Identification of *Klebsiella granulomatis* strain K22-14 16S rRNA gene and its primer design

N.Mehta[1], A.Raikwar[1], P.Chauhan[1], S.Kushwaha[1] and K.R.Pardasani[2]

[1]Department of Bioinformatics, Maulana Azad National Institute of Technology, Bhopal - 462051, Madhya Pradesh, India

[2] Department of Mathematics and Computing Applications, Maulana Azad National Institute of Technology, Bhopal - 462051, Madhya Pradesh, India.

Abstract:

Klebsiella granulomatis is gram-negative bacteria of the genus *Klebsiella* which causes sexually transmitted disease (STD) *Donovanosis* and urinary tract infection in older persons. In the present work, *in silico* approach for primer designing has been implemented; to gather more information about the bacterium. Primer plays an important role to initiate the process of Polymerase Chain Reaction (PCR), which amplifies DNA sequences. Proper primer designing is one of the most important parameter for successful DNA sequencing. Plethora of bioinformatics programs for PCR primer design reflects the central role of PCR experiments in modern molecular biology and in the –omics era. The work summarizes the flowchart approach to design primer for *Klebsiella Granulomatis* strain, to study its characteristics and generate specific primer sequences using bioinformatics tool. Result reveals a dire need to develop strategies through wet lab experiments to obtain effective drug target for curation of this deadly disease *Donovanosis.*

Keywords: *Donovanosis, Klebsiella Granulomatis*, *Polymerase Chain Reaction and Primer.*

Introduction:

Bioinformatics is an interdisciplinary research area which may be broadly defined as the interface between biological and computational sciences and the solutions of which are obtained using computational tools, software and systems. One of the most significant roles of Bioinformatics is to design and generate primer sequences. Primers are short oligonucleotides which are responsible for amplification of DNA sequences through polymerase chain reaction (PCR). Amplified DNA sequences are used during experimentation in biotechnology and molecular biology to find possible mutation detection, restriction fragment length polymorphism, develop strategies for drug designing, etc. *Klebsiella granulomatis* bacterium whose primer sequence of K22-14 16S rRNA strain is needed to be designed as it causes sexually transmitted disease (STD) *Donovanosis*. [1] Various Bioinformatics tools and software are available on World Wide Web for designing primer. [2] The motivation behind this work was to prevent the spread of *Donovanosis* by designing primer through *in silico* approach. It is mandatory to perform PCR wet lab experiments for evaluating the result predicted by various computational approaches for validation. [3]

Material and Methods:

Primer and approach for its design:

A primer is a single stranded short synthetic oligonucleotide which consists of 18 – 30 nucleotides and is used in molecular techniques right from PCR to DNA sequencing. Primers are to be designed in such a way that they should be the reverse complement of a region in a DNA template to which it should anneal as shown in figure 1. After the process of annealing, multiple copies of DNA template sequences are produced during PCR experiments. It is useful in DNA sequencing, cloning and construction of a suitable vector and for study of genetic diseases. [4] During PCR experiment, the performance and output mainly depends on the efficiency and the binding affinity of primers. The characteristics which play an important role for an oligonucleotide to act as a primer for PCR depends on various factors such as the duplex stability of mismatched nucleotides and their location, the efficiency with which the polymerase can recognize and extend a mismatched duplex, and the kinetics of association and dissociation of primer-template duplexes at the annealing and extension temperatures. A primer which is not designed properly can result in less or no product due to non-specific amplification or primer-dimer formation.

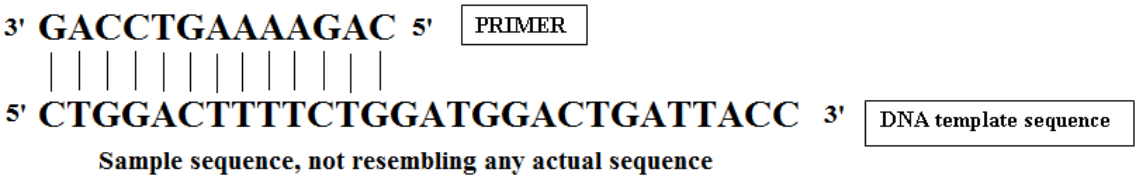

Figure 1: Annealing of primer with DNA template sequence.
The above figure pictorially shows the annealing process taking place between generated primer and the DNA template sequence

Table 1: Main factors and attributes affecting characteristics features while designing primer.

Sequence Feature	Characteristics
Sequence:	• Avoid runs of 3 or more G or C at the 3' end • Avoid a T at the 3' end • Avoid mismatches at the 3' end • Avoid complementary sequences within a primer and between primers • Avoid primer with hairpin loop • Avoid Dimers and false priming • Avoid repetitive sequence for proper specificity
Length:	Between 18–30 nucleotides
GC content:	Between 40–60%
Tm:	Tm = 2°C x (A+T) + 4°C x (C+G), Tm = 45 – 65 ^{0}C Use nearest-neighbour thermodynamic values for estimating range of melting temperature. [
Concentration:	Between 0.1–0.5 µM (0.2 µM)

Therefore following basic factors are to be considered while designing a specific primer as mentioned in Table 1. [5, 6, 7, 8]
Degeneracy in primer sequence should be taken into consideration. Degenerate primers based on the amino acid sequence of conserved regions are used to search for members of a gene family and computational methods are being developed specifically for degenerate primer design. Inosine should not be included in sequencing primers. They either do not work or give poor experimental sequencing result. [9] If mutations are to be introduced via the primer into the PCR product, it is important to leave at least three bases at the 3' end of the primer which are homologous to the template DNA. Mismatches at these sites will greatly reduce the efficiency of PCR experiment. [10]

Specification of primers

1) primer sequences should be unique
2) primers should flank the sequence of interest
3) primers that match multiple sequences will give multiple products
4) primers should not have self-annealing regions within each primer
5) pairs of primers should not anneal each other to form the dreaded "primer dimers"
6) It must be capable of extension by Taq DNA Polymerase.

Identification of bacterial organism Klebsiella granulomatis & details of Donovanosis

The flowchart approach was developed as shown in figure 2.

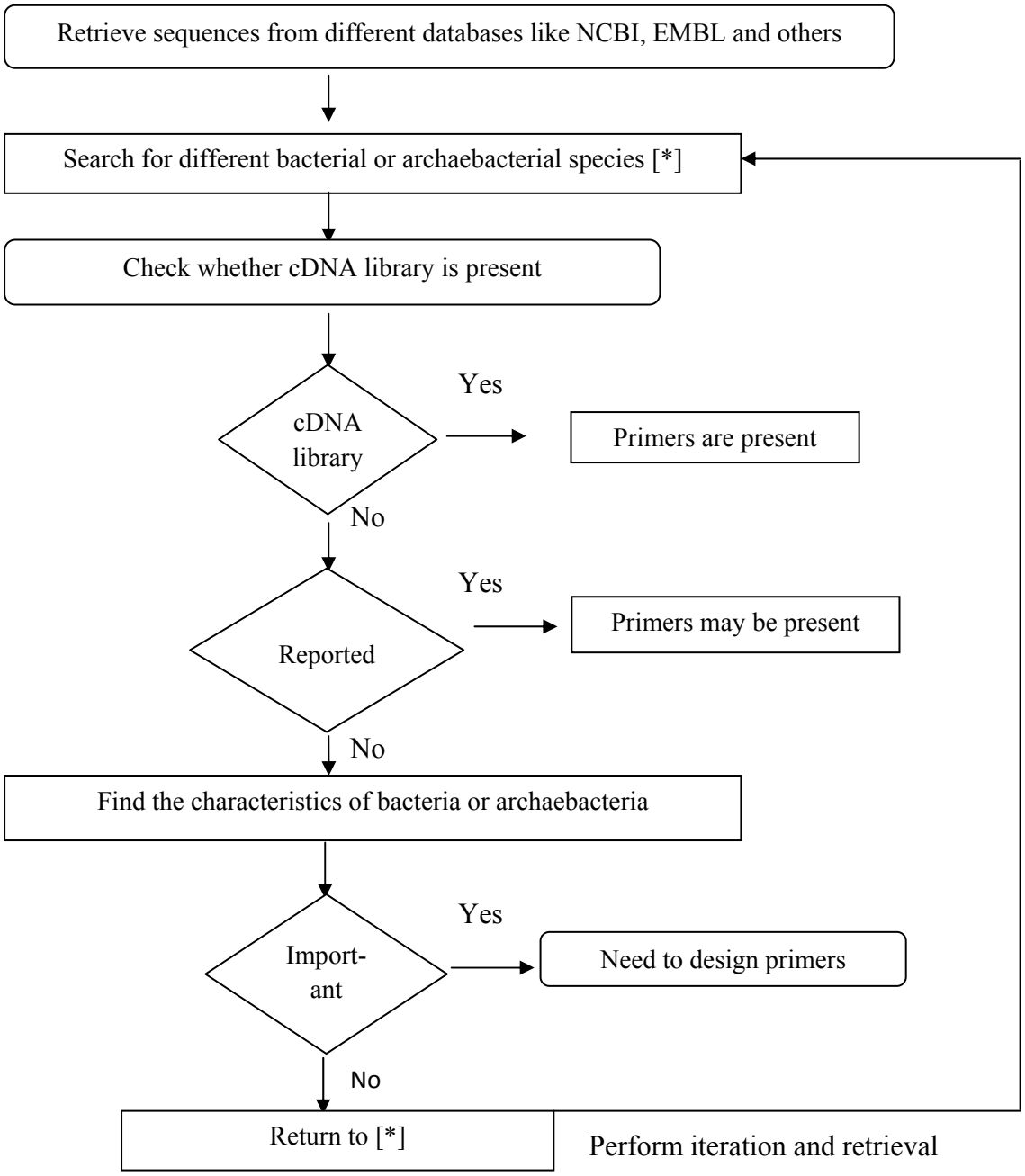

Figure 2: Flowchart approach developed for selection of bacterial organism for primer design

The NCBI database is used to retrieve sequences and finding out valid sequences which does not have cDNA library and whose complete sequence is present. Searching with the above mentioned flowchart approach; *Klebsiella granulomatis* strain K22-14 16S ribosomal RNA gene was obtained.[11]

General details shown on NCBI webpage:

1) GenBank ID - EU333881.1
2) cDNA library is absent
3) Length of complete sequence - 1471 base pairs.
4) Date of availability of sequence in NCBI - 13 Jan 2008
5) Isolation source: Soil from cold desert of Indian Himalayas

Primer3 Output

```
No mispriming library specified
Using 1-based sequence positions
OLIGO            start  len      tm     gc%   any   3'  seq
LEFT PRIMER        301   20   60.02   55.00  3.00  1.00 CAGCCACACTGGAACTGAGA
RIGHT PRIMER       504   20   60.02   55.00  4.00  1.00 GTTAGCCGGTGCTTCTTCTG
SEQUENCE SIZE: 1471
INCLUDED REGION SIZE: 1471
```

Figure 3: Output shown by Primer3

Note: None of the figure is copied from any research paper, book, websites, etc

6) Taxonomy Id – 39824

Website link:

www.ncbi.nlm.nih.gov/nuccore/164504839

Klebsiella granulomatis is gram-negative, rod shaped bacteria of the genus *Klebsiella* known to cause the sexually transmitted disease (STD) *Donovanosis*. [12] It is also called *Calymmatobacterium granulomatis* and it ranks 2nd to *E.Coli* for causing urinary tract infection in older persons. [13] It has cytoplasmic membrane, thin peptidoglycan layer and polysaccharide-based capsule and it is pathogenic for people with chronic pulmonary disease, enteric pathogenicity and rhinoscleroma. [14] Feces are the most significant source of patient infection, followed by contact with contaminated instruments. *Klebsiella* pneumonia is a necrotizing process with a predilection for debilitated people and it has a high mortality rate of approximately 50% even with antimicrobial therapy. The mortality rate approaches 100% for persons with alcoholism and bacteraemia. The proper clinical designation for *Donovanosis* is *granuloma inguinale* and it is result of infection by *Calymmatobacterium granulomatis*. Because of limited medical treatments available, the disease goes untreated and unnoticed. The destructive nature of *Donovanosis* also increases the risk of super infection by other pathogenic microbes. At least one person in India suffers from auto-amputation of genital organs every year because of *Donovanosis*. [15]

Need to design a primer for Klebsiella granulomatis through Primer3

The cDNA library ensures that the primer has been designed for a particular gene and has been used in PCR experiments, but cDNA library is absent for this mentioned gene. [16] *Donovanosis* caused by *Klebsiella granulomatis* is a STD and it affects the sensitive tissues in the genital parts. Wet-lab techniques require sophisticated instruments and large quantity of DNA sequences to perform a frequent experiment which is not easily available, as samples is to be taken from genital part of the infected person which may be against the ethics and may hurt their sentiments. Also the cost for performing wet lab experiments is very high when compared to *in silico* approaches.

Different primer designing tools are available for finding short oligonucleotide sequences i.e. primer. [17, 18] But for the present study, Primer3 was selected (http://frodo.wi.mit.edu/primer3/) as it suggests PCR primer for a variety of applications. For example: To create STSs (sequence tagged sites) for radiation hybrid mapping or to amplify sequences for single nucleotide polymorphism discovery. Primer3 can also select single primers for sequencing reactions and can design oligonucleotide hybridization probes. Primer3 consider many factors which include melting temperature, length, GC content, 3′ stability, estimated secondary structure, the likelihood of annealing to DNA template sequence or

amplifying undesirable sequences (for example interspersed repeats), the likelihood of primer–dimer formation between two copies of the same primer, and the accuracy of the source sequence. [19] Reasons to choose Primer3 are as follows. 1) It is the most popular used software for primer designing and developed at MIT, USA. 2) It provides sophistication in selection of parameters while designing primers. 3) When performed comparative analysis with other tools, it showed better results.

Primer3 accepts many options that specify which primers are acceptable and better than others. In WWW interface, the user selects these options through text boxes, check boxes, and pull-down menus. Primer3 examines all primer pairs that satisfy the constraints and finds pairs that are closest to the optimum. By default WWW interface tries to balance equally primer length, primer melting temperature, and product length. Primer3 never considers a primer that is unacceptable because of its position and if it seems as though very few primers are even being considered, the user might need to modify the maximum and minimum product size options, or expand the included region.

Result and Discussion:

- Left Primer: CAGCCACACTGGAACTGAGA
- Right Primer: GTTAGCCGGTGCTTCTTCTG
- Length: 20 base pair
- Tm : 60.00 degree C
- GC: 55 %
- Sequence Size: 1471, Included region size: 1471

Klebsiella granulomatis strain K22-14 16S ribosomal RNA gene complete sequence is inputted in Primer3 tool and with default parameters it is submitted on web. The result so obtained is mentioned below in figure3. Primer sequence so obtained have all its parameters in a verified range and this primer sequence must be used in PCR to amplify the infected *Klebsiella granulomatis* strain for wet lab experiments.

Conclusion:

The key to PCR lies in the design of the two oligonucleotide primers. Several parameters including the length of the primer, %GC content and the 3' sequence is need to be optimized for successful PCR. The increasing use of information from the internet and the sequences held in gene databases are practically the starting points while designing primers and reaction conditions for PCR. A number of *in silico* approaches such as Primer3 and others have speed up the process of primer design and to be less troublesome. *Klebsiella granulomatis* strain K22-14 16S ribosomal RNA gene has shown very severe impact affecting human health and this primer sequence should be used to generate multiple copies so that proper curation of *Donovanosis* can be done by developing a potential drug target through various chemo informatics and drug design strategies which will become boon for living world, at large.

Acknowledgement:

We are grateful to express our deep sense of gratitude and indebtedness to Prof. K.R. Pardasani (Former Head, Department of Mathematics and Computing Applications) for support and untiring supervision. Also, we are thankful to all members of Department of Bioinformatics and Department of Mathematics at Maulana Azad National Institute of Technology – Bhopal.

References:

[1] O'Farrell N. *Sex Transm Infect*, 2003; 78: 452 - 457.
[2] Binas M. *Bio techniques*, 2000; 29: 988-990
[3] He Q, Marjamaki M, Soini H, Mertsola J, Viljanen MK. *Bio techniques*, 1994; 17: 82-87
[4] Kamel A. Abd-Elsalam. *African Journal of Biotechnology*, 2003; Vol. 2 (5): 91-95
[5] Rychlik, W., W.J. Spencer, and R.E. Rhoads *Nucleic Acids Res*, 1990; 18: 6409-6412.

[6] Chen H, Zhu G. *Bio techniques*, 1997; 22: 1158-1160
[7] Linz U, Delling U, Rubsamem-Waigmann J. Clin. Chem. Clin. Biochem. 1990; 28: 5-13
[8] SantaLucia JJr, Allawi HAT, Seneviratne PA. *Biochemistry* 1996; 35: 3555-3562
[9] C W Dieffenbach, T M Lowe and G S Dveksler *Genome Res.*, 1993; 3: S30-S37
[10] Erlich HA, Gelfand D, Sninsky JJ *Science* 1991; 252: 1643-1651
[11] M. Drancourt, C. Bollet, A. Carta and P. Rousselier, *International Journal of Systematic and Evolutionary Microbiology*, 2001; 51: 925–932
[12] Ryan KJ; Ray CG. *Sherris Medical Microbiology* 2004 (4th ed.) McGraw Hill; 370
[13] Ian M. Mackay, Gerry Harnett, Neisha Jeoffreys, Ivan Bastian, Kadaba S Sriprakash, David Siebert and Theo P. Sloots. *Clinical Infectious Diseases* 2006; 42: 1431–1438
[14] R. Podschun and U. Ullmann, *Clinical Microbiology Reviews*, 1998; 11: 589–603
[15] TS Chandra Gupta, Thrimurthulu Rayudu, Srinivasa V Murthy *Indian Journal of Dermatology, Venereology and Leprology*, 2008; 74: 490 – 492
[16] Clifford F. Brunk, Jinliang Li, and Erik Avaniss-Aghajani, *Curr. Issues Mol. Biol*, 2002; 4: 13-18
[17] Gorelenkov V, Antipov A, Lejnine S, Darasella N, Yuryev A. *Bio techniques*, 2001; 31: 1326-1330
[18] Lowe T, Sharefkin J, Yang SQ, Dieffenbach CW. *Nucleic Acids Res*. 1990; 18: 1757-1761
[19] Steve Rozen and Helen Skaletsky, *Bioinformatics Methods and Protocols*, NJ, vol. 132: 365-386