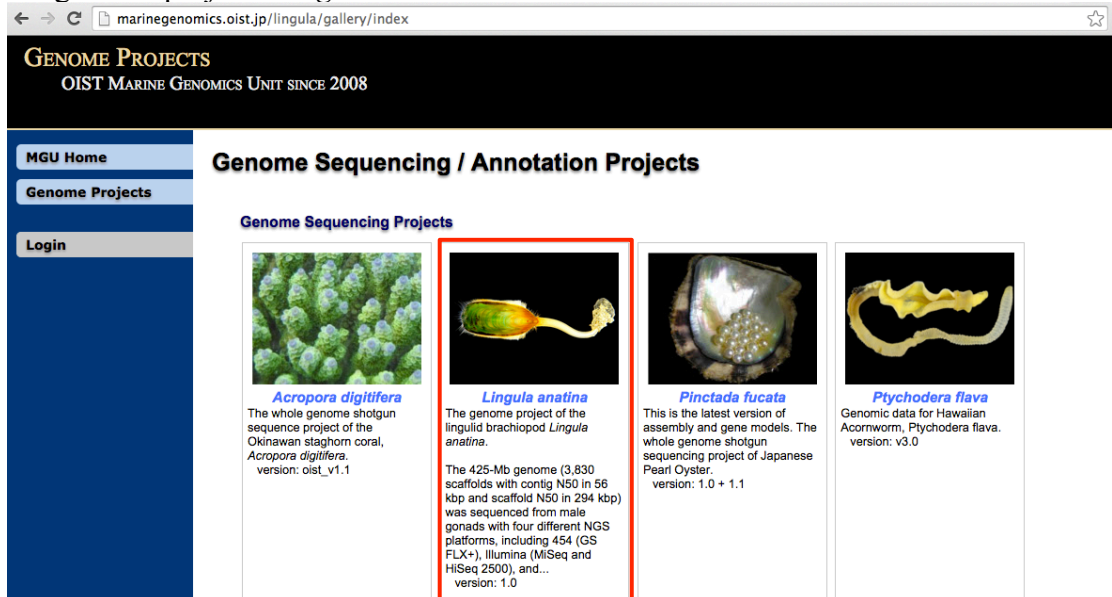


# Tutorial for checking expression profiles of *Lingula* transcripts

Yi-Jyun Luo and Kanako Hisata | 2015.11.25

1. Connect to the genome browser (<http://marinegenomics.oist.jp/>). And click on the genome project “*Lingula anatina*”.

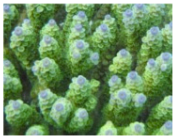

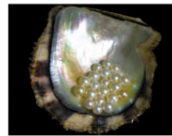
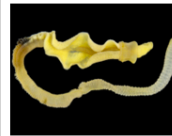


GENOME PROJECTS  
OIST MARINE GENOMICS UNIT SINCE 2008

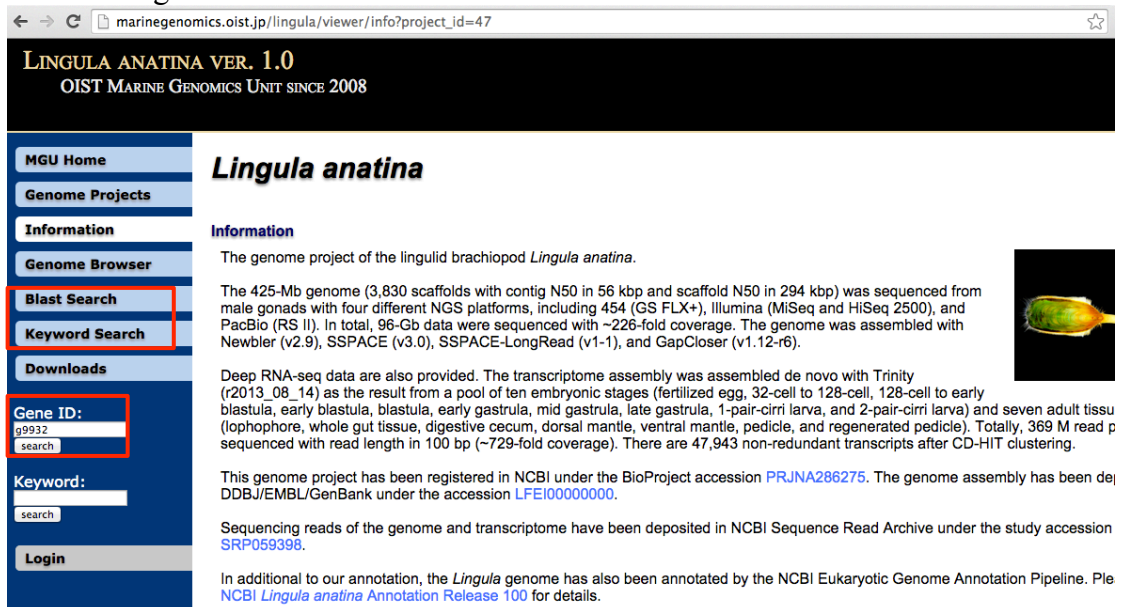
MGU Home  
Genome Projects  
Login

## Genome Sequencing / Annotation Projects

### Genome Sequencing Projects

 <p><b><i>Acropora digitifera</i></b> The whole genome shotgun sequence project of the Okinawan staghorn coral, <i>Acropora digitifera</i>. version: oist_v1.1</p>	 <p><b><i>Lingula anatina</i></b> The genome project of the lingulid brachiopod <i>Lingula anatina</i>. The 425-Mb genome (3,830 scaffolds with contig N50 in 56 kbp and scaffold N50 in 294 kbp) was sequenced from male gonads with four different NGS platforms, including 454 (GS FLX+), Illumina (MiSeq and HiSeq 2500), and... version: 1.0</p>	 <p><b><i>Pinctada fucata</i></b> This is the latest version of assembly and gene models. The whole genome shotgun sequencing project of Japanese Pearl Oyster. version: 1.0 + 1.1</p>	 <p><b><i>Ptychodera flava</i></b> Genomic data for Hawaiian Acornworm, <i>Ptychodera flava</i>. version: v3.0</p>
---	--	--	---

2. In the main page, use “Blast Search” or “Keyword Search” to find out genes of your interests. Or you can directly search the “Gene ID” if you already know one. We will search for g9932 in this example, which is a homolog of vertebrate BMP2 and BMP4 genes.



LINGULA ANATINA VER. 1.0  
OIST MARINE GENOMICS UNIT SINCE 2008

MGU Home  
Genome Projects  
Information  
Genome Browser  
Blast Search  
Keyword Search  
Downloads  
Gene ID:  
g9932  
search  
Keyword:  
search  
Login

## *Lingula anatina*

### Information

The genome project of the lingulid brachiopod *Lingula anatina*.

The 425-Mb genome (3,830 scaffolds with contig N50 in 56 kbp and scaffold N50 in 294 kbp) was sequenced from male gonads with four different NGS platforms, including 454 (GS FLX+), Illumina (MiSeq and HiSeq 2500), and PacBio (RS II). In total, 96-Gb data were sequenced with ~226-fold coverage. The genome was assembled with Newbler (v2.9), SSPACE (v3.0), SSPACE-LongRead (v1-1), and GapCloser (v1.12-r6).

Deep RNA-seq data are also provided. The transcriptome assembly was assembled de novo with Trinity (r2013\_08\_14) as the result from a pool of ten embryonic stages (fertilized egg, 32-cell to 128-cell, 128-cell to early blastula, early blastula, blastula, early gastrula, mid gastrula, late gastrula, 1-pair-cirri larva, and 2-pair-cirri larva) and seven adult tissue (lophophore, whole gut tissue, digestive cecum, dorsal mantle, ventral mantle, pedicle, and regenerated pedicle). Totally, 369 M read p sequenced with read length in 100 bp (~729-fold coverage). There are 47,943 non-redundant transcripts after CD-HIT clustering.

This genome project has been registered in NCBI under the BioProject accession [PRJNA286275](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA286275). The genome assembly has been deposited in DDBJ/EMBL/GenBank under the accession [LFEI00000000](https://www.ncbi.nlm.nih.gov/nuclseq/LFEI00000000).

Sequencing reads of the genome and transcriptome have been deposited in NCBI Sequence Read Archive under the study accession [SRP059398](https://www.ncbi.nlm.nih.gov/sra/SRP059398).

In addition to our annotation, the *Lingula* genome has also been annotated by the NCBI Eukaryotic Genome Annotation Pipeline. Please refer to [NCBI \*Lingula anatina\* Annotation Release 100](https://www.ncbi.nlm.nih.gov/genome/100) for details.

3. Then you will see the gene page, including the information such as 1) gene ID and locus, 2) Pfam domain composition, 3) best blast hit to Swiss-Prot, and 4) nucleotide and protein sequences. To check its transcripts, we click on “To GenomeBrowser”.

← → ↻ marinegenomics.oist.jp/lingula/genesearch/genemodel?utf8=√&gene\_modelid=g9932&project\_id=47

**LINGULA ANATINA VER. 1.0**  
OIST MARINE GENOMICS UNIT SINCE 2008

**MGU Home**  
**Genome Projects**  
**Information**  
**Genome Browser**  
**Blast Search**  
**Keyword Search**  
**Downloads**

Gene ID: g9932  
search

Keyword: search

Login

### Gene

Gene Model ID: g9932  
Locus: scaffold177 : 354144 ... 358644 : +  
To GenomeBrowser: scaffold177:354144..358644  
Genes list of scaffold: scaffold177

#### Hammer Search for Pfam

query	Total	ID	Target	From	To	Eval	Score	cEval	iEval	Score
g9932.t1	1	1	TGFb_propeptide	75	307	0.0	184.9	0.0	0.0	184.4
g9932.t1	1	1	TGF_beta	333	436	7.00649e-43	145.1	0.0	1.19951e-42	144.4

#### Blast Hit to nr / sp

query	Subject ID	Subject Name	evalue
g9932.t1	sp Q90752 BMP4_CHICK	Bone morphogenetic protein 4	0.0

4. You will then connect to a JBrowse-based genome browser. If you do not see any annotation, make sure that if the check boxes on the “Available Tracks” column are all checked. Gene models are shown in red, and transcripts from transcriptome are in blue. To check the expression level of the transcript, we click on the transcript ”comp125124\_c3\_seq14”. [NOTE: For the transcriptome assembly from older version of Trinity (such as r2013\_08\_14), each “transcript” is shown as compXXXXXX\_cXX, while seqXX means different isoforms.]

← → ↻ marinegenomics.oist.jp/lingula/viewer?project\_id=47&qname=scaffold177%3A354144..358644

**LINGULA ANATINA VER. 1.0**  
OIST MARINE GENOMICS UNIT SINCE 2008

**MGU Home**  
**Genome Projects**  
**Information**  
**Genome Browser**  
**Blast Search**  
**Keyword Search**  
**Downloads**

Gene ID: search  
Keyword: search

**Available Tracks**

filter by text

- Reference sequence 1
  - Reference sequence
- Gene Models 1
  - AUGUSTUS\_Gene\_Models
- Transcriptome 1
  - Trinity

**JBrowse** File View Help

50,000 100,000 150,000 200,000 250,000 300,000 350,000 400,000 450,000

scaffold177 scaffold177:347701..362780 (15.08 Kb) Go

Reference sequence Zoom in to see sequence Zoom in to see sequence Zoom

AUGUSTUS\_Gene\_Models

g9932.t1  
Bone morphogenetic protein 4

Trinity

comp125124\_c3\_seq18  
Bone morphogenetic protein 4

comp125124\_c3\_seq7

comp137949\_c1\_seq12

comp137949\_c1\_seq2

comp81547\_c0\_seq1

comp125124\_c3\_seq14  
Bone morphogenetic protein 4

5. You will then be redirected to the transcript page, including information of 1) gene ID and locus, 2) best blast hit to Swiss-Prot, and 3) expression profile 4) all isoforms with nucleotide and protein sequences. The expression profile contains expression levels in embryonic stages and adult tissues. Medium expressed stages or tissues (FPKM $\geq$ 10 $\leq$ 50) are labeled with blue, and highly expressed ones (FPKM $>$ 50) are labeled with pink.

FPKM, Fragments Per Kilobase of transcript per Million mapped reads.

**MGU Home**

**Genome Projects**

**Information**

**Genome Browser**

**Blast Search**

**Keyword Search**

**Downloads**

Gene ID:  
comp125124\_c3  
search

Keyword:  
search

Login

### Gene

Gene Model ID	comp125124_c3
Locus	scaffold177 : 342152 ... 354867 : +
To GenomeBrowser	<a href="#">scaffold177:342152..354867</a>
Genes list of scaffold	<a href="#">scaffold177</a>

#### Blast Hit to nr / sp

query	Subject ID	Subject Name	evalue
comp125124_c3_seq18	sp Q90752 BMP4_CHICK	Bone morphogenetic protein 4	0.0

#### Expression profile

(Color code: FPKM>10&<50, blue; FPKM>50, pink)

Libraries	FPKM
<b>&gt;Embryonic stages</b>	
Fertilized eggs	0.87
32-cell	4.42
128-cell	6.71
Early blastula	12.64