eISSN: 2231-0541 CAS CODEN: PHARN8 An ELSEVIER Covered Journal

PHARMANEST

An International Journal of Advances in Pharmaceutical Sciences

Volume 4 | Issue 6 | November-December 2013 | Pages 1409-1418

Original Research Article

GENOME ANALYSIS AND ANNOTATION OF *PKC* GENE IN BIPOLAR DISORDER

[a]S. SHANTHIPRIYA*, [b]VA DOSS

[a]Research scholar, Department of Bioinformatics, Karpagam University, Coimbatore, India.

[b]Associate professor, Department of Biochemistry, PSG College of Arts & Science, Coimbatore, India.

Author for Correspondence: shanthipriyaachu@gmail.com

Received: 03-10-2013 Revised: 21-10-2013

Accepted: 29-10-2013 Available online: 01-11-2013

ABSTRACT

Bipolar disorder (BPD) is a complex genetic disorder in which the core feature is pathological disturbance in mood ranging from extreme elation, or mania, to severe depression usually accompanied by disturbances in thinking and behavior. Recent evidence indicates that an alteration in PKC activity plays a significant role in pathophysiology of BPD. Protein kinase C (PKC) is a group of calcium and phospholipid – dependent enzymes, enriched in brain, where it plays a major role in regulating both pre-and postsynaptic aspects of neurotransmission. Inhibition of PKC plays an important role in neuroprotection against BPD. In this paper we have carried out the Gene location in chromosome, Analysis and Annotation. Coding and non-coding region was also identified to predict the Exon in the gene sequence. By using various Bioinformatics tools, the downstream genes were also identified. This analysis may help in further comparative analysis of genes in various organisms and to design a definite drug which inhibits the activity of the gene in Bipolar disorder.

Key words: *PKC* gene, Bipolar disorder, Genome Analysis.

INTRODUCTION

Bipolar disorder is a genetic disorder which shows pathological disturbance in mood ranging from extreme elation or mania to severe depression visually disturbances in thinking and behaviour. Certain genes play major role in bipolar disorder, but interaction of multiple genes (epistasis) or more complex genetic mechanisms (mutation) also play a vital role in BPD. BPD is sub classified into Bipolar I disorder (Clear cut Mania occur) and Bipolar II disorder (Milder forms of Mania called hypomania). Treatments are available for both manic and depressive phases of BPD in order to reduce the reoccurrence of acute episodes of depression and mania [1]

The main brain areas involved in BPD includes the frontal and temporal lobes of the fore brain, the prefrontal cortex, the basal ganglia and parts of limbic system[2]. Clinical levels indicate that PKC signalling may play an important role in pathophysiology and treatment of bipolar disorder. PKC is implicated in a diversity of cellular functions, including cell cycle progression, proliferation, differentiation and apoptosis. PKC has the major role in regulating survival signals in a variety of cell types, including neurons[3]. Various proofs identify that family of protein kinase C (PKC) can act as target in the treatment of BPD.

PKC is a group of calcium and phospholipid dependent enzymes, which plays a pivotal role in cell signalling systems and many evidences indicates that alterations in PKC activity plays significant role in the pathophysiology of Bipolar Disorder[4].

MATERIALS AND METHOD

Complete Gene sequences were retrieved from www.ncbi.nlm.nih.gov. Nucleotide-Nucleotide BLAST[5] was performed to obtain similar sequences like *pkc* gene sequence. Sequences were selected by observing the alignment criteria, bit score, expected value (E-value) and %identities. E-value varies from sequence to sequence depending upon their length; sequences with % identity were taken using CLC workbench software packages. Bioedit is a biological sequence editor that runs in windows (95/98/NT/2000/XP) and is intended to provide basic functions for protein and nucleic acid sequence editing, alignment, manipulation and analysis. In this study it was mainly used for local alignment[6], multiple sequence alignment[7] and identification of Exon. This software has the inbuilt functionality that it creates on its own local database[8]. CLC workbench was used for the prediction of Molecular properties, Atomic composition, Nucleotide distribution and Genome annotation[9].

RESULTS AND DISCUSSION

Genomics of *pkc*

Protein kinase C (*pkc*) is a family of serine and threonine specific protein kinases that can be activated by calcium and the second messenger diacylglycerol. *Pkc* family members phosphorylate a wide variety of protein targets and are known to be involved in diverse cellular signaling pathways. *Pkc* family members also serve as major receptors for phorbol esters, a class of tumor promoters. *Pkc* remains to be one of the most interesting target for mood disorders, especially Bipolar disorder and can be used as a long term drug target for Bipolar disorder[10] (*Pkc* signaling pathways plays a significant impact in the treatment of Bipolar disorder[11].

Genetic analysis of Protein kinase Epsilon of human genome provides a comprehensive analysis of protein in normal and disease states, as well as detailed view of the current state of human genome analysis[12]*Pkc's* nucleotide sequence was obtained and its Sequence information, Atomic composition, Nucleotide distribution table, Counts of di-nucleotides and codon statistics from coding regions were calculated.

Genome Organization of *pkc* in chromosome 2

Pkc gene is located in chromosome 2 and the entire chromosome structure is displayed (Figure 1). In chromosome 2 *pkc* genome content starts at 45,878,484 bp and ends at 46,415,129 bp. Total size are 536,646 bases. Total Number of genes in chromosome 2 is 7208.

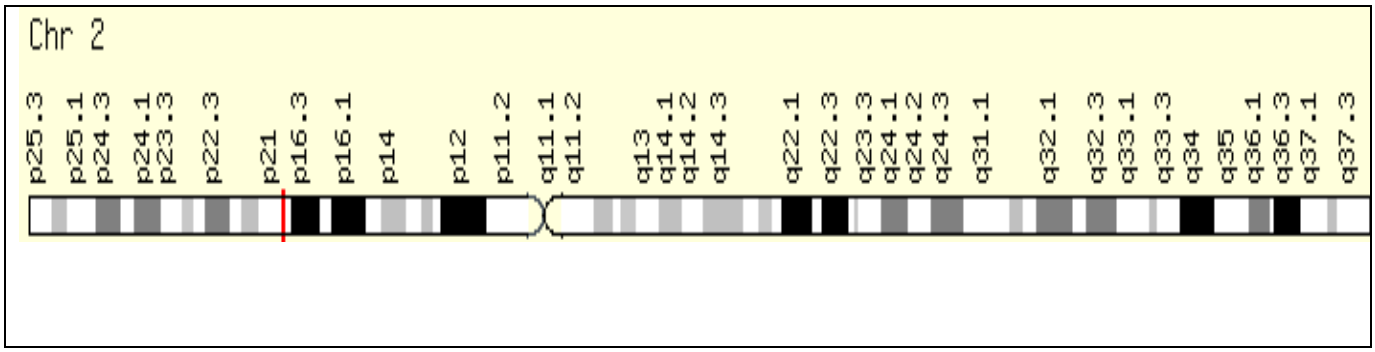

Fig.1. *Pkc* coding of Chromosome 2

■ - ***Pkc* coding**

Diagrammatic representation of the chromosome 2 and the location of entire genes that are present in it are shown Figure 2. In the figure, location of *pkc* gene is highlighted in purple colour.

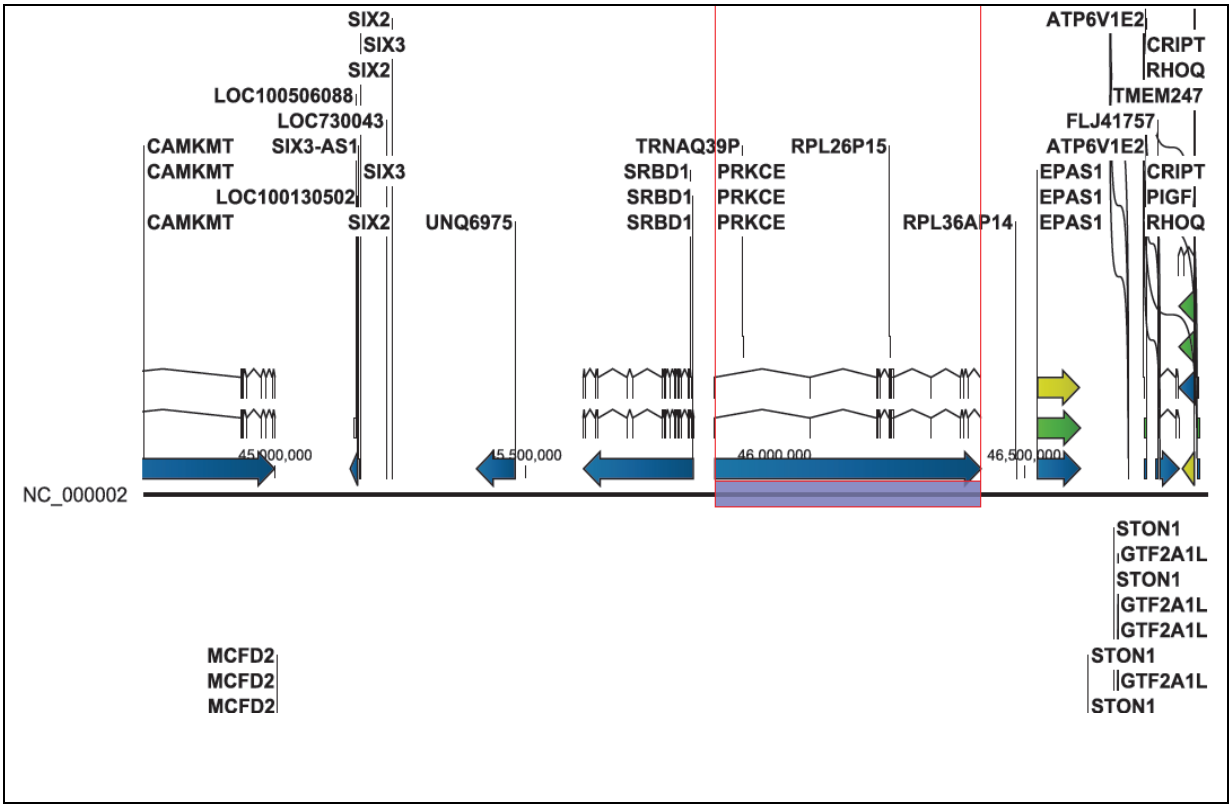

Fig.2. Location of Entire genes in Chromosome 2

- *Pkc* coding region in the entire chromosome 2

Pkc genome structure with different genes and their locations were displayed in Figure 2. In chromosome 2 *pkc* coding region was located in region between 45878484 to 46415129. It was found that *camkmt* gene was located in the first coding region, and *unq6975* next to it was located. The upstream gene for *pkc* was *srbd1* and the downstream gene was *epsa1* followed by *cript,pigf and rhoq* gene etc.

Molecular properties of *pkc*

Identification of molecular properties of *pkc* plays a major role in order to understand the stability of the gene and position of the coding region. To understand the *pkc's* stability it is important to study the nucleotide sequence statistics of pkc like identification of number of exons, polyA signal site, STS regions, miscellaneous features and CDS region, and the same were calculated as shown in Table 1. The length of the sequence was identified as 5,537bp.

Table.1. Molecular properties of *pkc* in Bipolar Disorder

Sequence type	RNA
Length	5,537bp
Organism	Homo sapiens
Weight (single-stranded)	1,779.809 kDa
Weight (double-stranded)	3,557.07 kDa
CDS	1
Exon	15
Gene	1
Misc feature	9
PolyA signal	1
STS	7
Source	1

The atomic composition of *pkc* gene for both Single and Double stranded DNA was calculated, when its strand differs its count and frequency changes. Atoms which show the highest: least, count: frequency in single: double strand are hydrogen and phosphorus.

Table.2a. Atomic composition of *pkc*

Atomic composition	**Single-stranded**		**Double-stranded**	
Atom	Count	Frequency	Count	Frequency
Hydrogen (H)	64,970	0.356	129,951	0.356
Carbon (C)	52,631	0.288	105,203	0.288
Nitrogen (N)	20,731	0.114	41,355	0.113
Oxygen (O)	38,771	0.212	77,520	0.212
Phosphorus (P)	5,537	0.030	11,074	0.030

Distribution of Adenine, Guanine, Cytosine and Uracil in nucleotide sequence plays a vital role to study its properties. Stability of the target depends upon the Guanine and Cytosine content (GC content) in the nucleotide sequence. *Pkc* gene contains GC content as 2,596 and 0.469 in count and frequency[13]. Since *pkc* have highest GC content it shows good stability and the value of other nucleotides are displayed in Table 2b.

Table.2b. Nucleotide distribution of *pkc*

Nucleotide	Count	Frequency
Adenine (A)	1,465	0.265
Cytosine (C)	1,263	0.228
Guanine (G)	1,333	0.241
Uracil (U)	1,476	0.267
C + G	2,596	0.469
A + U	2,941	0.531

A codon is a series of three nucleotides (triplets) that encodes a specific amino acid residue in a polypeptide chain or for the termination of translation (stop codons). Different factors have been proposed to be related to codon usage bias, including gene expression level, % G+C composition, GC skew, amino acid conservation, protein hydropathy, transcriptional selection, RNA stability, optimal growth temperature and hypersaline adaptation[14]. In computational biology, many statistical methods have been proposed and used to analyze codon usage bias [15]. Methods such as the 'frequency of optimal codons' (Fop)[16], the Relative Codon Adaptation (RCA) [17]or the 'Codon Adaptation Index' (CAI) [18] were used to the predict gene expression levels. Presence of GC formation libraries, AU pair libraries and triplets which shows strong and weak binding were identified in *pkc* gene [19]. The codon AAG counts for 35 times and codes for aminoacid Proline (K).

Genome annotation

Genome annotation was performed for *pkc* gene in order to identify its coding region, PolyAsite, Signal region and Restriction enzymes cutting region. The genome annotation identifies the coding region. So, *pkc* gene sequence was submitted to identify the start and end coding region of Protein kinase epsilon. When the sequence was analyzed it was identified that Protein kinase epsilon coding region starts and ends at 199th and 1983rd positions. Coding region starts with ATG residue which codes aminoacid methionine. In this annotation the restriction enzymes cutting regions was also identified. The enzymes PstI cuts at positions 517, 1291and 1975, EcorI at 1855, SmaI at 1795 and Hind III at 1665. PolyA signaling region and PolyA site was located at 3510 and 3534 position (Figure 3).

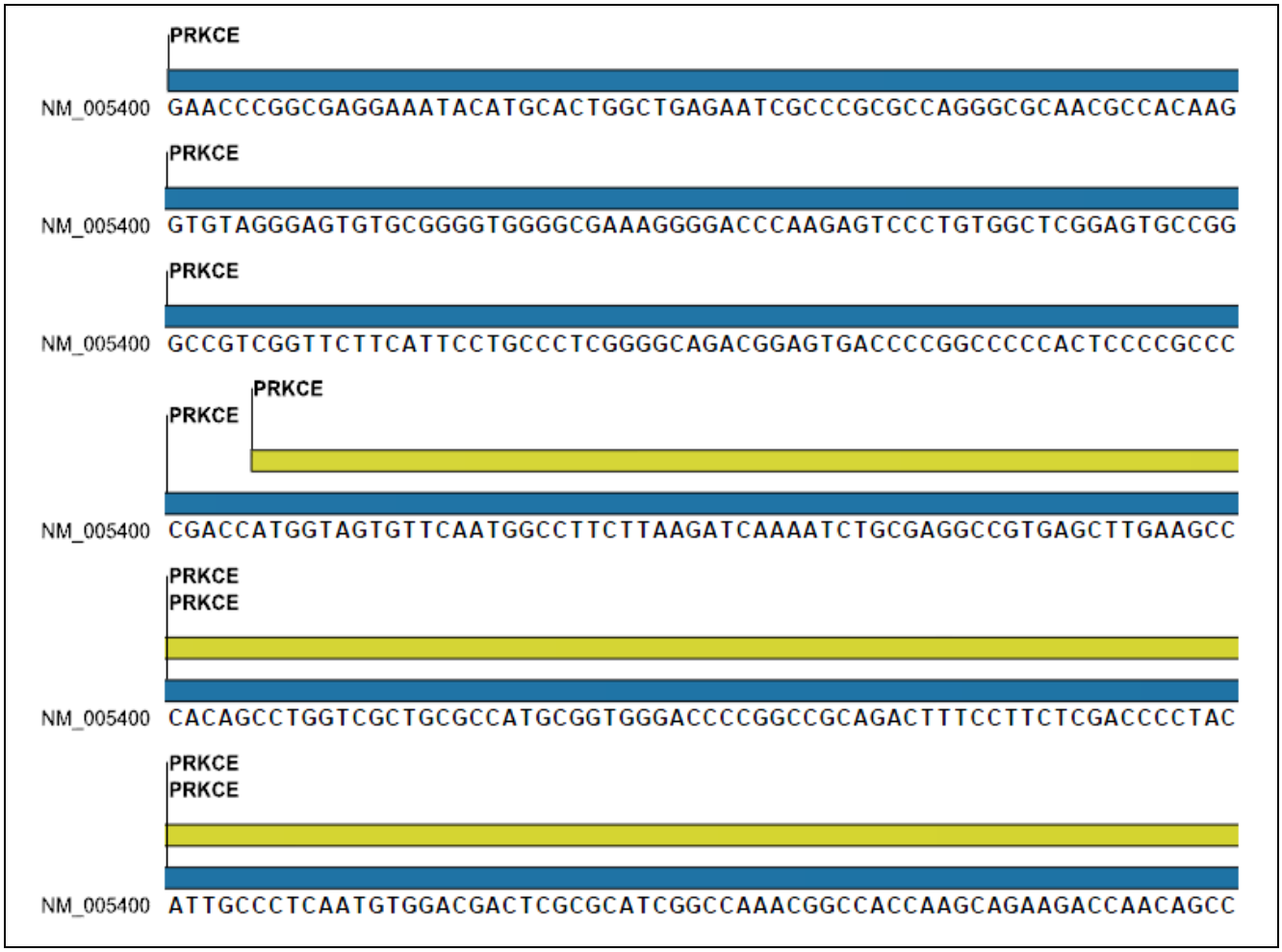

Fig.3. *pkc* genome annotation

Exon is a nucleotide sequence encoded by a gene that remains within the final mature RNA product of that gene after introns have been removed by RNA splicing. Exon refers to both the DNA sequence within a gene and to the corresponding sequence in RNA transcripts. In the *pkc* gene sequence the number of Exons was identified and displayed in Figure 4. Fifteen Exons were present and their position was also identified in chromosome assembly and mRNA (Table 3). Exon 1 position starts at 1 residue and Exon 15 ends at 5520 residue in *pkc* gene sequence[20].

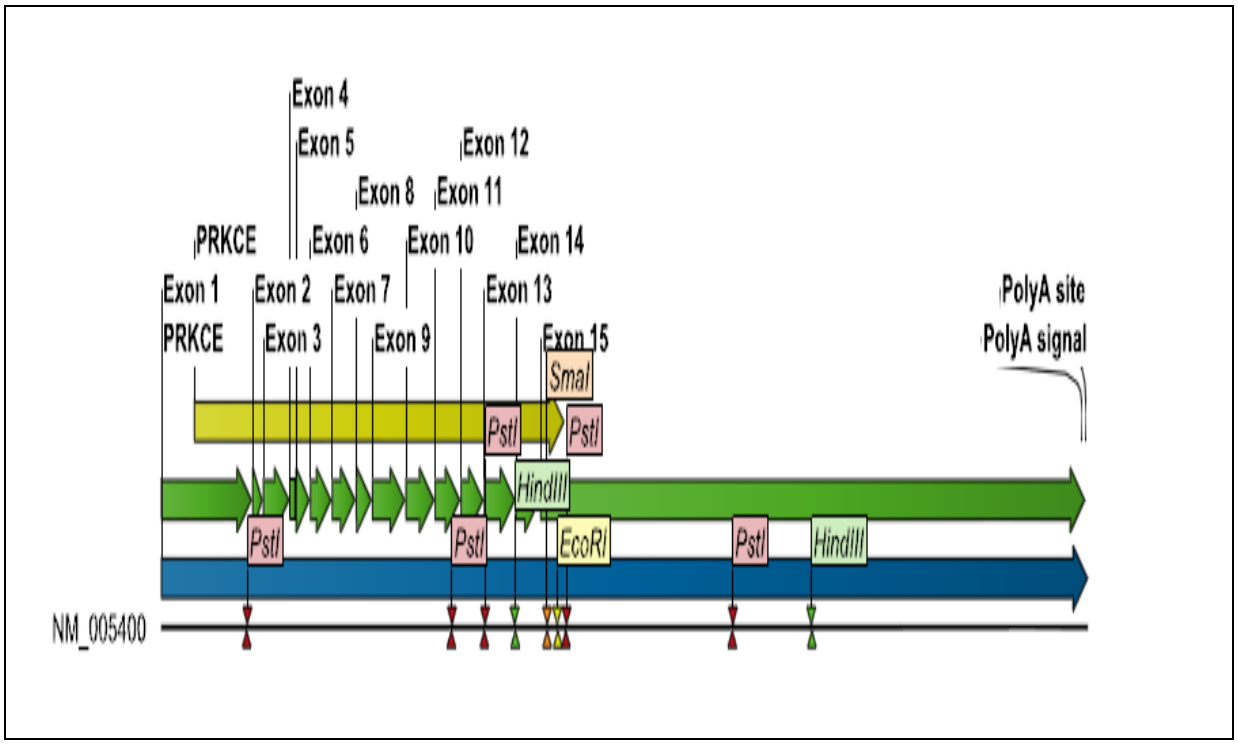

Fig.4.Location of Exon's and in *pkc* gene sequence

Table.3. Position of Exons

Exon No.	**Chromosome assembly (base pair)**	**mRNA (base pair)**
Exon1	45879240.. 45879587	1..545
Exon2	46070139.. 46070202	546..609
Exon3	46203568.. 46203727	610..769
Exon4	46206115.. 46206149	770..804
Exon5	46207435.. 46207520	805..890
Exon6	46211690.. 46211819	891..1020
Exon7	46228543.. 46228685	1021..1163
Exon8	46231681.. 46231777	1164..1260
Exon9	46234601.. 46234800	1261..1460
Exon10	46237483.. 46237656	1461..1634
Exon11	46313347.. 46313501	1635..1789
Exon12	46372232.. 46372370	1790..1928
Exon13	46378180.. 46378368	1929..2117
Exon14	46386745.. 46386891	2118..2264
Exon15	46411874.. 46412020	2265..5520

CONCLUSION

Bipolar disorder is one of the most increasing mood disorders in modern world and it affects most of the people in developing countries. It is evident that PKC plays an important role in signaling systems and alteration in its system plays a major role in pathophysiology of BPD. Hence PKC remains to be a promising target for BPD. Genomic study of *pkc* plays a role to understand its molecular properties and Exons position. In *pkc* gene location of polyAsite region and restriction enzymes positions were also identified. These studies are important for protein analysis and docking studies.

REFERENCES

1. Nick C. and Ian J (1999). Genetics of bipolar disorder. *J. Med Genet*, 36: 585-594.
2. Chandrasekar BVN., Rameshkumar N., Chakravarthy A. and Mukkanti K (2011). Novel synthetic approaches for the synthesis of anti psychotic drug olanzapine. *International Journal of pharma and Biosciences,* 2(3): 426-432.
3. Zarate CA. and Manji HK (2009). Protein kinase C Inhibitors: Rationale for use and potential in the treatment of Bipolar Disorder. *NIH Public Access*, 23(7): 569-582.
4. Nishizuka Y (1992). Intracellular signalling by hydrolysis of phospholipids and activation of protein kinase C. *Science*, 258: 607-614.
5. Altschul SF., Gish W., Miller W. Myers EW and Lipman DJ (1990). Basic local alignment search tool. J. Mol. Biol. 215: 403-410.
6. Thompson JD., Higgins DG and Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through matrix choice. Nucleic Acids Res. 22(22): 4673-4680.
7. Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis editor and analysis program for Windows 95/98/NT. Symp. Ser.41:95-98.
8. Khan MS., Zahra S. and Rashid H (2012). Bioinformatic analysis to discover putative drug targets against diarrheal causative agents. African Journal of Biotechnology. 11(8): 2058-2066.
9. Chen J., Zhang Y. and Shen P (2008). A protein kinase C activity localized to neuropeptide Y-like neurons mediates ethanol intoxication in Drosophila melanogaster. *Neuroscience*, 156:42–47.
10. Manji HK. and Lenox RH (2000). Signaling: cellular insights into the pathophysiology of bipolar disorder. *Biol Psychiatry*, 48:518–530.
11. Manning D., Whyte DB., Martinez R., Hunter T. and Sudarsanam S (2002). The Protein Kinase Complement of the Human Genome. *Science* ,298 (5600): 1912-1934.
12. Stegger M., Price LB., Larsen AR., Gillece JD., Waters AE., Skov R. and Andersen PS (2012). Genome Sequence of *Staphylococcus aureus* Strain 11819-97, an ST80-IV European Community-Acquired Methicillin-Resistant Isolate. *J. Bacteriol. 194(6):* 1625-1626.
13. Dupuis ME. and Moineau S (2009). Genome Organization and Characterization of the Virulent Lactococcal Phage 1358 and Its Similarities to *Listeria* Phages. *Applied and Environmental Microbiology,* 02173-09: 1623–1632.
14. Ermolaeva MD. (2001). "Synonymous codon usage in bacteria". *Curr Issues Mol Biol*, 3(4): 91–94.
15. Comeron JM., Aguadé M. (1998). "An evaluation of measures of

synonymous codon usage bias". *J. Mol. Evol* , 47 (3): 268–274.
16. Ikemura T. (1981). "Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the E. coli translational system". *J. Mol. Biol, 151*(3): 389–409.
17. Fox JM. and Erill I (2010). "Relative codon adaptation: a generic codon bias index for prediction of gene expression". *DNA Rev*, 17 (3): 185–196.
18. Sharp PM., Li WH. (1987). "The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications". *Nucleic Acids Res,* 15 (3): 1281–1295.
19. Genome analyzer information (2009) Illumina, Pub. No. 970- 2008- 039.
20. Davuluri RV., Grosse I. and Zhang MQ (2001). Computational identification of promoters and first exons in the human genome. DOI: 10.1038/ng780.