ISSN:1991-8178

Australian Journal of Basic and Applied Sciences

Journal home page: www.ajbasweb.com

Theoretically Construction Expression Vector Consist of Two Reporter Gene as a Fusion Protein

[1]Mona Al-Terehi, [2]Ali H. Al-Saadi, [2]Haider K. Zaidan, [2]Abbas Al-shariafe

[1]*University of Kufa -College of Science,* [2]*University of Babylon, College of Science. Iraq*

ARTICLE INFO

Article history:
Received 28 January 2015
Accepted 25 February 2015
Available online 2 April 2015

Keywords:
bioinformatics, DNA sequence, on line software

ABSTRACT

The study aim to implemented designs for primers and vectors using simple methods and online bioinformatics softwares, new vector design by two fluorescence protein gene, Turbo-GFP vector and RFP vector DNA sequence was used in design and analysis results show that online software were simple and handling with DNA data was easier than proprietary programs, also it offered more than one choosing of primers and data analysis, study conclude can design any vector pattern in simple approach.

To Cite This Article: Mona Al-Terehi, Ali H. Al-Saadi, Haider K. Zaidan, Abbas Al-shariafe., Theoretically Construction Expression Vector Consist of Two Reporter Gene as a Fusion Protein. ***Aust. J. Basic & Appl. Sci.,*** *9(7): 586-595, 2015*

INTRODUCTION

Bioinformatics of biological data has been the most important information's in computerized science, in last decades this field become completed all studies that deal with genomics and proteomics' by storage , analysis and used an algorithm to design primers, probes and vectors also it used in another application such as Evolutionary relationships between species, genetic similarity and differences (bioinformatics, Wikipedia, 2015).

Luscombe and others (2001) refer to the aims of bioinformatics at the beginning of the development this field in three categories First, it organizes data in a way that allows investigators to access existing information and to submit new entries as they are produced. The second aim is to develop tools and programs used in data analysis the third aim is how to use these programs in data analyses and clarified the results with biological systems.

In last decades many bioinformatics software's were created with different features and application in international network or by scientific company, the database of this software is special for the product of company if it is created by this company also it not free online and high cost,

Other Programmers created free online software's with updating database which can use with any database that upload by users, these software's are used for many application, depending on the aims of the study, raw data and its option. Many software for primer design it used users DNA data as database while other types have simple database and others have big database such as NCBI. Vectors analysis and graphic software have been used to crate physical map of vectors and gene carrier also it used to detection restriction site and type of DNA sequence segment.

The Present study aims to use simple methods and sequential software processing to design novel vectors by two fluorescence reporter gene GFP and RFP gene as a fusion protein with confirmatory test and PCR amplification products.

Subjected and software's:

Data and raw DNA sequence :Turbo-GFP vector and Turbo-RFP sequence from Evrogen company http://www.evrogen.com/products/vectors/pTurboGFP-N/pTurboGFP-N.shtml and http://www.evrogen.com/products/vectors/pTurboGFP-C/pTurboGFP-C.shtml, (2013), these sequences were processed in the following software's

Design vectors:

Vector was designed to consist two fluorescence protein genes GFP and RFP as a fusion protein with cytomegalovirus promoter by the following steps.

1- Choosing type of vectors: N-Turbo-GFP vector was chosen as blank vector and C-turbo-RFP as a source of red florescence protein gene.
2- Primer design, primer design using primer3 online software for amplification RFP gene then;
3- Primers passed in PCR primer states to test it if it is valid;

Corresponding Author: Mona Al-Terehi, University of Kufa -College of Science.Iraq
E-mail: Mona_diabetis@yahoo.com

4- Valid primers and DNA sequence of RFP- gene source vectors was passed through the PCR product in sequence manipulation site (primer states ,2014).

5- Turbo-GFP vector sequence was cut using double restriction enzyme NheI-ScaI by restriction digestion site (Sequence Manipulation Suite,2014)

Table 1: Online bioinformatics software's used in vector design.

Subject	Websites
PCR primer design	http://primer3.ut.ee/ http://www.biomol.unb.br/sms2/pcr_primer_stats.html
Vector graphic and design	http://www.biomol.unb.br/sms2/rest_digest.html http://www.addgene.org/analyze-sequence/
PCR amplification	http://www.biomol.unb.br/sms2/pcr_products.html

1- RFP gene amplicon also digested by the same restriction enzyme which cut GFP vector in the same software.
2- RFP gene was ligated in GFP- vector to create new vector.
3- New vector sequence was pass throw addgene – analysis sequence to graphic and physical map.
4- Confirmatory test was performed using the same primer and new vector sequence as raw data by PCR product software.

Results:

The results of primer design were a number of primers which have different annealing temperature ™ and different features, suitable primers set was chosen after it pass through PCR primer state. As fallow in the table (1)

Primer (1) was chosen to amplification RFP gene in PCR product in sequence manipulation site results show that amplicon has 1042 bases as in table (2 and 3), this sequence then digested by double digest restriction enzyme *NheI-ScaI* by manipulation site also,

Table 2: Primers sequences and features of RFP gene primer in Primer 3 software on line.

Design	Primer sequences 5‟→3‟	TM	GC%	GC clump	Self-annealing	Hairpin formation	Single base run
1	*F*- CGT GGA TAG CGG TTT GAC T	63.1	52.63	pass	pass	pass	Pass
	R- ACC TCT ACA AAT GTG GTA TGG C	62.1	45.45	pass	pass	pass	Pass
2	*F*- GCACCAAAATCAACGGGACT	64.2	50.00	pass	pass	pass	Pass
	R- GGTACCGTCGACTGCAGAAT more than 3 self-annealing bases in a row; GGTACCGTCGACTGCAGAAT |||||| TAAGACGTCAGCTGCCATGG	64.77°	50.00	pass	more than 3 self-annealing bases	pass	Pass
3	*F*- GGGAGTTTGTTTTGGCACCA	64.1	50.00	pass	pass	pass	pass
	R- GGGAGGTGTGGGAGGTTTT	64.87	57.89	pass	pass	pass	pass
4	*F*- GCACCAAAATCAACGGGACT	64.03	50.00	pass	pass	pass	pass
	R- GGTACCGTCGACTGCAGAAT more than 3 self-annealing bases in a row; GGTACCGTCGACTGCAGAAT |||||| TAAGACGTCAGCTGCCATGG	64.77	50.00	pass	more than 3 self-annealing bases	pass	pass
5	*F*- GGGAGTTTGTTTTGGCACCA	64.6°	50.00	pass	pass	pass	pass
	R- GGGAGGTGTGGGAGGTTTT	64.78	57.89	pass	pass	pass	Pass
6	*F* - GTAACAACTCCGCCCCATTG	64.14	55.00	pass	pass	pass	pass
	R- ACCACAACTAGAATGCAGTGA	62.18	42.86	pass	pass	pass	pass

Table 3: RFP gene sequence after amplification and double digested by restriction enzyme.

	Product sequences	**description**
1	CGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGG AGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCA TTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTA GTGAACCGTCAGATCCGCTAGCGCTACCGGTCGCCACCATGGTGAGCGAGCTGATTAAGG AGAACATGCCCATGAAGCTGTACATGGAGGGCACCGTGAACAACCACCACTTCAAGTGCA CATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAGACCATGAGAATCAAGGTCGTCG AGGGCGGCCCTCTCCCCTTCGCCTTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCA GAACCTTCATCAAGCACCCTCCGGGCATCCCCGACTTCTTTAAGCAGTCCTTCCCTGAGG GCTTCACATGGGAGAGAGTCACCACATACGAAGACGGGGGCGTGCTGACCGCTACCCAGG ACACCAGCCTCCAGGACGGCTGCCTCATCTACAACGTCAAGGTTAGAGGGGTGAACTTCC CAGCCAACGGCCCTGTGATGCAGAAGAAAACACTCGGCTGGGAGGCCTCCACCGAGACGA TGTACCCCGCTGACGGCGGCCTGGAAGGCGCATGTGACATGGCCCTGAAGCTCGTGGGCG GGGGCCACCTGATCTGCAACCTTGAGACCACATACAGATCCAAGAAACCCGCTACGAACC TCAAGATGCCCGGCGTCTACAACGTGGACCACAGACTGGAAAGAATCAAGGAGGCCGACG	1042 bp product from linear template Untitled, base 395 to base 1436 (D1 - D1)

	ATGAGACCTACGTCGAGCAGCACGAGGTGGCTGTGGCCAGATACTCTACTGGTGGCGCTG GTGATGGAGGTAAAGGTGGAGGAGGTTCCGGACTCAGATCTCGAGCTCAAGCTTCGAATT CTGCAGTCGACGGTACCGCGGGCCCGGGATCCACCGGATCTAGATAACTGATCATAATCA GCCATACCACATTTGTAGAGGT	
2	CTAGCGCTACCGGTCGCCACCATGGTGAGCGAGCTGATTAAGGAGAACATGCCCATGAAG CTGTACATGGAGGGCACCGTGAACAACCACCACTTCAAGTGCACATCCGAGGGCGAAGGC AAGCCCTACGAGGGCACCCAGACCATGAGAATCAAGGTCGTCGAGGGCGGCCCTCTCCCC TTCGCCTTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCAGAACCTTCATCAAGCAC CCTCCGGGCATCCCCGACTTCTTTAAGCAGTCCTTCCCTGAGGGCTTCACATGGGAGAGA GTCACCACATACGAAGACGGGGGCGTGCTGACCGCTACCCAGGACACCAGCCTCCAGGAC GGCTGCCTCATCTACAACGTCAAGGTTAGAGGGGTGAACTTCCCAGCCAACGGCCCTGTG ATGCAGAAGAAAACACTCGGCTGGGAGGCCTCCACCGAGACGATGTACCCCGCTGACGGC GGCCTGGAAGGCGCATGTGACATGGCCCTGAAGCTCGTGGGCGGGGGCCACCTGATCTGC AACCTTGAGACCACATACAGATCCAAGAAACCCGCTACGAACCTCAAGATGCCCGGCGTC TACAACGTGGACCACAGACTGGAAAGAATCAAGGAGGCCGACGATGAGACCTACGTCGAG CAGCACGAGGTGGCTGTGGCCAGATACTCTACTGGTGGCGCTGGTGATGGAGGTAAAGGT GGAGGAGGTTCCGGACTCAGATCTCGAGCT	750 bp linear fragment from linear parent RFP gene, base 198 to base 947 (NheI g|ctagc - SacI gagct|c)
3	CGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGG AGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCA TTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTA GTGAACCGTCAGATCCG	197 bp linear fragment from linear parent RFP gene was removed, base 1 to base 197 (sequence start - NheI g|ctagc).
4	CAAGCTTCGAATTCTGCAGTCGACGGTACCGCGGGCCCGGGATCCACCGGATCTAGATAA CTGATCATAATCAGCCATACCACATTTGTAGAGGT	95 bp linear fragment from linear parent RFP gene was removed, base 948 to base 1042 (SacI gagct|c - sequence end).

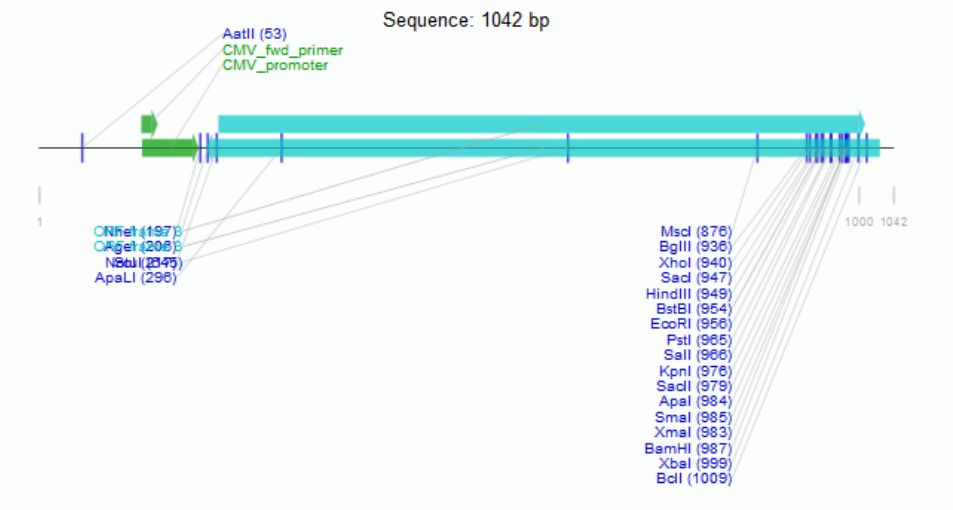

Fig. 1: Amplicon physical map in addgene analysis sequence.

GFP- vector figure(3,4) was digested by double digestion enzyme NheI-ScaI using sequence manipulation site, restriction digested as shown in table(4).

Table 4: Turbo-GFP vector digested by double digested restriction enzymes.

	Product sequences	**description**
1	CAAGCTTCGAATTCTGCAGTCGACGGTACCGCGGGCCCGGGATCCACCGGTCGCCACCATGGA GAGCGACGAGAGCGGCCTGCCCGCCATGGAGATCGAGTGCCGCATCACCGGCACCCTGAACGG CGTGGAGTTCGAGCTGGTGGGCGGCGGAGAGGGCACCCCCGAGCAGGGCCGCATGACCAACA AGATGAAGAGCACCAAAGGCGCCCTGACCTTCAGCCCCTACCTGCTGAGCCACGTGATGGGCT ACGGCTTCTACCACTTCGGCACCTACCCCAGCGGCTACGAGAACCCCTTCCTGCACGCCATCAA CAACGGCGGCTACACCAACACCCGCATCGAGAAGTACGAGGACGGCGGCGTGCTGCACGTGAG CTTCAGCTACCGCTACGAGGCCGGCCGCGTGATCGGCGACTTCAAGGTGATGGGCACCGGCTTC CCCGAGGACAGCGTGATCTTCACCGACAAGATCATCCGCAGCAACGCCACCGTGGAGCACCTG CACCCCATGGGCGATAACGATCGGATGGCAGCTTCACCCGCACCTTCAGCCTGCGCGACGGCG	4092 bp linear fragment from linear parent GFP-vector, base 621 to base 4712 (SacI gagct|c - sequence end).

	GCTACTACAGCTCCGTGGTGGACAGCCACATGCACTTCAAGAGCGCCATCCACCCCAGCATCCT GCAGAACGGGGGCCCCATGTTCGCCTTCCGCCGCGTGGAGGAGGATCACAGCAACACCGAGCT GGGCATCGTGGAGTACCAGCACGCCTTCAAGACCCCGGATGCAGATGCCGGTGAAGAATAAAG CGGCCGCGACTCTAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAA ACCTCCCACCCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTA TTGCAGCTTATAATGGTTACAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTT TCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCATGTATCTTAAGGCGTAAATTGTAAGCGT TAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTAACCAATAGGCCGA AATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCATTT GGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATC AGGGCGAGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAA AGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAA CGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAG CGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCAG GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAAT ATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGT CCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCT CCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGT CCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAG TCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCAT GGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGA AGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAGATCGATCAAGAGACAGGATGA GGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGA GGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCT GTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTG CAAGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTC GACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTC CTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGC ATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCAC GTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCG CGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGAGCATGCCCGACGGCGAGGATCTCGTCGTGA CCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGA CTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCT GAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATT CGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAA ATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATG AAAGGTTGGGCTTCGGAATCGTTTTCCGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTC ATGCTGGAGTTCTTCGCCCACCCTAGGGGGAGGCTAACTGAAACACGGAAGGAGACAATACCG GAAGGAACCCGCGCTATGACGGCAATAAAAAGACAGAATAAAACGCACGGTGTTGGGTCGTTT GTTCATAAACGCGGGGTTCGGTCCCAGGGCTGGCACTCTGTCGATACCCCACCGAGACCCCAT TGGGGCCAATACGCCCGCGTTTCTTCCTTTTCCCCACCCCACCCCCCAAGTTCGGGTGAAGGCC CAGGGCTCGCAGCCAACGTCGGGGCGGCAGGCCCTGCCATAGCCTCAGGTTACTCATATATACT TTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATC TCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACC GCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGC TTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCA AGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAG TGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCG GTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACT GAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACA GGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAAC GCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATG CTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGC CTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTAT TACCGCCATGCAT	
2	TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATT GACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCA ATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCC AAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTA CATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC CATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGG ATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG GGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCCG	591 bp linear fragment from linear parent GFP-vector base 1 to base 591 (sequence start - NheI g|ctagc)
3	CTAGCGCTACCGGACTCAGATCTCGAGCT	29 bp linear fragment from linear parent GFP-vector was removed, base 592 to base 620 (NheI g|ctagc - SacI gagct|c)

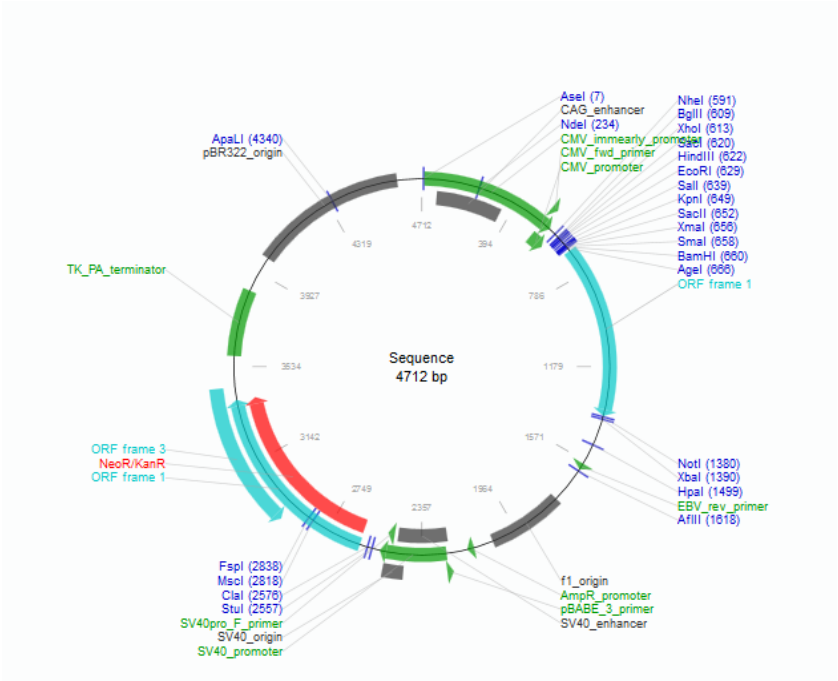

Fig. 2: Physical map of turbo GFP- vector.

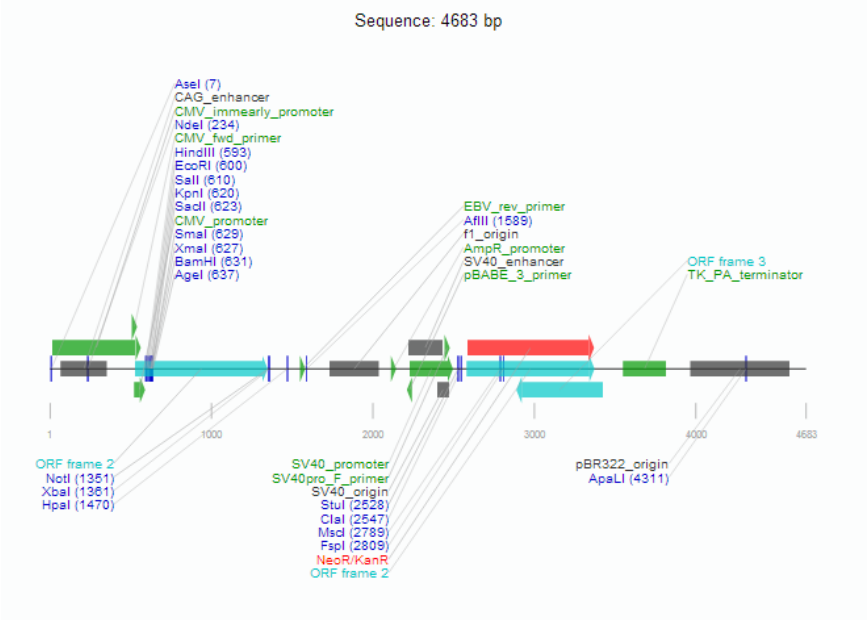

Fig. 3: Physical map of Turbo-GFP vector after double digested.

RFP- gene then ligated between 1 and 2 sequences in the table (3) to create D1 vector as follow sequence table (4) and maps figure (4). Also the features of D1open reading frame, restriction site and types of DNA sequences was at table (5) and (6) its created by analysis sequences in addgene site.

Table 5: D1 vector sequence.

	D1 vector design	
	TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGT AAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGC CAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTAT CATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTT ATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACAT CAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTT GGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT	

Table 5: D1 vector sequence.

ACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCCGCTAGCGCTACCGGTCGCCACCATGGT
GAGCGAGCTGATTAAGGAGAACATGCCCATGAAGCTGTACATGGAGGGCACCGTGAACAACCACCACTTCAAGTGC
ACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAGACCATGAGAATCAAGGTCGTCGAGGGCGGCCCTCTCC
CCTTCGCCTTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCAGAACCTTCATCAAGCACCCTCCGGGCATCCCCG
ACTTCTTTAAGCAGTCCTTCCCTGAGGGCTTCACATGGGAGAGAGTCACCACATACGAAGACGGGGGCGTGCTGACC
GCTACCCAGGACACCAGCCTCCAGGACGGCTGCCTCATCTACAACGTCAAGGTTAGAGGGGTGAACTTCCCAGCCA
ACGGCCCTGTGATGCAGAAGAAAACACTCGGCTGGGAGGCCTCCACCGAGACGATGTACCCCGCTGACGGCGGCCT
GGAAGGCGCATGTGACATGGCCCTGAAGCTCGTGGGCGGGGGCCACCTGATCTGCAACCTTGAGACCACATACAGA
TCCAAGAAACCCGCTACGAACCTCAAGATGCCCGGCGTCTACAACGTGGACCACAGACTGGAAAGAATCAAGGAGG
CCGACGATGAGACCTACGTCGAGCAGCACGAGGTGGCTGTGGCCAGATACTCTACTGGTGGCGCTGGTGATGGAGG
TAAAGGTGGAGGAGGTTCCGGACTCAGATCTCGAGCTCAAGCTTCGAATTCTGCAG
TCGACGGTACCGCGGGCCCGGGATCCACCGGTCGCCACCATGGAGAGCGACGAGAGCGGCCTGCCCGCCATGGAGA
TCGAGTGCCGCATCACCGGCACCCTGAACGGCGTGGAGTTCGAGCTGGTGGGCGGCGGAGAGGGCACCCCCGAGCA
GGGCCGCATGACCAACAAGATGAAGAGCACCAAAGGCGCCCTGACCTTCAGCCCCTACCTGCTGAGCCACGTGATG
GGCTACGGCTTCTACCACTTCGGCACCTACCCCAGCGGCTACGAGAACCCCTTCCTGCACGCCATCAACAACGGCGG
CTACACCAACACCCGCATCGAGAAGTACGAGGACGGCGGCGTGCTGCACGTGAGCTTCAGCTACCGCTACGAGGCC
GGCCGCGTGATCGGCGACTTCAAGGTGATGGGCACCGGCTTCCCCGAGGACAGCGTGATCTTCACCGACAAGATCAT
CCGCAGCAACGCCACCGTGGAGCACCTGCACCCCATGGGCGATAACGATCTGGATGGCAGCTTCACCCGCACCTTCA
GCCTGCGCGACGGCGGCTACTACAGCTCCGTGGTGGACAGCCACATGCACTTCAAGAGCGCCATCCACCCCAGCATC
CTGCAGAACGGGGGCCCCATGTTCGCCTTCCGCCGCGTGGAGGAGGATCACAGCAACACCGAGCTGGGCATCGTGG
AGTACCAGCACGCCTTCAAGACCCCGGATGCAGATGCCGGTGAAGAATAAAGCGGCCGCGACTCTAGATCATAATC
AGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAAT
GAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCAC
AAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTAAGGCGTAAATTGTAA
GCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCA
AAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTA
AAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT
AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGA
CGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAG
TGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCAGGTGGCACTT
TTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAAT
AACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTCCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGT
CAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAA
CCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAT
AGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAAT
TTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGG
CCTAGGCTTTTGCAAAGATCGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGC
AGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCG
CCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTG
CAAGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGA
AGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGA
AAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCG
AAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATC
AGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGAGCATGCCCGACGGCGAGGATCTCGTCGTGACCCA
TGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTG
TGGCGGACCGCTATCAGGACATAGCGTTGGCT
ACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGA
TTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAA
GCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTC
CGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCTAGGGGGAGGCTAAC
TGAAACACGGAAGGAGACAATACCGGAAGGAACCCGCGCTATGACGGCAATAAAAAGACAGAATAAAACGCACGG
TGTTGGGTCGTTTGTTCATAAACGCGGGGTTCGGTCCCAGGGCTGGCACTCTGTCGATACCCCACCGAGACCCCATT
GGGGCCAATACGCCCGCGTTTCTTCCTTTTCCCCACCCCACCCCCCAAGTTCGGGTGAAGGCCCAGGGCTCGCAGCC
AACGTCGGGGCGGCAGGCCCTGCCATAGCCTCAGGTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAA
TTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGA
GCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACA
AAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTT
CAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAC
CGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGG
ACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGA
GCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAG
GCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
TATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGC
CTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTG
CGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCATGCAT

Table 6: Some D1 vector features and its positions.

Features	Site position	
CMV_immearly_promoter	10	562
CAG_enhancer	65	352
CMV_fwd_primer	519	539
CMV_promoter	520	589
EBV_rev_primer	2312	2331
f1_origin	2787	2481

AmpR_promoter	2866	2894
pBABE_3_primer	2980	2960
SV40_enhancer	3181	2966
SV40_promoter	2978	3246
SV40_origin	3145	3222
SV40pro_F_primer	3207	3226
NeoR/KanR	3332	4120
TK_PA_terminator	4298	4567
pBR322_origin	4715	5334

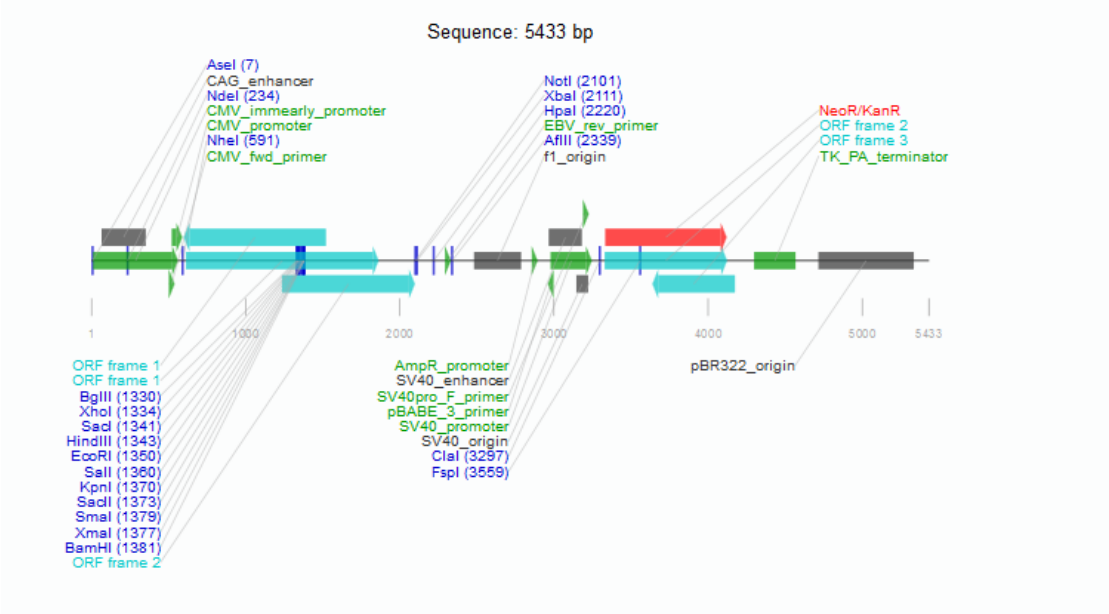

Fig. 4: Linear physical map of D1 vector.

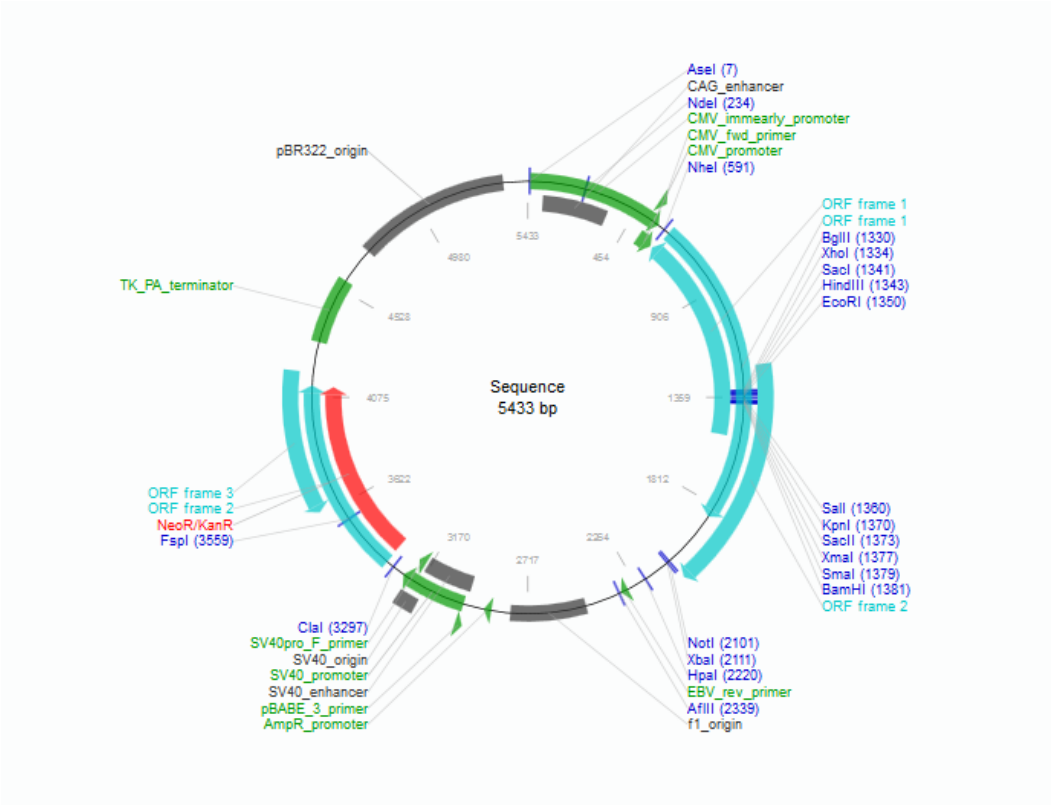

Fig. 5: Circle physical map of D1 vector.

Tables 7: Open reading frame position and unique restriction site in D1 vector.

Enzyme Name	Cut
AseI	7
NdeI	234
NheI	591
BglII	1330
XhoI	1334
SacI	1341
HindIII	1343
EcoRI	1350
SalI	1360
KpnI	1370
SacII	1373
SmaI	1379
XmaI	1377
BamHI	1381
NotI	2101
XbaI	2111
HpaI	2220
AflII	2339
ClaI	3297
FspI	3559

Open reading frame	Start	End
1	1521	598
1	613	1863
2	1235	2098
2	3329	4123
3	4174	3638

Confirmatory test was performed using the same primer that used in amplification RFP-gene; products were larger than the products before ligation as a following in the table (7);

Table 8: Confirmatory test of D1 vector and turbo GFP- vector using d1-d1 primers.

	Product sequence	Description
1	CGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGG AGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCA TTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTA GTGAACCGTCAGATCCGCTAGCGCTACCGGTCGCCACCATGGTGAGCGAGCTGATTAAGG AGAACATGCCCATGAAGCTGTACATGGAGGGCACCGTGAACAACCACCACTTCAAGTGCA CATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAGACCATGAGAATCAAGGTCGTCG AGGGCGGCCCTCTCCCCTTCGCCTTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCA GAACCTTCATCAAGCACCCTCCGGGCATCCCCGACTTCTTTAAGCAGTCCTTCCCTGAGG GCTTCACATGGGAGAGAGTCACCACATACGAAGACGGGGGCGTGCTGACCGCTACCCAGG ACACCAGCCTCCAGGACGGCTGCCTCATCTACAACGTCAAGGTTAGAGGGGTGAACTTCC CAGCCAACGGCCCTGTGATGCAGAAGAAAACACTCGGCTGGGAGGCCTCCACCGAGACGA TGTACCCCGCTGACGGCGGCCTGGAAGGCGCATGTGACATGGCCCTGAAGCTCGTGGGCG GGGGCCACCTGATCTGCAACCTTGAGACCACATACAGATCCAAGAAACCCGCTACGAACC TCAAGATGCCCGGCGTCTACAACGTGGACCACAGACTGGAAAGAATCAAGGAGGCCGACG ATGAGACCTACGTCGAGCAGCACGAGGTGGCTGTGGCCAGATACTCTACTGGTGGCGCTG GTGATGGAGGTAAAGGTGGAGGAGGTTCCGGACTCAGATCTCGAGCTCAAGCTTCGAATT CTGCAGTCGACGGTACCGCGGGCCCGGGATCCACCGGTCGCCACCATGGAGAGCGACGAG AGCGGCCTGCCCGCCATGGAGATCGAGTGCCGCATCACCGGCACCCTGAACGGCGTGGAG TTCGAGCTGGTGGGCGGCGGAGAGGGCACCCCCGAGCAGGGCCGCATGACCAACAAGATG AAGAGCACCAAAGGCGCCCTGACCTTCAGCCCCTACCTGCTGAGCCACGTGATGGGCTAC GGCTTCTACCACTTCGGCACCTACCCCAGCGGCTACGAGAACCCCTTCCTGCACGCCATC AACAACGGCGGCTACACCAACACCCGCATCGAGAAGTACGAGGACGGCGGCGTGCTGCAC GTGAGCTTCAGCTACCGCTACGAGGCCGGCCGCGTGATCGGCGACTTCAAGGTGATGGGC ACCGGCTTCCCCGAGGACAGCGTGATCTTCACCGACAAGATCATCCGCAGCAACGCCACC GTGGAGCACCTGCACCCCATGGGCGATAACGATCTGGATGGCAGCTTCACCCGCACCTTC AGCCTGCGCGACGGCGGCTACTACAGCTCCGTGGTGGACAGCCACATGCACTTCAAGAGC GCCATCCACCCCAGCATCCTGCAGAACGGGGGCCCCATGTTCGCCTTCCGCCGCGTGGAG GAGGATCACAGCAACACCGAGCTGGGCATCGTGGAGTACCAGCACGCCTTCAAGACCCCG GATGCAGATGCCGGTGAAGAATAAAGCGGCCGCGACTCTAGATCATAATCAGCCATACCA CATTTGTAGAGGT	1753 bp product from linear template D1, base 395 to base 2147 (d1 - d1).
2	CGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGG AGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCA TTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTA GTGAACCGTCAGATCCGCTAGCGCTACCGGACTCAGATCTCGAGCTCAAGCTTCGAATTC TGCAGTCGACGGTACCGCGGGCCCGGGATCCACCGGTCGCCACCATGGAGAGCGACGAGA GCGGCCTGCCCGCCATGGAGATCGAGTGCCGCATCACCGGCACCCTGAACGGCGTGGAGT TCGAGCTGGTGGGCGGCGGAGAGGGCACCCCCGAGCAGGGCCGCATGACCAACAAGATGA AGAGCACCAAAGGCGCCCTGACCTTCAGCCCCTACCTGCTGAGCCACGTGATGGGCTACG	1032 bp product from linear template GFP- vector, base 395 to base 1426 (d1 - d2).

	GCTTCTACCACTTCGGCACCTACCCCAGCGGCTACGAGAACCCCTTCCTGCACGCCATCA ACAACGGCGGCTACACCAACACCCGCATCGAGAAGTACGAGGACGGCGGCGTGCTGCACG TGAGCTTCAGCTACCGCTACGAGGCCGGCCGCGTGATCGGCGACTTCAAGGTGATGGGCA CCGGCTTCCCCGAGGACAGCGTGATCTTCACCGACAAGATCATCCGCAGCAACGCCACCG TGGAGCACCTGCACCCCATGGGCGATAACGATCTGGATGGCAGCTTCACCCGCACCTTCA GCCTGCGCGACGGCGGCTACTACAGCTCCGTGGTGGACAGCCACATGCACTTCAAGAGCG CCATCCACCCCAGCATCCTGCAGAACGGGGGCCCCATGTTCGCCTTCCGCCGCGTGGAGG AGGATCACAGCAACACCGAGCTGGGCATCGTGGAGTACCAGCACGCCTTCAAGACCCCGG ATGCAGATGCCGGTGAAGAATAAAGCGGCCGCGACTCTAGATCATAATCAGCCATACCAC ATTTGTAGAGGT	

Discussion:

As a result of great evolution in genetic engineering and biotechnology techniques which have main roles in researches and different applications such as industrial, agriculture, pharmacology and therapeutic applications, the present study was suggested to introduce new approaches in different sequence designs and manipulation.

Different software's were used to design new sequences such as primers and vectors, most of these software's are used for specific company productions because it has limited database, thus present study introduces simple and new methods for sequence design using online and simple software's. Also present study processed many problems that interference with a designer will be affected in gene cutting, insertion, amplification and ligation in addition of primers and linker design.

As intuitive in scientific methods vectors choosing in cloning was based on the aims of study, gene size and its types also host types have major roles. In present study vector was choses according to aims of the study then many basics were dangled, first when we search in a commercial company of vectors; large numbers and types of vectors were found, many of these were specific and others were blank vectors, which choosing to present study because the model of gene therapy would perform.

Tow vectors were chosen from Evrogen company production (http://www.evrogen.com/).Turbo green fluorescence protein (GFP) and Turbo Red florescence protein (RFP).

The important part of these designs were primers design, some basics were dependent on it, most important was TM of annealing temperature every primer set have approximate TM to employ accurate amplification it must be lower than 68C°. Also GC% must be lower than 60% (Guo and Bi, 2000**)**.

Self-complement was avoided and hairpin in primer sequence, in present study tow methods used in primer design, then every primer was passed in to the software to make sure of its ability to accurately complement in amplification.

Another company used special calculation of TM in primer design it depended on salt concentration and pH of master mix buffer (biolabbs, 2015) After primer design were finished virtual amplification performed using online software http://insilico.ehu.es/PCR/ using gene sequences available from Evrogene website. Then amplified sequence transfer to other software to analyses new amplification sequencehttp://www.addgene.org/analyze-sequence/ also http://www.bioinformatics.org/sms2/pcr_products.html was used.

The target gene must be inserted in suitable site, vector map gives all sites can be inserted in, such as a multiple cloning site, it is free sequences manipulation for insertion and deletion as researcher like, in order to obtain complement sequences between target gene and vector. Other site can be used but in present study reporter gene choses as model, thus any site can be inactivated it or causes knock out gene must be discarded. Also target gene can be inserted in other loci than the multiple cloning site, from vector map free sequences were inserted between gene (intragenic sequences) have restriction site can be used. In present study use tow enzyme to cut target gene and vector when gene transfer as a fusion protein in order to employ insertion target gene with same direction with reporter gene. The confirmatory test was designed to employ insertion sequence in the suitable site.

This design can be applied to create different patterns of primers and vectors for cloning in prokaryotic and eukaryotic cells, in vivo and in vitro.

REFERENCES

BioFreaks Biochemistry Blog (v.Beta)**,The** GGS LIVE - Making a fusion protein (2010)**,** http://bio-ggs.blogspot.com/2010/11/ggs-live-making-fusion-protein.html

Evrogen pTurboGFP-N vector (2013) http://www.evrogen.com/products/vectors/pTurboGFPN/pTurboGFP-N.shtml

Evrogen, pFusionRed-C vector (2013). http://www.evrogen.com/products/vectors/pFusionRedC/pFusionRed-C.shtml

Guo, B. and Bi, Y. (2000). Methods in molecular biology Cloning PCR products 192,111-119.

http://www.biomol.unb.br/sms2/pcr_primer_stats.html

http://www.evrogen.com/ (2013).

Luscombe, N. M. ; Greenbaum, D, and Gerstein, M. (2001). What is bioinformatics? An introduction and overview, Yearbook of Medical Informatics 83-100.

The European Molecular Biology Laboratory (EMBL) (2015) http://www.embl.de/pepcore/pepcore_services/cloning/pcr_strategy/primer_design/

The new England *biolabs* Inc, restriction endonuclease , (2014) https://www.neb.com/products/restriction-endonucleases/restriction-endonucleases.

The nonprofit plasmid repository addgene , analysis sequence (2015) http://www.addgene.org/plasmid_protocols/PCR_cloning/

The sequence manipulation site , PCR primer states (2014).

The sequence manipulation site, PCR product (2014) http://www.biomol.unb.br/sms2/pcr_products.html

The sequence manipulation site, restriction digested (2014) .http://www.biomol.unb.br/sms2/rest_digest.html

Bioinformatics , Wikipedia the free encyclopedia (2015), http://en.wikipedia.org/wiki/Bioinformatics