

University of Arkansas, Fayetteville

ScholarWorks@UARK

Biological Sciences Undergraduate Honors
Theses

Biological Sciences

5-2015

Evolutionary analysis of the genomic region of FOXP3 in vertebrates

Ethan G. Guffey

University of Arkansas, Fayetteville

Follow this and additional works at: <https://scholarworks.uark.edu/biscuht>

Citation

Guffey, E. G. (2015). Evolutionary analysis of the genomic region of FOXP3 in vertebrates. *Biological Sciences Undergraduate Honors Theses* Retrieved from <https://scholarworks.uark.edu/biscuht/4>

This Thesis is brought to you for free and open access by the Biological Sciences at ScholarWorks@UARK. It has been accepted for inclusion in Biological Sciences Undergraduate Honors Theses by an authorized administrator of ScholarWorks@UARK. For more information, please contact scholar@uark.edu.

**Evolutionary Analysis of the Genomic Region of FOXP3 in
Vertebrates**

An Honors Thesis submitted in partial fulfillment of the requirements of
Honors Studies in Biology

By

Ethan Guffey

Spring 2015

Biology

J. William Fulbright College of Arts and Sciences

The University of Arkansas

Table of Contents

Section	Page Number
Abstract	3
Introduction	4
Results	6
Discussion	10
References	14
Appendix	16

Abstract

The Human FOXP3 gene plays a role in immunosuppression, which includes the suppression of the maternal immune system to allow the growth of a semi-foreign fetus inside her uterus. This has lead scientists to hypothesize that FOXP3 gene played a crucial role in the development of viviparous species. FOXP3 is highly conserved across Eutherian mammals, but the degree of conservation in other viviparous species has not been investigated. This study found that FOXP3 orthologs are much less highly conserved in aplacental species and may be more closely related to other members of the FOXP gene family than to the FOXP3 found in mammals. Through the creation of phylogenetic trees, I found that all of the members of the Forkhead Box Protein Family were named correctly, but FOXP3 had the highest degree of substitutions, which may explain its relatively novel function in immunosuppression.

Keywords: FOXP3, forkhead box protein, viviparity, Eutheria, immunosuppression, orthologs

The Forkhead Box P3 gene, or FOXP3, belongs to a larger group called the forkhead box winged helix and more specifically the FOXP subfamily including FOXP1, FOXP2, FOXP3, and FOXP54, which are a group of transcription factors grouped together based DNA binding domain structure. The FOXP3 gene produces a protein by the same name that works as a transcription factor and serves as a “master regulator” in the immune system most importantly for CD4+ CD25+ T regulatory cell (T-reg) development and function (Andersen, Nissen, & Betz, 2012). It’s role as master regulator means that it works on multiple genes that activate or inhibit immune function given certain conditions. For example, it is important in preventing autoimmunity, or self-tolerance, by deleting self-reactive T cells in the thymus, which is the site of T-reg development (Samstein, Josefewicz, Arvey, Treuting, & Rudensky, 2012).

An additional function of the FOXP3 protein as the master regulator is the response to non-self antigens. During pregnancy, the mother constantly comes into contact with paternal antigens expressed by the fetus through the maternal-fetal barrier, known as the placenta. These non-self antigens trigger a response in the immune system of the mother that may result in a termination of the pregnancy. Obviously, termination of pregnancy would not allow for propagation of the species and the proliferation of the parental gene pool. For placental pregnancy to be viable, a mechanism had to evolve to allow maternal-fetal interaction and suppress these immune responses, and FOXP3 mediates this interaction (Andersen et al., 2012). Samstein et al. (2012) found that when female mice deficient in CNS1, an enhancer of FOXP3, were mated with males that were antigenically different from the female, they showed increased rates of spontaneous fetal absorption while the CNS1 positive individuals did not. This shows that sufficient expression of the FOXP3 gene plays a role in reducing the immune response of the mother to the growing fetus in the uterus.

FOXP3 is just one among many genes that have been recruited over the course of evolutionary time to slowly give rise to placentals. In mammalia major transitions exist in the form evolutionary novel phenotypes that are observable examples of the development of placentation (Lynch et al., 2015). For example, the monotremes are mammals that grow embryos in utero for 10 days transferring nutrients via a yolk-like placenta before laying thin, poorly mineralized eggs that hatch within 10 to 20 days. The next group of mammals appears to have embellished the yolk-sac into a structure more recognizable as the placenta. This evolutionary advancement marks the rise of viviparity, or live-birth, in a group called Theria, which includes both marsupials and Eutherians or “placental” mammals. While viviparity evolved only once in Eutherians, interestingly evidence suggests that it evolve over 100 times in squamous reptile species including lizards and snakes (Dykes & Beaupre, 2012). Marsupials have relatively short pregnancies of around 25 days, which is shorter than one estrous cycle. Placentals have evolved much longer pregnancies lasting on average 131 days but as much as 670 days. These longer pregnancies led to differentiation of cells in the uterus and induced large-scale gene regulation, which includes maternal immunotolerance of the semi-foreign fetus (Lynch et al., 2015; Murphy, Thompson, & Belov, 2009).

Due to the unique function of the FOXP3 gene in viviparity, studies have suggested that FOXP3 is only found in placentals. According to Andersen et al. (2012), the forkhead box subfamily can only be found in jawed vertebrates. They went further to say that FOXP3-like genes could be found in bony fish, amphibians, and reptiles, but not in bird species. Samstein et al. (2012) found that the FOXP3 enhancer, CNS1, may have played a role in the development of viviparity in mammals through regulation of the FOXP3 gene.

Andersen et al. (2012) specifically looked at the domains that make up FOXP3. These regions evolved in a step-wise fashion that allowed for the rise of placentals. They found that all

mammals contained the proline rich (ProR) region and the forkhead domain that functions in nuclear localization and DNA binding. They hypothesized that the very N-terminal tail of the ProR region, which was only found in placentals, may have increased the evolutionary differentiation from other FOXP3s because it may aid in binding to factors specifically associated with T cell activation.

The goal of this study was to analyze the genomic region surrounding FOXP3 in mammals and use that information to explore the genomes of mammals and other animals to find FOXP3 or FOXP3-like genes to determine its evolutionary origins, but the goal evolved into a search for orthologs of FOXP3 in other species in which it has not previously been found.

Results

FOXP3

In order to discover the origins of the FOXP3 gene and its role in placentation, sequences across all species included on NCBI blast were analyzed to determine which were most related to human FOXP3. A preliminary blastn of the NCBI NR nucleotide database with human FOXP3 DNA sequence (Accession: EF534714.1) returned different isoforms of human FOXP3. A follow-up blastn search was conducted excluding the human taxid to remove other isoforms of the same gene. The results were all FOXP3 isoforms of mammalian species.

To increase the likelihood of more comprehensive findings, blastp was used for the remainder of this study rather than blastn because amino acid sequences are more likely to be highly conserved across species than nucleotide sequences due to the degenerate nature of the genetic code. Additionally, substitution of an amino acid with similar properties would lead to

similar protein structures. For example, a non-polar amino acid such as alanine would likely offer very little change if substituted for another non-polar amino acid such as leucine. The same can be said for other groups of amino acids with acidic and basic sidechains.

A blastp of the NCBI protein database using the human FOXP3 protein (Accession: ABQ15210.1) showed high percent identity and high query coverage among mammals with the most highly conserved being among primate species including chimpanzee (*Pan troglodytes*), bonobo (*Pan paniscus*), gorilla (*Gorilla gorilla*), and macaque (*Macaca fascicularis*, Figure 1). All four of these species have a 99% identity and 100% query coverage, which suggests very similar sequences between these species and human FOXP3. Since the overall genomes of chimpanzee and these other species are the most closely related genomes to humans, these results seem reasonable. Cow (*Bos taurus*), cat (*Felis catus*), camel (*Camelus ferus*), armadillo (*Dasypus novemcinctus*), rhinoceros (*Ceratotherium simum*), and horse (*Equus caballus*) are all non-primates but have 100% query coverage and at least 90% identity. However, all of these species are mammals. It is likely that mammals would share more closely related sequences to each other than to non-mammalian vertebrates.

Figure 2 shows a visual representation of the blastp results with some additional information about taxonomy. All of these species mentioned above belong to a larger group called Eutheria, more commonly known as placentals. Considering the function of the products of the FOXP3 gene in live-bearing mothers, it stands to reason that this gene would be highly conserved across placentals. It is probable that the advantageous mutation occurred in a common ancestor of placentals that allowed for the successful development of viviparous species, because of the function of the FOXP3 gene in immunosuppression of the mother. The highly conserved nature of the FOXP3 gene across mammals leads to looking outside this group into other species including amphibians, fish, and birds to determine whether these contain

FOXP3. If a FOXP3 appears in these species but in a much less highly conserved form, it may suggest that the mutation entered the genome in a more distant common ancestor than previously thought.

A blastp of the NCBI protein database search was completed using human FOXP3 as the query but excluding the entire taxid of mammals. This allowed a closer look at species believed by previous studies not to contain FOXP3. The results showed much lower query coverage and percent identity than the search including mammals (Figure 3). These results include genes named for other members of the forkhead box family including FOXP1 and FOXP2 in addition to a select few named FOXP3.

In contrast with Figure 1, the results including non-mammalian vertebrates (notably they are aplacental) show query coverage from 50-60% and percent identity from 40-50%. Since the gene products of FOXP3 do not have a known equivalent in aplacental vertebrates, there appears to be a lack of fitness advantage and therefore no selection in these species, which may account for the varying degree of amino acid substitutions in non-Eutherian species. This suggests that the fitness advantage of FOXP3 in placentals serves to support the high similarity of the gene across the majority of mammals.

Forkhead Box Protein Family

Aside from the low coverage and identity, FOXP3 appears to be largely non-existent outside of mammalian species with a few exceptions. When FOXP3 is conserved in non-mammalian vertebrates, as in the case of some frogs (*Xenopus tropicalis*) and fish (*Danio rerio*), it is much less highly conserved than in mammals. Other than those few species where there is a counterpart, FOXP3 in fish, reptiles, and birds appears to be much more similar to other members of the forkhead box protein family, FOXP1, FOXP2, and FOXP4, than to forms of FOXP3

in other species as shown in Figure 3. We then considered whether some of the genes in these other lineages might be misnamed, which might account for “absence” of the gene named, FOXP3 in other species.

To determine whether the members of the forkhead box family are named properly, each member of the forkhead box family was analyzed, and a chromosomal region was obtained including each gene and their adjacent genes. It is likely that adjacent genes either upstream or downstream of the gene in question are likely to be linked, meaning that these genes are very unlikely to separate over the course of time in a species’ genome. Throughout the course of evolutionary time, these genes are likely to remain in proximity even if that region is translocated to another chromosome.

I therefore considered whether I could identify forkhead protein orthologs by their gene neighborhood. The hypothesis was that the true ortholog of FOXP3 would have the some of the same proximal genes in different lineages. The region of interest, or the FOXP3 gene region, is located on the X chromosome in humans. According to figure 4, genes within close proximity to FOXP3 are called PPP1R3F, which is a member of the protein phosphatase 1 family, and CCDC22, which is part of the coiled-coil domain family. In addition to FOXP3, the gene neighborhoods of FOXP1 (Figure 5), FOXP2 (Figure 6), and FOXP4 (Figure 7) were analyzed as well in case they are misnamed in other species. EIF4E3 is in close proximity to FOXP1. FOXP2 has two genes nearby called PPP1R3A and MDFIC. MDFI is in close proximity to FOXP4.

I used the MultiAlign program in the LaserGene software package (DNASStar, Madison, WI) to generate sequence alignments and distance trees for human (*Homo sapien*), mouse (*Mus musculus*), Chicken (*Gallus gallus*), American alligator (*Alligator mississippiensis*), western painted turtle (*Chrysemys picta bellii*), cobra (*Ophiophagus hannah*), python (*Python bivittatus*), and

western clawed frog (*Xenopus tropicalis*) to analyze the gene families of Forkhead Box Protein (FOXP_, Figure 8), Protein Phosphatase 1 (PPP1R3_, Figure 9), Coiled-Coil Domain Containing (CCDC_, Figure 10), and MyoD Family Inhibitor (MDFI, Figure 11). From Figure 8 the forkhead box protein family members are different from one another meaning they are evolutionarily distinct. A closer look at the forkhead box protein family tree reveals that clades of FOXP1, FOXP2, and FOXP4 differ by fewer than 30 amino acid substitutions. However, FOXP3 has many more substitutions with the fewest being 60 substitutions in both cobra (*Ophiophagus hannah*) and python (*Python bivittatus*). According to these trees, each separate gene forms its own clade, which implies that all of the genes are named correctly in the NCBI blast database.

The inconsistent nature of the amino acid sequence for FOXP3, as evidenced by the high degree of substitutions (Figure 8) and the low percent identity and query coverage (Figure 3), shows that this gene is the least highly conserved of the forkhead box family across species and possibly the newest paralog of this family that arose from another member of this family. In terms of evolution, FOXP3 appears to be the most diverse gene within the forkhead box protein family, which might suggest the reason for the relatively novel function in the gene of this family that may have aided in the rise of placentals and viviparity.

Discussion

The FOXP3 protein is highly conserved across mammalian species with highest identity and coverage among the various primate species. The FOXP3 gene is the most diverse paralog of the forkhead box protein family and, from these results, appears to be highly conserved in placentals with few clear orthologs in non-placental species. Representative orthologs could be

identified in some species of frogs and snakes. FOXP3 has more substitutions than any other member of this family, which may explain the reason for the relatively diverse gene products.

Interestingly there is a commonality between genes proximal to FOXP2 and those of FOXP3 and FOXP4. Both FOXP2 and FOXP3 regions have neighboring genes from the protein phosphatase 1 family (PPP1R3A and PPP1R3F, respectively). Also the FOXP2 region shares a gene neighbor with the FOXP4 family MyoD Family Inhibitor (MDFIC and MDFI, respectively). Because of the gene maps of the forkhead box family (Figures 4-7), FOXP2 appears to have the most similarity to the other family members. For example, the FOXP2 gene region includes genes from the Protein Phosphatase 1 family, which is also in the FOXP3 region, and the MyoD Inhibitor family, which is also in the FOXP4 region.

It is possible that the forkhead box family arose in a paralogous nature in which a whole or partial genome duplication gave rise to duplicate copies of the original ancestor of this gene family. The commonly shared genes in the region around FOXP2 with other forkhead box members may be evidence that FOXP2 is closest to the ancestor of this family, but it has since undergone its own evolutionary divergence. Of course more research needs to be conducted on all members of the forkhead box protein family with specific focus on the regions around the genes and how they relate to each other to determine the validity of this conjecture of the paralogous nature of the rise members of this family.

Due to the interests of my research laboratory, I specifically searched the chicken (*Gallus gallus*) genome for sequences similar to human FOXP3 but no such gene was found. The result of all searches of the chicken genome using the human FOXP3 as query resulted in different isoforms of FOXP2. After the majority of research was completed for this project, it was brought to my attention that the chicken genome is not complete. This genome is missing

sequences for chromosome 16 and many of the microchromosomes. It is possible that a FOXP3 ortholog does exist in chicken but that the region containing it has not yet been sequenced. This revealing gap in the knowledge about this genome brings concern to other genomes as well. During my research, I was unaware of these gaps, so future research should be done to determine which genomes used in these findings contain sequenced chromosomes and regions. Additionally, new methods of sequencing should also be researched allow the possibility to fill in these gaps.

Due to the function of the FOXP3 protein in suppressing the immune system, additional research may have practical applications in medicine. The autoimmune disorder IPEX syndrome, or Immunodysregulation Polyendocrinopathy Enteropathy X-linked syndrome, results from mutations in FOXP3 that result in non-conservative substitutions rendering the protein ineffective by changing the shape of the protein's structure or prematurely stopping translation creating a shorter, incomplete protein (Passerini, Santoni, Roncarolo, & Bacchetta, 2014). This ineffective or incomplete protein will not be able to bind correctly to DNA and serve as an enhancer for genes that lead to CD4+ CD25+ T regulator cell production. As a result there will be little or no regulatory response due to lower serum levels of T-regulator cells in the blood and no way to suppress immune responses including those against the individual's own body. This is known to cause diseases such as insulin dependent diabetes, eczema, and food allergies. This is a recessive, X-linked trait, therefore males are more commonly affected than females, which are more commonly carriers. Research into the origin of this gene may offer insight into the structure that may lead to potential gene therapies that bolster the levels of T-reg cells when needed to reduce the immune responses and promote self-tolerance.

Similarly women who have difficulty carrying their pregnancy to term as a result of spontaneous abortions may benefit from this research. Pregnant women generally have

elevated levels of T-reg cells in their blood than women who are not pregnant. However, this elevated level of T-reg cells was not found in women experiencing unexplained recurrent spontaneous abortions (URSA), which is defined as two or more consecutive losses of pregnancy before twenty weeks (Zaigui et al., 2012). With reduced T-reg counts, these women are unable to suppress the immune responses triggered by the introduction of the paternal antigens by the fetus. The lower level of T-reg in the blood of URSA patients is likely caused by reduced expression or functionality of FOXP3. Again it appears that mutations in FOXP3 are the cause of reduced functionality. Research needs to be done to determine more specifically which regions of this gene are prone to mutations that results in undesired loss of pregnancy in these women and to determine if any form of therapy can be implemented to raise the T-reg cell counts of these women during pregnancy to prevent spontaneous abortions. With continued research into the origins and sequence variations of FOXP3, it may be possible to increase the quality of life in people affected by mutations that result in deficient forms of this immunologically significant protein.

References

- Andersen, K. G., Nissen, J. K., & Betz, A. G. (2012). Comparative Genomics Reveals Key Gain-of-Function Events in Foxp3 during Regulatory T Cell Evolution. *Frontiers In Immunology*, 3113. doi:10.3389/fimmu.2012.00113
- Lynch, V. J., Nnamani, M. C., Kapusta, A., Brayer, K., Plaza, S. L., Mazur, E. C., & ... Wagner, G. P. (2015). Ancient Transposable Elements Transformed the Uterine Regulatory Landscape and Transcriptome during the Evolution of Mammalian Pregnancy. *Cell Reports*, doi:10.1016/j.celrep.2014.12.052
- Murphy, B. F., Belov, K., & Thompson, M. B. (2010). Evolution of viviparity and uterine angiogenesis: vascular endothelial growth factor (VEGF) in oviparous and viviparous skinks. *Journal Of Experimental Zoology. Part B, Molecular And Developmental Evolution*, 314(2), 148-156. doi:10.1002/jez.b.21317
- O'Leary, M. On the Trail of the First Placental Mammals. *American Scientist* [serial online]. May 2014; 102(3):190-197. Available from: Academic Search Complete, Ipswich, MA. Accessed March 18, 2015.
- Passerini, L., Santoni de Sio, F. R., Roncarolo, M. G., & Bacchetta, R. (2014). Forkhead box P3: The Peacekeeper of the Immune System. *International Reviews Of Immunology*, 33(2), 129-145. doi:10.3109/08830185.2013.863303
- Samstein, R. M., Josefowicz, S. Z., Arvey, A., Treuting, P. M., & Rudensky, A. Y. (2012). Extrathymic generation of regulatory T cells in placental mammals mitigates maternal-fetal conflict. *Cell*, 150(1), 29-38. doi:10.1016/j.cell.2012.05.031

Van Dyke, J. U., & Beaupre, S. J. (2012). Stable isotope tracer reveals that viviparous snakes transport amino acids to offspring during gestation. *The Journal Of Experimental Biology*, 215(Pt 5), 760-765. doi:10.1242/jeb.058644

Zaigui, W., Zeshan, Y., Cai, Z., Zhuyu, L., Xiumei, S., Xiuming, Z., & Yinguang, L. (2012). Association between Functional Polymorphisms of Foxp3 Gene and the Occurrence of Unexplained Recurrent Spontaneous Abortion in a Chinese Han Population. *Clinical & Developmental Immunology*, 1-7. doi:10.1155/2012/896458

Appendix

	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X1 (Homo sapiens)	884	884	100%	0.0	100%	XP_005272667.2
<input type="checkbox"/>	forkhead box protein P3 isoform a (Homo sapiens)	879	879	100%	0.0	100%	NP_054728.2
<input type="checkbox"/>	unnamed protein product (Homo sapiens)	878	878	100%	0.0	99%	BAF84741.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 (Pan troglodytes)	876	876	100%	0.0	99%	XP_009437357.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X1 (Pan paniscus)	874	874	100%	0.0	99%	XP_003807307.2
<input type="checkbox"/>	Foxp3 protein (Pan troglodytes)	874	874	100%	0.0	99%	AGK90310.1
<input type="checkbox"/>	Foxp3 protein (Gorilla gorilla)	868	868	100%	0.0	99%	AGK90312.1
<input type="checkbox"/>	RefName: Full=Foxhead box protein P3 (Macaca fascicularis)	868	868	100%	0.0	99%	G6UB07.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X3 (Macaca fascicularis)	871	871	100%	0.0	99%	XP_005593607.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X2 (Rhinopithecus roxellana)	867	867	100%	0.0	98%	XP_010360814.1
<input type="checkbox"/>	Foxp3 protein (Hylobates lar)	866	866	100%	0.0	98%	AGK90311.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X11 (Papio anubis)	870	870	100%	0.0	99%	XP_009195851.1
<input type="checkbox"/>	FOXP3 protein (Homo sapiens)	866	866	100%	0.0	95%	AA43787.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X3 (Homo sapiens)	870	870	100%	0.0	95%	XP_008724596.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X1 (Pongo abelii)	863	863	100%	0.0	98%	XP_009233122.1
<input type="checkbox"/>	forkhead box protein P3 (Macaca mulatta)	859	859	100%	0.0	98%	NP_001028090.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X10 (Chlorocebus sabaeus)	861	861	100%	0.0	98%	XP_007889882.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X7 (Macaca fascicularis)	854	854	100%	0.0	94%	XP_005593611.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X1 (Macaca fascicularis)	857	857	100%	0.0	94%	XP_005593605.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X8 (Papio anubis)	857	857	100%	0.0	94%	XP_009195848.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X1 (Rhinopithecus roxellana)	851	851	100%	0.0	93%	XP_010360913.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X9 (Chlorocebus sabaeus)	848	848	100%	0.0	93%	XP_007889881.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X2 (Saimiri boliviensis boliviensis)	842	842	100%	0.0	95%	XP_010348534.1
<input type="checkbox"/>	forkhead box protein P3 (Callithrix jacchus)	836	836	100%	0.0	96%	NP_001254669.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X4 (Saimiri boliviensis boliviensis)	830	830	100%	0.0	91%	XP_010348536.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X1 (Saimiri boliviensis boliviensis)	829	829	100%	0.0	91%	XP_010348533.1
<input type="checkbox"/>	FOXP3delta7 variant (Homo sapiens)	814	814	100%	0.0	94%	ACJ46653.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X2 (Galeopterus variegatus)	815	815	100%	0.0	92%	XP_008585353.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform 1 (Ceratotherium simum simum)	814	814	100%	0.0	92%	XP_004343973.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X2 (Vicugna pacos)	807	807	99%	0.0	91%	XP_006213548.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 (Cholemur garnettii)	806	806	100%	0.0	90%	XP_003799714.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X2 (Camelus ferus)	804	804	99%	0.0	91%	XP_006195300.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform 1 (Dasypus novemcinctus)	803	803	100%	0.0	91%	XP_004465077.1
<input type="checkbox"/>	forkhead box protein P3 (Equus caballus)	802	802	100%	0.0	91%	NP_001156744.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X4 (Bubalus bubalis)	802	802	100%	0.0	90%	XP_006073710.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X1 (Upotes vexillifer)	802	802	99%	0.0	91%	XP_007450879.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform X2 (Tarsius syrichta)	801	801	100%	0.0	91%	XP_008072845.1
<input type="checkbox"/>	PREDICTED: forkhead box protein P3 isoform 2 (Dasypus novemcinctus)	802	802	100%	0.0	91%	XP_004465078.1
<input type="checkbox"/>	forkhead box protein P3 (Felis catus)	801	801	100%	0.0	91%	NP_001077421.1
<input type="checkbox"/>	forkhead box protein P3 (Bos taurus)	800	800	100%	0.0	90%	NP_001030398.1

Figure 1 Results of Blastp of human FOXP3 (ABQ15210) on NR protein database

Lineage Report			
root			
.. Eutheria [placentals]			
.. .. Boreoeutheria [placentals]			
.. .. Euarctontoglires [placentals]			
.. Primates [primates]			
.. Simiiformes [primates]			
.. Catarrhini [primates]			
.. Hominoidea [primates]			
.. Hominiidae [primates]			
.. Homininae [primates]			
.. Homo sapiens (man)	884	17 hits	[primates]
.. Pan paniscus (bonobo)	877	2 hits	[primates]
.. Pan troglodytes	874	1 hit	[primates]
.. Gorilla gorilla (gorilla)	868	1 hit	[primates]
.. Gorilla gorilla gorilla (lowland gorilla)	781	1 hit	[primates]
.. Pongo abelii (orang utan)	863	3 hits	[primates]
.. Pongo pygmaeus (orang utan)	863	1 hit	[primates]
.. Hylobates lar (white-handed gibbon)	866	1 hit	[primates]
.. Homaeus leucogenys (White-cheeked Gibbon)	760	1 hit	[primates]
.. Macaca fascicularis (long-tailed macaque)	871	9 hits	[primates]
.. Macaca mulatta (rhesus monkeys)	859	2 hits	[primates]
.. Papio anubis (baboon)	782	1 hit	[primates]
.. Callicebus jacchus (white ear-tufted marmoset)	836	2 hits	[primates]
.. Saimiri boliviensis boliviensis	756	1 hit	[primates]
.. Ootolemur garnettii	806	1 hit	[primates]
.. Oryctolagus cuniculus (rabbits)	800	1 hit	[rabbits & hares]
.. Lepus timidus (rodents)	788	2 hits	[rodents]
.. Tupia chinensis	787	2 hits	[placentals]
.. Ochotona princeps (southern American pika)	778	2 hits	[rabbits & hares]
.. Heterocephalus glaber (naked mole rat)	777	3 hits	[rodents]
.. Microtus ochropaster (prairie voles)	776	1 hit	[rodents]
.. Cricetus griseus (Chinese hamsters)	774	3 hits	[rodents]
.. Mesocricetus auratus (Syrian hamsters)	768	1 hit	[rodents]
.. Rattus norvegicus (rats)	768	11 hits	[rodents]
.. Mus musculus (mouse)	766	16 hits	[rodents]
.. Chinchilla lanigera	761	11 hits	[rodents]
.. Cavia porcellus (guinea pig)	744	1 hit	[rodents]
.. Octodon degus	742	1 hit	[rodents]
.. Jaculus jaculus	732	1 hit	[rodents]
.. Ceratotherium simum simum (southern square-lipped...)	814	2 hits	[odd-toed ungulates]
.. Vicugna pacos	807	4 hits	[even-toed ungulates]
.. Camelus ferus	804	5 hits	[even-toed ungulates]
.. Equus caballus (equine)	802	3 hits	[odd-toed ungulates]
.. Bos taurus (domestic cow)	802	7 hits	[even-toed ungulates]
.. Felis catus (cats)	801	2 hits	[carnivores]
.. Odobenus rosmarus diversus	800	2 hits	[carnivores]
.. Bos mutus (wild yak)	800	2 hits	[even-toed ungulates]
.. Ovis aries (wild sheep)	798	2 hits	[even-toed ungulates]
.. Capra hircus (goats)	798	3 hits	[even-toed ungulates]
.. Myotis brandtii	798	2 hits	[bats]
.. Orcinus orca (Orca)	797	1 hit	[whales & dolphins]
.. Myotis davidii	796	2 hits	[bats]
.. Canis lupus familiaris (dogs)	796	4 hits	[carnivores]
.. Sus scrofa (wild boar)	795	4 hits	[even-toed ungulates]
.. Condylyura cristata	794	2 hits	[insectivores]
.. Pandtholops hodgsonii	794	3 hits	[even-toed ungulates]
.. Pteropus alecto	791	1 hit	[bats]
.. Mustela putorius furo (ferret)	787	2 hits	[carnivores]
.. Alluoropoda melanocephala	781	1 hit	[carnivores]
.. Sorex araneus (Eurasian shrew)	753	1 hit	[insectivores]
.. Dasypus novemcinctus	803	2 hits	[placentals]
.. Trichechus manatus latirostris	796	3 hits	[placentals]
.. Loxodonta africana (African savannah elephant)	768	1 hit	[placentals]
.. Echinops telfairi (lesser hedgehog tenrec)	766	1 hit	[placentals]
.. synthetic construct	879	3 hits	[other sequences]
			PREDICTED: forkhead box protein P3 isoform X1 [Homo sapiens]
			PREDICTED: forkhead box protein P3 isoform 1 [Pan paniscus]
			Foxp3 protein [Pan troglodytes]
			Foxp3 protein [Gorilla gorilla]
			PREDICTED: forkhead box protein P3 [Gorilla gorilla gorilla]
			PREDICTED: forkhead box protein P3 isoform 2 [Pongo abelii]
			PREDICTED: forkhead box protein P3 isoform 2 [Pongo abelii]
			Foxp3 protein [Hylobates lar]
			PREDICTED: LOW QUALITY PROTEIN: forkhead box protein P3 [No
			PREDICTED: forkhead box protein P3 isoform X3 [Macaca fasci
			forkhead box protein P3 [Macaca mulatta] >gi 52354003 gb AA
			PREDICTED: forkhead box protein P3 [Papio anubis]
			forkhead box protein P3 [Callicebus jacchus] >gi 254763667/
			PREDICTED: forkhead box protein P3 [Saimiri boliviensis bol
			PREDICTED: forkhead box protein P3 [Ootolemur garnettii]
			PREDICTED: forkhead box protein P3 isoform 1 [Oryctolagus cuniculus]
			PREDICTED: forkhead box protein P3 isoform X1 [Lepus timidus tr
			PREDICTED: forkhead box protein P3 isoform X1 [Tupaia chine
			PREDICTED: forkhead box protein P3 isoform X1 [Ochotona pri
			PREDICTED: forkhead box protein P3 isoform X1 [Heterocephal
			PREDICTED: forkhead box protein P3 [Microtus ochropaster]
			PREDICTED: forkhead box protein P3 [Cricetus griseus] >gi
			PREDICTED: forkhead box protein P3 [Mesocricetus auratus]
			forkhead box protein P3 [Rattus norvegicus] >gi 564398711/
			forkhead box protein P3 [Mus musculus] >gi 1135635001 ref NP
			PREDICTED: forkhead box protein P3 isoform X10 [Chinchilla
			PREDICTED: forkhead box protein P3 isoform 1 [Cavia porcell
			PREDICTED: forkhead box protein P3 isoform X1 [Octodon degu
			PREDICTED: forkhead box protein P3 isoform X2 [Jaculus jacu
			PREDICTED: forkhead box protein P3 isoform 1 [Ceratotherium
			forkhead box protein P3 [Felis catus] >gi 126093301 gb BAB7
			PREDICTED: forkhead box protein P3 isoform 2 [Odobenus rosm
			PREDICTED: forkhead box protein P3 isoform X1 [Bos mutus]
			forkhead box protein P3 [Ovis aries] >gi 1548531175 ref XP_0
			forkhead box protein P3 [Ovis aries] >gi 1548531175 ref XP_0
			PREDICTED: forkhead box protein P3 isoform X2 [Myotis brand
			PREDICTED: forkhead box protein P3 isoform 1 [Orcinus orca]
			forkhead box protein P3 [Myotis davidii]
			forkhead box protein P3 [Canis lupus familiaris] >gi 1545557
			PREDICTED: forkhead box protein P3 isoform X1 [Sus scrofa]
			PREDICTED: forkhead box protein P3 isoform X1 [Condylyura cr
			PREDICTED: forkhead box protein P3 isoform X3 [Pandtholops h
			forkhead box protein P3 [Pteropus alecto]
			PREDICTED: forkhead box protein P3 [Mustela putorius furo]
			PREDICTED: forkhead box protein P3-like [Alluoropoda melano]
			PREDICTED: forkhead box protein P3 isoform X1 [Sorex araneu
			PREDICTED: forkhead box protein P3 isoform 1 [Dasypus novem
			PREDICTED: forkhead box protein P3 isoform 1 [Trichechus na
			PREDICTED: forkhead box protein P3-like isoform 1 [Loxodonte
			PREDICTED: forkhead box protein P3 isoform X1 [Echinops tel
			forkhead box protein P3 isoform 3 [Homo sapiens] >gi 1145480

Figure 2 Taxonomy report of blastp of human FOXP3.

Description	Max score	Total score	Query cover	E value	Ident	Accession
FOXP3 protein [Homo sapiens]	879	879	100%	0.0	100%	AAH11854.1
transcription factor foxp3 [Xenopus laevis]	273	273	98%	9e-83	38%	NP_001121199.1
PREDICTED: forkhead box protein P3 [Xenopus (Silurana) tropicalis]	254	254	82%	9e-76	42%	XP_002941017.2
PREDICTED: forkhead box protein P2-like IL episcosteus oculatus]	244	244	57%	8e-72	49%	XP_006625452.1
PREDICTED: forkhead box P1a isoform X2 [Danio rerio]	242	242	56%	4e-70	47%	XP_005162017.1
forkhead box P1a [Danio rerio]	244	244	56%	8e-70	47%	NP_001071032.2
PREDICTED: forkhead box P1a isoform X1 [Danio rerio]	243	243	56%	9e-70	47%	XP_005162016.1
PREDICTED: forkhead box protein P1-B isoform X3 [Danio rerio]	239	239	57%	5e-69	47%	XP_009300836.1
forkhead box protein P1-B [Danio rerio]	241	241	57%	7e-69	48%	NP_001034726.1
PREDICTED: forkhead box protein P1-like isoform X1 IL episcosteus oculatus]	242	242	57%	9e-69	47%	XP_006630670.1
PREDICTED: forkhead box protein P1-B isoform X1 [Danio rerio]	240	240	57%	2e-68	47%	XP_005166115.1
PREDICTED: forkhead box protein P1-B isoform X2 [Danio rerio]	240	240	57%	3e-68	47%	XP_005166114.1
PREDICTED: forkhead box protein P1 [Phalacrocorax carbo]	233	233	54%	3e-68	47%	XP_009500234.1
PREDICTED: forkhead box protein P1-B-like isoform X6 [Astranax mexicanus]	237	237	59%	5e-68	45%	XP_007256410.1
PREDICTED: forkhead box protein P1-B-like isoform X5 [Astranax mexicanus]	237	237	59%	5e-68	45%	XP_007256409.1
PREDICTED: forkhead box protein P1-B-like isoform X4 [Astranax mexicanus]	238	238	59%	5e-68	45%	XP_007256408.1
PREDICTED: forkhead box protein P1-B-like isoform 1 [Takifugu rubripes]	239	239	57%	6e-68	46%	XP_003973806.1
PREDICTED: forkhead box protein P1-B-like isoform X3 [Astranax mexicanus]	238	238	59%	1e-67	45%	XP_007256407.1
PREDICTED: forkhead box protein P1-B-like isoform X1 [Astranax mexicanus]	238	238	59%	1e-67	45%	XP_007256406.1
foxp1-like protein [Tetraodon nigroviridis]	231	231	54%	3e-67	48%	AC070857.1
PREDICTED: forkhead box protein P1-like isoform X3 [Python bivittatus]	238	238	57%	4e-67	46%	XP_007421370.1
PREDICTED: forkhead box protein P1 isoform X4 [Balearica pavonina alberticeps]	235	235	57%	4e-67	46%	XP_010300441.1
PREDICTED: LOW QUALITY PROTEIN: forkhead box protein P1 [Columba livia]	237	237	57%	4e-67	47%	XP_005514946.1
PREDICTED: forkhead box protein P1 [Cariama cristata]	230	230	53%	4e-67	48%	XP_009708153.1
PREDICTED: forkhead box protein P1 isoform X5 [Mesitornis unicolor]	234	234	57%	7e-67	46%	XP_010178252.1
Forkhead box protein P1 [Phalacrocorax carbo]	229	229	54%	7e-67	47%	KFF089536.1
PREDICTED: forkhead box protein P1 isoform X2 [Larimichthys crocea]	234	234	57%	8e-67	46%	KFF20700.1
PREDICTED: forkhead box protein P1 [Tauraco erythrophus]	231	231	57%	1e-66	46%	XP_009888198.1
PREDICTED: forkhead box protein P1-B [Coryxias latipes]	234	234	56%	1e-66	47%	XP_004070878.1
Forkhead box protein P1 [Tauraco erythrophus]	231	231	57%	1e-66	46%	XP_009888198.1
unnamed protein product [Oncorhynchus mykiss]	229	229	55%	1e-66	47%	CD061679.1
PREDICTED: forkhead box protein P1 isoform X5 [Calypste anna]	236	236	57%	1e-66	46%	XP_008497865.1
PREDICTED: forkhead box protein P1 isoform X6 [Pseudopodoces humilis]	236	236	57%	1e-66	46%	XP_005522043.1
PREDICTED: forkhead box protein P1 isoform X1 [Cuculus canorus]	236	236	57%	1e-66	46%	XP_009561416.1
Forkhead box protein P1 [Cariama cristata]	229	229	54%	1e-66	48%	KFF61865.1
PREDICTED: forkhead box protein P1 isoform X6 [Pseudopodoces humilis]	236	236	57%	1e-66	46%	XP_005522043.1

Figure 3 Blastp of human FOXP3 (ABQ15210) excluding mammals

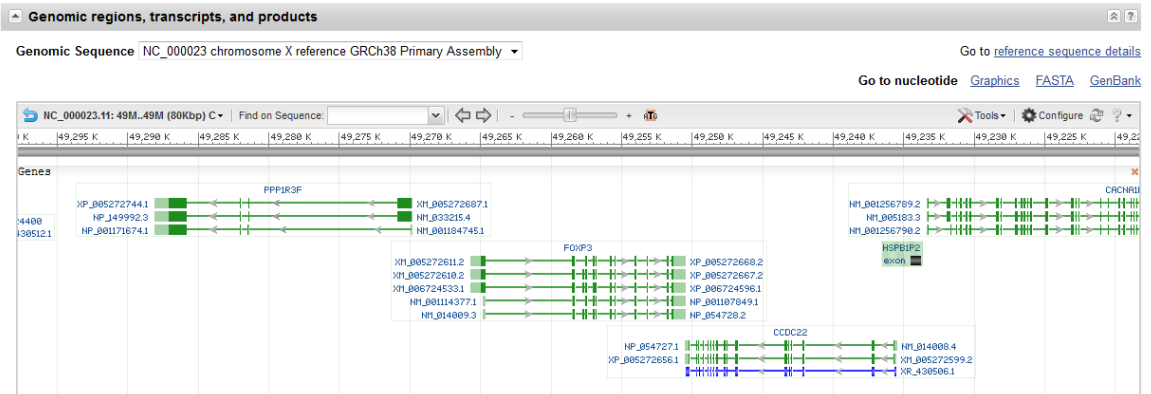


Figure 4 Map of human FOXP3 gene region



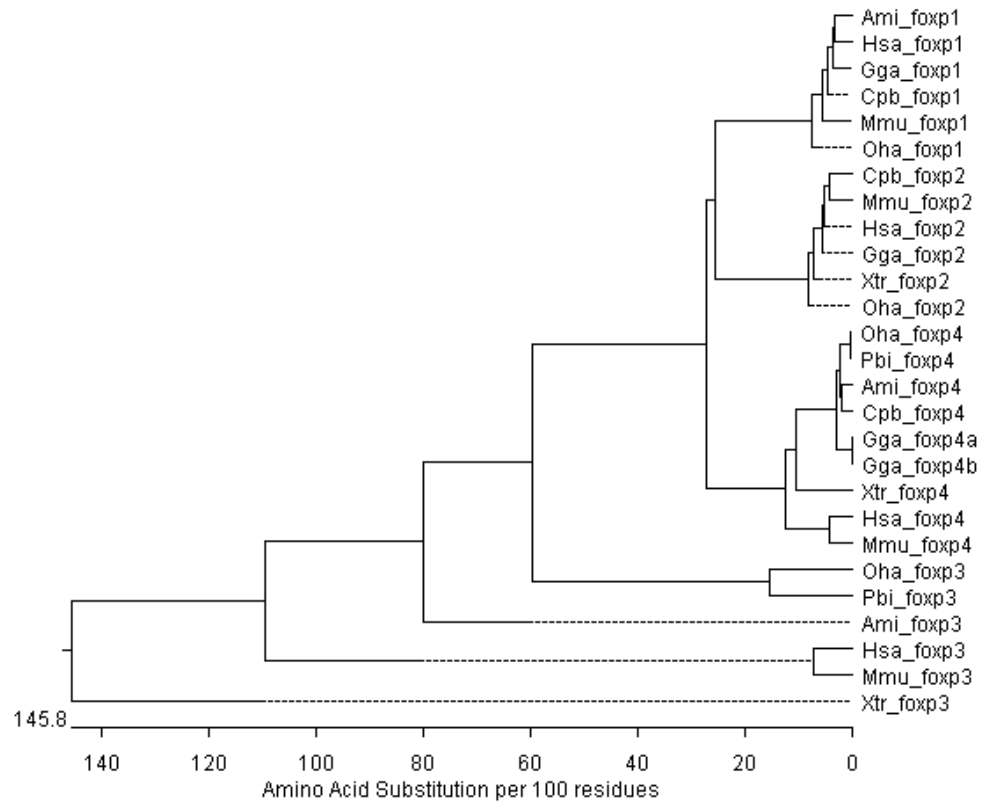


Figure 8 Distance tree of vertebrate Forkhead Box family. [Key: human (Homo sapien, "Hsa"), mouse (Mus musculus, "Mmu"), Chicken (Gallus gallus, "Gga"), American alligator (Alligator mississippiensis, "Ami"), western painted turtle (Chrysemys picta bellii, "Cpb"), cobra (Ophiophagus hannah, "Oha"), python (Python bivittatus, "Pbi"), and western clawed frog (Xenopus tropicalis, "Xtr")]

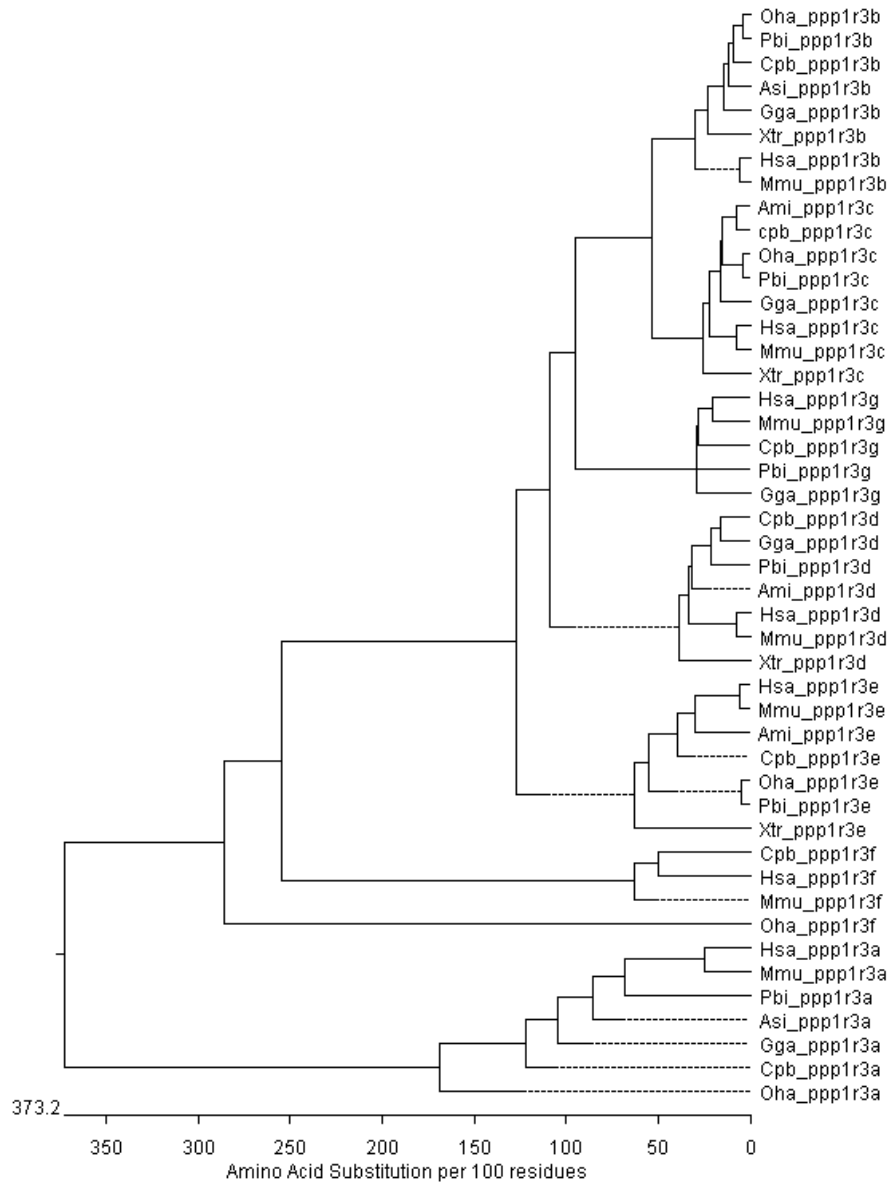


Figure 9 Distance Tree of Protein Phosphatase 1 family. [Key: same as Figure 8]

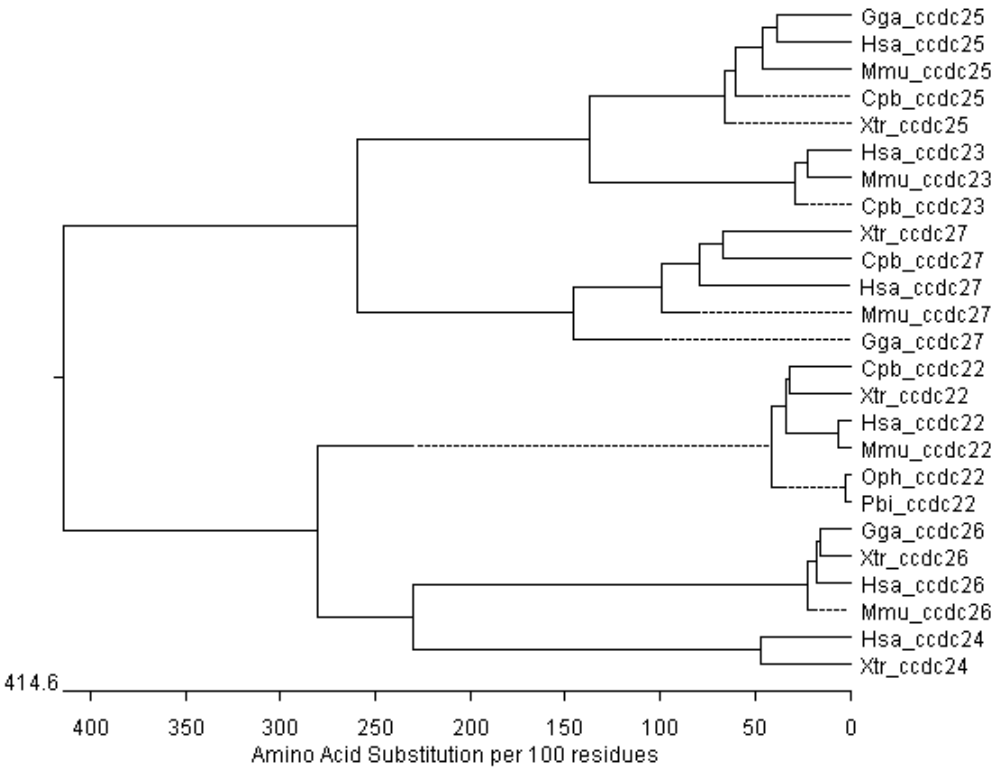


Figure 10 Distance tree of Coiled-Coil Domain family. [Key: same as Figure 8]

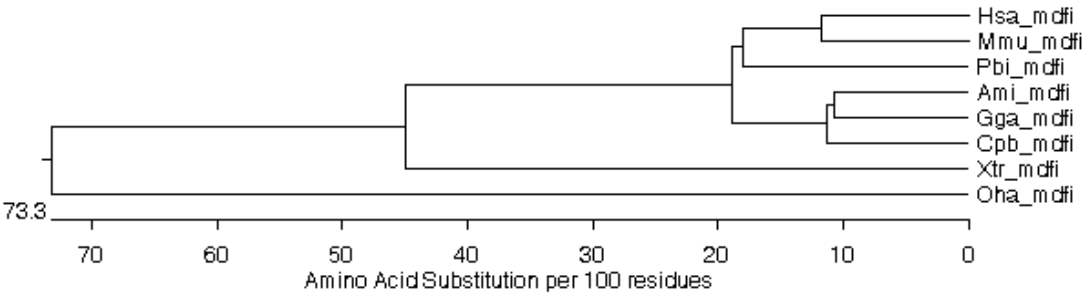


Figure 11 Distance Tree of MyoD Family Inhibitor [Key: same as Figure 8]