

รายงาน  
หัวข้อ NCBI database and BLAST

จัดทำโดย  
นายพุทธิพน สวัสดิ์  
รหัสนักศึกษา 665020084-2

รายงานนี้เป็นส่วนหนึ่งของวิชา Practical Skills in Biology  
สาขาชีววิทยาสำหรับครู  
คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

## *Bmp2* gene (Bone Morphogenetic Protein 2)

### ที่มาและความสำคัญ

ยีน *Bmp2* (Bone Morphogenetic Protein 2) เป็นยีนที่เกี่ยวข้องกับการควบคุมการเจริญเติบโต การเปลี่ยนแปลงของเซลล์ และการพัฒนาเนื้อเยื่อต่าง ๆ โดยเฉพาะการพัฒนาโครงสร้างกระดูกและกระดูกอ่อน (Chen et al., 2004) โดยบทบาทสำคัญของ *Bmp2* คือความสามารถในการกระตุ้นให้เซลล์ต้นกำเนิดชนิด mesenchymal stem cells เปลี่ยนไปเป็นเซลล์สร้างกระดูก (osteoblasts) และกระดูกอ่อน (chondrocytes) ซึ่งเป็นกระบวนการที่สำคัญทั้งในช่วงพัฒนาการของตัวอ่อนและการซ่อมแซมเนื้อเยื่อในผู้ใหญ่ (Chen et al., 2004) ซึ่งในทางคลินิก *Bmp2* ถูกนำมาใช้เป็นโปรตีนสังเคราะห์ (recombinant human BMP2 – rhBMP2) เพื่อรักษาภาวะกระดูกหักที่ไม่สามารถเชื่อมติดเองได้ โดยใช้ร่วมกับวัสดุปลูกกระดูกในผู้ป่วยที่มีบาดแผลกระดูกหักรุนแรงหรือกรณีผ่าตัดกระดูกสันหลัง ซึ่งแสดงผลการรักษาที่ดีขึ้นเมื่อเทียบกับวิธีแบบดั้งเดิม (Govender et al., 2002)

จำนวนลำดับเบสทั้งหมดของยีนนี้ : 3545 bp

จำนวนลำดับกรดอะมิโนของยีนนี้ : 396

## ขั้นตอนการดำเนินงาน

### 1. พิมพ์ชื่อที่ต้องการศึกษา ในการค้นหา All Databases และกด Search

NIH National Library of Medicine  
National Center for Biotechnology Information

All Databases Bmp2 Search

NCBI Home  
Resource List (A-Z)  
All Resources  
Chemicals & Bioassays

Welcome to NCBI  
The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.  
[About the NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News & Blog](#)

Popular Resources  
PubMed  
Bookshelf  
PubMed Central

### 2. เลือกหัวข้อ Gene

Search NCBI Bmp2 Search

Results found in 28 databases

GENE  
**BMP2 - bone morphogenetic protein 2**  
*Homo sapiens (human)*  
Also known as: BDA2, BMP2A, SSFSC, SSFSC1  
Gene ID: 650

RefSeq products Orthologs Genome Data Viewer

New - Visualize gene across multiple species

RefSeq Sequences +

Literature  
Bookshelf 110  
MeSH 5  
NLM Catalog 2  
PubMed 12,820

Genes  
Gene 1,916  
GEO DataSets 2,693  
GEO Profiles 28,683  
PopSet 0

Proteins  
Conserved Domains 9  
Identical Protein Groups 441  
Protein 4,645  
Protein Family Models 1

### 3. เลือกยีนในสิ่งมีชีวิตที่สนใจ ที่ต้องการศึกษา

**Search results**  
Items: 1 to 20 of 1916  
<< First < Prev Page 1 of 96 Next > Last >>  
See also 21 discontinued or replaced items.

Name/Gene ID	Description	Location	Aliases	MIM
<input type="checkbox"/> BMP2 ID: 650	bone morphogenetic protein 2 [ <i>Homo sapiens</i> (human)]	Chromosome 20, NC_000020.11 (6767686..6780246)	BDA2A, SSFSC, SSFSC1, BMP2	112261
<input type="checkbox"/> Bmp2 ID: 12156	bone morphogenetic protein 2 [ <i>Mus musculus</i> (house mouse)]	Chromosome 2, NC_000068.8 (133394079..133404816)	Bmp2a	
<input type="checkbox"/> Bmp2 ID: 29373	bone morphogenetic protein 2 [ <i>Rattus norvegicus</i> (Norway rat)]	Chromosome 3, NC_086021.1 (141264648..141275416)		
<input type="checkbox"/> BMP2 ID: 378779	bone morphogenetic protein 2 [ <i>Gallus gallus</i> (chicken)]	Chromosome 3, NC_052534.1 (14796478..15040250, complement)		
<input type="checkbox"/> BMP2 ID: 615037	bone morphogenetic protein 2 [ <i>Bos taurus</i> (domestic cattle)]	Chromosome 13, NC_037340.1 (49152163..49163417)		
<input type="checkbox"/> BMP2 ID: 100157103	bone morphogenetic protein 2 [ <i>Sus scrofa</i> (pig)]	Chromosome 17, NC_010459.5 (15750487..15762982)		

4. เมื่อได้ยีนที่สนใจแล้ว กดเลือกหัวข้อ NCBI Reference Sequences (RefSeq) บริเวณแถบทางขวามือ

NIH National Library of Medicine  
National Center for Biotechnology Information

Gene   Advanced Help

Full Report

**BMP2 bone morphogenetic protein 2 [ *Homo sapiens* (human) ]**

Gene ID: 650, updated on 3-May-2025

**Summary**

**Official Symbol** BMP2 provided by HGNC  
**Official Full Name** bone morphogenetic protein 2 provided by HGNC  
**Primary source** HGNC:HGNC:1069  
**See related** [Ensembl:ENSG00000125845](#) [MIM:112281](#) [AllianceGenome:HGNC:1069](#)  
**Gene type** protein coding  
**RefSeq status** REVIEWED  
**Organism** [Homo sapiens](#)  
**Lineage** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorhini; Catarrhini; Hominidae; Homo  
**Also known as** BDA2; BMP2A; SSFSC; SSFSC1  
**Summary** This gene encodes a secreted ligand of the TGF-beta (transforming growth factor-beta) superfamily of proteins. Ligands of this family bind various TGF-beta receptors leading to recruitment and activation of SMAD family transcription factors that regulate gene expression. The encoded preproprotein is proteolytically processed to generate each subunit of the disulfide-linked homodimer, which plays a role in bone and cartilage development. Duplication of a regulatory region downstream of this gene causes a form of brachydactyly characterized by a malformed index finger and second toe in human patients. [provided by RefSeq, Jul 2016]  
**Expression** Broad expression in fat (RPKM 13.9), colon (RPKM 12.1) and 21 other tissues [See more](#)  
**Orthologs** [mouse](#) [all](#)  
 [Try the new Gene table](#)  
[Try the new Transcript table](#)

**Table of contents**

- Summary
- Genomic context
- Genomic regions, transcripts, and products
- Expression
- Bibliography
- Phenotypes
- Variation
- HIV-1 interactions
- Pathways from PubChem
- Interactions
- General gene information
  - Markers, Homology, Gene Ontology
- General protein information
- NCBI Reference Sequences (RefSeq)
- Related sequences
- Additional links
  - Locus-specific Databases

5. คลิกที่ NP\_001191.1 เพื่อค้นหาลำดับกรดอะมิโนของยีนที่สนใจในสิ่งมีชีวิตที่สนใจ

NCBI Reference Sequences (RefSeq)  [Try the new Transcript table](#)

☐ RefSeqs maintained independently of Annotated Genomes

These reference sequences exist independently of genome builds: [Explain](#)

**Genomic**

1. **NG\_023233.1 RefSeqGene**

Range 4589..17149  
Download [GenBank](#) [FASTA](#) [Sequence Viewer \(Graphics\)](#)

**mRNA and Protein(s)**

1. **NM\_001200.4 → NP\_001191.1 bone morphogenetic protein 2 preproprotein**

[See identical proteins and their annotated locations for NP\\_001191.1](#)

**Status:** REVIEWED

Source sequence(s) [AA489383](#), [AL035668](#), [BC069214](#), [BM670141](#), [M22489](#)  
Consensus CDS [CCDS13099.1](#)  
UniProtKB/Swiss-Prot [P12643](#)  
UniProtKB/TrEMBL [C8C060](#), [C8C061](#)  
Related [ENSP00000368104.3](#), [ENST00000378827.5](#)

Conserved Domains (2) [summary](#)

<a href="#">smart00204</a>	TGFB; Transforming growth factor-beta (TGF-beta) family
Location:296 → 396	
<a href="#">pfam00688</a>	TGfb_propeptide; TGF-beta propeptide
Location:37 → 267	

6. คลิกที่ FASTA เพื่อดูลำดับกรดอะมิโนของยีนที่สนใจในสิ่งมีชีวิตที่สนใจ

NIH National Library of Medicine  
National Center for Biotechnology Information

Protein   GenPept

**bone morphogenetic protein 2 preproprotein [Homo sapiens]**

NCBI Reference Sequence: NP\_001191.1

[Identical Proteins](#) [FASTA](#) [Graphics](#)

7. copy ลำดับกรดอะมิโนของยีนที่สนใจในสิ่งมีชีวิตที่สนใจ

FASTA ▾

## bone morphogenetic protein 2 preproprotein [Homo sapiens]

NCBI Reference Sequence: NP\_001191.1

[GenPept](#) [Identical Proteins](#) [Graphics](#)

>NP\_001191.1 bone morphogenetic protein 2 preproprotein [Homo sapiens]  
MVAGTRCLLALLLPQVLLGGAAGLVPELGRRKFAAASSGRPSSQPSDEVLSEFELRLLSMFGLKQRTPS  
RDVAVPPYMLDLRRHSGQPGSPAPDHLERAASRANTVRSFHHEESLEELPETSCKTTRRFFNLSSIP  
TEEFITSAELQVFREQMDALGNSSFHHRINIYEIIPATANSKFVTRLLDTRLVNQNASRWESFDVT  
PAVMRWTAQGHANHGFFVEVAHLEEKQGVSKRHVRISRLHQDEHSWSQIRPLLVTFGHDGKGHPLHKRE  
KRQAKHKQRKRLKSSCKRHPLYVDFSDVGWNDWIVAPPGYHAFYCHGECPPFLADHLNSTNHAIVQTLVN  
SVNSKIPKACCVPTLSAISMLYLDENEKVVLKNYQDMVVEGCGCR

8. เข้าสู่ BLAST แล้วเลือก Protein BLAST

NIH National Library of Medicine  
National Center for Biotechnology Information

BLAST® Home Recent Results Saved Strategies Help

**Basic Local Alignment Search Tool**

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

Mon, 17 Mar 2025  
Improvements include upgrading to GCP Artifact Registry and better handling of job completion status in kubernetes version 1.30+.  
ElasticBLAST 1.4.0 is now available! [More BLAST news...](#)

**Web BLAST**

**Nucleotide BLAST**  
nucleotide → nucleotide

**blastx**  
translated nucleotide → protein

**tblastn**  
protein → translated nucleotide

**Protein BLAST**  
protein → protein

**BLAST Genomes**

Enter organism common name, scientific name, or tax id [Search](#)

Human Mouse Rat Microbes

9. วางลำดับกรดอะมิโนที่ได้ทำการ copy ไว้

BLAST® » blastp suite

blastn **blastp** blastx tblastn tblastx

Standard Protein BLAST

BLASTP programs search protein databases using a protein query

**Enter Query Sequence**

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

KRQAKHKQRKRLKSSCKRHPLYVDFSDVGWNDWIVAPPGYHAFYCHGECPPFLADHLNSTNHAIVQTLVNSVNSKIPKACCVPTLSAISMLYLDENEKVVLKNYQDMVVEGCGCR

Query subrange [?](#)

From

To

Or, upload file   [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

10. เลือก Database = refseq\_protein และพิมพ์ชื่อกลุ่มสิ่งมีชีวิตที่สนใจต้องการศึกษา และเลือก Exclude ทั้งหมดยกเว้น Models (XM/XP)

**Choose Search Set**

**Database** ☒ Standard databases (nr etc.): ☐ Experimental databases

Reference proteins (refseq\_protein) ?

**Organism** Optional Mammalia (taxid:40674) ☐ exclude [Add organism](#)

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown ?

**Exclude** Optional ☐ Models (XM/XP) ☒ Non-redundant RefSeq proteins (WP) ☒ Uncultured/environmental sample sequences

11. กำหนด Max target sequences = 50 จากนั้นกด BLAST

**Algorithm parameters**

**General Parameters**

**Max target sequences** 50 ?  
Select the maximum number of aligned sequences to display ?

**Short queries** ☒ Automatically adjust parameters for short input sequences ?

**Expect threshold** 0.05 ?

**Word size** 5 ?

**Max matches in a query range** 0 ?

**Scoring Parameters**

**Matrix** BLOSUM62 ?

**Gap Costs** Existence: 11 Extension: 1 ?

**Compositional adjustments** Conditional compositional score matrix adjustment ?

**Filters and Masking**

**Filter** ☐ Low complexity regions ?

**Mask** ☐ Mask for lookup table only ?  
☐ Mask lower case letters ?

**BLAST** Search database nr using **Blastp (protein-protein BLAST)**  
☐ Show results in a new window

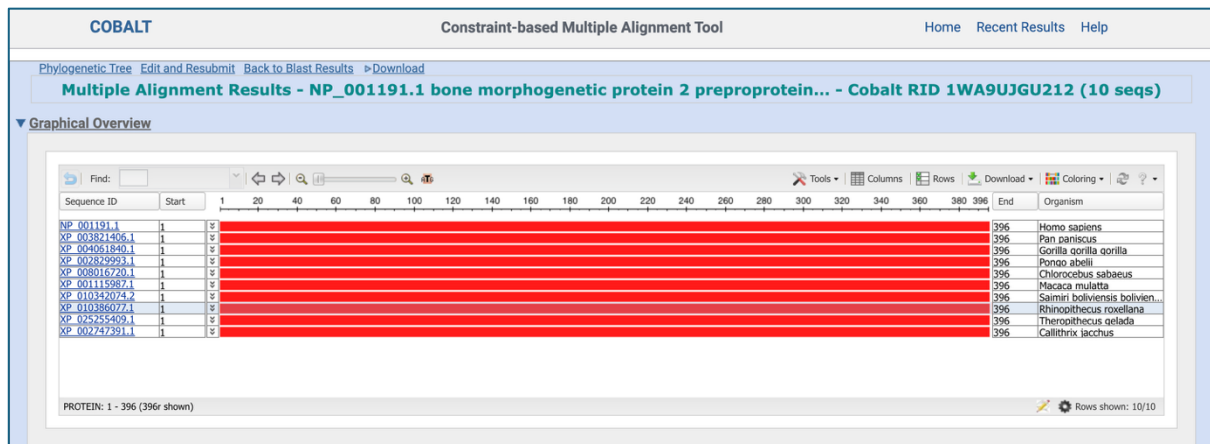
12. เลือกศึกษาในสิ่งมีชีวิตที่สนใจ จากนั้นกด Multiple alignment

Sequences producing significant alignments Download Select columns Show 50 ?

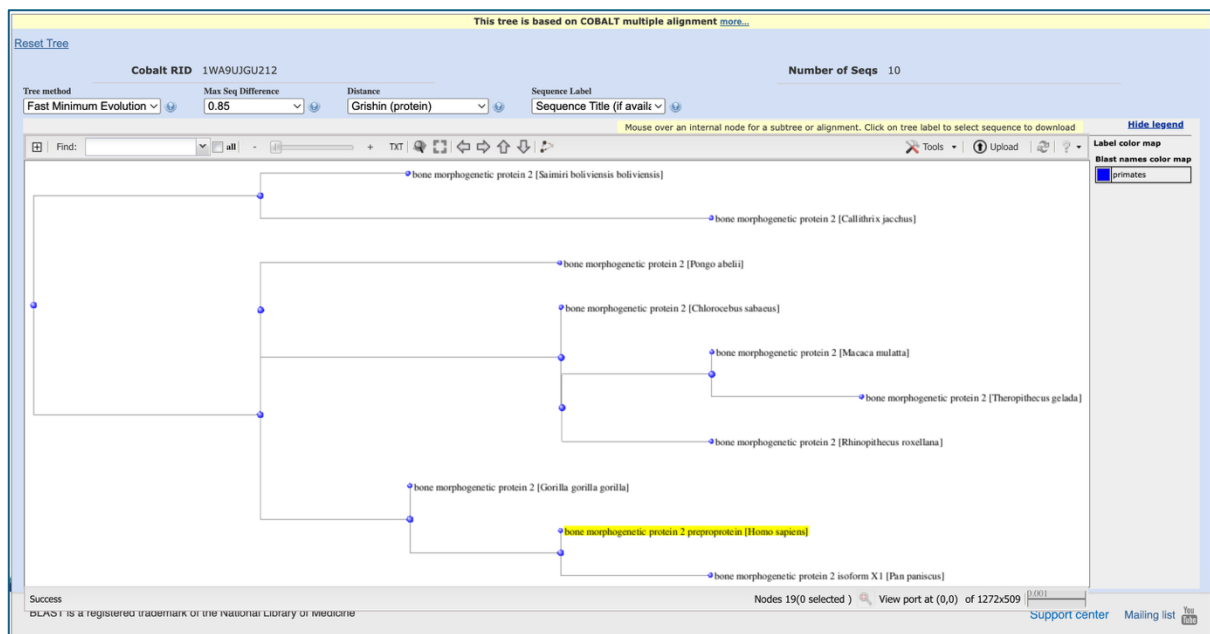
☐ select all 61 sequences selected

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	bone morphogenetic protein 2 (oreoprotein) [Homo sapiens]	Homo sapiens	825	825	100%	0.0	100.00%	396	NP_001191.1
<input checked="" type="checkbox"/>	bone morphogenetic protein 2 (isoform X1) [Pan paniscus]	Pan paniscus	825	825	100%	0.0	99.75%	396	XP_003821406.1
<input checked="" type="checkbox"/>	bone morphogenetic protein 2 (Gorilla gorilla gorilla)	Gorilla gorilla gorilla	824	824	100%	0.0	99.75%	396	XP_004061840.1
<input type="checkbox"/>	bone morphogenetic protein 2 (Nomascus leucogenys)	Nomascus leucogenys	820	820	100%	0.0	99.24%	396	XP_003252536.1
<input type="checkbox"/>	bone morphogenetic protein 2 (Hylobates moloch)	Hylobates moloch	819	819	100%	0.0	99.24%	396	XP_031986602.1
<input checked="" type="checkbox"/>	bone morphogenetic protein 2 (Pongo abelii)	Pongo abelii	819	819	100%	0.0	98.99%	396	XP_002829993.1
<input type="checkbox"/>	bone morphogenetic protein 2 (Cebus imitator)	Cebus imitator	818	818	100%	0.0	98.74%	396	XP_017379463.1
<input checked="" type="checkbox"/>	bone morphogenetic protein 2 (Chlorocebus sabaeus)	Chlorocebus sabaeus	818	818	100%	0.0	98.99%	396	XP_008016720.1
<input checked="" type="checkbox"/>	bone morphogenetic protein 2 (Macaca mulatta)	Macaca mulatta	818	818	100%	0.0	98.74%	396	XP_001115987.1
<input type="checkbox"/>	bone morphogenetic protein 2 (Symphalangus syndactylus)	Symphalangus syndactylus	818	818	100%	0.0	98.99%	396	XP_055122105.1
<input checked="" type="checkbox"/>	bone morphogenetic protein 2 (Saimiri boliviensis boliviensis)	Saimiri boliviensis boliviensis	818	818	100%	0.0	98.48%	396	XP_010342074.2
<input checked="" type="checkbox"/>	bone morphogenetic protein 2 (Rhinopithecus roxellana)	Rhinopithecus roxellana	816	816	100%	0.0	98.74%	396	XP_010386077.1
<input type="checkbox"/>	PREDICTED: bone morphogenetic protein 2 (Mandrillus leucorhæus)	Mandrillus leucorhæus	816	816	100%	0.0	98.48%	396	XP_011853912.1
<input checked="" type="checkbox"/>	bone morphogenetic protein 2 (Theropithecus gelada)	Theropithecus gelada	816	816	100%	0.0	98.48%	396	XP_025255409.1
<input type="checkbox"/>	PREDICTED: bone morphogenetic protein 2 (Cercopithecus aethiops)	Cercopithecus aethiops	815	815	100%	0.0	98.48%	396	XP_011907497.1
<input type="checkbox"/>	bone morphogenetic protein 2 (Pan troglodytes)	Pan troglodytes	815	815	100%	0.0	98.74%	396	XP_014508.2
<input checked="" type="checkbox"/>	bone morphogenetic protein 2 (Callithrix jacchus)	Callithrix jacchus	814	814	100%	0.0	97.98%	396	XP_002747391.1
<input type="checkbox"/>	PREDICTED: bone morphogenetic protein 2 (Colobus angolensis palliatus)	Colobus angolensis palliatus	813	813	100%	0.0	98.23%	396	XP_011796396.1

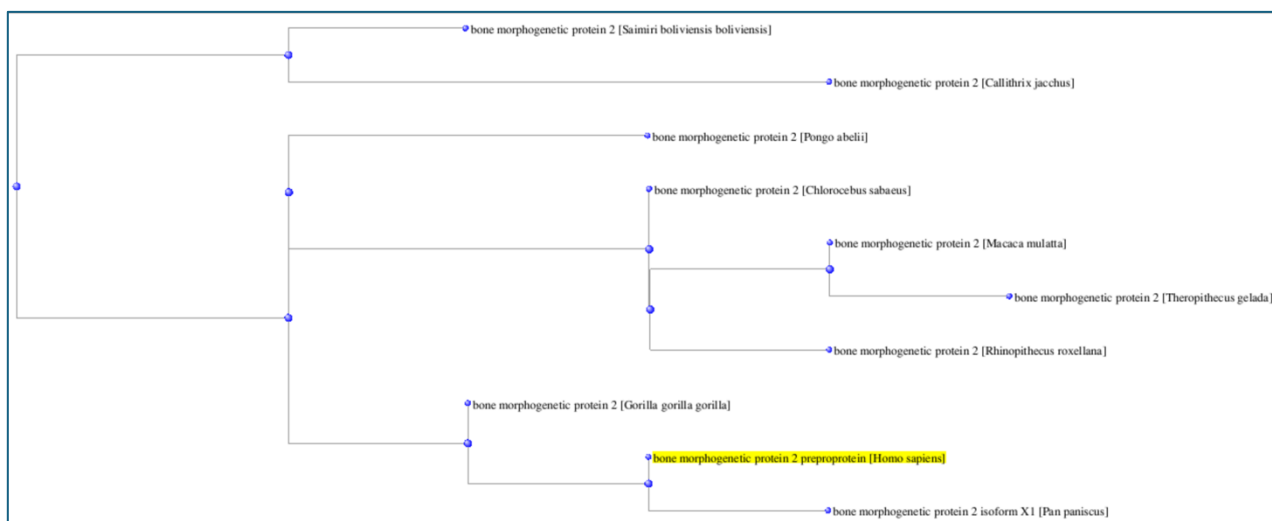
### 13. กัดไปที่ Phylogenetic Tree



### 14. ผลการสร้าง Phylogenetic Tree



## ผลการทดลอง



## การสร้าง Phylogenetic Tree จากผลของ blastp

จากผลการวิเคราะห์พบว่า *BMP2* ของมนุษย์ (*Homo sapiens*) อยู่ใกล้กับลิงโบนโบ (*Pan paniscus*) มากที่สุด นอกจากนี้กลุ่มลิงต่าง ๆ เช่น Gorilla, Rhinopithecus, และ Macaca ยังถูกจัดให้อยู่ในกลุ่มเดียวกัน แสดงถึงการมีลำดับกรดอะมิโนของโปรตีน *BMP2* ที่คล้ายคลึงกันและบ่งบอกถึงบรรพบุรุษร่วมที่ใกล้ชิดกัน ขณะที่สายพันธุ์ลิงใหม่บางชนิดอย่าง *Saimiri boliviensis* ถูกแยกออกไปอยู่ในกิ่งที่แสดงถึงความแตกต่างของลำดับกรดอะมิโนที่มากกว่า

ภาคผนวก

## NCBI accession number

- NP 001191.1

## ลำดับกรดอะมิโนของยีนที่สนใจ

>NP\_001191.1 bone morphogenetic protein 2 preproprotein [Homo sapiens]  
MVAGTRCLLALLLPQVLLGGAAGLPVELGRRKFAAASSGRPSSQPSDEVLSEFELRLLSMFGLKQRPTPS  
RDAVPPPYMLDLYRRHSGQPGSPAPDHRLERAASRANTVRSFHHEESLEELPETSGKTTTRRFFFNLS  
SIPTEEFITSAELQVFREQMDALGNSSSFHHRINIYEIIKPATANSKFPVTRLLDTRLVNQNASRWESFDVT  
PAVMRWTAQGHANHGFFVEVAHLEEKQGVSKRHVIRSRSLHQDEHSWSQIRPLLVTFGHDGKGHPLHKRE  
KRQAKHKQRKRLKSSCKRHPLYVDVSDVGWNDWIVAPPGYHAFYCHGECPPFLADHLNSTNHAIVQTLVN  
SVNSKIPKACCVPTLSAISMLYLDENEKVVLKNYODMVVEGCGCR