

Phylogenetic Analysis of the Genes Responsible for Albinism: Albinism Phylogenetic Analysis

Shikha Pandey, Kumud Sarin, Sumer Singh

Department of Bioinformatics, Singhania University, Rajasthan, India

Email: shikhapandey23@yahoo.com

Abstract

The aims of the research were to identify the genes responsible for Albinism. The genes have been identified and analyzed using BLAST. The genes responsible for causing albinism are OCA and OA. Albinism is caused by a mutation in one of several genes. Four forms of OCA are now recognized – OCA1, OCA2, OCA3 and OCA4; some are further divided into subtypes. Each of these genes provides instructions for making one of several proteins involved in the production of melanin. Melanin is produced by cells called melanocytes, which are found in the skin and eyes. A mutation may result in no melanin at all or a significant decline in the amount of melanin thus resulting in albinism. The phylogenetic relation of the genes responsible albinism has been established using various bioinformatics tools like Clustal Omega, MUSCLE, PHYLM, MEGA X and BioEdit. This phylogenetic study has helped us analysis how these genes have evolved with time and what level similarity exists between these sequences. This study paves path for the future analysis of the genes and drug designing for albinism.

Keywords: Albinism, Amino acid, Oculocutaneous, OccularOCA OA, Phylogeny, ClustalW, MUSCLE, PHYLM, MEGA X and BioEdit

INTRODUCTION

Amino acid disorders are major category of the inborn metabolic disorders. Amino acids are the building blocks of proteins. There are twenty amino acids present amongst which nine are the ones which are not synthesized by the body and can only be supplied through food, they are involved in metabolism.

Albinism is a congenital disorder categorised under amino acid metabolic disorder. The major characteristic feature of this disorder is complete or partial absence of pigment in the skin, hair and eyes, caused by the absence or defective tyrosinase. Tyrosinase is a copper containing enzyme responsible for the production of melanin.

There are several genes responsible for providing instructions to the several proteins involved in the production of

melanin. Mutation in any one of these several genes leads to albinism. Melanin is a pigment produced by melanocyte cells that provides colour to the skin and eyes. To have albinism and some of its type a person has to inherit two copies of a mutated gene i.e. one copy from each parent. This type of mutation leads to a significant drop or complete absence of the pigment melanin.

OCA genes are responsible for the Oculocutaneous albinism and OA genes are responsible for Ocular albinism. OCA genes located on the “autosomal” chromosome while the OA1 gene is located on the X chromosome. Autosomal chromosomes have genes responsible for the general body characteristics in contrast to the sex chromosomes. A normal individual has two copies of these chromosomes and genes one inherited from each of the parent i.e one from the father and the other from the mother.

Albinism is a recessive disorder, so in a person with albinism neither of the gene copies is functional, while if only one copy is also functional then a person can make melanin but will act as a carrier for the trait.

GPR143 gene plays an essential signalling role important for the pigmentation in the eye. Any defect in this gene causes Ocular albinism (OA1). OA1 gene is located on the X chromosome thus a male inherits it easily as they require just one copy of the defective chromosome from their mother while the females have two XX chromosomes, they would require both the defective copies for the disease to express itself. In females with one defective copy they would be normal but act as carrier.

People around the globe have different types of albinism. At times it also happens that people are not even aware that they have albinism. Various myths are prevalent in the society with respect to the colour of eyes in people with albinism, whereas the fact is there are various types of albinisms and with each type the eye pigment varies. Most frequently the eye colour varies from reddish to violet eyes majority being blue, but it might further range to hazel or brown also. It is basically the type and quantity of melanin produced by the body which determines the skin, hair and eyes colour.

As per the current situation there is no cure for albinism, one can only take protective measures to protect the skin and enhance vision. Over the period of time researchers have categorised albinism under various types and now studies are underway to find some suitable treatment for the same

Oculocutaneous albinism

Oculocutaneous (ok-u-low-ku-TAY-nee-us) albinism is caused by a mutation in one of four genes. People with oculocutaneous albinism (OCA) type 1 have milky white skin, white hair and blue eyes at birth. Some people with OCA type 1 never

experience an increase in pigmentation, but others begin to produce melanin during early childhood. Their hair may become a golden blond, brown or red, and their irises may change color and lose some translucence. Oculocutaneous albinism type 1 (OCA1 or tyrosinase-related albinism) results from a genetic defect in an enzyme called tyrosinase (hence 'ty' above). This enzyme helps the body to change the amino acid tyrosine into pigment. (An amino acid is a "building block" of protein.) There are two subtypes of OCA1. In OCA1A, the enzyme is inactive and no melanin is produced, leading to white hair and very light skin. In OCA1B, the enzyme is minimally active and a small amount of melanin is produced, leading to hair that may darken to blond, yellow/orange or even light brown, as well as slightly more pigment in the skin.

Oculocutaneous albinism type 2

(OCA2 or P gene albinism) results from a genetic defect in the P protein that helps the tyrosinase enzyme to function. Individuals with OCA2 make a minimal amount of melanin pigment and can have hair color ranging from very light blond to brown. OCA type 2 is most common in sub-Saharan Africans, African-Americans and Native Americans. The hair may be yellow, auburn, ginger or red, the eyes can be blue-gray or tan, and the skin is white at birth. With sun exposure, the skin may, over time, develop freckles, moles or lentigines.

Oculocutaneous albinism type 3

(OCA3) is rarely described and results from a genetic defect in TYRP1, a protein related to tyrosinase. Individuals with OCA3 can have substantial pigment. People with OCA type 3, mainly found in black South Africans, usually have reddish-brown skin, ginger or reddish hair, and hazel or brown eyes. OCA type 4

looks similar to type 2 and is most often found in people of East Asian descent.

Oculocutaneous albinism type 4

(OCA4) results from a genetic defect in the SLC45A2 protein that helps the tyrosinase enzyme to function. Individuals with OCA4 patients have similarity with OCA2 patients in the scenario that in both the conditions the patients make bare minimum quantities of melanin.

X-linked ocular albinism

This type of albinism occurs more or less exclusively in males and the gene responsible for the same is located on the X chromosome. People with this disorder have problems with vision whereas otherwise though on a lighter shade as compared to other family members their skin, hair and eye colour is usually in the normal range.

Oculocutaneous albinism (OCA) is a set of autosomal recessive disorders characterized by hypopigmented hair, skin and eyes. Homozygous or compound heterozygous mutations in the tyrosinase (TYR) gene can cause OCA1, which is the most common and severe subtype of albinism. A study stated here helps to describe the situation better: “17 patients with non-syndromic OCA were enrolled from eight provinces of China and were non-consanguineous. Total genomic DNA was isolated from peripheral blood. Screening was performed for the whole exons and their flanking regions of the TYR gene using Sanger sequencing and the pathogenicity of variants was predicted using in silico analysis. In total, 12 TYR mutations were identified in 10 patients, respectively. Of these, two patients carried homozygous mutations and eight patients carried compound heterozygous mutations. Among the 12 TYR mutations, two missense mutations” c.1198T>G

(p.W400G) and c.819G>T (p.Q273H) were novel [1].

TYR gene mutations have a more severe effect on pigmentation than mutations in OCA2 and the GPR143 gene. Nevertheless, mutations in these genes affect the development of visual function either directly or by interaction with other genes like MC1R, which can be deduced from a frequent association of MC1R variants with p.R305W or p.R419Q in OCA2 [2].

Objective: Phylogenetic analysis of the genes and their variants responsible for Albinism

Aims

- To identify the genes responsible for Albinism with the variants.
- To carry out the phylogenetic analysis of the genes and their variants responsible for Albinism.

Methodology

The gene sequences and their variants responsible for albinism have been identified and downloaded from the NCBI site in the Fasta format. These sequences have further been analyzed using the BLAST for their similarity with other related organisms. The sequences and their variants are further analyzed by performing multiple sequence alignment using software like CLUSTALW and MUSCLE for the similarity between these sequences [3]. Finally the phylogenetic relationship between these sequences has been established using various phylogenetic tools like PHLYM, MEGA X, and BioEdit. Studying the phylogenetic analysis of the sequences mentioned help us to predict how these sequences have evolved with time in different species and as per the current stage how closely the sequences are related and how they differ from each other [4].

Tools Used

BLASTn

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s)

Clear

Query subrange

From

To

Or, upload file

Job Title

Browse...

Enter a descriptive title for your BLAST search

☐ Align two or more sequences

Choose Search Set

Database

Human genomic + transcript Mouse genomic + transcript Others (nr etc.):

Nucleotide collection (nr/nt)

Organism

Optional

Enter organism name or id--completions will be suggested

Exclude

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

Exclude

Optional

Models (XM/XP) Uncultured/environmental sample sequences

Entrez Query

Optional

CLUSTAL Omega

Clustal Omega

Input form Web services Help & Documentation Bioinformatics Tools FAQ

Tools > Multiple Sequence Alignment > Clustal Omega

Multiple Sequence Alignment

Clustal Omega is a new multiple sequence alignment program that uses seeded guide trees and HMM profile profile techniques to generate alignments between **three or more** sequences. For the alignment of two sequences please instead use our pairwise sequence alignment tools.

Important note: This tool can align up to 4000 sequences or a maximum file size of 4 MB.

STEP 1 - Enter your input sequences

Enter or paste a set of

DNA

sequences in any supported format.

Or, upload a file Choose File No file chosen

See example inputs

Mega X



PhyML

← → http://www.phylogeny.fr/version2.cgi/one_task.cgi?task_type=phyml 🔍 📄 🔄 🗑️ Phylogeny.fr: PhyML

File Edit View Favorites Tools Help

🔍 Search **FANATIC™** 🎬 Movie Trailers 👍 Movie Reviews

🌟 hjcdgeszardfg, - Google S... 🌐 Google 📁 Suggested Sites 🌐 Internet Explorer cannot d... 🌐 Web Slice Gallery

Méthodes et Algorithmes pour la Bio-informatique LIRMM

Home Phylogeny Analysis Blast Explorer Online Programs Your Workspace Documentation Downloads Contacts

PhyML 3.0 aLRT (doc + aLRT)

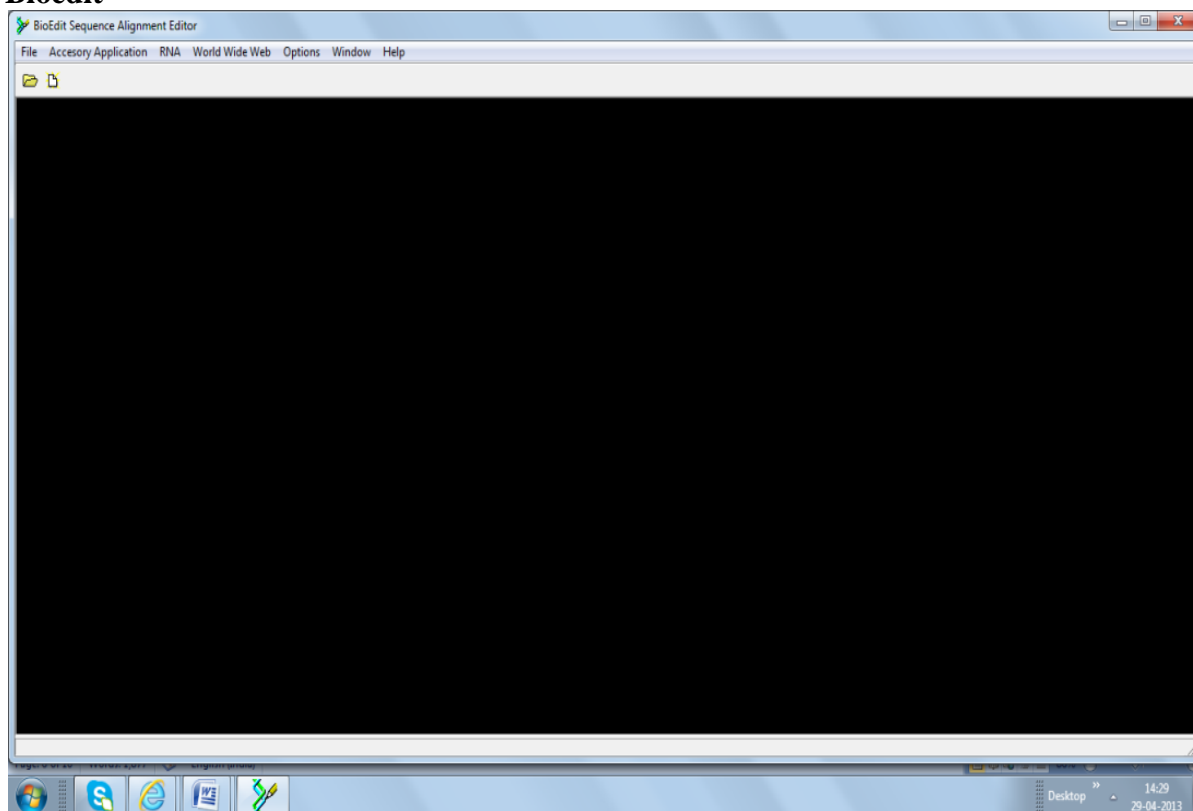
1. Overview 2. Data & Settings 3. Results

Datatype: ☒ auto-select ☐ protein ☐ DNA/RNA

Upload your alignment (FASTA, Phylip, Clustal, EMBL or NEXUS format) from a file:

Or paste it here (load example of alignment)

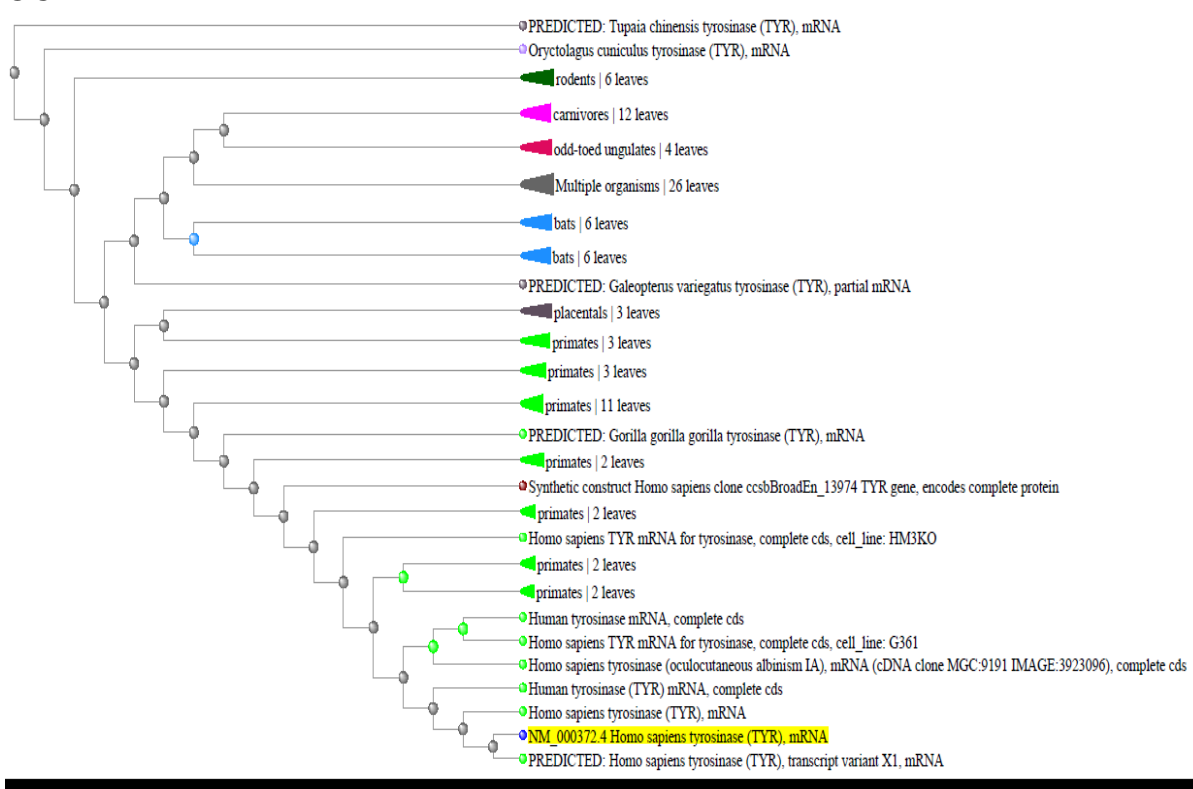
Bioedit



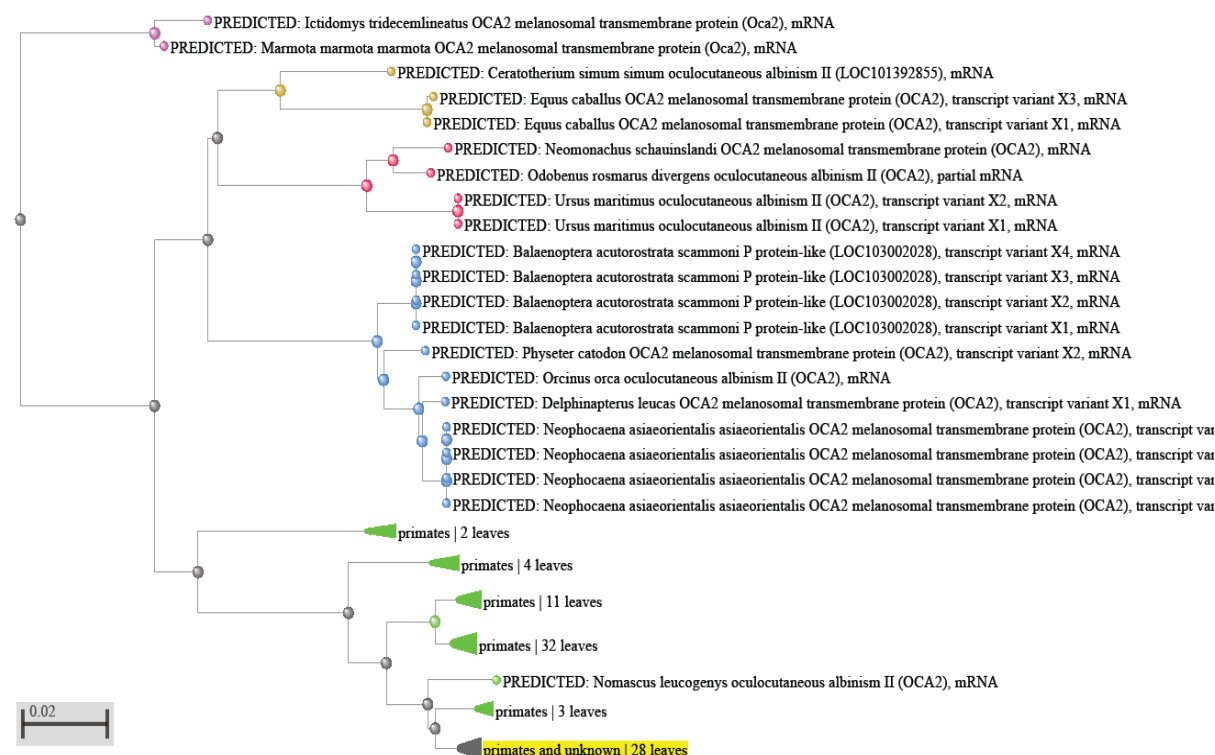
Results

Blast Results

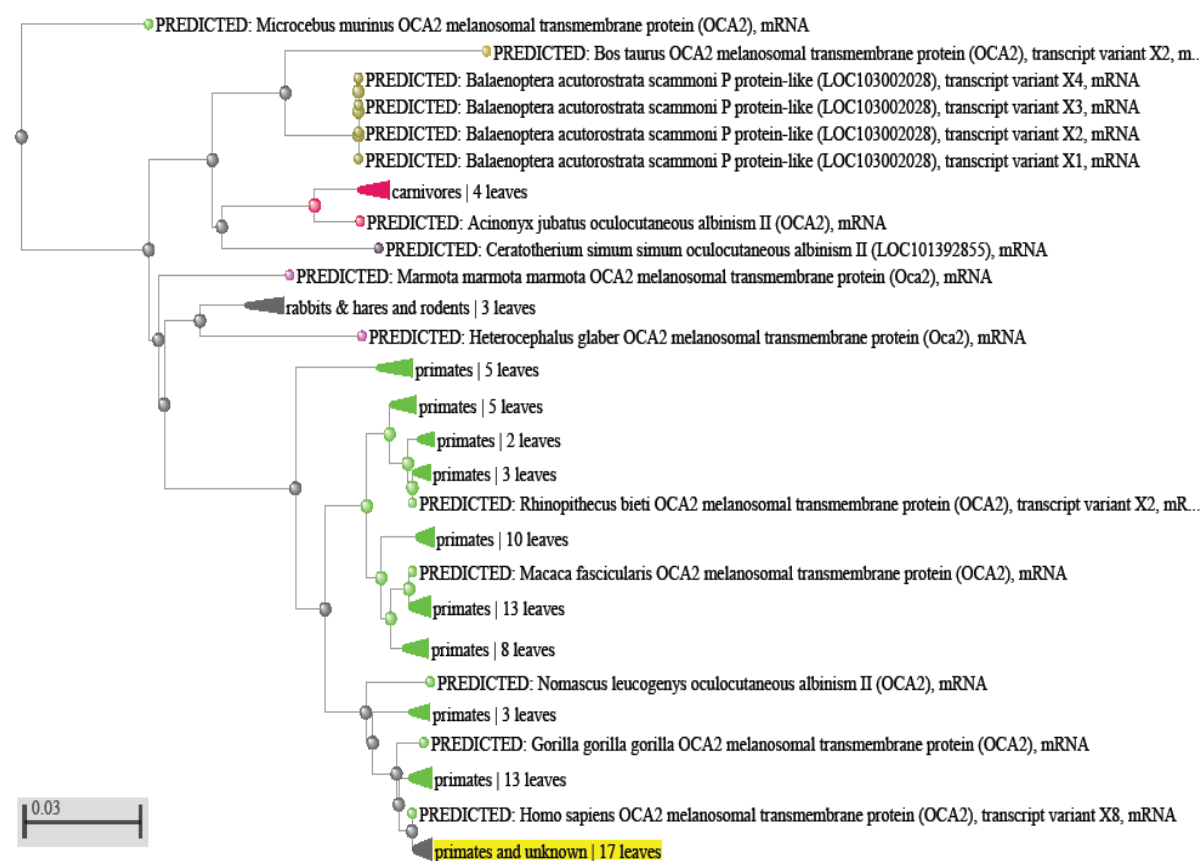
OCA1



OCA 2 Variant 1



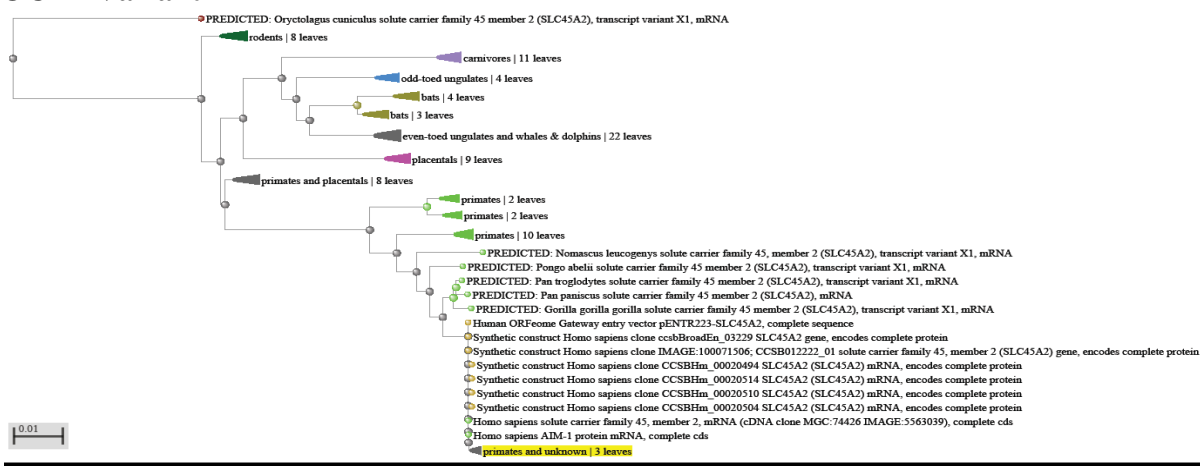
OCA 2 Variant 2



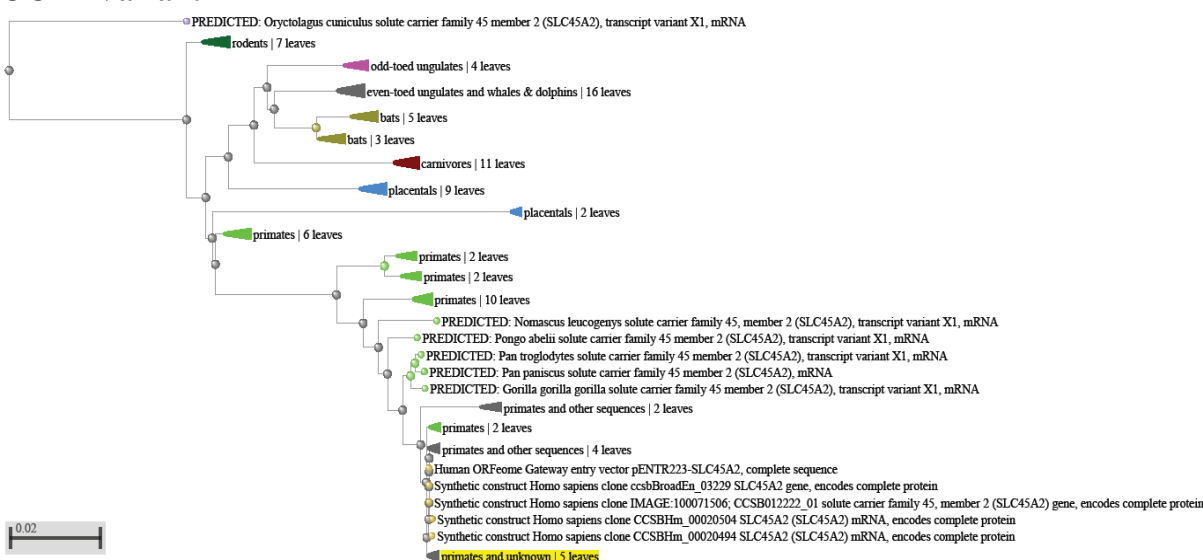
OCA3



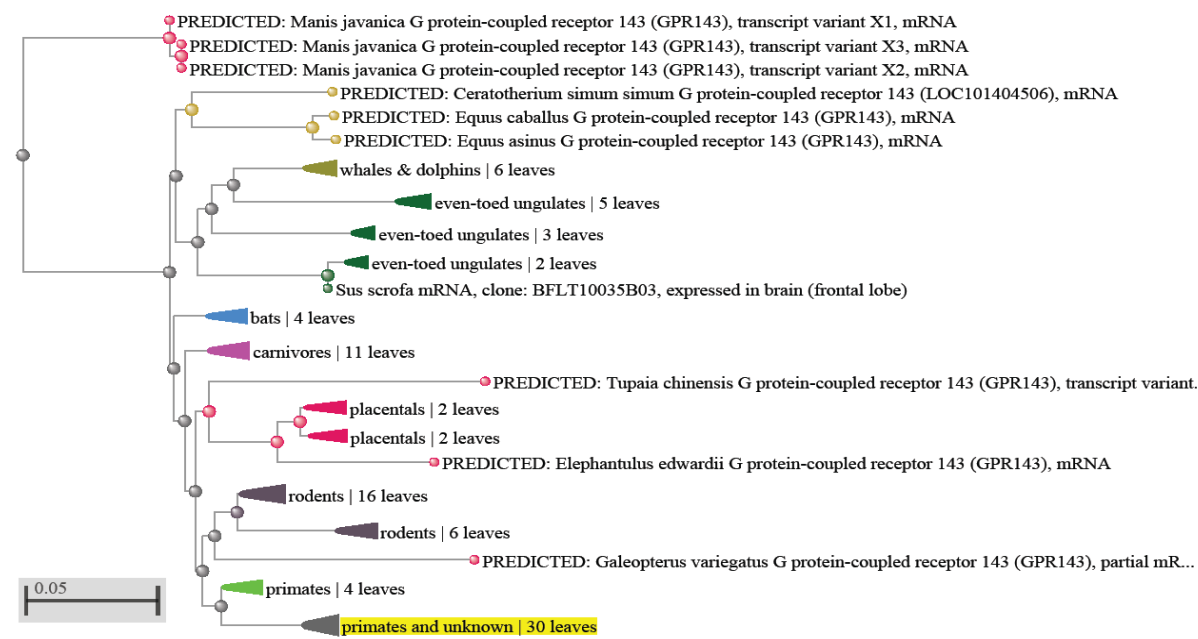
OCA4 Variant 1



OCA4 Variant 2



OA1



CLUSTAL W RESULTS

Table: 1. Percent Identity Matrix - created by Clustal2.1

1	NM_000273.2	100.00	44.81	45.02	34.54	34.11	31.93	33.44
2	NM_016180.4	44.81	100.00	91.21	38.30	38.11	36.62	37.41
3	NM_001012509.3	45.02	91.21	100.00	38.64	38.52	36.13	39.08
4	NM_001300984.1	34.54	38.30	38.64	100.00	100.00	39.65	41.85
5	NM_000275.2	34.11	38.11	38.52	100.00	100.00	39.63	42.00
6	NM_000372.4	31.93	36.62	36.13	39.65	39.63	100.00	50.32
7	NM_000550.2	33.44	37.41	39.08	41.85	42.00	5.32	100.00

MUSCLE RESULTS

```

NM_000372.  ATcACTgtagtagtagCtGG-----AaAGaGAAAtctgtGACTCca-----
NM_000550.  AcagtTctGAcGtgagaGGAcaagAaaCACAAGCaAataTaAAACAttcaAttCTAaGag
NM_000273.  -----caCCGaGC-----CtGgCtCTA-----
NM_0013009  cTTACTtCGAAG--GctGtGCTCCgCtCACcATCCAGAGcGgAggtGCgGACctTA-----
NM_000275.  cTTACTtCGAAG--GctGtGCTCCgCtCACcATCCAGAGcGgAggtGCgGACctTA-----
NM_016180.  -----GCCaGGCTCCACGtcaAATCCAGttTGAAACA-CaGACCCTAGG--
NM_0010125  -----GCCaGGCTCCACGtcaAATCCAGttTGAAACA-CaGACCCTAGG--

NM_000372.  -----AttAgcCAgtTccTGCAgacCTTGtgAG-----
NM_000550.  aagttcAtCAgAgACaTccTtCAGGA-TTgtgAGctGGAAttTCctCtAcgtGCTtcaGt
NM_000273.  -----cTGCAgGcgGtGggGttGG-----GGtgggGGA
NM_0013009  -----AACtCACTCcTGGaGaaAaGAtcTGCAAGTgcG-----
NM_000275.  -----AACtCACTCcTGGaGaaAaGAtcTGCAAGTgcG-----
NM_016180.  -----AcCACgCAggaGGTggtGGgCTcGCAgGaAGGTTCCTCTCCCAgTGGCAtGGG
NM_0010125  -----AcCACgCAggaGGTggtGGgCTcGCAgGaAGGTTCCTCTCCCAgTGGCAtGGG

NM_000372.  ---gActAGAGgaAga-----ATgctCCTgGCTG--
NM_000550.  cttCtCtAcAcAaAgaGCTGcaAaCCAaggTCTttgTtTtGcaCtCtTATttCaAGCaGaA
NM_000273.  gAGgcCCAGgGcacAtGaTGcCgcCCcCAgCc---cgcccAGCACATgaCCCagGCAGGc
NM_0013009  -----CAGAGAGAAGaCTGGCAGtgagcaTGCATcTgGAGggCA-----gAGacGGC
NM_000275.  -----CAGAGAGAAGaCTGGCAGtgagcaTGCATcTgGAGggCA-----gAGacGGC
NM_016180.  tAGCAaCAGtGgGcAGGCTGGCcGCCACATCT-----AtAaATCCCTAGCTG-A
NM_0010125  tAGCAaCAGtGgGcAGGCTGGCcGCCACATCT-----AtAaATCCCTAGCTG-A

```

```

NM_000372. -----tttTGtActGCCTGtGTGGAGtT-----
NM_000550. TGA--GTgCtCctaAaCtCCtctctctGgGCTGtaTcttcttCCC-----CTt
NM_000273. cGgCGGggtCCTGGcaCaCCc-----GAGCcGCGtCCgGcgAaCa-----CAGC-----
NM_0013009 aGgCGGTACCCcGGcgCGCCgGCgGTGGAGCTcCtgCAGAcgtCCgTGCCAGCgGACTC
NM_000275. aGgCGGTACCCcGGcgCGCCgGCgGTGGAGCTcCtgCAGAcgtCCgTGCCAGCgGACTC
NM_016180. TGAAtGGccCCtTtGACT-----CtGTGGAGCcGCcTaAaAgACCCA---CCAGCAGACTC
NM_0010125 TGAAtGGccCCtTtGACT-----CtGTGGAGCcGCcTaAaAgACCCA---CCAGCAGACTC

NM_000372. -----TCcaGaCCtcCgCtgGCCAT-TTCCctAGAGccTgTgTCTcCtCtaaG-----
NM_000550. GCTacttTttcaGCaGGCcCGgGctcaaTTCCCAAGAcAGTGtGcCactGttGaGgcttt
NM_000273. -----cCATGGCCtcCcCGCGC-----CtaGG-----
NM_0013009 GCTGAACttgTGGCCGGCAaGCGCaggCTTCC-----tcGGG-----
NM_000275. GCTGAACttgTGGCCGGCAaGCGCaggCTTCC-----tcGGG-----
NM_016180. a-----TCATGcaCaGCAAtG-GCCATgTTCggAAGAGAGTtCtgCTACGCgGtG-----
NM_0010125 a-----TCATGcaCaGCAAtG-GCCATgTTCggAAGAGAGTtCtgCTACGCgGtG-----

NM_000372. --aAcCtGaTGGAGaaGg-----AaTgCTGtCCacCG---TGGagCgGGGaCAGGa
NM_000550. gaGAagtGGTatgtGTtg-CC---CAgaCCTGtCCCctGtgtcTGGGcCTGGGaCAGaC
NM_000273. --GA-----CC---ttCTgCTGCCCCACG---cGGGaCgcaGCCAcGC
NM_0013009 --GAGCCGGTGGAGcTGACCCCTCGCACTCCTGCCCCAgG---GGGgCTG---CCgGGC
NM_000275. --GAGCCGGTGGAGcTGACCCCTCGCACTCCTGCCCCAgG---GGGgCTG---CCgGGC
NM_016180. --GAGgCaGcGtAtGTGACCC---CagTCCTGctCagCG---TaGGtCTG-----
NM_0010125 --GAGgCaGcGtAtGTGACCC---CagTCCTGctCagCG---TaGGtCTG-----

NM_000372. gtCcCTGTGGccagCtTtCAGGCaGaGgtT-----CCTGT---CAGaAAta
NM_000550. --CGCTGTGGctCatCatCAGGgAgGGGca-----gaTGTgAgGCAG--tG
NM_000273. AGCtC-GTGctGagCtTcCAGcCgcGGGccTTCcCacgcgctctgCCTGggcAGCgG----
NM_0013009 AGaGCTcTtGgGCTCCTGCAGGCcAGGagTTTgCTTCATT---CCTcacAAaagGGAGG
NM_000275. AGaGCTcTtGgGCTCCTGCAGGCcAGGagTTTgCTTCATT---CCTcacAAaagGGAGG
NM_016180. -----cCCaGCAG-----CCTGTAcAGCAAt----
NM_0010125 -----cCCaGCAG-----CCTGTAcAGCAAt----

NM_000372. TCctTCtgTCCaaTGCaCCA-----CttGGgCCTcAatTTCCc-----
NM_000550. aCTgcaGactCcCgGCCCCA-----caGcCCTcAGtaTCCc-----
NM_000273. -----cggGctCCg-----CttGGcgCTGgGccTtCTG-----
NM_0013009 TCTcaCtcTtCTtTGCCCCAGATGTCCAGCTCCAGGTcTaaAGATTCCCTGCTttACAgAA
NM_000275. TCTcaCtcTtCTtTGCCCCAGATGTCCAGCTCCAGGTcTaaAGATTCCCTGCTttACAgAA
NM_016180. TgTgTgGtTCCTCaGCCCCA-----TCCTGgGATTCCCTGCTGCAgcCcg
NM_0010125 TgTgTgGtTCCTCaGCCCCA-----TCCTGgGATTCCCTGCTGCAgcCcg

NM_000372. -----ttcaCaGG-----GgtgGAtGaCGGGagtCgTggCct
NM_000550. -----CaTGatGG-----cAGAGAtGatCGGGaggtCTggCcC
NM_000273. -----CaGCTGctGcCCggCcCGCGG---cccGcGggccCGGGtcCCCcgCgaC
NM_0013009 AACACtCcTttGCTGAgGAatTCCTtaCaG---GAGAAgGGtCacGGtgCaTACCTg
NM_000275. AACACtCcTttGCTGAgGAatTCCTtaCaG---GAGAAgGGtCacGGtgCaTACCTg
NM_016180. tggTcGgATCGGCcagcGACC-aCTGCCGGTCCAgtGtGgGgccGCCGGAgaCCCTACaTC
NM_0010125 tggTcGgATCGGCcagcGACC-aCTGCCGGTCCAgtGtGgGgccGCCGGAgaCCCTACaTC

NM_000372. TcCgtCtTttAtAATaGGACctgcCAgtgCtctGGCAActTCaTGGGAt-tCAACTGTGg
NM_000550. TTgcgCtTcttCAATaGGACatgTCActgCaaCGGCAATtTCTcaGGAC-ACAAGTGTGg
NM_000273. gTCCcCGgcCggcctcGGtcCgCATCtTgCgCGtGCGctgCTGcGAC--CttCTcgGC
NM_0013009 TttACCaTCcAgAgT-----TCATcaCtGctGaAgaGTCTTGGGAagACAgCTcTGC
NM_000275. TttACCaTCcAgAgT-----TCATcaCtGctGaAgaGTCTTGGGAagACAgCTcTGC
NM_016180. cTCACCc-----TGGGAgtCATgATgCtCGtGGgcATGgCTctGtACctCAAtgGgGC
NM_0010125 cTCACCc-----TGGGAgtCATgATgCtCGtGGgcATGgCTctGtACctCAAtgGgGC

NM_000372. aAACTG-----CaagTtTGGCttttGGgGACCaaaCTGcacagAgAgacgaC
NM_000550. gAcgTG-----CcGtccTGGCtGGAGagGAgCtGcCTGTgacCAGAgGgTTC
NM_000273. TgcCTGGGtatGgtGAT--CcGgTccacCGtGtGGT-----TAGGATTC

```

NM_0013009	TgACTGGGAGCGAAGATACTGcTaaGcaGGGAGGTGtCtGGTCTGTCTgCATcTGccTC
NM_000275.	TgACTGGGAGCGAAGATACTGcTaaGcaGGGAGGTGtCtGGTCTGTCTgCATcTGccTC
NM_016180.	TA-----CTGtTgTaGCaGc-----TtTGATTg
NM_0010125	TA-----CTGtTgTaGCaGc-----TtTGATTg
NM_000372.	tcttggtGAGAAGaAA--CaTcTTcGAtt-----TgAGTgcCcCagAGAA
NM_000550.	tcAtAgtcAGgAGaAA--tCTTcTGGACt-----TAAGTaaagaagAaAa
NM_000273.	CcAAAttttGttGacAG-CgTcTcGGAtA-----TgAacCaCACggAaA-
NM_0013009	CT---CCGAGAAGGgAGACCTTcTGGACAgCCCGCACATCCGACTccGTCTttCCAAGcT
NM_000275.	CT---CCGAGAAGGgAGACCTTcTGGACAgCCCGCACATCCGACTccGTCTttCCAAGcT
NM_016180.	CTAAcCCaAGgAGGAAGctggTTTGGgCcA-----TAAGTgTCACCATGAT
NM_0010125	CTAAcCCaAGgAGGAAGctggTTTGGgCcA-----TAAGTgTCACCATGAT
NM_000372.	GgACaAaTtTtTtgccTaCCTcActtTagCAAAGCAtACCaTcaGCtcagAcTaTGtcaT
NM_000550.	GAACCACTtTGTcGgGgcCCTGgAtaTggCAAAGCGcACaacTcACCCtTtTaTtTGtcaT
NM_000273.	-----TTTggccTGCTGc
NM_0013009	GAggCgCTGTGTgCaGTggCTGAAAGTCAtg---GgcCtGTTtGCCTTTgTggTGCTGT
NM_000275.	GAggCgCTGTGTgCaGTggCTGAAAGTCAtg---GgcCtGTTtGCCTTTgTggTGCTGT
NM_016180.	ag-----GtgtCGTTcTcTTgATTtTGCTGc
NM_0010125	ag-----GtgtCGTTcTcTTgATTtTGCTGc
NM_000372.	CccCaTaggga-----CCTatGgCCAAATGAAaaAtggatcAaCaCCaTgT
NM_000550.	tgcCacCAggagatcaGAagaaAataCTGGGgCCAGATGgca-----AcaCgCCacaAT
NM_000273.	ttTCTgCgTgG---gGAgTgCgatgTGGatCCAGcTGttGtA---cAGtgCctgctT
NM_0013009	gtTCTattTTG-----TtCAGCCTatatCCgGATcAAGGA-----AAGC-----T
NM_000275.	gtTCTattTTG-----TtCAGCCTatatCCgGATcAAGGA-----AAGC-----T
NM_016180.	CgaCTTCATTG-----aTGGGcCCAtca-----AAGCctaCtTAT
NM_0010125	CgaCTTCATTG-----aTGGGcCCAtca-----AAGCctaCtTAT
NM_000372.	TTaACGaCatcaatatTTATGAC-----ctCTttGTctGgaTgCattTattATGt
NM_000550.	TTGAgaaCatTTCCatTTATAAC-----TACTttGTttGgacaCACTattAct-
NM_000273.	cTGgtGGCTGTT-----tTgCtATG-
NM_0013009	cTGgCaGCTGTTggCtTTATcACcgCtgGAGAACTACTCCGTgAaCCTtagCagCCACG-
NM_000275.	cTGgCaGCTGTTggCtTTATcACcgCtgGAGAACTACTCCGTgAaCCTtagCagCCACG-
NM_016180.	TTGAtGtCTGcTCCATcAgGACAAgGAGAagg-----GCCTcCACTACCATG-
NM_0010125	TTGAtGtCTGcTCCATcAgGACAAgGAGA-Agg-----GCCTcCACTACCATG-
NM_000372.	gtCAaTggAtgCACtgCTTGGGGG-----atCTGAAatCTgGaGaGaCA-----TtGAt
NM_000550.	--CAGTCAaAaagacTtTccttGGgGtaggacagGAAagCTTtGGtGaag-----TgGAt
NM_000273.	--CAGTggAtgCttaTCTgGtGatccggagatCgGcAGgacTGAaGcaCCAtcctGctGta
NM_0013009	--tgGaCTcCACgCtgCTgcaGGtgGACCTGGCaGggGcCcTaGtgGCCAGTgGGccGAg
NM_000275.	--tgGaCTcCACgCtgCTgcaGGtgGACCTGGCaGggGcCcTaGtgGCCAGTgGGccGAg
NM_016180.	--CccTCTtCACAggTtTtGGAAGGtGcCCTGGgTtAccttTTGGGtGCTA-----TaGAc
NM_0010125	--CccTCTtCACAggTtTtGGAAGGtGcCCTGGgTtAccttTTGGGtGCTA-----TaGAc
NM_000372.	TttgCCCATGAaGcAcCAGctttTCTgCtTGGcataGaCtCttCtTgTgCGGgTgGGAA
NM_000550.	TtctCtCATGAGGGAcCAGctttTCTcaCaTGGcaCaGGtACcACCTccTGCGTcTGGAG
NM_000273.	TCacatCATG-----GCgTGGgGCctGgcCacCCTGcTctGTgTGGAG
NM_0013009	TCGtCCtggGAGGGAAgAGCACATCGtG-gTGGAGCTGaCcCAGgCTGAcGCtTTgGG--
NM_000275.	TCGtCCtggGAGGGAAgAGCACATCGtG-gTGGAGCTGaCcCAGgCTGAcGCtTTgGG--
NM_016180.	TgGgCCCAT-----cTGGAGCTGGgAagACt-----gTTGGgt
NM_0010125	TgGgCCCAT-----cTGGAGCTGGgAagACt-----gTTGGgt
NM_000372.	cAAGAAaTCCAGaaGctGaCaGgAgAtGaAaacTtCaCTatTCcatAtTGGGAcTggcgg
NM_000550.	AAAGAcATgCAGGaaatGttGcaAgAgcCtTcTtTCTCcCtTCcttACTGGaAtTTTgcA
NM_000273.	ggAGccgc-----CATGcTCTaCTaCcCT-----
NM_0013009	-----cTCCAGGTGgCGGCGGCcAcAGCAGGTcaCTCaCaaCTGgACgGtGtATTTaaA
NM_000275.	-----cTCCAGGTGgCGGCGGCcAcAGCAGGTcaCTCaCaaCTGgACgGtGtATTTaaA
NM_016180.	AcAGAAaTCCAGGT-----CATGTTCTTCTtCTCTGcAtTGG-----
NM_0010125	AcAGAAaTCCAGGT-----CATGTTCTTCTtCTCTGcAtTGG-----

```

NM_000372. gatGcaGAAAAAG---TGtGacaTtTGcacaGAtGaGTaCATGGGAGgtCAGcACCcCacA
NM_000550. aCgGgGaAAAAatgtcTGtGataTCTGcacgGAtGacTTgATGGGAtccagaagCaaCTtt
NM_000273. -----TCCGTGTccAGGTGTgagcGGGgc--CtGgACCAcgcc
NM_0013009 tCcGaGGAgAAG---cGAGCACTCaGTGaTGAG-----CAGgACCTtTGA
NM_000275. tCcGaGGAgAAG---cGAGCACTCaGTGaTGAG-----CAGgACCTtTGA
NM_016180. -----TGctCACTtTGTGTTttacTGTTTCATctGtG--CAGtAtCTCTGA
NM_0010125 -----TGctCACTtTGTGTTttacTGTTTCATctGtG--CAGtAtCTCTGA

NM_000372. AaTCCtAACTtActc---aGc--CCAgCATCATTC-----
NM_000550. gaTtCCAActctAAAtA---aGc--CCAAactCtgTC-----
NM_000273. AtcCCCcACTA-----tGTcaCCAtgtaCcTgC-----
NM_0013009 gGTaCtgACcAgAGAGACGGTgtCCAtCAGCATcCGGGCCTCCCtgCAGCAgACCCAGgc
NM_000275. gGTaCtgACcAgAGAGACGGTgtCCAtCAGCATcCGGGCCTCCCtgCAGCAgACCCAGgc
NM_016180. AGcCCCActTAcAGA---GGTtGCaAAggGCATTC-----CCCCACAGCAAACCCctCA
NM_0010125 AGcCCCActTAcAGA---GGTtGCaAAggGCATTC-----CCCCACAGCAAACCCctCA

NM_000372. -----tTCTccTctTGGCAGATtGTCTGTAg-----CcgattGGAgGAGTAcAAc
NM_000550. -----tTtTcTcaATGGCgagTgGTCTGTga-----CTCCttGGAAGAtTAgtAt
NM_000273. -----CcCTGcTGcTGGt-----
NM_0013009 tGtCCCTCTtTTGATGGC---TcaTcAGTAC-----CTCCgcGGAAGtGTAgAAA
NM_000275. tGtCCCTCTtTTGATGGC---TcaTcAGTAC-----CTCCgcGGAAGtGTAgAAA
NM_016180. GGACCTCcaTTGtcatCAGATgGaaTGTACGAGTATGGTTCTatcgaGAAagtTAaAAA
NM_0010125 GGACCTCcaTTGtcatCAGATgGaaTGTACGAGTATGGTTCTatcgaGAAagtTAaAAA

NM_000372. AgCcatCagtCtTtatGCaAtGGaAcgCCcgAgGGACCttTACGgcGTAATC---CTGG-
NM_000550. ACCcTGgGAACActtttGtaACaGCaccgaggAtGGgCCAATtAGgAGaAATCagCTGG-
NM_000273. -----TCtcGtGgCGaaccCaTCctGttCCAAaAgaCAGTgActgcAgTGGc
NM_0013009 cCCAgGtGAcCATC---GCGACGGCcATCCTCgcGGgCgtcTAcGCgcTgATCATAtTtGA
NM_000275. cCCAgGtGAcCATC---GCGACGGCcATCCTCgcGGgCgtcTAcGCgcTgATCATAtTtGA
NM_016180. tggTtAcGtAaATCCAGAGctGGCAATgC--AgGGAgCAaAaAaCAaaAATCATgCTGa-
NM_0010125 tggTtAcGtAaATCCAGAGctGGCAATgC--AgGGAgCAaAaAaCAaaAATCATgCTGa-

NM_000372. -----
NM_000550. -----
NM_000273. CtcttTaCtt-----
NM_0013009 g-----
NM_000275. gatcgTgCacagaactctggcgccatgctgggttcccttgcagcactggcagcactggc
NM_016180. -----
NM_0010125 -----

NM_000372. -----AAACCatGaCAaAtCcA-----GAacCccA
NM_000550. -----AAAtgtgGCCAGACCAA-----TGgtgC-AA
NM_000273. -----AAAggaAGaCAaggCAt-----TtACACggA
NM_0013009 tgtgattggcgatAgACCCAGCctGACCcATGTGGTGGAGTGGATTGATTTTGAACgct
NM_000275. tgtgattggcgatAgACCCAGCctGACCcATGTGGTGGAGTGGATTGATTTTGAACgct
NM_016180. -----AcAgaCtcgCAGggCAA-----TGACAttAA
NM_0010125 -----AcAgaCtcgCAGggCAA-----TGACAttAA

NM_000372. AGgCtcCCctcTTCagcTGATGTAGAattttTGcctg-----aGtTTgacCc
NM_000550. cGTCtTCCTGaacCaCAGGATGTcGctCAGTGcTtg-----gAAgttGgTTTAttt
NM_000273. gaaC-----gagaGgAgGATG-gGAGCcTGaTCA-----AgATcCGATTTTTTCa
NM_0013009 gGcCCTgCTGtTTgGCATGATG-AtcttAGTaGcCATATTTTCAGAAAcGgGATTTTTTCg
NM_000275. gGcCCTgCTGtTTgGCATGATG-AtcttAGTaGcCATATTTTCAGAAAcGgGATTTTTTCg
NM_016180. AGTC-----acTGcTG-AGAGCAcTGGTgA-----AcATGCctccTcaCt
NM_0010125 AGTC-----acTGcTG-AGAGCAcTGGTgA-----AcATGCctccTcaCt

NM_000372. AAAtATgAatCTG---GttcCA-----tGGATa-----
NM_000550. gACAc-----GCCTC-----
NM_000273. AA-ATcATgCTG---GttTtAattAtTTgtTgGttGtCGaATaTcatCaATGA-----

```

NM_0013009 AttATTgTgCTGTAAAGgCatACCggtTctcccCGGGACGGgTGTgGGCCATGATCATCA
NM_000275. AttATTgTgCTGTAAAGgCatACCggtTctcccCGGGACGGgTGTgGGCCATGATCATCA
NM_016180. AcCgcTAcctTt----GCaTCAGCcACcTcAT-----tGGATGgA-----
NM_0010125 AcCgcTAcctTt----GCaTCAGCcACcTcAT-----tGGATGgA-----

NM_000372. -----aAGCtgcCaatTtCagCTTtAgAAAtAcaCTGg
NM_000550. -----CtTTtTatTCCAActctACAAACAgtttTcC
NM_000273. -----aAGCCTTtTatTctAtCTTgAgAtgCAaaCaGa
NM_0013009 TGCTCTGTCTCATCGCGCCGTCCTCTCtGCCTTCTTGgaCAACgTcACcAcCATGCTcC
NM_000275. TGCTCTGTCTCATCGCGCCGTCCTCTCtGCCTTCTTGgaCAACgTcACcAcCATGCTcC
NM_016180. -----CAGCCTTcTGTCcAAC-----ATGCTGt
NM_0010125 -----CAGCCTTcTGTCcAAC-----ATGCTGt

NM_000372. -----AAGGAtTtTgctA---GtCCAct-----
NM_000550. gaaaCACagtG-----AAGGTTacAgtg---aCCC-----
NM_000273. TaTcaAtG-----gAGGTTcTtTGA---aaCctgtCagaActgcAGcCAAGaCC
NM_0013009 TCTTCACGCCTGTGACCATAAGGTTgTgTGAGGTGCTCAaCCTTGATCCAAGACAAGtCC
NM_000275. TCTTCACGCCTGTGACCATAAGGTTgTgTGAGGTGCTCAaCCTTGATCCAAGACAAGtCC
NM_016180. TCTTCA-----cAGaTTtcATGg-----gCC
NM_0010125 TCTTCA-----cAGaTTtcATGg---G-----CC

NM_000372. -----tActGGgATAgCgGAtgCctCtcaaaGcAGCAtgC
NM_000550. -----CACGGGAagtaTGACcCTGctgTtcGaAGtCttC
NM_000273. AcATgGTtT-----AttAtGGGAATcCtgaAtcC-----AGCCCAg
NM_0013009 tGATTGcAgAAGTGATCTTCACAAACAttGGAggAgCTGcCACTGCCaTCgGgaCCCTc
NM_000275. tGATTGcAgAAGTGATCTTCACAAACAttGGAggAgCTGcCACTGCCaTCgGgaCCCTc
NM_016180. AGATTGTgT-----AcCgCGGGgATcCCTa-----tAGtgCAC
NM_0010125 AGATTGTgT-----AcCgCGGGgATcCCTa-----tAGtgCAC

NM_000372. ACAATgCC-----TTgCaCATCTAtA--TGAATGGAAcAatGtCCcaGgtaCAGGGaT
NM_000550. ACAATT-----TggCtCATCTAttcCTGAATGGAAcAggGGgacaaAccCAtttgT
NM_000273. ggAtTTTctTGTctTTggCtTCTACgGCTGgAcaGgAtgcAGcCtgGGt-----
NM_0013009 caAATgtCATTaTtGTTCCaAcCaAgAGCTGA--GGAAgAtgGGCCTGGA-----
NM_000275. caAATgtCATTaTtGTTCCaAcCaAgAGCTGA--GGAAgAtgGGCCTGGA-----
NM_016180. ACAAcTCCAcagGag-TTTCTCATCTACgaaaGA--GGAgtcgAGGttgGatgTtgGGGCT
NM_0010125 ACAAcTCCAcagGag-TTTCTCATCTACgaaaGA--GGAgtcgAGGttgGatgTtgGGGCT

NM_000372. CtgCCAACgAtCCTATcTTccTtcttCacCATgC--ATTtgtTGacAGTattTTTGAgCa
NM_000550. CtcCaAAtgAtCCTATtTTtGTCctcCTGCACACcTtcaCagatGcAGT--CTTTGATga
NM_000273. -----TtTcaGTCTCcCaGgAaggAgATcCAGTGGGAaT--CacTGAcCa
NM_0013009 -----CtTtTGccgGatTCACTGCACACATgTTCAtTGGGAtTTGCcTTGtTCT
NM_000275. -----CtTtTGccgGatTCACTGCACACATgTTCAtTGGGAtTTGCcTTGtTCT
NM_016180. tCtgCatCAACtCcgTGTTtctCTACTtttAttCtTAcTttcaGaaAGT---TTTGgTaT
NM_0010125 tCtgCatCAACtCcgTGTTtctCTACTtttAttCtTAcTttcaGaaAGT---TTTGgTaT

NM_000372. -----GtgGCTccGaAGGcACcgTCCTctTcaAgaagttTaTCCAgaagGCcAAtGCacCC
NM_000550. -----atgGCTgAGGAGaTACAaTgCTgaTatAtccacaTtTCCAattgGAAAAAGCCcCt
NM_000273. CCTcGGCTGCTgAGGgGGctCACcCaT-----Cccc
NM_0013009 CCTgGtCTGCTt-----TcCgCTCCT-----CAGAcTCTTt
NM_000275. CCTgGtCTGCTt-----TcCgCTCCT-----CAGAcTCTTt
NM_016180. -----CCTaC
NM_0010125 -----CCTaC

NM_000372. ATTGGACAtAAccGGgaaTcCtACATGGttCctTtTaTaCcaCT--GtacaGAaAtGGtG
NM_000550. ATTGGACAtAAAGaGaaTACAACATGGtgCCaTtctGgCccCcaGtCaccaAcAcaGAa
NM_000273. AcTG-----aTgCccCATGaAAaC-----CcTGcttccGgGAAGGtG
NM_0013009 AcTGGA-AcAgaAaGCTTTAtAACAAAGGAACCCagTgaGaTtgTTGaaCTGaAGcAcGAG
NM_000275. AcTGGA-AcAgaAaGCTTTAtAACAAAGGAACCCagTgaGaTtgTTGaaCTGaAGcAcGAG
NM_016180. ATTGGAAttAAAaggGtCTTTACttCAcGGgA---TATTTGCTgtTTGGCCTGGgGAcGGga
NM_0010125 ATTGGAAttAAAaggGtCTTTACttCAcGGgA---TATTTGCTgtTTGGCCTGGgGAcGGga


```

NM_000372.    ATTtCtTt-TatTtCatCCaAagatC-----
NM_000550.    ATgttTgTtTaCTgCtCCaGACaaCCTgGgatacacttatgaaattCAatgGCCAaGtCgG
NM_000273.    tcTCAaG-TGggtgGgCaGAC-----ttCtgaCgAaGCC--
NM_0013009    ATTCaCg-TctggCGCCTGACTGtCAGCG-----CATCAGCCcGGCCAG
NM_000275.    ATTCaCg-TctggCGCCTGACTGtCAGCG-----CATCAGCCcGGCCAG
NM_016180.    tTTatTG-gGCTcttCCCGAaTGtCTA-----CtCCAcCCTGGtC--
NM_0010125    tTTatTG-gGCTcttCCCGAaTGtCTA-----CtCCAcCCTGGtC--

NM_000372.    -----TGggCtaTGaCTataGtATctaCAAGaTtCaGacCCaGa-----
NM_000550.    gaG-----TttAgtgTaCCTGagatAATtGCCAtaG--CaGTAgttGG-----
NM_000273.    CTG-----aGCAtgCTGtCTG-----AAGGTtCTGatgCCaG-----
NM_0013009    CcGCGAGGAGACAGCTGtgCgCcGCCTGcTGC--TGGggAAGGTgCTGgcaCtGGAGCAC
NM_000275.    CcGCGAGGAGACAGCTGtgCgCcGCCTGcTGC--TGGggAAGGTgCTGgcaCtGGAGCAC
NM_016180.    CTG-----TGCAGcCTGttTGgTGtAATGtCCAgcaccCTGTACaCtG-----
NM_0010125    CTG-----TGCAGcCTGttTGgTGtAATGtCCAgcaccCTGTACaCtG-----

NM_000372.    -ctCTtTT-----
NM_000550.    -cGCTtTgttactGGtTgcACTCATTtTtTggGACtGcTtcttatctgattcgtgccagaC
NM_000273.    -caCaaTT-----gAAATTCaCaCtgCAagtgaATCc-----t
NM_0013009    CTGCTCgcCCGgAGGCTgcACaCcTTCCaCAGACAGATCt-----C
NM_000275.    CTGCTCgcCCGgAGGCTgcACaCcTTCCaCAGACAGATCt-----C
NM_016180.    -TGCcCTT-----TAAcCTCATTaCTgAGtaccAcCg-----C
NM_0010125    -TGCcCTT-----TAAcCTCATTaCTgAGtaccAcCg-----C

NM_000372.    -----caAgacTa-----CATtAAGtCCTAT
NM_000550.    GCAGtAtggAtGAAGcTa---ACCAGccTCtCctcACT-----gATCAgtAtCaAT
NM_000273.    GCAacAaAAAtGAAGgTG---ACCctgcTCtCccAACc-----CATggAGACCTAT
NM_0013009    aCAGgAGgAcAaAAAtTGGGAGACCAATATCCaAgAACT-----CcaaAAaAagCAT
NM_000275.    aCAGgAGgAcAaAAAtTGGGAGACCAATATCCaAgAACT-----CcaaAAaAagCAT
NM_016180.    GagGaAGAAAAGgAGagG-----CAGCAGGcCCCA-
NM_0010125    GagGaAGAAAAGgAGgTGTGctgtCatTgaaCCtctgCcttgggtaaCATCcAGACTCAT

NM_000372.    ttGgaAcaaGcGaGtcg-----
NM_000550.    gctATgctGaAGaatATgaaaaaCtCCAgaatCCTAAtcagtCTGtGgTcTAAcAaaTgC
NM_000273.    -----GaAGGGGAT-----
NM_0013009    agGATAtCtGAcGGGATTCTGCTCGCCAAATGCCtgACAGtGtTGGGATTTgttAtCTtC
NM_000275.    agGATAtCtGAcGGGATTCTGCTCGCCAAATGCCtgACAGtGtTGGGATTTgttAtCTtC
NM_016180.    -----GGAGGG-----
NM_0010125    ctGAaggCtGgGGAgaTtTGgTaGCaAtgTGtagAgCAGaGCTGGaAaaTAAgggtTc-

NM_000372.    -----
NM_000550.    cctacTcTCtTatgcattagTaTCAcaaaaaccaCctggTtgaaTataATaGATTGAgTATA
NM_000273.    -----
NM_0013009    ATGTTTTTCCTCAATTCGTTTGTCCTGGCATTcATCTTGATCTGGATGGATTGctATt
NM_000275.    ATGTTTTTCCTCAATTCGTTTGTCCTGGCATTcATCTTGATCTGGATGGATTGctATt
NM_016180.    -----
NM_0010125    -----TTGAaAcA

NM_000372.    -----
NM_000550.    tTAACTGtatTtTctTTcACttTaTTAcCTtctttctaatacaagcatatgttaGcATTa
NM_000273.    -----GTgcTg-----
NM_0013009    cTgggTGCCATCTGgTTgctAATTTTAAGCTG-----aTATTC
NM_000275.    cTgggTGCCATCTGgTTgctAATTTTAAGCTG-----aTATTC
NM_016180.    -----
NM_0010125    gcAACaGtCAaaTGcccaACcAacT-----GcAacC

NM_000372.    --GATCTgGt-----CATggctCcttGgGgcGgCgAtgggtAgGGGCcgtCCTcactgCcc
NM_000550.    AaGtTCTAGg-----CATACTtttCAaAGctGggaAGaCCcTttcagAatCTTtTcaaTg
NM_000273.    ggGgTCcAGAccc-----CATATtCtCAGAcTcaACaA-----
NM_0013009    ATGATtTtGAgAT--aAtTcTACACAGAGTgGA-----ATGGGCaACCCTTCTgTtTT

```

NM_000275. ATGATtTtGAgAT--aATtCTACACAGAGTgGA-----ATGGGCaACCCTTCTgTtTT
NM_016180. --GAcCcAGA-----CA-----ACAGcGTgag-----
NM_0010125 ATGATCTgGtacaggCATgCTcacCAGcccaGcCcAGcCCA---GctgtCCTgtccTCcT

NM_000372. TGcTgGCAGggCTTGT-----GAGC-----
NM_000550. gGtTttaAtTTtcaGTTCTaTTtaaaatggtgaATGacaCTAaActcCatGAataTtTaag
NM_000273. -----TTCTTGTTCT-----tTaGAaCTg-----
NM_0013009 TtgcaGCgcTctTTGTTCTg-----ATGGAGgcA-----
NM_000275. TtgcaGCgcTctTTGTTCTg-----ATGGAGgcA-----
NM_016180. -----
NM_0010125 TatTctCAacTCcTGcTCacTT-----AgGGAGCTgtAgggCAaGAagTgT---

NM_000372. -----TTGctgtgTCgTC-----acAAGAGAAag
NM_000550. gatagtgtgaagatctTTGgCATgatTTaaaggTTgagTAgtGAAGATatAAGtaAgtg
NM_000273. -----TgttCtCAcCTTC-----ccaA
NM_0013009 -----TTGgCACATCTcC-----ActTAAtAGAATATGttgGAGAAcA
NM_000275. -----TTGgCACATCTcC-----ActTAAtAGAATATGttgGAGAAcA
NM_016180. -----AgGgAAGgGcAtg
NM_0010125 -----TactCcCATtTTtctaaTTAgaaAAagGctTggtgAgGctgAgA

NM_000372. cAgcttCCTGa-----AGAAAAGCAGCcACTCcT-----
NM_000550. AACTaCCaTGC--ttTgtttAcGtgtaaAGgAAAatAatGttTgATAGtaaaTGTcCaCtT
NM_000273. cACTGCaCTGC-----cGAagtGtAGCGgCcCccAaaCctTGcTCTC--
NM_0013009 AACTGCttTGCTAATAAAgATGGtCCCAGAggAGCAGCGcCTCATAGCCgCcaTTgTCCT
NM_000275. AACTGCttTGCTAATAAAgATGGtCCCAGAggAGCAGCGcCTCATAGCCgCcaTTgTCCT
NM_016180. gACTGCgCc-----
NM_0010125 AACcagCCCaCaAtTAcAcAaaGaC--AaAAAAaggGgGACcCAGgtCttCTGTaacCac

NM_000372. --CATgGagaAAGaGG-----
NM_000550. aaaATAcATGAAtgGG---CAtttCTAaaATgtTaAaAcATaaACAcattTtCCaTTCAt
NM_000273. -----
NM_0013009 -----GGTGgtGtGGGTCTCAgCCCTgGCgTcGTCcCtgATTGACAACATcCCGTTTCAC
NM_000275. -----GGTGgtGtGGGTCTCAgCCCTgGCgTcGTCcCtgATTGACAACATcCCGTTTCAC
NM_016180. -----
NM_0010125 ctCATAGGTctccctGGTgcCAcCtgcAGCATttctAgAtAccGACAAaAagggtcTCcC

NM_000372. -----AtTACCAcAgcT-----TGtAtCAGA-GCcAtTtataaAAGG-----
NM_000550. ggatatttgTcAaCAGAttTAaAGaaaA-----CCAcA-GtTatTAattaAAGa-----
NM_000273. -----AtcACCAG-----CtAGA-GCTtCTtCCCCAAGG-----
NM_0013009 T-----gCTACCAtgAtTcCcGTGctCCTGAACCTGA-GCCACgACCctgAGGTTGGC
NM_000275. T-----gCTACCAtgAtTcCcGTGctCCTGAACCTGA-GCCACgACCctgAGGTTGGC
NM_016180. -----ACcctCAcAtgcAtgGTGCAgCTGgctCAGA---TcCTggtCGgAGG-----
NM_0010125 TatccccaACTttCAgAggTACATAgtggaatttCttGgcatTATtTgCtGAGGGTAgGA

NM_000372. -----cTtAGGcAAAtAGa
NM_000550. -----AaAttaattaTGTGtaGTtAtAaa
NM_000273. -----
NM_0013009 CTGCCCCACCGCCGCTCATgTATGCCCTGGCCTTCgGtgCTTGCCCTGgGAGGTAACgGG
NM_000275. CTGCCCCACCGCCGCTCATgTATGCCCTGGCCTTCgGtgCTTGCCCTGgGAGGTAACgGG
NM_016180. -----
NM_0010125 aaGtCCtaggCaatGgagtccagattttaGtaagTCAGAcCcatgtGTGAGaaggaAaG

NM_000372. gtAGgGcCAaaaaGcCTGAcCtcactcTaacTcAAaGTAAATGTccag---GtTccCAGAg
NM_000550. cCAatGaaATtttGatTaAcCtt-----tTcAAatTAaTGTtcCa---GtTTGaAGAc
NM_000273. -----GCCtTtAGgA
NM_0013009 ACA-----CTGATtggCgCgTCGgcAAACGTcGTGTGtgCAGGGaTTgCAGAA
NM_000275. ACA-----CTGATtggCgCgTCGgcAAACGTcGTGTGtgCAGGGaTTgCAGAA
NM_016180. -----TGgCCtGGGCTTtCtGgt
NM_0010125 ctgGaataCATgggtggTaATCccCaCtTtGcTgcccCtgAGcGgGaCCtGtGCcatttaAA

```

NM_000372.      aAt-----aTcTgCtgGTAtTttTctg-----
NM_000550.      CAatcaAAtatAttattTAGTcaACataTACtatTtag-----
NM_000273.      tAGgAgAAaGg-----GTtcAtGcacACaGgTgtg-----
NM_0013009      CAGCAt-----GgATAtGgGTtCTCcTtCA-----
NM_000275.      CAGCAt-----GgATAtGgGTtCTCcTtCA-----
NM_016180.      CAaCacAgC-----
NM_0010125      CcttggaACGcAa---cTAGTgaACaaGcgCctaTgCAGtgcttataacagaggttggtgt

NM_000372.      -----TaAagAccATTtGCAaaTTgTaaCTaaTAcAA
NM_000550.      -----TCtcAggTTcaAGGCTa---caACaAAAATcacCATtCTTgTcaAactTTgGAg
NM_000273.      -----AGAA
NM_0013009      -----TgGAAattTTTCAGGCTgGGCTTcCCAAatgaTGGttgtGTcCTgCACTGTTgGgA
NM_000275.      -----TgGAAattTTTCAGGCTgGGCTTcCCAAatgaTGGttgtGTcCTgCACTGTTgGgA
NM_016180.      -----
NM_0010125      ttccacTCtAAaaTTaCcacCTcttCTTAtCAAAATTtGCAActTTCTtCctgtgaAaAt

NM_000372.      AGtGTAgcCTtCtTCcAaCTCagGtAG-----
NM_000550.      AGGGaaAatcttCaCTttCTtAaGCAacaATGGA-----tatTg
NM_000273.      tGgaagAgCcCCTCAGacCActC-----
NM_0013009      tGtGTTATCTCCTtGtGtGCTCAtGtgGTGgTGGgATGGAATTAATAGACATCCATCtATT
NM_000275.      tGtGTTATCTCCTtGtGtGCTCAtGtgGTGgTGGgATGGAATTAATAGACATCCATCtATT
NM_016180.      -----
NM_0010125      AGGcTTATagttCctTAACTggaGCAaTGATatAAaaGctacAgagataActCcaagtTT

NM_000372.      -----aAcAcAcCTgtCtTtGt
NM_000550.      cCTgtgtttgccActGtggtTcCtGcTCtCAAattCgCTgAAaAaggAACTACctatcC
NM_000273.      -----TAcAgCTgCtCTaGC
NM_0013009      gCTCgAAgACTaAAGGAAACTtCATCCATCAACAACCAAtTAgtCATAAaAACTACCCTgaC
NM_000275.      gCTCgAAgACTaAAGGAAACTtCATCCATCAACAACCAAtTAgtCATAAaAACTACCCTgaC
NM_016180.      -----cgggACcgttgTcGt
NM_0010125      aaTCcAgctCagcAtGAAACTgCA-----AAaaCACTgAACATAtAACcAtCagact

NM_000372.      CTTgCTGTT-----ttcActcAgccctTtTAacAttTtccCctAaGCcCA-----
NM_000550.      tTacaTtTc-----AcctAcTAaTgtCTCTtctAacatcttAGAGgtcCA-----
NM_000273.      CTTAgTtgc-----CaCTAGgAAgTtTtCtGAGGctgg-----
NM_0013009      CccACTGTTTgAAGAAGAAAAggTgcTTaCcCTgGAgAtgCTaCAGAGaCACA----GTG
NM_000275.      CccACTGTTTgAAGAAGAAAAggTgcTTaCcCTgGAgAtgCTaCAGAGaCACA----GTG
NM_016180.      CgTggTGaT-----cAcAgcgTCTgCgGtGGCACt-----
NM_0010125      aTaAtTcTggccAtgttttttAtcAtTTgaTCTgGcAgcagggCAGAGgtgCAccagGaG

NM_000372.      -----
NM_000550.      -----
NM_000273.      -----
NM_0013009      GAATAGACCTTgACACTAACACTCTAATTCAAAGCgAATgTTGGAACACCATGACCTCCTC
NM_000275.      GAATAGACCTTgACACTAACACTCTAATTCAAAGCgAATgTTGGAACACCATGACCTCCTC
NM_016180.      -----
NM_0010125      tAAaAacCtTTcAaAaaAgaggTgTggcTttcctcAtcccctGgcaAaCATtgtCaCCaa

NM_000372.      -----TATgtCTAAGGAaAggATG
NM_000550.      -----tggAgaAggCaTATGGAgAAcATG
NM_000273.      -----CTgtaaAgtAAgTG
NM_0013009      TGTGTGTCCCTT-----TcTCCCcAAGGAcAAAAATG
NM_000275.      TGTGTGTCCCTT-----TcTCCCcAAGGAcAAAAATG
NM_016180.      -----GatAggcTG
NM_0010125      gGTGgGatgTTgctgctaattgggagagtcataaaaaAcTgTCCCTAcaaggAgAgTG

NM_000372.      cTAttTGgTaaTgaggAACTgtTaTt-----TgtaTgTGaaTTA
NM_000550.      TTttATacTGcTctATAAaTAGTaTt-----CcaAtcAcTgtGctTaaTTTA
NM_000273.      T-----
NM_0013009      TagAAaGATGTgAGATAACT--TACTCAAGATTCCCTCCAGAAAAATACGTaTGtTTaA

```

NM_000275. TagAAaGATGTgAGATAACT--TACTCAAGATTCCCTCCAGAAAAATACGTaTgTtTaA
NM_016180. T-----TGcTTTg
NM_0010125 TTAcATGgTGTtGgGgAagAaaACctAcGcTTCaggctCAGAgAA----accTGgaATTc

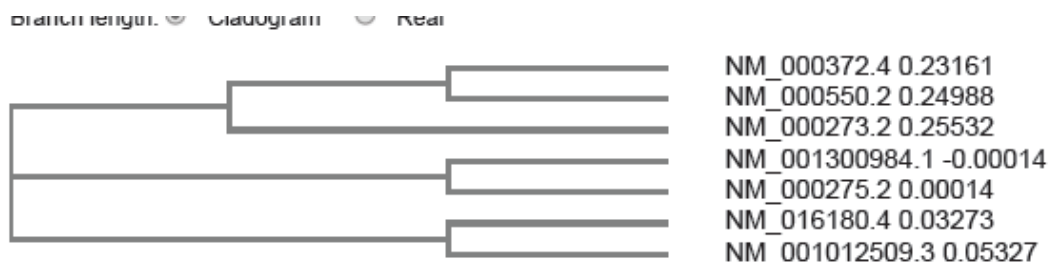
NM_000372. AAgtgCTcttatTT-----tAAAAAATtgAaAT
NM_000550. AAtagCaTtaTcTTATcatTtatcAgccttttAtgTatttttCaAgtAAAAATtAAcAT
NM_000273. -----
NM_0013009 AAACCCTTCCTgcTATACATAGGAAaAGAcACACAT-----CCACctAAAAATGAActgT
NM_000275. AAACCCTTCCTgcTATACATAGGAAaAGAcACACAT-----CCACctAAAAATGAActgT
NM_016180. tcgCtCTcttTgTTAgA-----
NM_0010125 AAAcCCTgCCTcTctcActTgGaAgctGtggttACAc-----TGgggAT

NM_000372. AaTtTTgAtTTtTgc-----cTtcTgaTTAttTaAaGaTcTAT
NM_000550. AtTATTTcATTGgtc-----TTCtTtTTTATCTG--GtTcTAT
NM_000273. -----AaGgTccAc
NM_0013009 ACTgTTTAAcTGTcA-----ATTctcCTGAGGcTAaAc
NM_000275. ACTgTTTAAcTGTcA-----ATTctcCTGAGGcTAaAc
NM_016180. -----TAT
NM_0010125 gtTATTTAcTTGTttattagtttagcacttagaatgagTTCaTtTTTAaggTGAGataATAT

NM_000372. ATAtgT-----TTTATTggcCCctT
NM_000550. ATgaAT-----gcTATTTTTtCCcT
NM_000273. ATcctT-----
NM_0013009 AcAgtT-----TgTTTTTCTtgT
NM_000275. AcAgtT-----TgTTTTTCTtgT
NM_016180. gTggAT-----
NM_0010125 ATtcAcaatgatacacaaaagtgcacactgtgtgctccctttaccaTTTAcTaTctCCcT

NM_000372. ctT-----
NM_000550. tcTCttCTAACATGaaAtAtattTtctCtTTttgATctTgTGCTaT--GAAacaAtctTc
NM_000273. -----
NM_0013009 AATCACtTttCATGTTAAAATAATCAgCATTCAAATtGTATGCTTCTGAATATAGACTT
NM_000275. AATCACtTttCATGTTAAAATAATCAgCATTCAAATtGTATGCTTCTGAATATAGACTT
NM_016180. -----
NM_0010125 ccagcCCcAACgTGTTAccActcTtActAgTttttTgtTAatCTTTC-----
NM_000372. -----TAttTtAATAAAAcAgtgA-----gAAATc-----
NM_000550. caaaGaActgtaTaaggTGGTCaTAAGTgAATAtttTAatTA-----AAAtTG-----
NM_000273. -----ggGGaaGTAgTAAAAATAAAATAgTTA-----tgACTG-----
NM_0013009 TCTGGGAAAAGGTTTACTGcTCGTAAGgAAAcAtttTatgTATTAAAAATAAACTGTTTCCT
NM_000275. TCTGGGAAAAGGTTTACTGcTCGTAAGgAAAcAtttTatgTATTAAAAATAAACTGTTTCCT
NM_016180. -----TAGGTcAATAAAgagacaA-----tgACccTaaCc
NM_0010125 -----GgCaTtttTAAATAAAaccaaag-----cAAATG-----

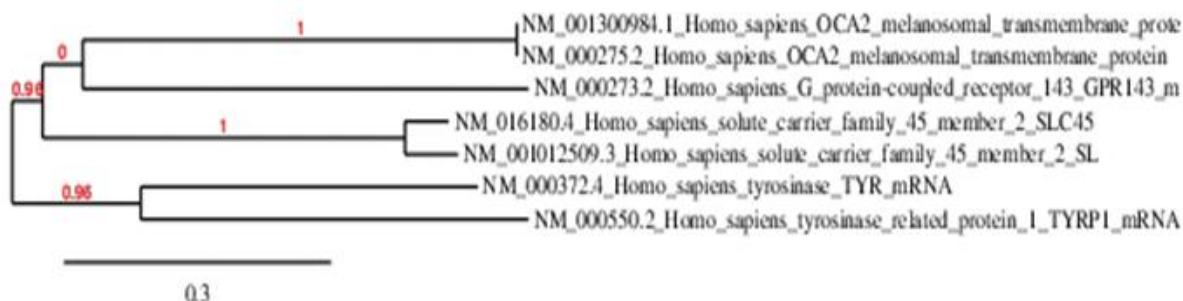
NM_000372. ---TAAAAAAAAAAAAAAAAA
NM_000550. ---gtAAAAAtAAAtAAtAAcA
NM_000273. ---aAAAAAAAAAAAAAAAAA---
NM_0013009 TgATAAAAAAAAAAAAAAAAAA
NM_000275. TgATAAAAAAAAAAAAAAAAAA
NM_016180. TcAgAAAAAAAAAAAAAAAAA
NM_0010125 ---TAAAAAAAAAAAAAAAAA---



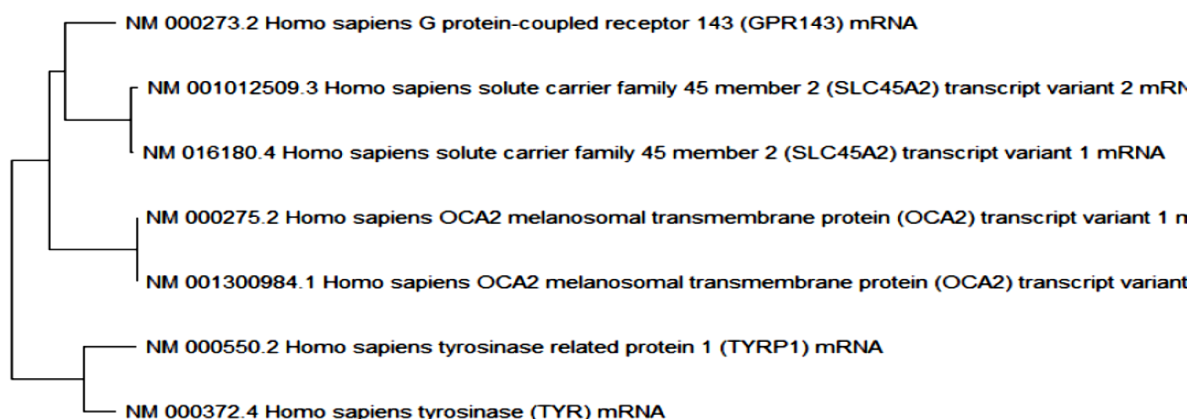
Tree Data

```
(
(
(
NM_000372.4:0.23161,
NM_000550.2:0.24988)
:0.04651,
NM_000273.2:0.25532)
:0.00514,
(
(
NM_001300984.1:-0.00014,
NM_000275.2:0.00014)
:0.27160,
(
NM_016180.4:0.03273,
NM_001012509.3:0.05327)
:0.22222);
```

PhyML RESULTS



Mega X RESULTS

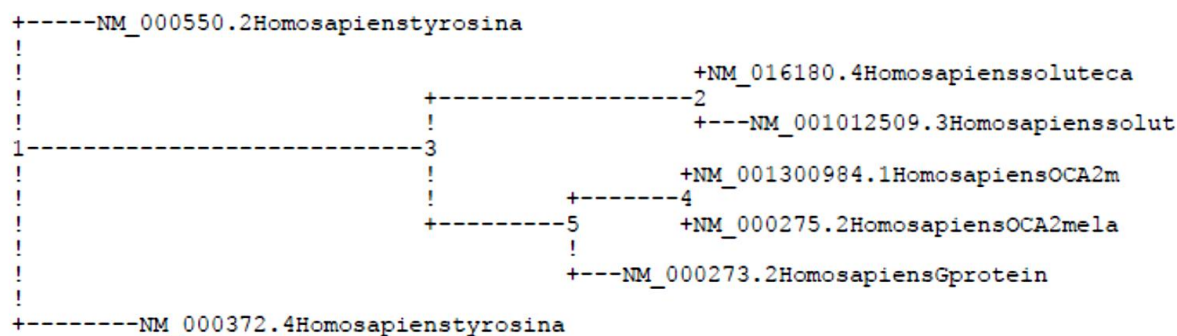


7 Populations

Neighbor-Joining/UPGMA method version 3.6a2.1

Neighbor-joining method

Negative branch lengths allowed



remember: this is an unrooted tree!

Between	And	Length	
1	NM_000550.2	Homo sapiens tyrosinase	0.32134
1	3	1.45397	
3	2	0.93395	
2	NM_016180.4	Homo sapiens solute carrier family 12 member 1	-0.13791
2	NM_001012509.3	Homo sapiens solute carrier family 12 member 1	0.20801
3	5	0.47644	
5	4	0.41271	
4	NM_001300984.1	Homo sapiens OCA2	-0.00053
4	NM_000275.2	Homo sapiens OCA2	0.00053
5	NM_000273.2	Homo sapiens G protein-coupled receptor 1	0.19789
1	NM_000372.4	Homo sapiens tyrosinase	0.47246

Albinism is a typical genetic disease. The only cause for this disease is the changed genetic code of the genes. This is the major reasons why till date no medicine for disease has been developed. The study of its genetic code in humans and other related species with its similarity to the related genes may help us know more details about the disease and in the development of genetic medicine or treatment of these diseases. Genetic treatment is the only way to cure these genetic diseases [5].

equivalent genes coming from several species for reconstructing the genealogic tree of the genes and finding out who is the closest relative of whom in the family. The phylogenetic analysis of the genes reveals the relationship of the genes and their variants responsible for phenylketonuria and hyperphenylalanemia. The purpose of the phylogeny is to study how the environmental conditions and other factors have led to the diversification of the sequences [6-9]. The studies also help us to analyse where mutation have taken place in the nucleotide sequences and thus leading to change in the amino acids. This analysis groups the sequences according to level of their similarity.

As the first step towards phylogenetic analysis the genes and their variants responsible for the cause of albinism have been identified. The genes OCA1, OCA2, OCA3, OCA 4, OA1 and its variants were first identified and selected from NCBI. The similarity of these sequences in related organisms has been studied by using BLAST for the selected gene sequences. The similar sequences responsible for the concerned disease are selected for the further analysis and constructing their phylogeny.

The further analysis of the genes was done by conducting multiple sequence alignment of the selected sequences using bioinformatics tools like Clustal Omega, MUSCLE and Mega X. These alignments showed the pattern of how the gene codes evolved or changed by mutation with time and other environmental factors.

To study the phylogenetic relationship of the genes responsible for oculocutaneous and ocular albinism and to check the validity of the results received phylogenetic trees have been generated using various bioinformatics tools like PHYLIP, MEGA X, BioEdit. The results obtained show that evolution and time all the responsible genes have undergone change in their genetic code due to mutations. These mutations may have been caused by the changing environmental conditions, the living and the food habits. All the sequences analysed show significant divergence from the node of the tree. All the sequences OCA1, OCA2, OCA3, OCA 4, OA1 and its variants have shown great similarity with each other. The OCA2 and the OA1 genes have shown the maximum divergence from the tree node showing that these genes have undergone the maximum number of mutations with time. The present phylogenetic studies have opened a new path for the further analysis of the gene sequences and to develop new genetic

medicines for Albinism and its various types.

REFERENCES

1. Sun W, Shen Y, Shan S, Han L, Li Y, Zhou Z, Zhong Z, Chen J Identification of TYR mutations in patients with oculocutaneous albinism. *Mol Med Rep.* 2018 Jun;17(6):8409-8413. doi: 10.3892/mmr.2018.8881. Epub 2018 Apr 13.
2. Markus N. Preising,^{1, 2} Hedwig Forster,¹ Miriam Gonser,¹ Birgit Lorenz^{1,2}; *Molecular Vision* 2011; 17:939-948
<http://www.molvis.org/molvis/v17/a104> Received 17 December 2011 | Accepted 5 April 2011 | Published 15 April 2011.
3. KAZUTOMO TOYOFUKU, IKUO WADA,* JULIO C. VALENCIA,† TSUNETO KUSHIMOTO, VICTOR J. FERRANS,† AND VINCENT J. HEARING¹, Laboratory of Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA; *Department of Biochemistry, Sapporo Medical University, Sapporo, Japan; and †Pathology Section, National Oculocutaneous albinism types 1 and 3 are ER retention diseases: mutation of tyrosinase or Tyrp1 can affect the processing of both mutant and wildtype proteins, 0892-6638/01/0015-2149 © FASEB Vol. 15 October 2001.
4. Karen Grønskov, Jakob Ek and Karen Brøndum-Nielsen* Address: Kennedy Center. National Research Center for Genetics, visual Impairment and Mental Retardation, Gl. Landevej 7, 2600 Glostrup, Denmark. Review Open Access Oculocutaneous albinism* Corresponding author Published: 2 November 2007 *Orphanet Journal of Rare Diseases* 2007, 2:43 doi:10.1186/1750-1172-2-43.

5. Kazutomo TOYOFUKU*, Ikuo WADA., Richard A. SPRITZ. and Vincent J. HEARING*¹*Laboratory of Cell Biology, National Cancer Institute, National Institutes of Health, Building 37, Room 1B25, Bethesda, MD 20892, U.S.A., .Department of Biochemistry, Sapporo Medical University, Chu-ou-ku, Minami 1, Nishi 17, Sapporo, Hokkaido 060-8556, Japan, and .Human Medical Genetics Program, University of Colorado Health Sciences Center, Denver, CO 80262, U.S.A. The molecular basis of oculocutaneous albinism type 1 (OCA1): sorting failure and degradation of mutant tyrosinases results in a lack of pigmentation, *Biochem. J.* (2001) 355, 259±269
6. Ray K ,Chaki M, Sengupta M. Tyrosinase and ocular diseases: some novel thoughts on the molecular basis of oculocutaneous albinism type 1. *Prog Retin Eye Res.* 2007 Jul;26(4):323-58. Epub 2007 Jan 17.
7. Richard Alan Lewis, MD, MS Departments of Ophthalmology, Medicine, Pediatrics, and Molecular and Human Genetics /Huffington Center on Aging Cullen Eye Institute Baylor College of Medicine Houston, Texas. Ocular Albinism, X-Linked Synonyms: Nettleship-Falls Ocular Albinism, OA1, Ocular Albinism Type 1, XLOA. Initial Posting: March 12, 2004; Last Update: April 5, 2011.
8. Tamio Suzuki, MD, PhD and Masahiro Hayashi, MD, PhD. Oculocutaneous Albinism Type 4. Synonym: OCA4, Initial Posting: November 17, 2005; Last Revision: September 15, 2011.
9. Richard Alan Lewis, MD, MS, Cullen Eye Institute, Baylor College of Medicine, Houston, Texas. Oculocutaneous Albinism Type 2 Synonym: OCA2. Includes: Brown OCA. Initial Posting: July 17, 2003; Last Update: August 16, 2012.

Acknowledgement: I would take this opportunity to extend my heartiest thanks to my reverent guide, and management for their genuine support and kind help without which this study would not have completed.