EUROPEAN JOURNAL OF BIOMEDICAL AND PHARMACEUTICAL SCIENCES
http://www.ejbps.com

ISSN 2349-8870
Volume: 6
Issue: 3
391-404
Year: 2019

FUNCTIONAL AND STRUCTURAL CHARACTERIZATION OF MECP2 GENE WHICH IS RESPONSIBLE FOR RETT SYNDROME

***P. Divya and K. Shoba**

Department of Biochemistry, D.K.M College for Women, Vellore, Tamil Nadu.

***Corresponding Author: P. Divya**
Department of Biochemistry, D.K.M College for Women, Vellore, Tamil Nadu.

Article Received on 09/01/2019 **Article Revised on 29/01/2019** **Article Accepted on 19/02/2019**

ABSTRACT
Rett syndrome is dueto a genetic mutation of the MECP2 gene. This gene occurs on the x chromosome typically it developed as a new mutation, with less than one percent of cases being inherited from a person parent. If occur almost exclusively in girls. Rett syndrome is a genetic brain disorder which typically becomes apparent after 6 to 18 month of age in female. MECP2 (methyl CPG binding protein 2) is a gene that encodes the protein MECP2.MECP2 appears to be essential for the function of nerve cell. The sequence of protein retrieved from national Centre for biotechnology information in fasta format. The structural analysis of MECP2 were carried out by using bioinformatics tools like scanprosite, blast-p,T -coffee, gene card, MHC, Aceview…etc.

KEYWORDS: Rett syndrome, methyl CPG binding protein 2, NCBI, scanprosite, blast-p, T coffee, gene card, MHC, Aceview.

INTRODUCTION

Rett syndrome (RTT) is a genetic brain disorder which typically becomes apparent after 6 to 18 months of age in females. Symptoms include problems with language, coordination, and repetitive movements. Often there is slower growth, problems walking, and a smaller head size. Complications can include seizures, scoliosis, and sleeping problems. Those affected, however, may be affected to different degrees.

Rett syndrome is due to a genetic mutation of the MECP2 gene. This gene occurs on the X chromosome. Typically it develops as a new mutation, with less than one percent of cases being inherited from a person's parents. It occurs almost exclusively in girls. Boys who have a similar mutation typically die shortly after birth. Diagnosis is based on symptoms and can be confirmed with genetic testing.

There is no known cure for Rett syndrome. Treatment is directed at improving symptoms. Anticonvulsants may be used to help with seizures. Special education, physiotherapy, and braces may also be useful. Many people with the condition live into middle age.

The condition affects about 1 in 8,500 females. Andreas Rett, a pediatrician in Vienna, first described the condition in 1966. As his writings were in German, they did not become widely known in the English-speaking world. Bengt Hagberg, a Swedish pediatrician, published an English article in 1983 and named the condition after Rett. In 1999, Lebanese-American physician Huda Zoghbi discovered the mutation that causes the condition.

METHODOLOGY

Target protein sequenceof MECP2 were retrieved from NCBI data base. The retrieved sequence is submitted to the following server and bioinformatics tools for functional and structural characterization of scanprosite MECP2 protein. The retrieved sequence is submitted to blast-p server similarity search of MECP2 protein. The retrieved sequence is submitted to T-COFFEE server for evolutionary analysis of MECP2 protein. The retrieved sequence is submitted to GVS server for motif analysis of MECP2 protein. The retrieved sequence is submitted to Interproscan sequence search for domainanalysis in MECP2 protein. The retrieved sequence is submitted to palm server for phylogenetic analysis of MECP2 protein. The retrieved sequence is submitted to Aceview server for presence of redundant sequence of MECP2 protein. The retrieved sequence is submitted to gene card toolfor chromosomal mapping position of MECP2 protein. The retrieved sequence is submitted to Annie server for functional analysis of MECP2 protein. The retrieved sequence is submitted to MHC server for antigen binding region of MHCP2 protein.

RESULT AND DISCUSSION

1. SEQUENCE RETRIEVEL

NCBI

PROTEIN

>NP_001104262.1 methyl-CpG-binding protein 2 isoform 2 [Homo sapiens]
maaaaaaapsgggggeeerleeksedqdlqglkdkplkfkkvkkdkkeekegkhepvqpsahhsaepae
agkaetsegsgsapavpeasaspkqrrsiirdrgpmyddptlpegwtrklkqrksgrsagkydvylinpq
gkafrskveliayfekvgdtsldpndfdftvtgrgspsrreqkppkkpkspkapgtgrgrgrpkgsgttr
pkaatsegvqvkrvlekspgkllvkmpfqtspggkaegggattstqvmvikrpgrkrkaeadpqaipkkr
grkpgsvvaaaaaeakkkavkessirsvqetvlpikkrktretvsievkevvkpllvstlgeksgkglkt
ckspgrkskesspkgrsssassppkkehhhhhhhhsespkapvpllpplpppppepessedptsppepqdl
sssvckeekmprggslesdgcpkepaktqpavataataaekykhrgegerkdivsssmprpnreepvdsr
tpvtervs

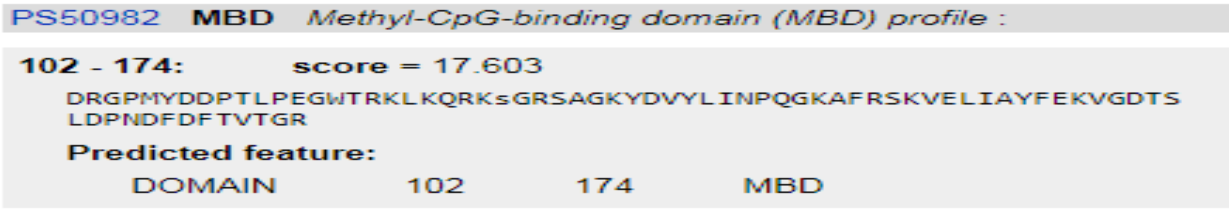
NUCLEOTIDE

>NM_001110792.1:67-1563 Homo sapiens methyl-CpG binding protein 2 (MECP2), transcript variant 2, Mrna
ATGGCCGCCGCCGCCGCCGCCGCGCCGAGCGGA
GGAGGAGGAGGAGGCGAGGAGGAGAGACTGGA
AGAAA
AGTCAGAAGACCAGGACCTCCAGGGCCTCAAGG
ACAAACCCCTCAAGTTTAAAAAGGTGAAGAAAG
ATAA
GAAAGAAGAGAAAGAGGGCAAGCATGAGCCCG
TGCAGCCATCAGCCCACCACTCTGCTGAGCCCG
CAGAG
GCAGGCAAAGCAGAGACATCAGAAGGGTCAGG
CTCCGCCCCGGCTGTGCCGGAAGCTTCTGCCTCC
CCCA
AACAGCGGCGCTCCATCATCCGTGACCGGGGAC
CCATGTATGATGACCCCACCCTGCCTGAAGGCT
GGAC
ACGGAAGCTTAAGCAAAGGAAATCTGGCCGCTC
TGCTGGGAAGTATGATGTGTATTTGATCAATCCC
CAG
GGAAAAGCCTTTCGCTCTAAAGTGGAGTTGATT
GCGTACTTCGAAAAGGTAGGCGACACATCCCTG
GACC
CTAATGATTTTGACTTCACGGTAACTGGGAGAG
GGAGCCCCTCCCGGCGAGAGCAGAAACCACCTA
AGAA
GCCCAAATCTCCCAAAGCTCCAGGAACTGGCAG
AGGCCGGGGACGCCCCAAAGGGAGCGGCACCA
CGAGA
CCCAAGGCGGCCACGTCAGAGGGTGTGCAGGTG
AAAAGGGTCCTGGAGAAAAGTCCTGGGAAGCTC
CTTG
TCAAGATGCCTTTTCAAACTTCGCCAGGGGGCA
AGGCTGAGGGGGGTGGGGCCACCACATCCACCC
AGGT
CATGGTGATCAAACGCCCCGGCAGGAAGCGAAA
AGCTGAGGCCGACCCTCAGGCCATTCCCAAGAA
ACGG
GGCCGAAAGCCGGGGAGTGTGGTGGCAGCCGCT
GCCGCCGAGGCCAAAAAGAAAGCCGTGAAGGA
GTCTT
CTATCCGATCTGTGCAGGAGACCGTACTCCCCAT
CAAGAAGCGCAAGACCCGGGAGACGGTCAGCA
TCGA
GGTCAAGGAAGTGGTGAAGCCCCTGCTGGTGTC
CACCCTCGGTGAGAAGAGCGGGAAAGGACTGA
AGACC
TGTAAGAGCCCTGGGCGGAAAAGCAAGGAGAG
CAGCCCCAAGGGGCGCAGCAGCAGCGCCTCCTC
ACCCC
CCAAGAAGGAGCACCACCACCATCACCACCACT
CAGAGTCCCCAAAGGCCCCCGTGCCACTGCTCC
CACC
CCTGCCCCCACCTCCACCTGAGCCCGAGAGCTC
CGAGGACCCCACCAGCCCCCCTGAGCCCCAGGA
CTTG
AGCAGCAGCGTCTGCAAAGAGGAGAAGATGCC
CAGAGGAGGCTCACTGGAGAGCGACGGCTGCCC
CAAGG
AGCCAGCTAAGACTCAGCCCGCGGTTGCCACCG
CCGCCACGGCCGCAGAAAAGTACAAACACCGA
GGGGA
GGGAGAGCGCAAAGACATTGTTTCATCCTCCAT
GCCAAGGCCAAACAGAGAGGAGCCTGTGGACA
GCCGG
ACGCCCGTGACCGAGAGAGTTAGCTGA

The above result shows the fasta format of protein and nucleotide sequence of Methyl CPS binding protein 2.

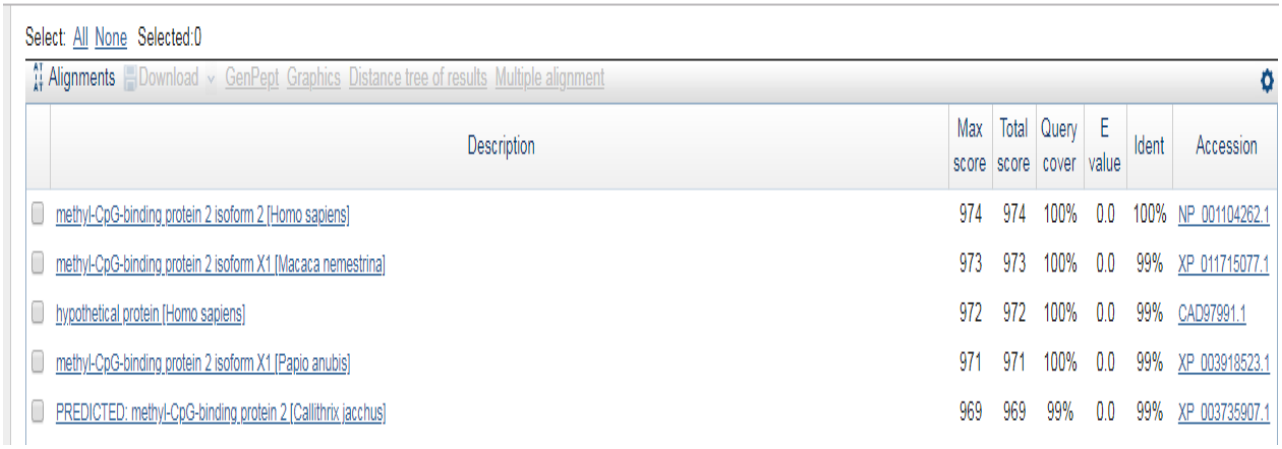
2. MOTIF SEARCH

SCANPROSITE



The above results shows the motif region in yellow colour and conserved domain region in green colour of MECP2 protein, here.

3. SIMILARITY SEARCH
BLAST P



Select: All None Selected:0

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession
methyl-CpG-binding protein 2 isoform 2 [Homo sapiens]	974	974	100%	0.0	100%	NP_001104262.1
methyl-CpG-binding protein 2 isoform X1 [Macaca nemestrina]	973	973	100%	0.0	99%	XP_011715077.1
hypothetical protein [Homo sapiens]	972	972	100%	0.0	99%	CAD97991.1
methyl-CpG-binding protein 2 isoform X1 [Papio anubis]	971	971	100%	0.0	99%	XP_003918523.1
PREDICTED: methyl-CpG-binding protein 2 [Callithrix jacchus]	969	969	99%	0.0	99%	XP_003735907.1

The above result shows the similarity of the MECP2 protein.

4. MULTIPLE ANALYSIS
T COFFEE

T-COFFEE, Version_11.00.d625267 (2016-01-11 15:25:41 - Revision d625267 - Build 507)
Cedric Notredame
SCORE=992
*
BAD AVG GOOD
*
NP_001104262.1 : 98
XP_024843940.1 : 98
NP_073164.2 : 98
XP_024843940.1_ : 98
AKP92847.1 : 98
AKP92847.1_1 : 98
cons : 99

NP_001104262.1
MAAAAAAAPSGGGGGGEEERLEEKSEDQDLQGL
KDKPLKFKKVKKDKKEEKEGKHEPVQPSAH
XP_024843940.1
MAAAAAAAPSGGGGGGEEERLEEKSEEQDLQGL
KDKPLKFKKVKKDKKEDKEGKHEPLQPAAH
NP_073164.2 MVAGML----------
GLREEKSEDQDLQGLKEKPLKFKKVKKDKKEDKE
GKHEPLQPSAH
XP_024843940.1_
MAAAAAAAPSGGGGGGEEERLEEKSEEQDLQGL
KDKPLKFKKVKKDKKEDKEGKHEPLQPAAH
AKP92847.1 MAAA-
AAAPSGGGGGGEEERLEEKSEDQDLQGLKDKPLK
FKKVKKDKKEDKEGKHEPVQPPAH
AKP92847.1_1 MAAA-
AAAPSGGGGGGEEERLEEKSEDQDLQGLKDKPLK
FKKVKKDKKEDKEGKHEPVQPPAH
cons *.*.
*****:*******:**************:*******:**.**
NP_001104262.1
HSAEPAEAGKAETSEGSGSAPAVPEASASPKQRRSI
IRDRGPMYDDPTLPEGWTRKLKQRKSG
XP_024843940.1
HSAEPAEAGKAETSEGSGSAPAVPEASASPKQRRSI
IRDRGPMYDDPTLPEGWTRKLKQRKSG
NP_073164.2
HSAEPAEAGKAETSESSGSAPAVPEASASPKQRRSI
IRDRGPMYDDPTLPEGWTRKLKQRKSG
XP_024843940.1_
HSAEPAEAGKAETSEGSGSAPAVPEASASPKQRRSI
IRDRGPMYDDPTLPEGWTRKLKQRKSG
AKP92847.1
HSAEPAEAGKAETSEGSGSAPAVPEASASPKQRRSI
IRDRGPMYDDPTLPEGWTRKLKQRKSG
AKP92847.1_1
HSAEPAEAGKAETSEGSGSAPAVPEASASPKQRRSI
IRDRGPMYDDPTLPEGWTRKLKQRKSG
cons
**********.**********
****
NP_001104262.1
RSAGKYDVYLINPQGKAFRSKVELIAYFEKVGDTS
LDPNDFDFTVTGRGSPSRREQKPPKKPK
XP_024843940.1
RSAGKYDVYLINPQGKAFRSKVELIAYFEKVGDTS
LDPNDFDFTVTGRGSPSRREQKPPKKPK
NP_073164.2
RSAGKYDVYLINPQGKAFRSKVELIAYFEKVGDTS
LDPNDFDFTVTGRGSPSRREQKPPKKPK
XP_024843940.1_
RSAGKYDVYLINPQGKAFRSKVELIAYFEKVGDTS
LDPNDFDFTVTGRGSPSRREQKPPKKPK
AKP92847.1
RSAGKYDVYLINPQGKAFRSKVELIAYFEKVGDTS
LDPNDFDFTVTGRGSPSRREQKPPKKPK
AKP92847.1_1
RSAGKYDVYLINPQGKAFRSKVELIAYFEKVGDTS
LDPNDFDFTVTGRGSPSRREQKPPKKPK
cons
****
NP_001104262.1
SPKAPGTGRGRGRPKGSGTTRPKAATSEGVQVKR
VLEKSPGKLLVKMPFQTSPGGKAEGGGAT
XP_024843940.1

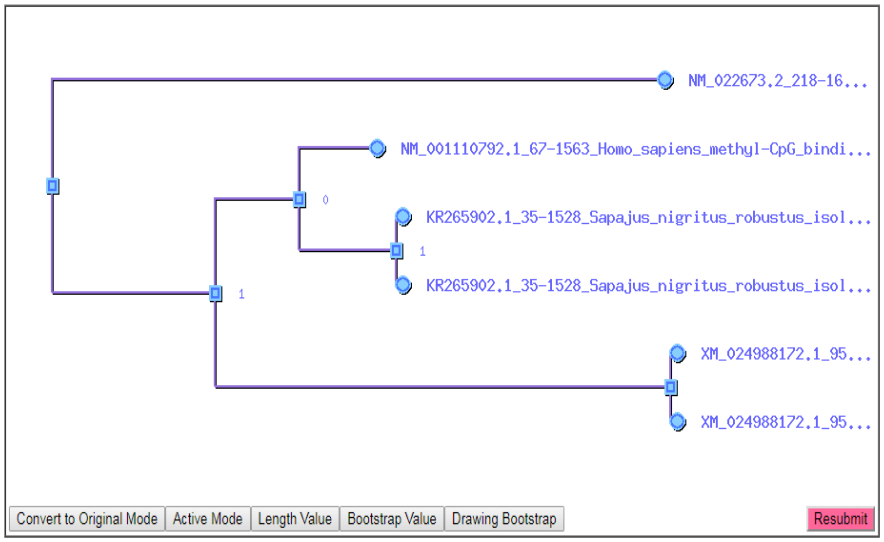
```
SPKAPGTGRGRGRPKGSGTTRPKAAASEGVQVKR
VLEKSPGKLLVKMPFQAAPGSKAEGGGAT
NP_073164.2
SPKAPGTGRGRGRPKGSGTGRPKAAASEGVQVKR
VLEKSPGKLLVKMPFQASPGGKGEGGGAT
XP_024843940.1_
SPKAPGTGRGRGRPKGSGTTRPKAAASEGVQVKR
VLEKSPGKLLVKMPFQAAPGSKAEGGGAT
AKP92847.1
SPKAPGTGRGRGRPKGSGTTRPKAATSEGVQVKR
VLEKSPGKLLVKMPFQTSPGGKAEGGGAT
AKP92847.1_1
SPKAPGTGRGRGRPKGSGTTRPKAATSEGVQVKR
VLEKSPGKLLVKMPFQTSPGGKAEGGGAT
cons
******************** *****:**********************
***::**.*.******
NP_001104262.1
TSTQVMVIKRPGRKRKAEADPQAIPKKRGRKPGS
VVAAAAAEAKKKAVKESSIRSVQETVLPI
XP_024843940.1
TSAQVMVIKRPGRKRKAEADPQAIPKKRGRKPGS
VVAAATAEAKKKAVKESSIRSVQETVLPI
NP_073164.2
TSAQVMVIKRPGRKRKAEADPQAIPKKRGRKPGS
VVAAAAAEAKKKAVKESSIRSVQETVLPI
XP_024843940.1_
TSAQVMVIKRPGRKRKAEADPQAIPKKRGRKPGS
VVAAATAEAKKKAVKESSIRSVQETVLPI
AKP92847.1
TSTQVMVIKRPGRKRKAEADPQAIPKKRGRKPGS
VVAAAAAEAKKKAVKESSIRSVQETVLPI
AKP92847.1_1
TSTQVMVIKRPGRKRKAEADPQAIPKKRGRKPGS
VVAAAAAEAKKKAVKESSIRSVQETVLPI
cons
**:**************************************:******
******************
NP_001104262.1
KKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKT
CKSPGRKSKESSPKGRSSSASSPPKKEH
XP_024843940.1
KKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKT
CKSPGRKSKESSPKGRSGSASSPPKKEH
NP_073164.2
KKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKT
CKSPGRKSKESSPKGRSSSASSPPKKEH
XP_024843940.1_
KKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKT
CKSPGRKSKESSPKGRSGSASSPPKKEH
AKP92847.1
KKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKT
CKSPGRKSKESSPKGRSSSASSPPKKEH
AKP92847.1_1
KKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKT
CKSPGRKSKESSPKGRSSSASSPPKKEH
cons
**********************************************
******.**********
NP_001104262.1
HHHHHHSESPKAPVPLLPPLPPPPPEPESSEDPTSPP
EPQDLSSSVCKEEKMPRGGSLESDGC
XP_024843940.1
HHHHHHVEPPKAPAPLLLPPPPPPPEPQSSEDPASP
PEPQDLSSSVCKEEKMPRAGSLESDGC
NP_073164.2 HHHHHHAESPKAPMPLLP--
PPPPPEPQSSEDPISPPEPQDLSSSICKEEKMPRAGSL
ESDGC
XP_024843940.1_
HHHHHHVEPPKAPAPLLLPPPPPPPEPQSSEDPASP
PEPQDLSSSVCKEEKMPRAGSLESDGC
AKP92847.1
HHHHHHSESPKAPVPLLPPLPPPPPEPESSEDPTSPP
EPQDLSSSVCKEEKMPRGGSLESDGC
AKP92847.1_1
HHHHHHSESPKAPVPLLPPLPPPPPEPESSEDPTSPP
EPQDLSSSVCKEEKMPRGGSLESDGC
cons ****** *.**** ***
*******:***** **********:********.********
NP_001104262.1 PKEPAKTQPAVATAAT--------
AAEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTP
VTER
XP_024843940.1 PKEPAKTQPALATAAP--------
ATEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTP
VTER
NP_073164.2
PKEPAKTQPMVAAAATTTTTTTTTVAEKYKHRGE
GERKDIVSSSMPRPNREEPVDSRTPVTER
XP_024843940.1_ PKEPAKTQPALATAAP--------
ATEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTP
VTER
AKP92847.1 PKEPAKTQPAVATAAT--------
AAEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTP
VTER
AKP92847.1_1 PKEPAKTQPAVATAAT--------
AAEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTP
VTER
cons ********* :*:**.
.:**************************************
NP_001104262.1 VS
XP_024843940.1 VS
NP_073164.2 VS
XP_024843940.1_ VS
AKP92847.1 VS
AKP92847.1_1 VS
cons **
```

The above result shows the evolutionary relationship of MECP2 protein.

5. PHYLOGENTIC ANALYSIS

PALM

Tree Image Area



Note: If the tree is oversized and would not be displayed properly, we recommend users to download the Newick format tree file and view the phylogenetic tree with other tools. (e.g. ATV or FigTree tree viewer)

The above result shows the phylogenetic analysis of MECP2 protein.

6. MUTATION STUDIES

A.GVS

Set up parameters for display and analysis

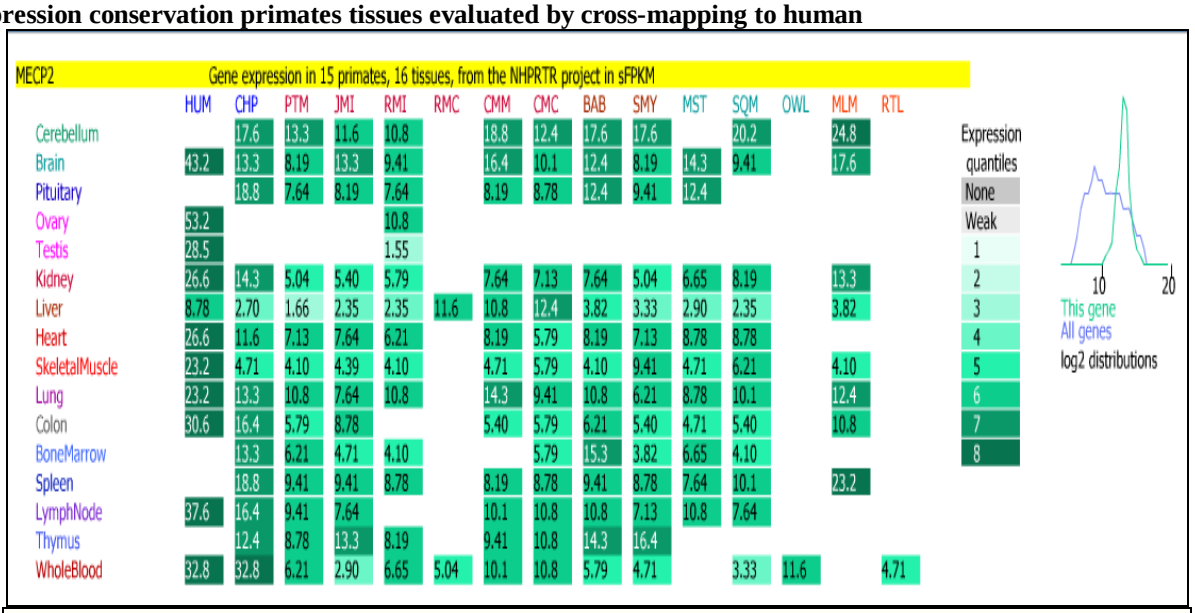
Merging Data Sets			
Merge Samples and Variations For tag SNPs, choice of B or C may invoke the MultiPop algorithm.	C - combined samples with combined variations		
Data Output and Display			
Output SNPs By	rs ID and Position in graph	Display SNPs By	Table/Image
Filtering SNPs			
Allele Frequency Cutoff (%)	0	No Monomorphic Sites	☐
Clustering in Graphic Display			
Cluster SNPs	☐	Cluster Samples	☐
Selecting Tag SNPs			
r^2Threshold (0.0-1.0)	0.80	Data Coverage (%) for Tag SNPs	14
		Data Coverage (%) for Clustering	12
Show More Parameters			

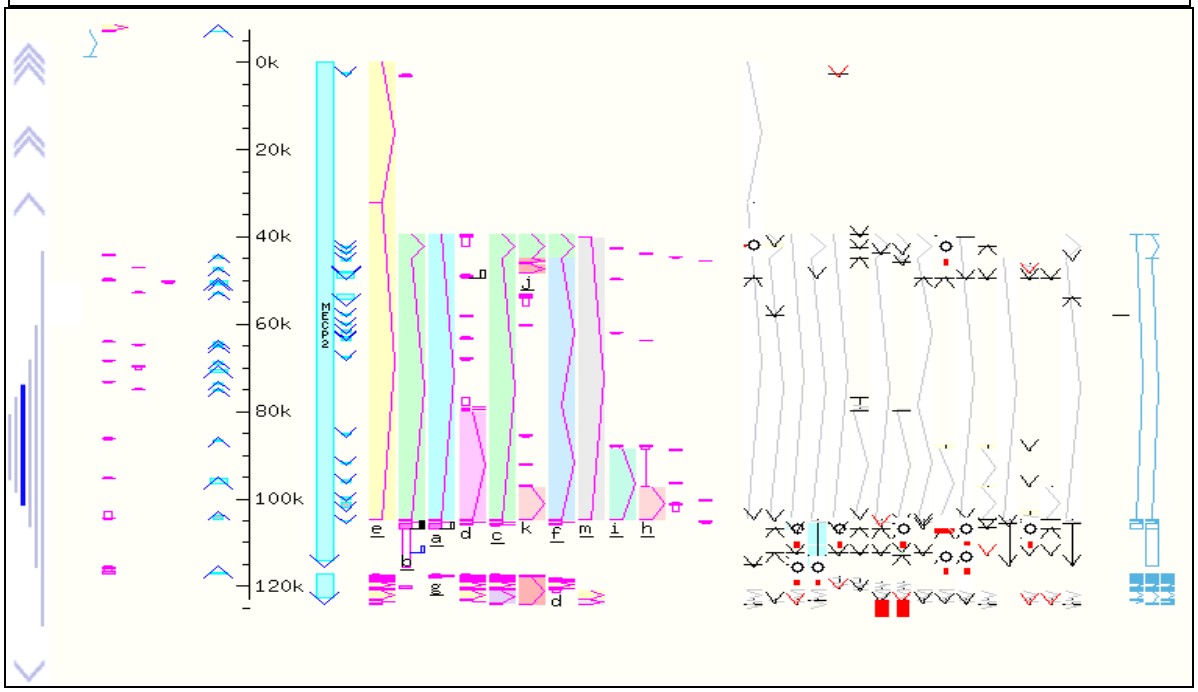
The above result shows the motif region of MECP2 protein.

7. ANNOTATION STUDIES

A. ACEVIEW

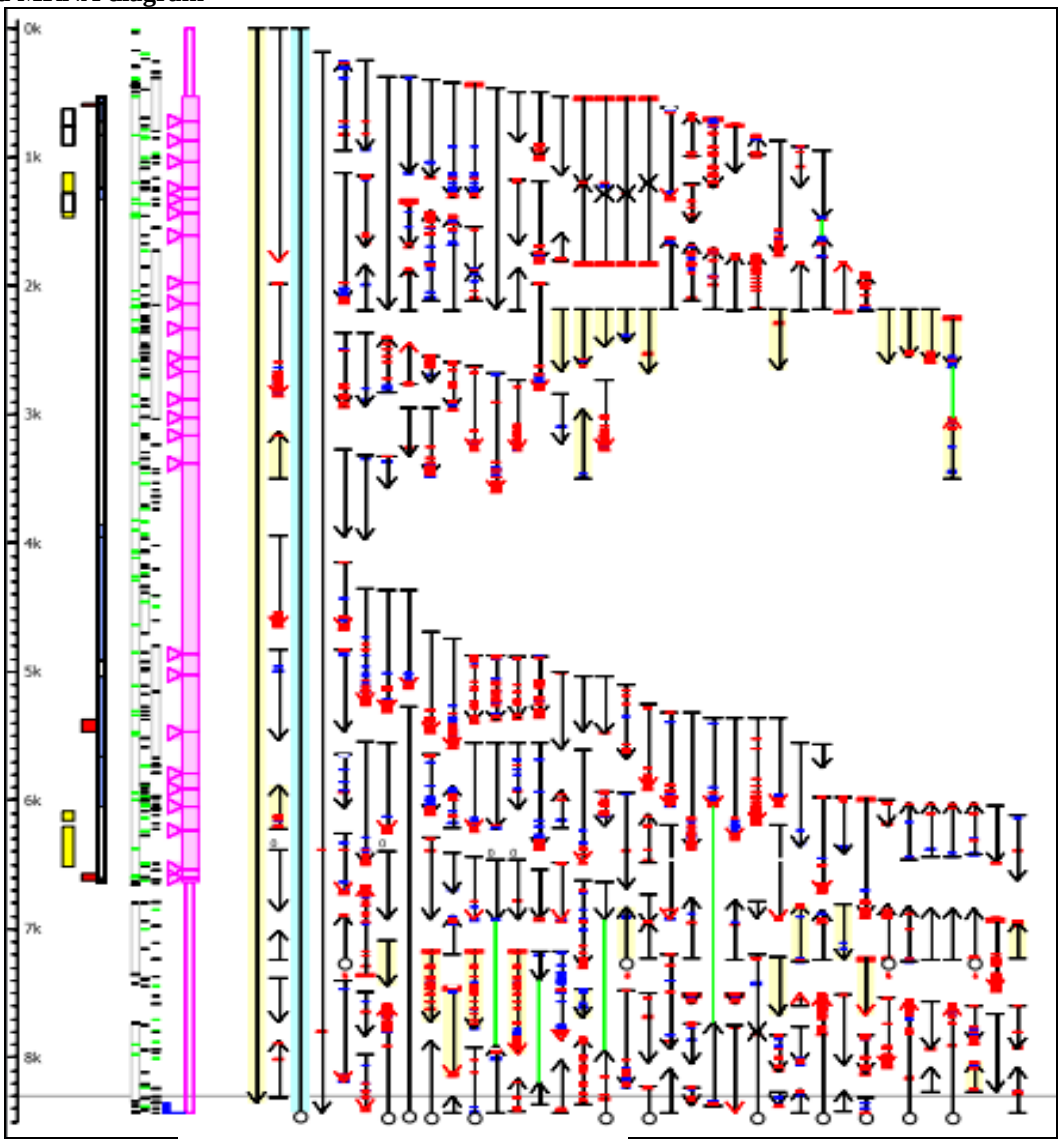
Expression conservation primates tissues evaluated by cross-mapping to human





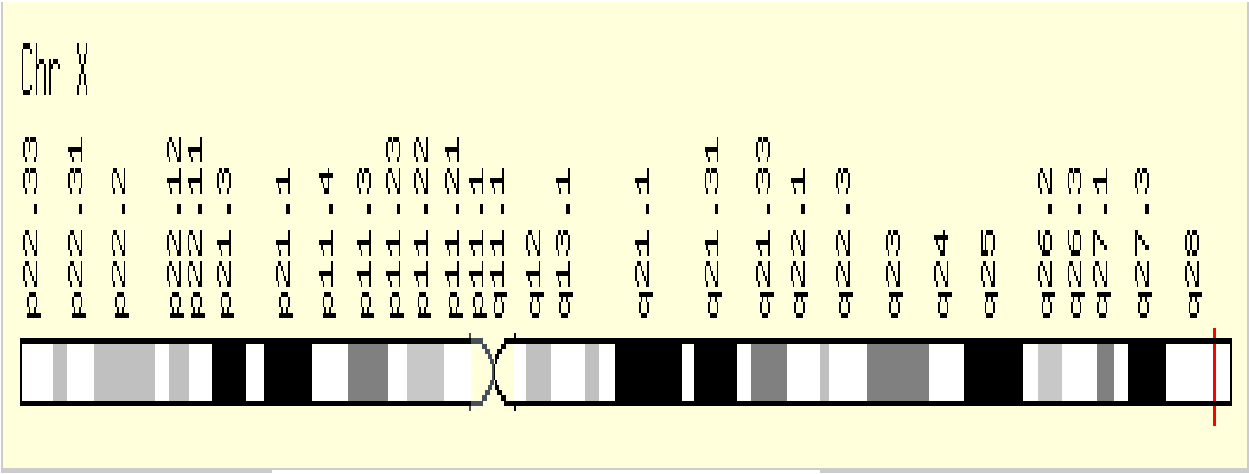
mRNA variant	mRNA matching the genome	Best predicted protein	5' UTR	3' UTR	uORF	Upstream sequence	Transcription unit pre-mRNA	Downstream sequence
aAug10	1693 bp	520 aa		130 bp		2kb	67501 bp	1kb
bAug10	10502 bp	486 aa	242 bp	8799 bp	102 bp	2kb including Promoter	76186 bp	1kb
cAug10	878 bp	172 aa	186 bp	173 bp		2kb possibly including promoter	66671 bp	1kb
dAug10	2890 bp	172 aa	2373 bp			2kb possibly including promoter	29073 bp	1kb
eAug10	490 bp	163 aa				2kb	104755 bp	1kb
fAug10	786 bp	144 aa	352 bp		102 bp	2kb possibly including promoter	66302 bp	1kb
gAug10-unspliced	2727 bp	132 aa	211 bp	2117 bp		2kb possibly including promoter	2727 bp	1kb
hAug10	1198 bp	71 aa		981 bp		2kb	17045 bp	1kb
iAug10	1088 bp	72 aa		869 bp		2kb	17047 bp	1kb
jAug10	769 bp	60 aa	461 bp	125 bp		2kb	8923 bp	1kb
kAug10	684 bp	118 aa	102 bp	225 bp		2kb possibly including promoter	7918 bp	1kb
lAug10-unspliced	442 bp	40 aa		317 bp		2kb	442 bp	1kb
mAug10	342 bp	34 aa		237 bp		2kb	64604 bp	1kb
nAug10-unspliced	269 bp	41 aa	118 bp	25 bp		2kb	269 bp	1kb

Annotated MRNA diagram



The above result show the of redundant sequence of MECP2 protein.

B.GENE CARD



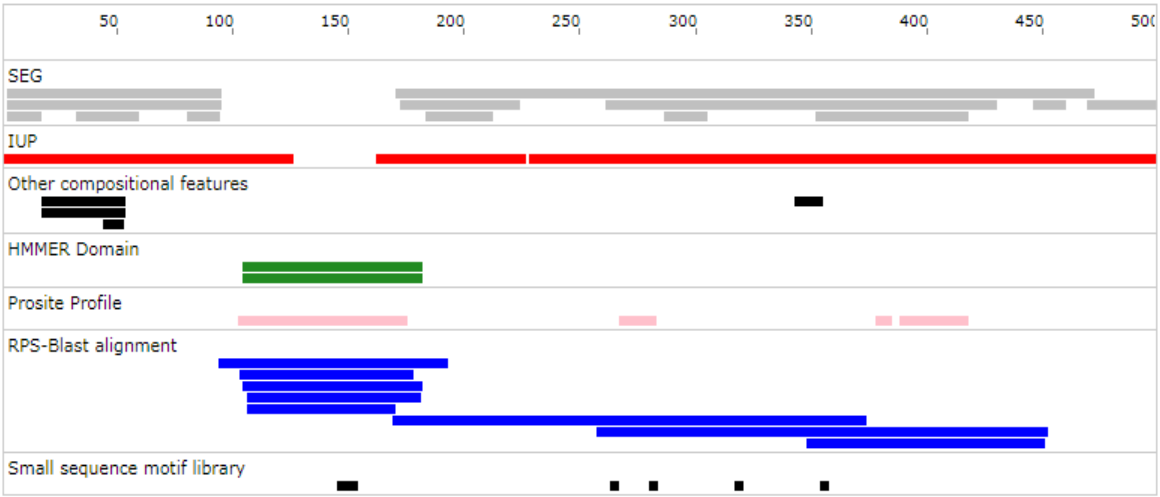
The above result show the chromosomal mapping position of MECP2 protein.

8. FUNCTIONAL ANALYSIS STUDIES

A.ANNIE

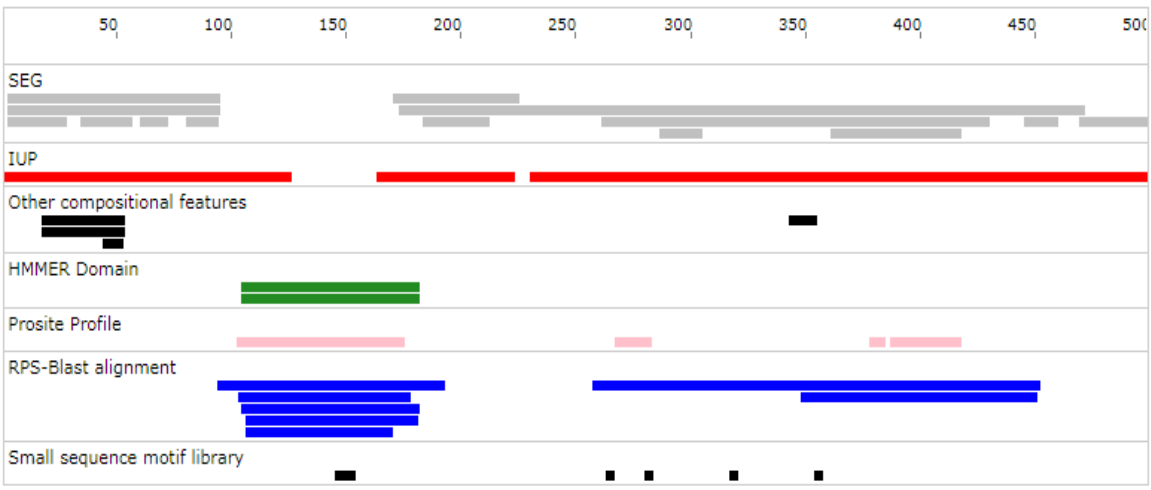
Description: NP_001104262.1 methyl-CpG-binding protein 2 isoform 2 [Homo sapiens]

Length: 498



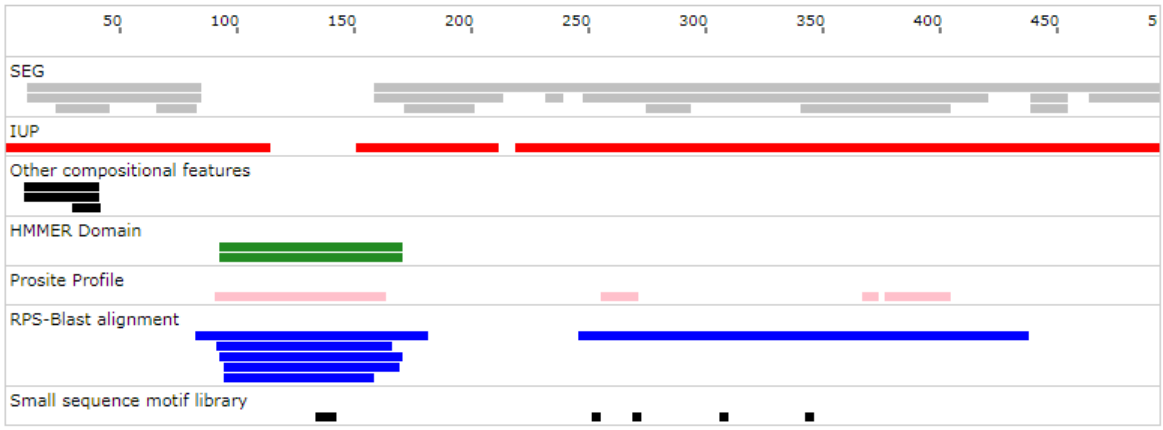
Description: XP_024843940.1 methyl-CpG-binding protein 2 isoform X1 [Bos taurus]

Length: 498



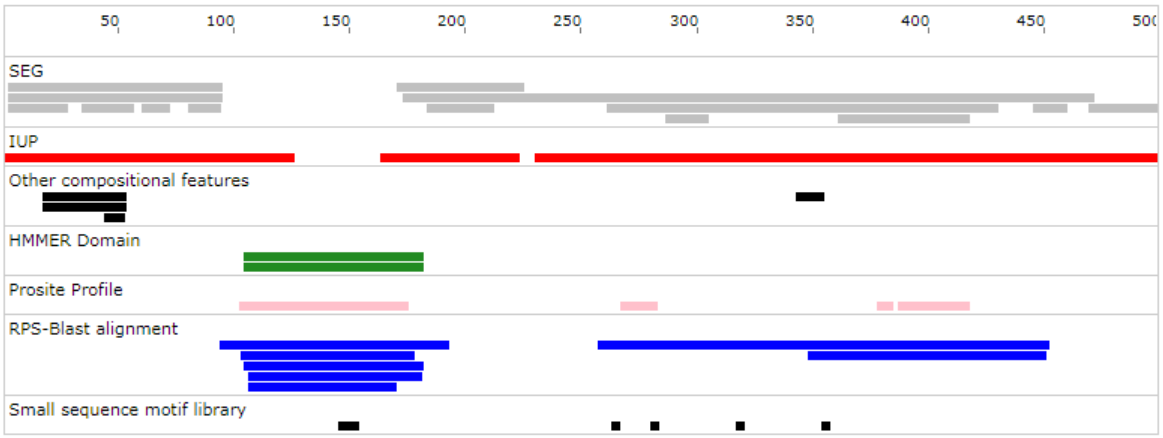
Description: NP_073164.2 methyl-CpG-binding protein 2 [Rattus norvegicus]

Length: 492



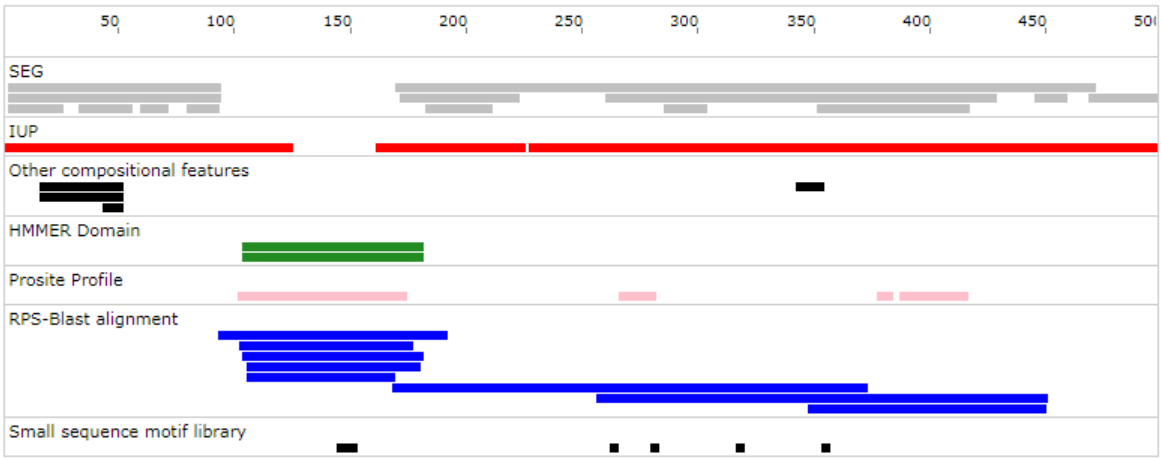
Description: XP_024843940.1 methyl-CpG-binding protein 2 isoform X1 [Bos taurus]

Length: 498



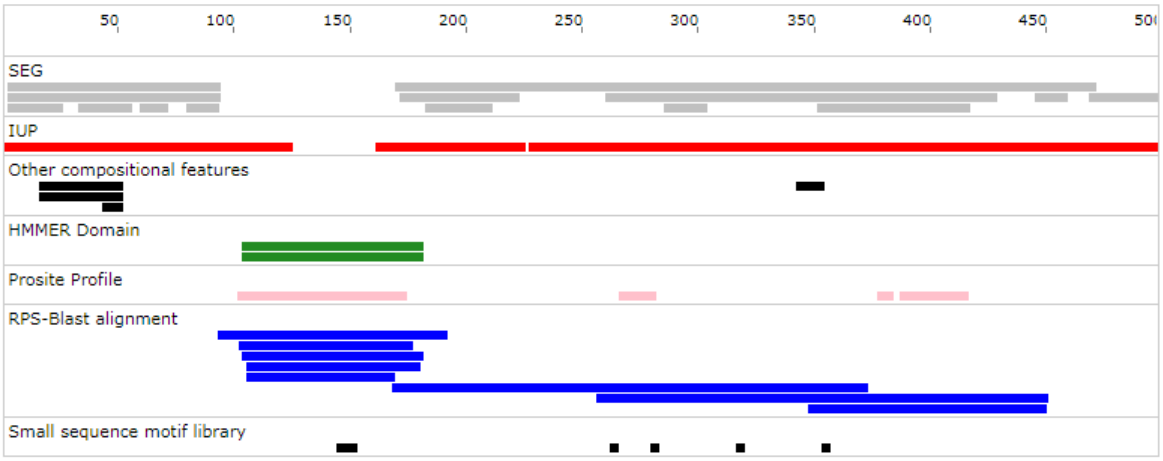
Description: AKP92847.1 methyl CpG binding protein 2 transcript variant 2 [Sapajus nigritus robustus]

Length: 497



Description: AKP92847.1 methyl CpG binding protein 2 transcript variant 2 [Sapajus nigritus robustus]

Length: 497



The above show thefunctional analysis of MECP2 protein.

B.MHC

#	Name	Sequence
1	sequence 1	MAAAAAAAPSGGGGGGEEERLEEKSEDQDLQGLKDKPLKFKKVKKDKKEE KEGKHEPVQPSAHHSAEPAEAGKAETSEGSGSAPAVPEASASPKQRRSII RDRGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSKVEL IAYFEKVGDTSLDPNDFDFTVTGRGSPSRREQKPPKKPKSPKAPGTGRGR GRPKGSGTTRPKAATSEGVQVKRVLEKSPGKLLVKMPFQTSPGGKAEGGG ATTSTQVMVIKRPGRKRKAEADPQAIPKKRGRKPGSVVAAAAAEAKKKAV KESSIRSVQETVLPIKKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKT CKSPGRKSKESSPKGRSSSASSPPKKEHHHHHHHSESPKAPVPLLPPLPP PPPEPESSEDPTSPPEPQDLSSSVCKEEKMPRGGSLESDGCPKEPAKTQP AVATAATAAEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTPVTERVS

Allele	#	Start	End	Peptide Length	Sequence	Method used	Percentile Rank
HLA-A*11:01	1	367	375	9	SSSASSPPK	Consensus (*ANN, SMM*)	0.10
HLA-A*11:01	1	453	461	9	ATAATAAEK	Consensus (*ANN, SMM*)	0.30
HLA-A*11:01	1	368	376	9	SSASSPPKK	Consensus (*ANN, SMM*)	0.35
HLA-A*11:01	1	349	357	9	KTCKSPGRK	Consensus (*ANN, SMM*)	0.60
HLA-A*11:01	1	253	261	9	TSTQVMVIK	Consensus (*ANN, SMM*)	0.65
HLA-A*11:01	1	338	346	9	STLGEKSGK	Consensus (*ANN, SMM*)	0.65
HLA-A*11:01	1	214	222	9	ATSEGVQVK	Consensus (*ANN, SMM*)	0.75
HLA-A*11:01	1	254	262	9	STQVMVIKR	Consensus (*ANN, SMM*)	0.75
HLA-A*11:01	1	258	266	9	MVIKRPGRK	Consensus (*ANN, SMM*)	0.75
HLA-A*11:01	1	223	231	9	RVLEKSPGK	Consensus (*ANN, SMM*)	0.85
HLA-A*11:01	1	421	429	9	SSSVCKEEK	Consensus (*ANN, SMM*)	0.85

#	Name	Sequence
1	sequence 1	MAAAAAAAPSGGGGGGEEERLEEKSEEQDLQGLKDKPLKFKKVKKDKKED KEGKHEPLQPAAHHSAEPAEAGKAETSEGSGSAPAVPEASASPKQRRSII RDRGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSKVEL IAYFEKVGDTSLDPNDFDFTVTGRGSPSRREQKPPKKPKSPKAPGTGRGR GRPKGSGTTRPKAAASEGVQVKRVLEKSPGKLLVKMPFQAAPGSKAEGGG ATTSAQVMVIKRPGRKRKAEADPQAIPKKRGRKPGSVVAAATAEAKKKAV KESSIRSVQETVLPIKKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKT CKSPGRKSKESSPKGRSGSASSPPKKEHHHHHHHVEPPKAPAPLLLPPPP PPPEPQSSEDPASPPEPQDLSSSVCKEEKMPRAGSLESDGCPKEPAKTQP ALATAAPATEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTPVTERVS

Allele	#	Start	End	Peptide Length	Sequence	Method used	Percentile Rank
HLA-A*11:01	1	453	461	9	ATAAPATEK	Consensus (*ANN, SMM*)	0.30
HLA-A*11:01	1	253	261	9	TSAQVMVIK	Consensus (*ANN, SMM*)	0.35
HLA-A*11:01	1	349	357	9	KTCKSPGRK	Consensus (*ANN, SMM*)	0.60
HLA-A*11:01	1	338	346	9	STLGEKSGK	Consensus (*ANN, SMM*)	0.65
HLA-A*11:01	1	368	376	9	GSASSPPKK	Consensus (*ANN, SMM*)	0.70
HLA-A*11:01	1	258	266	9	MVIKRPGRK	Consensus (*ANN, SMM*)	0.75
HLA-A*11:01	1	223	231	9	RVLEKSPGK	Consensus (*ANN, SMM*)	0.85
HLA-A*11:01	1	421	429	9	SSSVCKEEK	Consensus (*ANN, SMM*)	0.85
HLA-A*11:01	1	367	375	9	SGSASSPPK	Consensus (*ANN, SMM*)	1.10

#	Name	Sequence
1	sequence 1	MVAGMLGLREEKSEDQDLQGLKEKPLKFKKVKKDKKEDKEGKHEPLQPSA HHSAEPAEAGKAETSESSGSAPAVPEASASPKQRRSIIRDRGPMYDDPTL PEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSKVELIAYFEKVGDTSL DPNDFDFTVTGRGSPSRREQKPPKKPKSPKAPGTGRGRGRPKGSGTGRPK AAASEGVQVKRVLEKSPGKLLVKMPFQASPGGKGEGGGATTSAQVMVIKR PGRKRKAEADPQAIPKKRGRKPGSVVAAAAAEAKKKAVKESSIRSVQETV LPIKKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKTCKSPGRKSKESS PKGRSSSASSPPKKEHHHHHHHAESPKAPMPLLPPPPPPEPQSSEDPISP PEPQDLSSSICKEEKMPRAGSLESDGCPKEPAKTQPMVAAAATTTTTTTT TVAEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTPVTERVS

Allele	#	Start	End	Peptide Length	Sequence	Method used	Percentile Rank
HLA-A*11:01	1	355	363	9	SSSASSPPK	Consensus (*ANN, SMM*)	0.10
HLA-A*11:01	1	241	249	9	TSAQVMVIK	Consensus (*ANN, SMM*)	0.35
HLA-A*11:01	1	356	364	9	SSASSPPKK	Consensus (*ANN, SMM*)	0.35
HLA-A*11:01	1	447	455	9	TTTTTVAEK	Consensus (*ANN, SMM*)	0.45
HLA-A*11:01	1	337	345	9	KTCKSPGRK	Consensus (*ANN, SMM*)	0.60
HLA-A*11:01	1	326	334	9	STLGEKSGK	Consensus (*ANN, SMM*)	0.65
HLA-A*11:01	1	246	254	9	MVIKRPGRK	Consensus (*ANN, SMM*)	0.75
HLA-A*11:01	1	211	219	9	RVLEKSPGK	Consensus (*ANN, SMM*)	0.85
HLA-A*11:01	1	407	415	9	SSSICKEEK	Consensus (*ANN, SMM*)	0.95

#	Name	Sequence
1	sequence 1	MAAAAAAAPSGGGGGGEEERLEEKSEEQDLQGLKDKPLKFKKVKKDKKED KEGKHEPLQPAAHHSAEPAEAGKAETSEGSGSAPAVPEASASPKQRRSII RDRGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSKVEL IAYFEKVGDTSLDPNDFDFTVTGRGSPSRREQKPPKKPKSPKAPGTGRGR GRPKGSGTTRPKAAASEGVQVKRVLEKSPGKLLVKMPFQAAPGSKAEGGG ATTSAQVMVIKRPGRKRKAEADPQAIPKKRGRKPGSVVAAATAEAKKKAV KESSIRSVQETVLPIKKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKT CKSPGRKSKESSPKGRSGSASSPPKKEHHHHHHHVEPPKAPAPLLLPPPP PPPEPQSSEDPASPPEPQDLSSSVCKEEKMPRAGSLESDGCPKEPAKTQP ALATAAPATEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTPVTERVS

Allele	#	Start	End	Peptide Length	Sequence	Method used	Percentile Rank
HLA-A*11:01	1	453	461	9	ATAAPATEK	Consensus (*ANN,SMM*)	0.30
HLA-A*11:01	1	253	261	9	TSAQVMVIK	Consensus (*ANN,SMM*)	0.35
HLA-A*11:01	1	349	357	9	KTCKSPGRK	Consensus (*ANN,SMM*)	0.60
HLA-A*11:01	1	338	346	9	STLGEKSGK	Consensus (*ANN,SMM*)	0.65
HLA-A*11:01	1	368	376	9	GSASSPPKK	Consensus (*ANN,SMM*)	0.70
HLA-A*11:01	1	258	266	9	MVIKRPGRK	Consensus (*ANN,SMM*)	0.75
HLA-A*11:01	1	223	231	9	RVLEKSPGK	Consensus (*ANN,SMM*)	0.85
HLA-A*11:01	1	421	429	9	SSSVCKEEK	Consensus (*ANN,SMM*)	0.85

#	Name	Sequence
1	sequence 1	MAAAAAAPSGGGGGGEEERLEEKSEDQDLQGLKDKPLKFKKVKKDKKEDK EGKHEPVQPPAHHSAEPAEAGKAETSEGSGSAPAVPEASASPKQRRSIIR DRGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSKVELI AYFEKVGDTSLDPNDFDFTVTGRGSPSRREQKPPKKPKSPKAPGTGRGRG RPKGSGTTRPKAATSEGVQVKRVLEKSPGKLLVKMPFQTSPGGKAEGGGA TTSTQVMVIKRPGRKRKAEADPQAIPKKRGRKPGSVVAAAAAEAKKKAVK ESSIRSVQETVLPIKKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKTC KSPGRKSKESSPKGRSSSASSPPKKEHHHHHHHSESPKAPVPLLPPLPPP PPEPESSEDPTSPPEPQDLSSSVCKEEKMPRGGSLESDGCPKEPAKTQPA VATAATAAEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTPVTERVS

Allele	#	Start	End	Peptide Length	Sequence	Method used	Percentile Rank
HLA-A*11:01	1	366	374	9	SSSASSPPK	Consensus (*ANN,SMM*)	0.10
HLA-A*11:01	1	452	460	9	ATAATAAEK	Consensus (*ANN,SMM*)	0.30
HLA-A*11:01	1	367	375	9	SSASSPPKK	Consensus (*ANN,SMM*)	0.35
HLA-A*11:01	1	348	356	9	KTCKSPGRK	Consensus (*ANN,SMM*)	0.60
HLA-A*11:01	1	252	260	9	TSTQVMVIK	Consensus (*ANN,SMM*)	0.65
HLA-A*11:01	1	337	345	9	STLGEKSGK	Consensus (*ANN,SMM*)	0.65
HLA-A*11:01	1	213	221	9	ATSEGVQVK	Consensus (*ANN,SMM*)	0.75
HLA-A*11:01	1	253	261	9	STQVMVIKR	Consensus (*ANN,SMM*)	0.75
HLA-A*11:01	1	257	265	9	MVIKRPGRK	Consensus (*ANN,SMM*)	0.75

#	Name	Sequence
1	sequence 1	MAAAAAAPSGGGGGGEEERLEEKSEDQDLQGLKDKPLKFKKVKKDKKEDK EGKHEPVQPPAHHSAEPAEAGKAETSEGSGSAPAVPEASASPKQRRSIIR DRGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSKVELI AYFEKVGDTSLDPNDFDFTVTGRGSPSRREQKPPKKPKSPKAPGTGRGRG RPKGSGTTRPKAATSEGVQVKRVLEKSPGKLLVKMPFQTSPGGKAEGGGA TTSTQVMVIKRPGRKRKAEADPQAIPKKRGRKPGSVVAAAAAEAKKKAVK ESSIRSVQETVLPIKKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKTC KSPGRKSKESSPKGRSSSASSPPKKEHHHHHHHSESPKAPVPLLPPLPPP PPEPESSEDPTSPPEPQDLSSSVCKEEKMPRGGSLESDGCPKEPAKTQAV ATAATAAEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTPVTERVS

The above result show the antigen binding region of MHCP2 protein.

CONCLUSION

Rett syndrome is a neuron development disorder that affects girls almost exclusively. It is characterized by normal early growth and development, loss of purposeful use of the hands, distinctive hand movement, slowed brain head growth and intellectual disability. DNA methylation is a major modification of eukaryotic genomes and plays an essential role in mammalian development. Chromosomal protein that bind to methylated DNA. It can bind specifically to a singal methyl-CPG pair. Human protein MECP2 (this protein)MBD1, MBD2, MBD3 and MBD4 comprise a family of nuclear protein related by the persence in each if a methyl –CPG binding domain(MBD). The protein sequence of MECP2 was retrieved from NCBI data base. Functional and structural characterization of MECP2 protein were done through using bioinformatics software's and tools. Hence by way our project we get the results like functional and structural characterization of target protein, which is acts as a potential therapeutic agents for clinical and pharmacy informatics studies.

REFERENCES

1. Abbeduto, Leonard; Ozonoff, Susan; Thurman, Angela John; McDuffie, Angela; Schweitzer, Julie. Hales, Robert; Yudofsky, Stuart; Robert, Laura Weiss, eds. Chapter 8. Neurodevelopmental Disorders, The American Psychiatric Publishing Textbook of Psychiatry (6 ed.). Arlington, VA: American Psychiatric Publishing. ISBN 978-1-58562-444-7. Retrieved 11 March 2015.
2. Acampa, M.; Guideri, F. (May 2006). "Cardiac disease and Rett syndrome". Archives of Disease in Childhood, 91(5): 440–443. doi:10.1136/adc.2005.090290. ISSN 1468-2044. PMC 2082747. PMID 16632674.
3. Amir, Ruthie; Van den Veyver, Ignatia; Wan, Mimi; Tran, Charles; Francke, Uta; Zoghbi, Huda (1999). "Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2". Nature Genetics, 23(2): 185–8. doi:10.1038/13810. PMID 10508514.
4. Ariani, Francesca; Hayek, Giuseppe; Rondinella, Dalila; Artuso, Rosangela; Mencarelli, Maria Antonietta; Spanhol-Rosseto, Ariele; Pollazzon, Marzia; Buoni, Sabrina; Spiga, Ottavia; Ricciardi, Sara; Meloni, Ilaria; Longo, Ilaria; Mari, Francesca; Broccoli, Vania; Zappella, Michele; Renieri, Alessandra (11 July 2008). "FOXG1 is Responsible for the Congenital Variant of Rett Syndrome". The American Journal of Human Genetics, 83(1): 89–93. doi:10.1016/j.ajhg.2008.05.015. PMC 2443837. PMID 18571142.
5. Berridge, Craig W; Waterhouse, Barry D (2003). "The locus coeruleus–noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes". Brain Research Reviews, 42(1): 33–84. doi:10.1016/S0165-0173(03)00143-7. PMID 12668290.
6. Björklund, A.; Lindvall, O (1984). "Dopamine-containing systems in the CNS". In Björklund, A.; Hökfelt, T. Handbook of Chemical Neuroanatomy. Classical Transmitters in the CNS, Part l. 2. New York: Elsevier. pp. 55–122.
7. Björklund, Anders; Dunnett, Stephen B. (2007). "Dopamine neuron systems in the brain: An update". Trends in Neurosciences, 30(5): 194–202. doi:10.1016/j.tins.2007.03.006. PMID 17408759.
8. Buoni, Sabrina; Zannolli, Raffaella; De Felice, Claudio; De Nicola, Anna; Guerri, Vanessa; Guerra, Beatrice; Casali, Stefania; Pucci, Barbara; Corbini, Letizia; Mari, Francesca; Renieri, Alessandra; Zappella, Michele; Hayek, Joseph (May 2010). "EEG features and epilepsy in MECP2-mutated patients with the Zappella variant of Rett syndrome". Clinical Neurophysiology, 121(5): 652–7. doi:10.1016/j.clinph.2010.01.003. PMID 20153689.
9. Chen, Richard Z.; Akbarian, Schahram; Tudor, Matthew; Jaenisch, Rudolf (2001). "Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice". Nature Genetics, 27(3): 327–31. doi:10.1038/85906. PMID 11242118.
10. Cheval, H; Guy, J; Merusi, C; De Sousa, D; Selfridge, J; Bird, A (2012). "Postnatal inactivation reveals enhanced requirement for MeCP2 at distinct age windows". Human Molecular Genetics, 21(17): 3806–14. doi:10.1093/hmg/dds208. PMC 3412380. PMID 22653753.

11. Dauer, William; Przedborski, Serge (2003). "Parkinson's Disease". Neuron, 39(6): 889–909. doi: 10.1016/S0896-6273(03)00568-3. PMID 12971891.
12. Ehrhart, Friederike; Coort, Susan L. M.; Cirillo, Elisa; Smeets, Eric; Evelo, Chris T.; Curfs, Leopold M. G. (25 November 2016). "Rett syndrome – biological pathways leading from MECP2 to disorder phenotypes". Orphanet Journal of Rare Diseases, 11(1). doi:10.1186/s13023-016-0545-5.
13. Fitzgerald, Patricia M.; Jankovic, Joseph; Percy, Alan K. (1990). "Rett syndrome and associated movement disorders". Movement Disorders, 5(3): 195–202. doi: 10.1002/mds.870050303. PMID 2388636.
14. Shoba.k and Dr. Mazher sultana, Three - dimensional structure and motif prediction studies on collagenase protein in fiddler crab, International journal of novel trends in pharmaceutical sciences, Issn: 2277 – 2782, volume 6, issue 4, pages 79 – 83.
15. Shoba K., Manjuladevi M, Dr. Mazher sultana, Biochemical analysis and gene expression profiling on collagenase protein in fiddler crab, World journal of pharmacy and pharmaceutical sciences, issn 2278 – 4357, volume 6, issue 3, 747-756.
16. Shoba K., Sowmiya S and Dr. Mazher sultana, World Journal of Pharmaceutical and Life Sciences, ISSN 2454-2229, Vol. 3, Issue 1, 427-436.
17. Shoba.K, Hebsibahelsie.B, insilico homology modeling ofribulose-1, 5-bisphosphate carboxylase protein in gracilariaedulis, world journal of pharmacy and pharmaceutical sciences, 2017, volume 6, issue 8, 396-406, issn 2278 – 4357.
18. Shoba.K, Lavanya.G, Identification Of De Novo Peptide And Motif Prediction On Porphyria Protein (Hmbs) Using Insilico Tools, Universal Journal Of Pharmacy, 2018, Volume 8, Issue 1, Issn2320-303x.
19. Shoba.K, Lavanya.G, Tertiary Structural Prediction And Drug Binding Studies On Mutated Gene (Hmbs) In Human Porphyria, International Journal Of Novel Trends In Pharmaceutical Science, 2018, Issn 2277 -2782, Volume 8, Issue 1.
20. Shoba K., Hebsibah Elsie B. And Bavyasri S. Insilico Peptide Modeling Stuidies And Structural Analysis On Ribulose -1, 5 Bisphosphate Carboxylase In Gracilaria Edulis, World Journal Of Pharmacy And Pharmaceutical Sciences, 2018, Volume 7, Issue 3, 1086-1095, Issn 2278 – 4357.
21. Shoba. K, Hebsibah Elsie. B and Jayakumari. S. Sathya. R. (2018); Insilico Structural Analysis and Drug Docking Stuidies On Ribulose -1, 5 Bisphosphate Carboxylase In *Gracilaria Edulis*. International journal of advanced research, 6(9): 159-165] (ISSN 2320-5407).
22. Hebsibah Elsie B, Subashini.G, Nithya.G and Shoba.K. (2018); Purification and Identification of Antioxidant Peptides from the Skin Protein Hydrolysate of Marine Fish (*Aurigequula Fasciata*). European journal of Pharmaceutical and Medical Research, 5(10): 371 – 378. Issn No 2394-3211.
23. Shoba. K, Nithy.G and Deepa.L. (2018); Biochemical Analysis and Peptide Modeling of Lysozyme in Indian *Fenneropenaeus indicus* shrimp species. International journal of advanced research, 6(9): 159-165] (ISSN 2320-5407).
24. Kalpana K., Manjuvani S., Shoba K. In Silico Comparative Modeling of Maturase K Protein in Cymbopogon martinii Plant. *Research & Reviews: A Journal of Bioinformatics*, 2018; 5(3): 30–36p.
25. Shoba K., Kalpana K., Protein Modeling and Drug Docking Studies on Potential Protein Target (*E. coli*–dosP) and Compound Aldehyde (Sumatriptan) using Bioinformatics Tools. *Research & Reviews: A Journal of Bioinformatics*, 2018; 5(3): 9–18p.