

Variations in *Serratia marcescens* differentiation using different primers

Hassan Mohabatkar, MS, PhD, Anabita Hamidi, MS, Reza Mohammadzadegan, MS.

ABSTRACT

Objective: To study primer sequences (1060, 1247, 1254, 1281, 1283, and 1290) for random amplified polymorphic DNA polymerase chain reaction (RAPD-PCR).

Methods: Twenty-four clinical *Serratia marcescens* (*S. marcescens*) isolates were obtained over a 6-month period from April 2002 to September 2002 from hospitals in the Fars province of Iran. Six primers were used due to *S. marcescens* genome properties, and RAPD-PCR was carried out. The results were subjected to unweighted pair-group method analysis using NTSYSpc 2.02. The primers were blasted with the *S. marcescens* genome, and the primers efficiency was estimated.

Results: The results of blast primers with *S. marcescens* genome sequence showed that primer 1283 had the highest homology and primer 1290 had the lowest homology. Comparing the resulted dendrograms showed that the pattern of the primers to separate isolates was closely related to their sequence homology with the genome and their amount of guanine and cytosine nucleotide content.

Conclusion: There are clear differences in RAPD-PCR results when different primers are used, and it is recommended to consider genomic properties of an organism to design a primer for RAPD-PCR.

Saudi Med J 2007; Vol. 28 (10): 1520-1524

From the Department of Biology, College of Sciences, Shiraz University, Shiraz, Iran.

Received 10th December 2006. Accepted 30th May 2007.

Address correspondence and reprint request to: Dr. Hassan Mohabatkar, Department of Biology, College of Sciences, Shiraz University, Shiraz, Iran. Tel. +98711-2273900 Ext. 7426. Fax. +98711-2280916. E-mail: mohabatkar@susc.ac.ir

Serratia marcescens (*S. marcescens*) is an opportunistic Gram-negative, facultative anaerobic rod belonging to the family *Enterobacteriaceae*. This nosocomial pathogen is often associated with a characteristic red pigment, which has made it a part of history and religious folklore

for appearing as blood.^{1,2} This bacterium causes infections in many immunocompromised hosts.^{2,3} Nowadays, *S. marcescens* has become detected in nosocomial outbreaks in neonatology units, intensive care units, and some other hospital services.⁴⁻⁷ Sensitive and efficient methods for typing pathogenic microbes are important for tracing routes of infections, and for understanding the spread of antibiotic resistance and the evolution of virulence.⁸ Recently, many typing systems have been used to study and detect the *S. marcescens* population's behavior, and this may have caused scientists to adopt pertinent measures to control bacterial populations.^{9,10} Typing methods included both phenotypic and genotypic characterization, and are based on the assumption of relationships between closely related organisms that are determined by a unique characteristic distinguishing them from unrelated isolates. Phenotypic approaches that are based on metabolic or biological characteristics, or both, have included biotyping,¹¹ serotyping,¹² phage typing,¹² bacteriocin typing,¹³ antibiograms,¹⁴ and whole cell fingerprinting,¹⁵ however, their applicability has been limited due to insufficient discrimination between closely related organisms. Recently, genotyping methods such as pulsed field gel electrophoresis,¹⁶ plasmid profiling,¹³ restriction fragment length polymorphisms,¹⁷ ribotyping,¹⁵ and various polymerase chain reaction (PCR) approaches¹⁸ have gained popularity due to their increased sensitivity. The random amplified polymorphic DNA (RAPD) fingerprinting method, a molecular biological technique based on PCR, uses a single oligonucleotide of arbitrary chosen sequence to prime DNA synthesis from pairs of sites, to which it is fortuitously or partially matched.¹⁹ The RAPD method can be useful for whole microbial community monitoring, and has been reported as a reproducible tool for monitoring complex microbial communities.²⁰ The RAPD analysis has the high discriminatory capacity for typing *S. marcescens*

isolates and can be used for assessment of strain relationship.²¹ Now, RAPD is known as a method that uses an arbitrary primer. The main aim of this study is to determine if there is any relationship between a primer's isolates discretion pattern and its sequence similarity with genome sequence, and if it is important to design a primer based on genome properties. Briefly, our aim in this study was to determine if primer properties could affect RAPD results.

Methods. Twenty-four clinical *S. marcescens* isolates were obtained over a 6-month period from hospitals in the Fars province of Iran (from April 2002 to September 2002). The names of the patients were kept secret. The bacteria were isolated from blood, urine, and sputum. The organisms were routinely propagated on Luria-Bertani agar plates or in liquid media at 37°C.²² Genomic DNA was extracted from the various pure bacterial isolates using the CinnaGen Inc. (DNP™ High yield DNA Purification, Iran) kit. The PCR reactions were carried out in a DNA thermal cycler (Technique, United Kingdom). Thirty cycles were used. All PCR conditions included an initial 4 cycles of 94°C for 5 minutes, 36°C for 5 minutes, 72°C for 5 minutes, 94°C for one minute, 36°C for one minute, 72°C for 2 minutes, and at the end of all cycles 72°C for 10 minutes. Electrophoresis of 15 µl of the mixture was performed using Agarose 0.8% gel in 0.5 Tris-Borate-EDTA (TBE) buffer.²³ A 100bp DNA ladder (5 µl, Fermentas, MBI) was used as a size marker. The DNA was detected by staining with ethidium bromide and gels were photographed under ultraviolet light. Six primers were used due to *S. marcescens* genome properties. Those were 1060 (5'-CCCGGGATAA),²² 1247 (5'-AAGAGCCCGT),²⁴ 1254 (5'-CCGCAGCCAA),²² 1281 (5'-AACGCGCAAC),²⁵ 1283 (5'-GCGATCCCCA),²² and 1290 (5'-GTGGATGCGA).²⁵ Two primers, 1254 and 1283, consisted of 70% guanine and cytosine (G+C), while the others are consisted of just 60% G+C. Each band on the gel was considered a DNA marker and was scored across all samples. Bands were recorded as present (1) or absent (0). Very faint bands were not considered for final scoring. Molecular weights of the bands were estimated using the Gene Ruler 1 kb DNA ladder (Fermentas, UK) as standard. All amplifications were repeated at least twice, and only reproducible bands were considered for analyses. The data set of isolates and reproducible bands were used to calculate pair-wise similarity coefficients following Jaccard.²⁶ This matrix of similarity of coefficients was subjected to unweighted pair-group method analysis (UPGMA) to generate dendrograms using the average linkage procedure. The standardized data matrix was used to calculate correlations among

variables. These correlations were subjected to Eigen vector analysis to extract the first 3 most informative principal components. These 3 principal components were plotted in the 3 possible combinations to study the pattern of variations observed among the isolates. All the numerical taxonomic analyses were conducted using the computer program NTSYSpc, version 2.02 (Exeter Software, New York).

The primers were blasted with *S. marcescens* genome, and the primers efficiency was estimated. Firstly, the complete *S. marcescens* genome sequence was fetched from the Sanger Institute web site (http://www.sanger.ac.uk/Projects/S_marcescens/Sma.art), and then BioEdit software (version 12.0) was used to blast primers. A local nucleotide database file was created, and subsequently a local blast was used to blast the primers. At NCBI local BLAST window, Tblastn program was used, and match and identify matrixes were run separately. A local alignment without gaps simply consists of a pair of equal length segments, one from each of the 2 sequences being compared. A modification of the Smith-Waterman or Sellers algorithms finds all segment pairs whose scores cannot be improved by extension or trimming. These are called high scoring segment pairs (HSPs). In the limit of sufficiently large sequence lengths m and n , the statistics of HSP scores are characterized by 2 parameters, κ and λ . Most simply, the expected number of HSPs with score at least S is given by the formula:²⁷

$$E = Kmn e^{-\lambda S}$$

We call it E-value for the score S . The BLAST programs take this approach in calculating the database E-value. Notice that for DNA sequence comparisons, the length of database records is largely arbitrary, and therefore this is the only real tenable method for estimating statistical significance. The E-value is the number of hits that would be expected to have a score equal or better than this by chance alone. A good E-value is much less than one, around one is what we expect just by chance.²⁷

Results. In the *S. marcescens* genome, there were 15 sequences corresponding to primer 1254, 10 to 1283, 3 to 1281, and 1290, and only 2 to 1247. There was no corresponding sequence to primer 1060. The E-values of primer 1290 were 1.4 and 1.3 in match and identify matrixes. This primer scores were 23 in match matrix and 24 in identify matrix. These numbers indicated the lower specificity of this primer against other primers. There were different E-values and scores for primers 1254 and 1283, as shown in **Table 1**. Results indicate that the sequence of primer 1283 has the best coordination with the *S. marcescens* genome sequence. All *S. marcescens* genomes were amplified with the primers used, and

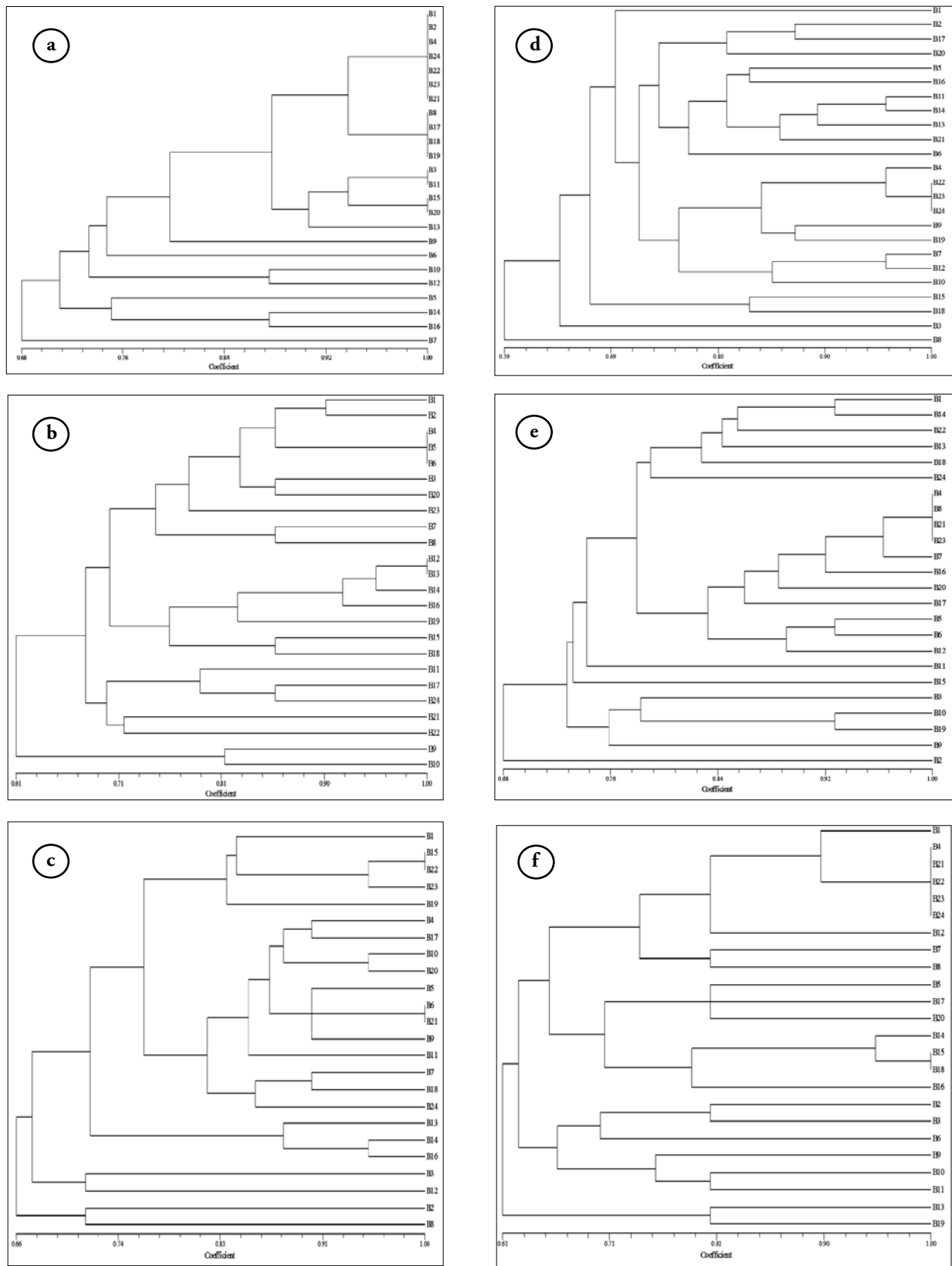


Figure 1 - Dendrograms obtained using each primer, a) primer 1060, b) primer 1247, c) primer 1254, d) primer 1281, e) primer 1283, f) primer 1290.

Table 1 - Homology of primer sequences *Serratia marcescens* genome sequences.

Reaction (primer)	G+C (%)	E-value*		Scores*	
		Match	Identify	Match	Identify
1060	60	0.034	0.032	29	29
1247	60	0.56	0.56	25	25
1254	70	0.012	0.009	30	31
1281	60	0.38	0.39	25	25
1283	70	0.009	0.007	31	31
1290	60	1.4	1.3	23	24

*Calculations were done in 2 cases, match and identify matrix.

we drew dendrograms of these primers. There were 6 primers, and so 6 discrete dendrograms were drawn (Figure 1). Also, a mixed dendrogram of primers 1283 and 1254, showing the most similarity with the *S. marcescens* genome (Figure 2) was drawn, and a mixed dendrogram of all primers (Figure 3) was also drawn. The dendrogram of primer 1290 also separated isolates, however there were many differences between this dendrogram and other dendrograms. The dendrograms of other primers showed intermediate patterns of isolate separating.

Discussion. There is a reverse relation between E-value and matching percentage of both blasted sequences, however, the scores are directly related to matching percentage of both blasted sequences.²⁷ If the E-value decreases, the matching percentage of blasted sequences will increase, however the higher the scores, the more phylogeny between the sequences. Also, it is recommended to use higher amounts of G+C in primer sequence.²⁸ The genome of *S. marcescens* includes approximately 59.5% of G+C nucleotides, and it is logical that the higher percentage of primers G+C, the higher possibility of primers matching with the genome. It was very clear that both primers 1254 and 1283 consisted of a higher percentage of G+C. It seemed that the higher percentage of G+C could increase complementation chance of primer to the *S. marcescens* genome. A special characteristic of RAPD-PCR is the arbitrary designing of primers, and it seems that different primers separate isolates differently. The primers with the lowest E-values are more similar to the genome sequence, and the primer with the highest E-value is different to the genome sequence. The results gained by using primers 1283 and 1254, with E-values less than one, and the results gained by using primer 1290, with the least percentage of G+C nucleotides content and higher E-values are different. By studying

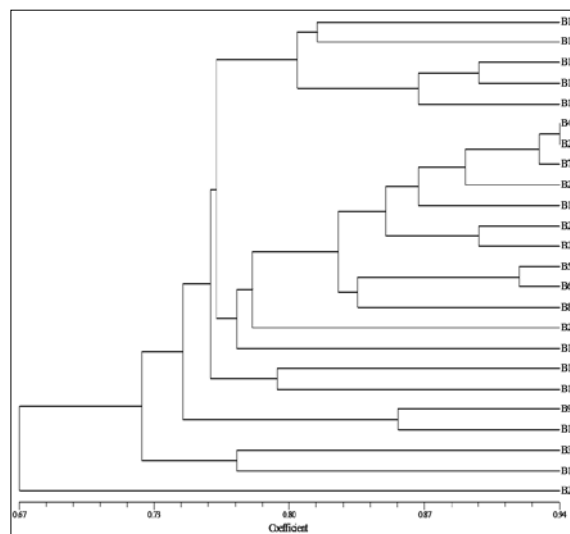


Figure 2 - The mixed dendrogram of primers 1283 and 1254, which had shown the best coordination with *Serratia marcescens* genome.

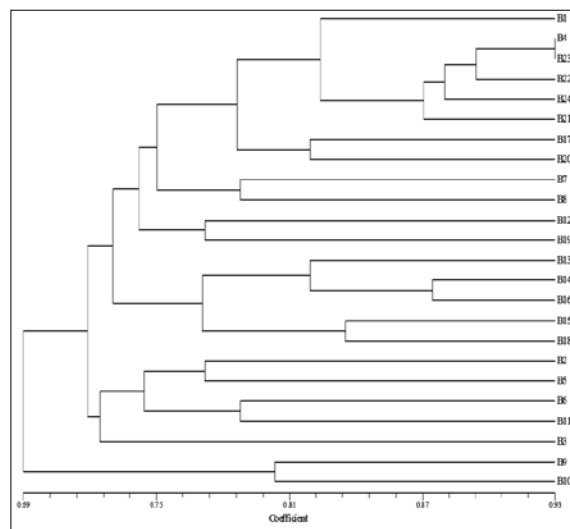


Figure 3 - The mixed dendrogram of all primers.

dendrograms accurately, one can recognize that there are so many differences in the isolates separating patterns; even primers 1254 and 1283 showed different patterns of isolate separating. Thus, it is very clear that different primers made different patterns. This is logical due to the different homology of primers with the *S. marcescens* genome sequence. It seems that primer sequence is an important element to gain RAPD-PCR results.

In conclusion, although RAPD-PCR is a very good method to detect and identify some organism sources even by non-designed primers, however, we recommend accurate design of the primer for optimal results.

