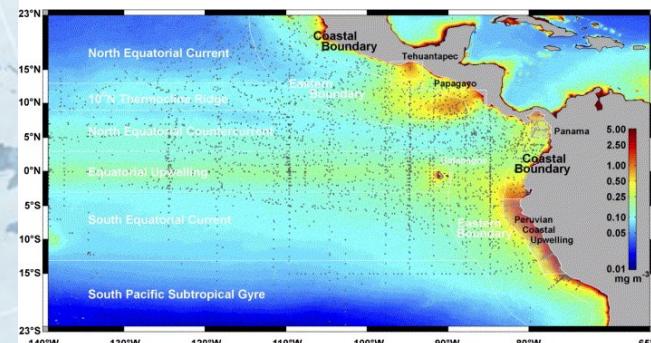
Adaptations to Oceanic Life

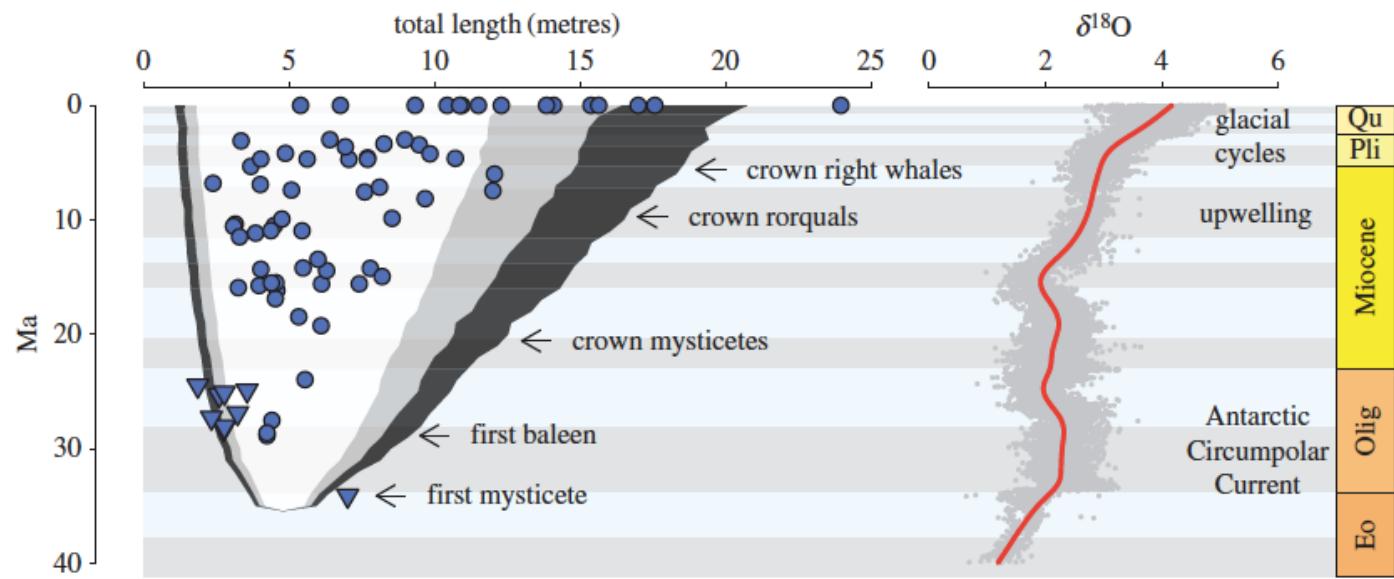
- Locomotion
- Gravity effects
- Alternative sensory systems
- O₂ capture
- Diving physiology
- Heat conservation
- Migrations & Ecology



SOUTH FOR THE WINTER

North Pacific blue whales were known to winter at the Costa Rica Dome, a nutrient-rich upwelling that supports abundant krill. But its location shifts, so researchers aboard R.V. *Pacific Storm* used satellite-tagged whales to lead the way. Three of 15 whales tagged off Santa Barbara made the migration in early 2008. Others wintered off Mexico. Further tagging at the dome could confirm a hunch that blue whales from as far south as Antarctica gather at this upwelling too.





Oceanography & top predators

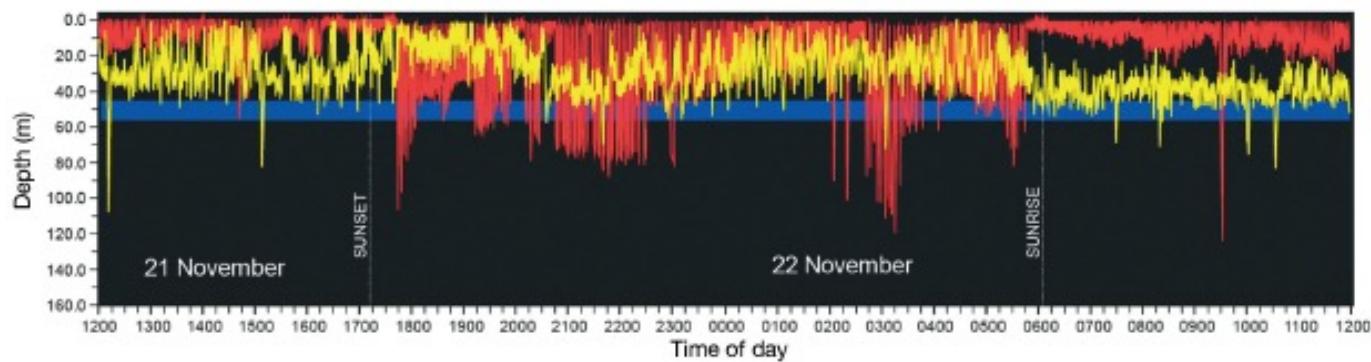
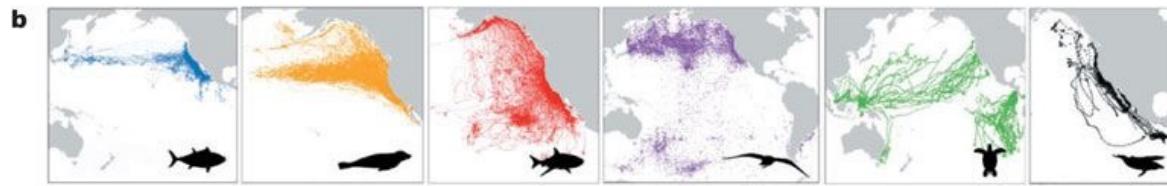
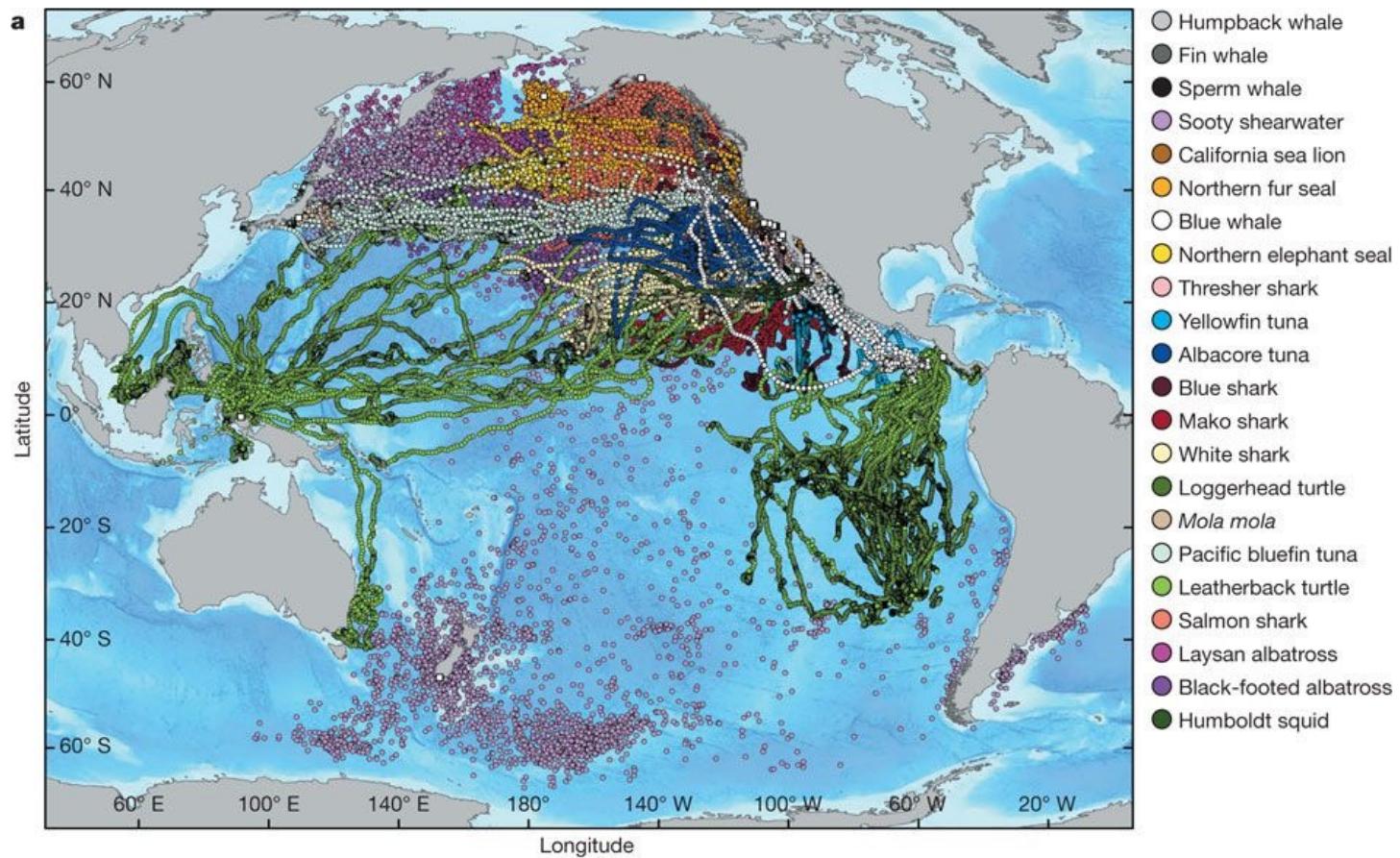
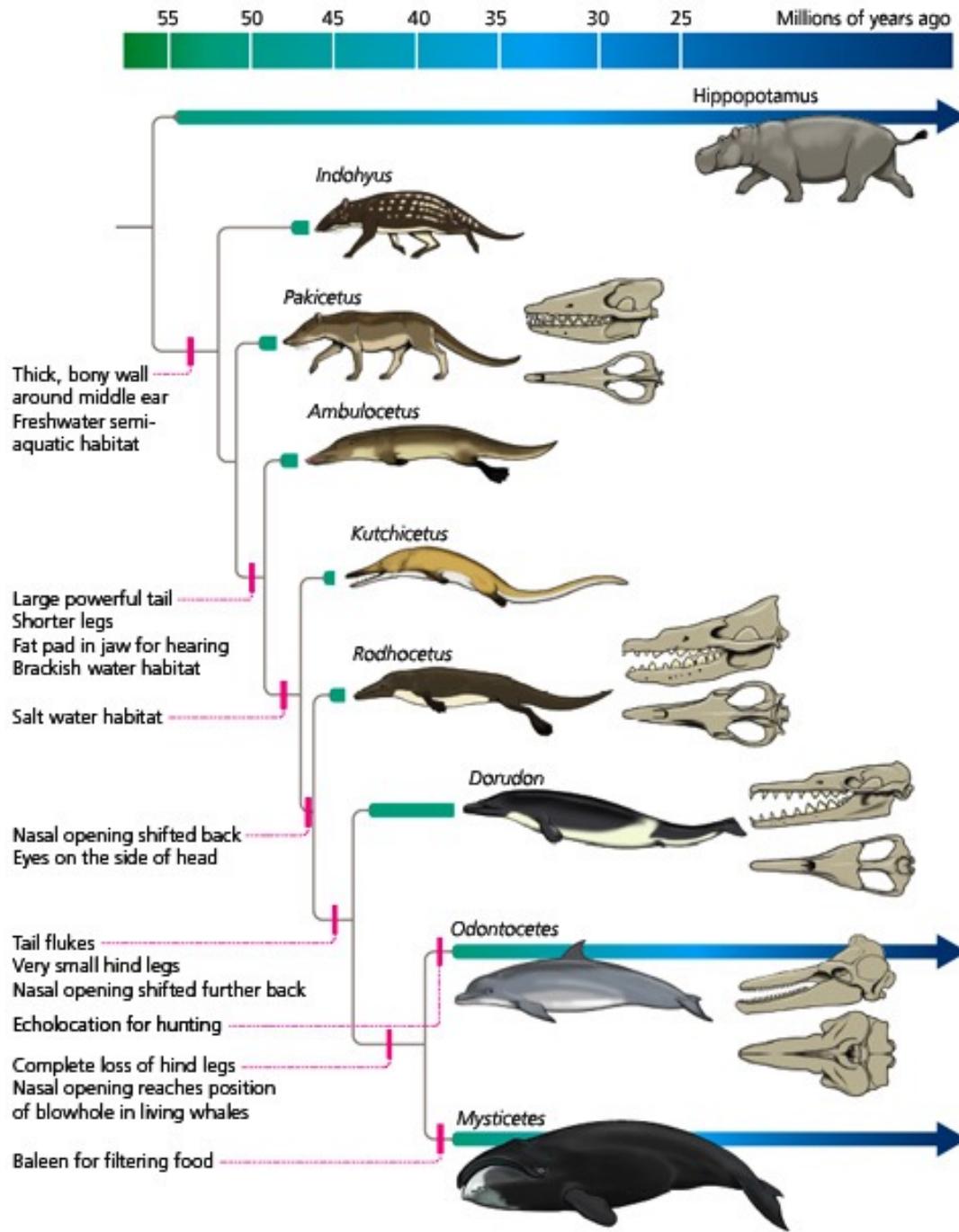
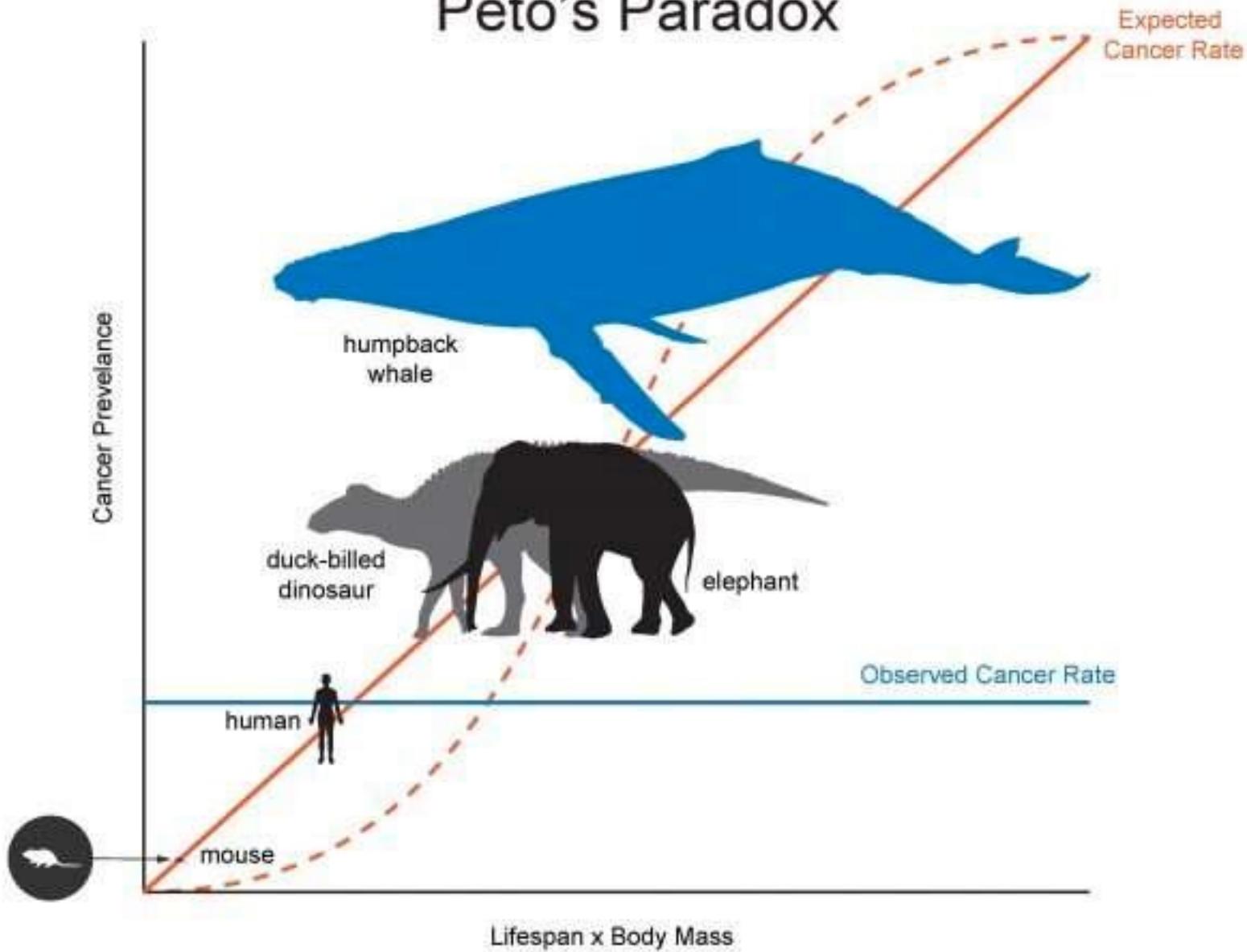


Fig. 2. *Stenella attenuata* and *Thunnus albacares*. Sample of vertical movements of Tuna T1 (yellow) and Dolphin D8 (orange) simultaneously tracked during 21 to 22 November 1993. The depth of the thermocline is represented as a blue band





Peto's Paradox



"SALT"



She returned to New England coast every year since 1975
13 calves, 14 grandcalves, 2 great-grandcalves
1st whale with tracking name instead of #

TP53

- Tumor suppressor
- Mutations found in up to half of human cancers
- Cell cycle arrest, DNA repair, apoptosis
- Washing machine analogy:
 - Start the machine, notice water leaking out
 - Stop the cycle!
 - Call the repair agent
 - If necessary, throw away the washer

Return to the Sea, Get Huge, Beat Cancer: An Analysis of Cetacean Genomes Including an Assembly for the Humpback Whale (*Megaptera novaeangliae*)

Marc Tollis,^{*,1,2,3} Jooke Robbins,⁴ Andrew E. Webb,⁵ Lukas F.K. Kuderna,⁶ Aleah F. Caulin,⁷ Jacinda D. Garcia,² Martine Bérubé,^{4,8} Nader Pourmand,⁹ Tomas Marques-Bonet,^{6,10,11,12} Mary J. O'Connell,¹³ Per J. Palsbøll,^{†,4,8} and Carlo C. Maley^{†,1,2}

Molecular Biology and Evolution, 2019

- They sequenced Salt
- Depth analysis indicated possibility of multiple TP53 genes.

Depth analysis

- When you sequence a genome, you:
 - Amplify all the DNA.
 - Fragment amplified DNA at random places → reads.
 - Determine sequence of each reads.
 - Hire a bioinformatician.
- Sequencing depth = average number of times a nucleotide in the genome is represented in a read
- There's always fluctuation, but depth anywhere is roughly within 20% of the average.

If you know the genome, it's easy to analyze depth

- Align each read to the genome, count how many time a position in the genome is hit.
- But we don't know the genome! That's why Salt was sequenced.
- We can get good information by using nearest relative with known annotated genome as a “reference genome”

Reference genome: fictitious example

- The Green Martian genome has been fully sequenced and annotated.



- The Blue Martian Genome Project has a partial assembly. Before they annotate (find and determine function of each gene), they hire you to analyze their reads.



- You decide that Green Martian is a close enough relative to use as a reference genome. You blast all your reads against the Green Martian genome

Reference genome: fictitious example

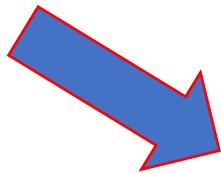
- Average read depth over the entire reference genome is 50.
- For the references genome's COI sequence, average read depth is 48.
 - Yes, Martians have COI.
 - 48 is within 4% of the overall average → nothing unusual here.
- For the references genome's Photosystem II sequences, average read depth is .05.
 - Yes, Green Martians photosynthesize. That's why they're green.
 - Conclude that Blue Martians don't have Photosystem II. Mabe they are related to UCYN-A?
- For the references genome's TP53 sequences, average read depth is 96.
 - Nearly 2x average depth.
 - Hard to explain if Blue Martians only have 1 TP53 gene.
 - Expected depth if Blue Martians have 2 TP53 genes.

They say that whales have multiple copies of TP53. What about Great-grandmother SALT?

National Library of Medicine
National Center for Biotechnology Information

Search NCBI

Megaptera novaeangliae TP53 x Search



- No nucleotide
- No gene
- No genome
- Just an assembly

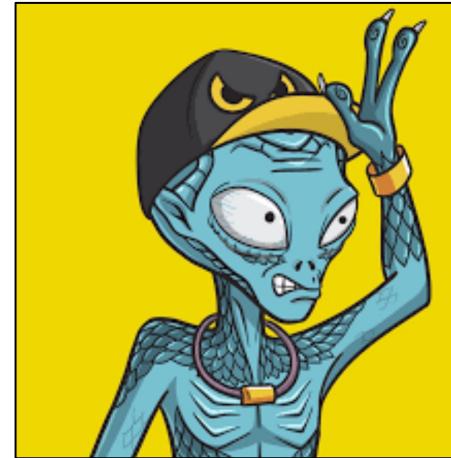
Literature	
Bookshelf	0
	0
	0
	0
	0
PubMed Central	31
Genes	
Gene	0
GEO DataSets	0
GEO Profiles	0
HomoloGene	0
PopSet	0
Genomes	
Assembly	1
BioCollections	0
BioProject	0
BioSample	0
Genome	0
Nucleotide	0
SRA	0
Taxonomy	0
Clinical	
ClinicalTrials.gov	0
ClinVar	0
dbGaP	0
dbSNP	0
dbVar	0
GTR	0
MedGen	0
OMIM	0

For the Salt assembly, reference genome was Blue Whale

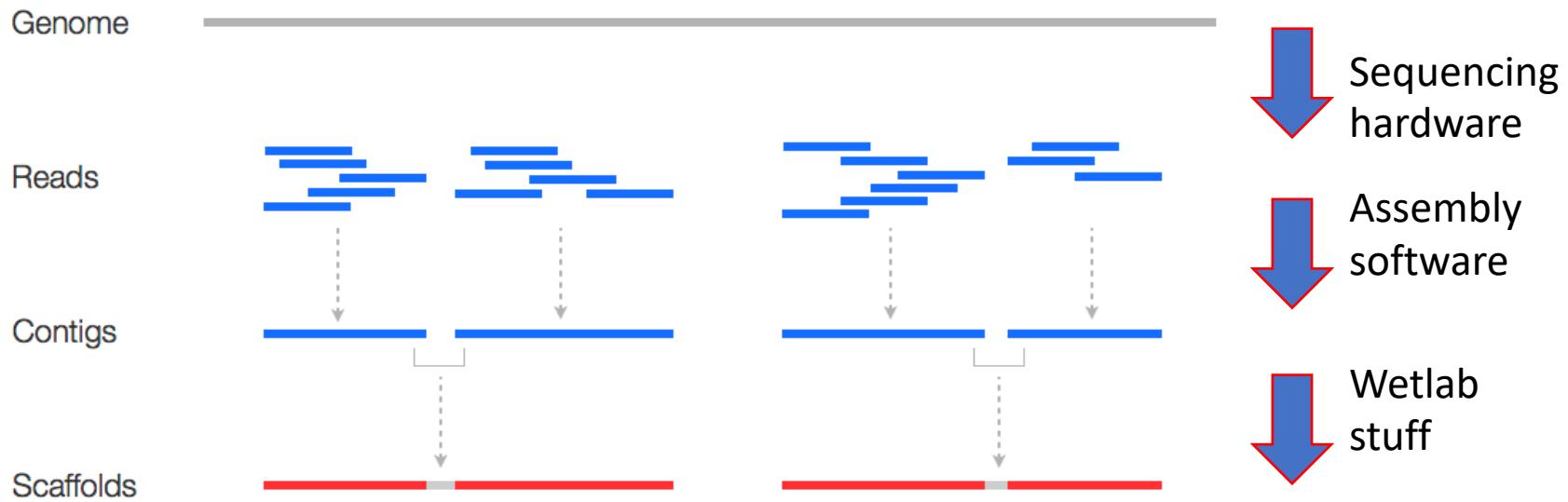
Reference



New assembly



Defining contigs & scaffolds



Collections of multiple scaffolds per chromosome are submitted to GenBank as Assemblies

The Salt assembly

- It isn't certain that this is a complete genome (all nucleotides of all chromosomes).
- “Assembly” is everything that's been sequenced, without annotation. It's a step below “Genome”.
- A genome would have 1 sequence per chromosome.
- An assembly has 1 sequence per scaffold (or per contig, if scaffolds weren't built).

So I clicked on the “Assembly” button

Literature	
Bookshelf	0
MeSH	0
NLM Catalog	0
PubMed	0
PubMed Central	31

Genes	
Gene	0
GEO DataSets	0
GEO Profiles	0
HomoloGene	0
PopSet	0

Genomes	
Assembly	1
BioCollections	0
BioProject	0
BioSample	0
Genome	0
Nucleotide	0
SRA	0
Taxonomy	0

Clinical	
ClinicalTrials.gov	0
ClinVar	0
dbGaP	0
dbSNP	0
dbVar	0
GTR	0
MedGen	0
OMIM	0

... and got to here:

Assembly Assembly ▾ | Advanced Browse by organism Search Help

⚠ In June 2023, Assembly record pages will be redirected to the new [Datasets Genome page](#). [Learn more](#)

Full Report ▾ Send to: ▾

megNov1

Organism name: [Megaptera novaeangliae \(humpback whale\)](#)

Infraspecific name: Breed: North Atlantic

Isolate: ASU/GRON-1

Sex: female

BioSample: [SAMN10585801](#)

BioProject: [PRJNA509641](#)

Submitter: Arizona Cancer Evolution Center (ACE)

Date: 2019/03/06

Assembly level: Scaffold

Genome representation: full

RefSeq category: representative genome

GenBank assembly accession: GCA_004329385.1 (latest)

RefSeq assembly accession: n/a

RefSeq assembly and GenBank assembly identical: n/a

WGS Project: [RYZJ01](#)

Assembly method: Meraculous + HiRise v. Feb-2016

Expected final version: yes

Genome coverage: 102.0x

Sequencing technology: Illumina HiSeq

IDs: 2336551 [UID] 8711508 [GenBank]

See [Genome](#) Information for **Megaptera novaeangliae**

Then I clicked here

Download Assembly

Access the data

BLAST the assembly

Run Primer-BLAST

Full sequence report

Statistics report

FTP directory for GenBank assembly

NCBI Datasets **NEW**

Assembly Information

Assembly Help

Assembly Basics

NCBI Assembly Data Model

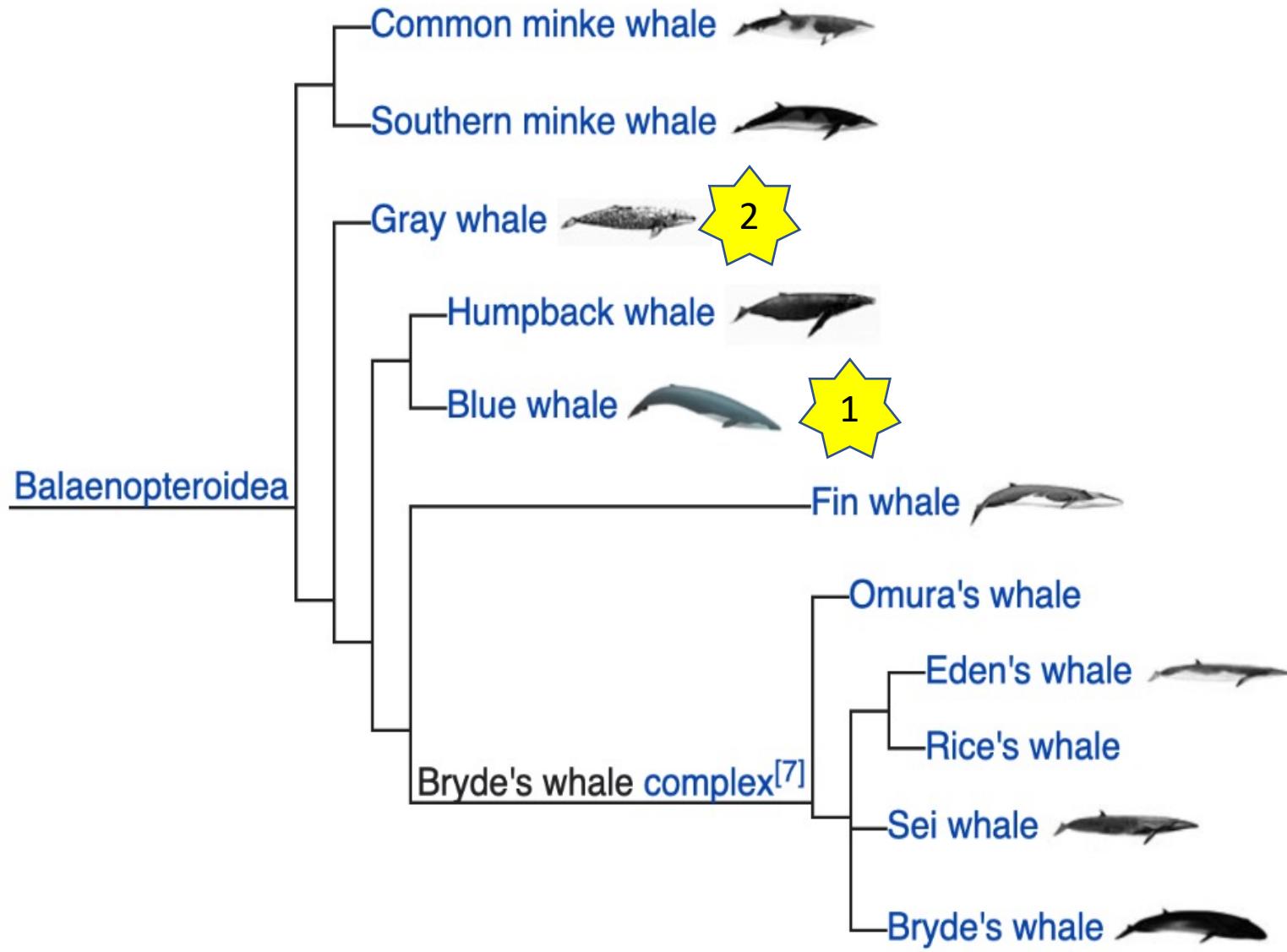
Related Information

BioProject

BioSample

Which downloaded GCA_004329385.1_megNov1_genomic.fna

- “.fna” means nucleotide fasta
- 28 million lines
- 2558 records ... apparently the scaffolds
- ~ 2 billion nucleotides
- How do I count the TP53 sequences?
- If I had a Humpback TP53 sequence, that would be a good start ... but I just showed that I don’t have one.
- Aha! I’ll get TP53 from a close relative of Humpback, and do a local alignment against the Humpback assembly.



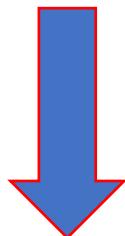


Nucleotide

Nucleotide

blue whale

[Create alert](#) [Advanced](#)



GENE

Was this helpful?



TP53 – tumor protein p53

Balaenoptera musculus (Blue whale)

Gene ID: 118886638

[RefSeq transcripts \(1\)](#) [RefSeq proteins \(1\)](#)

Yay!

[Orthologs](#)

[Genome Data Viewer](#)

[BLAST](#)

[Download](#)

Careful!

Look at the lengths!

- [**Balaenoptera musculus** isolate JJ_BM4_2016_0621 chromosome 4, mBalMus1.pri.v3, whole genome shotgun sequence](#)
144,968,589 bp linear DNA
Accession: NC_045788.1 GI: 1919258342
[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [**Balaenoptera musculus** isolate JJ_BM4_2016_0621 chromosome 6, mBalMus1.pri.v3, whole genome shotgun sequence](#)
116,510,015 bp linear DNA
Accession: NC_045790.1 GI: 1919258340
[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [**Balaenoptera musculus** isolate JJ_BM4_2016_0621 chromosome 7, mBalMus1.pri.v3, whole genome shotgun sequence](#)
113,414,938 bp linear DNA
Accession: NC_045791.1 GI: 1919258339
[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [**Balaenoptera musculus** isolate JJ_BM4_2016_0621 chromosome 10, mBalMus1.pri.v3, whole genome shotgun sequence](#)
104,744,437 bp linear DNA
Accession: NC_045794.1 GI: 1919258336
[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [**Balaenoptera musculus** isolate JJ_BM4_2016_0621 chromosome 14, mBalMus1.pri.v3, whole genome shotgun sequence](#)
90,457,838 bp linear DNA
Accession: NC_045798.1 GI: 1919258331
[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [**Balaenoptera musculus** isolate JJ_BM4_2016_0621 chromosome 15, mBalMus1.pri.v3, whole genome shotgun sequence](#)
88,470,553 bp linear DNA
Accession: NC_045799.1 GI: 1919258330
[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

Careful!
Look at the lengths!

These are entire
chromosomes that
contain blue whale
TP53

Sort this page by
sequence length and
look for something
reasonable

-
- [Balaenoptera musculus isolate JJ_BM4_2016_0621 chromosome 4, mBalMus1.pri.v3, whole genome shotgun sequence](#)
144,968,589 bp linear DNA
Accession: NC_045798.1 GI: 191958342
[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [Balaenoptera musculus isolate JJ_BM4_2016_0621 chromosome 6, mBalMus1.pri.v3, whole genome shotgun sequence](#)
116,510,015 bp linear DNA
Accession: NC_045790.1 GI: 191958340
[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [Balaenoptera musculus isolate JJ_BM4_2016_0621 chromosome 7, mBalMus1.pri.v3, whole genome shotgun sequence](#)
113,414,938 bp linear DNA
Accession: NC_045791.1 GI: 191958339
[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [Balaenoptera musculus isolate JJ_BM4_2016_0621 chromosome 10, mBalMus1.pri.v3, whole genome shotgun sequence](#)
104,744,437 bp linear DNA
Accession: NC_045794.1 GI: 191958336
[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [Balaenoptera musculus isolate JJ_BM4_2016_0621 chromosome 14, mBalMus1.pri.v3, whole genome shotgun sequence](#)
90,457,838 bp linear DNA
Accession: NC_045798.1 GI: 191958331
[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [Balaenoptera musculus isolate JJ_BM4_2016_0621 chromosome 15, mBalMus1.pri.v3, whole genome shotgun sequence](#)
88,470,553 bp linear DNA
Accession: NC_045799.1 GI: 191958330
[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

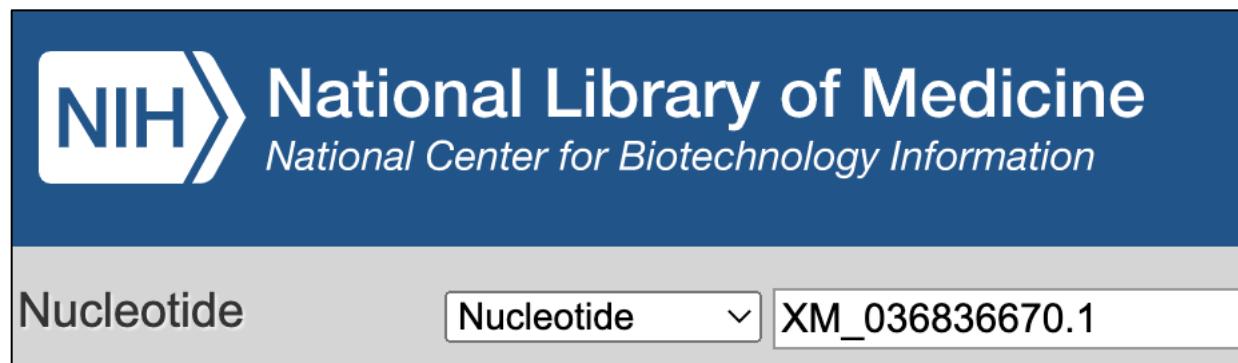
Further down on the same page...

Yes! “Predicted” means by software, not verified in wetlab.

No! “TP53” is in the name, but you have to read the entire annotation.

- [PREDICTED: **Balaenoptera musculus** tumor protein p53 \(**TP53**\), mRNA](#)
12. 2,210 bp linear mRNA
Accession: XM_036836670.1 GI: 1920199934
[BioProject](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [PREDICTED: **Balaenoptera musculus** WD repeat containing antisense to **TP53** \(**WRAP53**\), transcript variant X1, mRNA](#)
13. transcript variant X1, mRNA
2,177 bp linear mRNA
Accession: XM_036836660.1 GI: 1920199916
[BioProject](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [PREDICTED: **Balaenoptera musculus** WD repeat containing antisense to **TP53** \(**WRAP53**\), transcript variant X2, mRNA](#)
14. transcript variant X2, mRNA
1,942 bp linear mRNA
Accession: XM_036836661.1 GI: 1920199918
[BioProject](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [PREDICTED: **Balaenoptera musculus** WD repeat containing antisense to **TP53** \(**WRAP53**\), transcript variant X3, mRNA](#)
15. transcript variant X3, mRNA
1,856 bp linear mRNA
Accession: XM_036836662.1 GI: 1920199920
[BioProject](#) [Protein](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

Ok, let's retrieve that one...



Looks normal

GenBank ▾ Se

PREDICTED: Balaenoptera musculus tumor protein p53 (TP53), mRNA

NCBI Reference Sequence: XM_036836670.1

[FASTA](#) [Graphics](#)

Go to:

LOCUS XM_036836670 2210 bp mRNA linear MAM 28-OCT-2020

DEFINITION PREDICTED: Balaenoptera musculus tumor protein p53 (TP53), mRNA.

ACCESSION XM_036836670

VERSION XM_036836670.1

DBLINK BioProject: [PRJNA607322](#)

KEYWORDS RefSeq.

SOURCE Balaenoptera musculus (Blue whale)

ORGANISM [Balaenoptera musculus](#)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Artiodactyla; Whippomorpha;
Cetacea; Mysticeti; Balaenopteridae; Balaenoptera.

COMMENT MODEL [REFSEQ](#): This record is predicted by automated computational analysis. This record is derived from a genomic sequence ([NC_045804.1](#)) annotated using gene prediction method: Gnomon, supported by mRNA evidence.

So I made a blastable database from the humpback assembly...

Activity	Motrin Score
Download and install the blast executables.	 A grey rectangular box containing two orange oval-shaped Motrin IB tablets.
Convert the fasta to a blastable database:	Zero

```
makeblastdb -in  
GCA_004329385.1_megNov1_genomic.fna  
-out salt_the_humpback -dbtype nucl
```

Then I blasted Blue Whale TP53 against Humpback assembly

Activity	Motrin Score
Blast the blue whale TP53 sequence against the humpback assembly:	Zero
<pre>blastn -query BlueWhaleTP53.fa -db salt_the_humpback -out blue.html -html</pre>	Zero

Database: GCA_004329385.1_megNov1_genomic.fna
2,558 sequences; 2,265,788,366 total letters

Query= XM_036836670.1 PREDICTED: Balaenoptera musculus tumor protein p53
(TP53), mRNA

Length=2210

Sequences producing significant alignments:	Score (Bits)	E Value
RYZJ01000410.1 Megaptera novaeangliae breed North Atlantic isolat...	1725	0.0
RYZJ01000218.1 Megaptera novaeangliae breed North Atlantic isolat...	104	2e-19

(more ...)

- 2 scaffolds in the humpback assembly were hit at least once by the query, with E-value that blast considers “significant”.

The first hit (E-value = 0)

Query length was 2243.

This match starts at query location 1271.

Not a problem, just a thing.

RYZJ01000410.1 Megaptera novaeangliae breed North Atlantic isolat... [1725](#) 0.0
RYZJ01000218.1 Megaptera novaeangliae breed North Atlantic isolat... [104](#) 2e-19

> RYZJ01000410.1 Megaptera novaeangliae breed North Atlantic isolate ASU/GRON-1 ScjyzU5_270, whole genome shotgun sequence
Length=2192838

Score = 1725 bits (934), Expect = 0.0
Identities = 938/940 (99%), Gaps = 0/940 (0%)
Strand=Plus/Plus

Query	1271	AGCCACCTGAAGTCTAAGAAGGGCAGTCTCCCTCCGCCATAAAAACTGATGTTCAAG	1330
Sbjct	2008127	AGCCACCTGAAGTCTAAGAAGGGCAGTCTCCCTCCGCCATAAAAACTGATGTTCAAG	2008186
Query	1331	AGAGAAGGGCCTGACTCAGACTGACATCTCCTGCTTCCTGTCCTGTGACACCCCTAC	1390
Sbjct	2008187	AGAGAAGGGCCTGACTCAGACTGACATCTCCTGCTTCCTGTCCTGTGACACCCCTAC	2008246
Query	1391	CCTCGTTCCCCCTTCCCTGCCATTGGATCTCAGGTGCTTAAACCTCTGCTTGGTG	1450
Sbjct	2008247	CCTCGTTCCCCCTTCCCTGCCATTGGATCTCAGGTGCTTAAACCTCTGCTTGGTG	2008306
Query	1451	CAGGTGTGCCCTCAGAACACCCCTGGAATTCTTCCGTTGCTGGCCTGGGCTCTGCTAA	1510
Sbjct	2008307	CAGGTGTGCCCTCAGAACACCCCTGGAATTCTTCCGTTGCTGGCCTGGGCTCTGCTAA	2008366
Query	1511	AGAGGCTGGCCTGCACTCGTGGTTTTGGGGAGGTAGCTGGGCCTTCCAGCTTAGCTT	1570
Sbjct	2008367	AGAGGCTGGCCTGCACTCGTGGTTTTGGGGAGGTAGCTGGGCCTTCCAGCTTAGCTT	2008426
Query	1571	TTAAGGTTTTACTGTGGGAGAATTGAGAGAGGTAGGAAACGTTCTGCATGTGAGG	1630
Sbjct	2008427	TTAAGGTTTTACTGTGGGAGAATTGAGAGAGGTAGGAAACGTTCTGCATGTGAGG	2008486
Query	1631	AACATTTATAACCAGCCATACACTGGTTGGAAGCCCAGTTCTCCACCATAGCCAG	1690
Sbjct	2008487	AACATTTATAACCAGCCATACACTGGTTGGAAGCCCAGTTCTCCACCATAGCCAG	2008546

(more ...)

More hits to the same scaffold

- Much shorter.
 - Near the 1st hit.
 - These are at 2,004,963 and 2,005,664.
 - Top hit was at 2,008,127.
 - Would be strange in a completed genome.
 - But Humpback assembly is a work in progress, possibly with errors that will be corrected in the future.

Score = 457 bits (247), Expect = 1e-125
 Identities = 255/259 (98%), Gaps = 0/259 (0%)
 Strand=Plus/Plus

Query	286	GTCCTCTGAGCTCTCCCCGGCCGTGGATGACCTGCTGCTGTCCCCAGAAGACGTCGCAAA	345
Sbjct	2004963	GTCCTCTGAGCTCTCCCCGGCCGTGGATGACCTGCTGCTGTCCCCAGAAGACGTCGCAAA	2005022
Query	346	CTGGCTGGATGAACGTCCAGATGAagccccccagatgcctgacccactgcgcaggctgc	405
Sbjct	2005023	CTGGCTGGATGAACGTCCAGATGAAGCCCCCAGATGCCGGAGCCTACTGCGCCAGCCGC	2005082
Query	406	ccccccccagccgccccagcaccagccaCCTCCTGGCCCCGTGCGTCCCTTGTCCTTC	465
Sbjct	2005083	CCCCGCCCCAGCCGCCAGCACGCCACCTCTGGCCCCGTGCGTCCCTCGTCCCTTC	2005142
Query	466	CCAGAAGACCTACCCCTGGCAGCTATGGGTTCCGTCTAGGTTCCCTACATTCTGGAACTGC	525
Sbjct	2005143	CCAGAAGACCTACCCCTGGCAGCTATGGGTTCCGTCTAGATTCCCTACATTCTGGAACTGC	2005202
Query	526	CAAGTCTGTAACCTGCACG 544	
Sbjct	2005203	CAAGTCTGTAACCTGCACG 2005221	
Score = 346 bits (187), Expect = 3e-92 Identities = 194/197 (98%), Gaps = 1/197 (1%) Strand=Plus/Plus			
Query	537	CCTGCACGTATTCCCCCTGCCCTCAACAAGCTGTTTGCCAGCTGCCAAGACCTGCCGG	596
Sbjct	2005664	CCTCCA-GTATTCCCCCTGCCCTCAACAAGCTGTTTGCCAGCTGCCAAGACCTGCCGG	2005722
Query	597	TGCAGCTGTGGGTCAAGCTCACCAACCCCGCCTGGCACCCGGTCCGCGCCATGCCATCT	656
Sbjct	2005723	TGCAGCTGTGGGTCAAGCTCACCAACCCCGCCGGCACCCGGTCCGCGCCATGCCATCT	2005782
Query	657	ACAAGAAAGTCAGAGTACATGACGGAGGTTGTGAGGCCTGTCCCCACGAGCGCTGCT	716
Sbjct	2005783	ACAAGAAAGTCAGAGTACATGACGGAGGTTGTGAGGCCTGTCCCCACGAGCGCTGCT	2005842
Query	717	CTGACTATAGCGACGGT 733	
Sbjct	2005843	CTGACTATAGCGACGGT 2005859	

What we've seen

- Salt was the first Humpback Whale to be sequenced.
- Depth analysis, using Blue Whale genome as a reference, suggested that the Humpback has (at least) 2 copies of TP53.
- Thank you, Salt.

