Published in the Slovak Republic
European Journal of Molecular Biotechnology
Has been issued since 2013.
ISSN: 2310-6255
E-ISSN: 2409-1332
2017, 5(1): 4-10

DOI: 10.13187/ejmb.2017.1.4
www.ejournal8.com

Articles and Statements

A High-Throughput PCR-Amplification of GC-Rich DNA Sequences

Alexey S. Afoshin [a, *], Philipp V. Kochetkov [a], Zhanna I. Andreeva-Kovalevskaya [a], Zhanna I. Budarina [a], Marina V. Zakharova [a], Alexander V. Lisov [a], Andrey M. Shadrin [a, b], Alexey A. Leontievsky [a, b]

[a] G. K. Skryabin Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, Russian Federation
[b] Pushchino State Institute of Life Sciences, Russian Federation

Abstract

The PCR amplification of DNA molecules with a high GC content is a complex procedure and often requires optimisation. In last few years several improvements have been developed to facilitate the optimisation of PCR steps. In this work, five high-fidelity commercial DNA-polymerases for the amplification of GC-rich genes were tested. Using the most effective DNA polymerase, we were able to PCR amplify 161 of 187 (87 %) of target genes with a GC content from 57 to 77 % mostly up to 2500 bp long without an optimisation steps. The final yield of specific PCR products was more than 150 ng from 100 µL PCR reactions mix after the purification of specific DNA products by agarose gel electrophoresis.

Keywords: Actinomyces, PCR, GC-rich, high-throughput, DNA polymerase, dimethyl sulfoxide.

1. Introduction

The microorganisms of the *Actinomycetales* order are able to metabolize hard-to-reach substrates such as chitin (Schrempf, 2001), cellulose (Wilson, 1992), oil (Jirasripongpun, 2002), *etc*. This ability makes actinomycetes a valuable source of the genetic material encoding enzymes applicable in food industry, medicine, agriculture, and different areas of biotechnology. Most of A*ctinomycetales* genomes have a high content of guanine–cytosine pairs (GC), which leads to severe problems with successful PCR amplifications of target genes. There a many protocols based on optimisation of PCR thermal conditions (Orpana et al., 2012), primer design strategy (Li et al., 2011), reaction buffer composition and using special additives (Sahdev et al., 2007; Horakova et al., 2011) have been developed for amplification of GC-rich DNA (Strien et al., 2013; Naz, Fatima, 2013). However, currently there is no sufficient data characterizing the PCR amplification of GC-rich genes for large scale cloning using high-fidelity DNA polymerases. In this work, we test five

* Corresponding authors
E-mail addresses: alex080686@mail.ru (A.S. Afoshin), kochetkov-f@bk.ru (P.V. Kochetkov), hemolysin@rambler.ru (Zh.I. Andreeva-Kovalevskaya), budarina@ibpm.pushchino.ru (Zh.I. Budarina), zemskovam@mail.ru (M.V. Zakharova), ssl208@rambler.ru (A.V. Lisov), andrey2010s@gmail.com (A.M. Shadrin), leont@ibpm.pushchino.ru (A.A. Leontievsky)

high-fidelity DNA-polymerases and describe a system approach on application to the high-throughput PCR-amplification of hundreds of target genes with a high GC content.

2. Materials and Methods

Bacterial strains

Nocardiopsis synnemataformans VKM Ac-2518; *Streptomyces avermitilis* VKM Ac-1301; *Streptomyces mobaraensis* VKM Ac-928; *Saccharopolyspora erythraea* VKM Ac-1189; *Saccharopolyspora rectivirgula* VKM Ac-810; *Saccharothrix espanaensis* VKM Ac-1969; *Meiothermus ruber* VKM B-1258; *Thermomonospora curvata* VKM Ac-1241; *Nocardiopsis alba* VKM Ac-1883 strains with sequenced genomes obtained from All-Russian Collection of Microorganisms (VKM).

Dendrogram construction

The dendrogram was based on 16S ribosomal RNA genes sequences and constructed via BLAST (http://blast.ncbi.nlm.nih.gov) using the fast minimum evolution method for clustering (Desper, Gascuel, 2004).

DNA extraction

For the extraction of total DNA, 1–2 g of the biomass of the bacterial strain was homogenised with 2 ml of phenol and liquid nitrogen in the mortar box. The mixture was transferred into 20 ml of buffer containing 20 mM Tris-HCl, pH 8.0, and 20 mM EDTA and extracted four times with one volume of phenol–chloroform–isoamyl alcohol mixture (25 : 24 : 1). After the extraction, the residual amounts of phenol were removed by chloroform extraction. Purified DNA was precipitated with ethanol and dissolved in 200 μL of mQ water. From one to ten nanograms of DNA was used for a single PCR reaction (50 μL).

Primers Design

Primers were designed in accordance with the following requirements: (i) the Tm value determined by OligoAnalyzer 3.1 software (http://eu.idtdna.com/calc/analyzer) must be in the range from +55°C to +60°C; (ii) the melting temperatures of hairpin loops must be lower than +40°C; (iii) the oligonucleotide length must be at least 15 bp or longer. If one of the oligonucleotide parameters did not comply with the above requirements, the annealing region was shifted 10–20 bp further from the open reading frame of the target gene. This procedure can be automated using a Primer3 oligonucleotide design tool (http://bioinfo.ut.ee/?page_id=163&lang=en) (Untergasser et al., 2012, Koressaar, Remm, 2007). DNA oligonucleotides were synthesized by the Evrogen company service.

PCR conditions

DNA amplification was carried out using a C1000Touch PCR machine (BioRad). For each DNA-polymerase a specific program was used. For Q5 polymerase (cat. no. #M0491S, New England Biolabs): initial denaturation at 98°C (30 s); 35 cycles: 98°C (10 s), 60°C (20 s), 72°C (45 s); final elongation 3 min. For Phusion (cat. no. #M0530S, New England Biolabs): initial denaturation at 98°C (30 s); 30 cycles: 98°C (10 s), 55°C (30 s), 72°C (45 s); final elongation 7 min. For TaqSE (cat. no. E314, SibEnzyme): initial denaturation at 95°C (3 min); 35 cycles: 93°C (30 s), 60°C (30 s), 72°C (1.5 min); final elongation 3 min. For Herculase II (cat. no. 600675, Agilent Technologies): initial denaturation at 98°C (2 min); 35 cycles: 98°C (20 s), 60°C (20 s), 72°C (45 s); final elongation 3 min. For PfuUltra II Fusion HS (cat. no. 600670, Agilent Technologies): initial denaturation at 98°C (2 min); 35 cycles: 98°C (20 s), 60°C (20 s), 72°C (30 s); final elongation 3 min. 0.5 μM oligonucleotide primers used for all DNA-polymerases.

3. Results and discussion

As a source of GC-rich templates, we used a collection of highly diverse microorganisms consisting of eight representatives of the *Actinomycetales* order and one strain of the *Thermales* order with available genome sequences in NCBI database (Fig. 1).

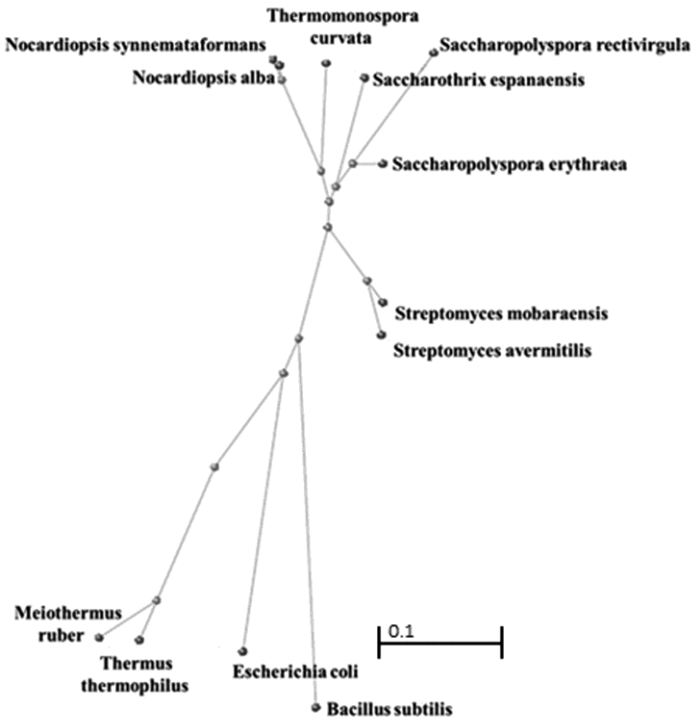

Fig. 1. Radial dendrogram showing the phylogenetic distance between the strains based on nucleotide sequences of 16 rRNA genes.

Table 1. Properties of tested high-fidelity DNA polymerases

Polymerase	Additive	Positive reactions	Error rate*	Price per reaction
Q5	Q5® High GC Enhancer	14 (88%)	100x	$$
Phusion	GC® Reaction Buffer	10 (63%)	50x	$$
TaqSE	No Q5® High GC Enhancer 3% DMSO	2 (13%) 8 (50%) 8 (50%)	n.d.	$
Herculase II Fusion	No Q5® High GC Enhancer 3% DMSO	2 (13%) 15 (94%) 15 (94%)	7x	$$
PfuUltra II Fusion HS	No Q5® High GC Enhancer 3% DMSO	1 (6%) 7 (44%) 7 (44%)	20x	$$$

Information taken from vendors web sites.
*) Error rate shown, relative to Taq-polymerase. n.d. – no data.

The most effective DNA polymerase was determined by testing five commercially available high-fidelity DNA polymerases on 16 target genes of the *Saccharopolyspora erythraea* (VKM Ac-1189) and *Nocardiopsis alba* (VKM Ac-1883) strains (Fig. 2). Q5 DNA polymerase in the presence of the *High GC enhancer* amplifies 14 of 16 target genes (Fig. 2A), whereas Phusion DNA polymerase in GC-reaction buffer amplifies only 10 of 16 genes with a greater amount of nonspecific products (Fig. 2B). TaqSE, Herculase II, and PfuUltra II Fusion HS DNA polymerases supplied without any special buffer or a reagent for amplification of a GC-rich DNA. In the PCRs with the standard buffers the specific products were detected only for one or two of 16 tested genes (Fig. 2C-E). These results are in agreement with the previously reported observation for Taq DNA-polymerase (Naz, Fatima, 2013).

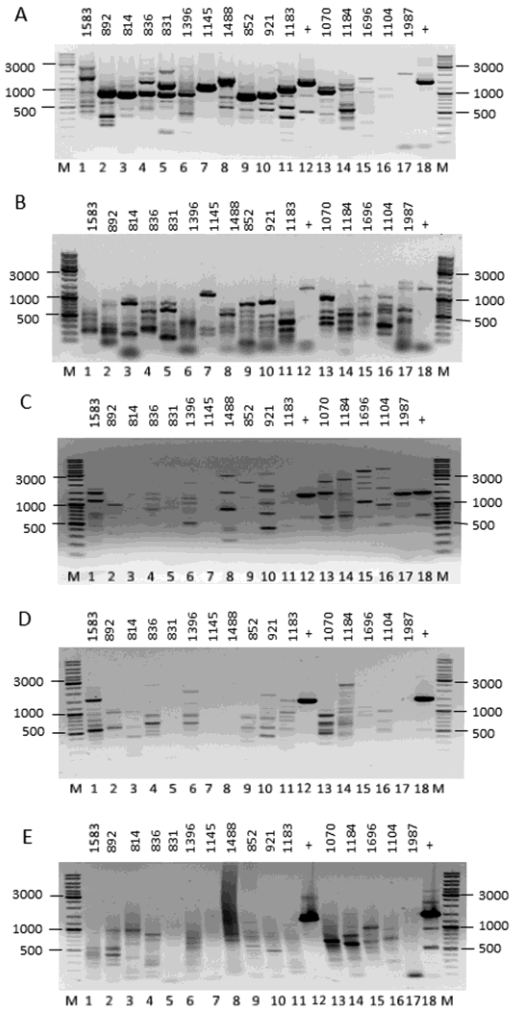

Fig. 2. Results of testing five different high-fidelity DNA polymerases on GC-rich templates. (A) Q5 polymerase; (B) Phusion; (C) TaqSE; (D) Herculase II; (E) PfuUltra II Fusion HS. M – DNA ladder (cat. no. SM0331; Thermo Scientific); tracks from 1 to 11 contain reactions for amplicons of the *Saccharopolyspora erythraea* VKM Ac-1189 strain; tracks from 13 to 17 contain reactions for amplicons from 183 to 187 (sup. Table 1) of the *Nocardiopsis Alba* VKM Ac-1883 strain. Tracks 12 and 18 contain the products of 16S ribosomal RNA gene amplification as positive controls for *Saccharopolyspora erythraea* VKM Ac-1189 and *Nocardiopsis Alba* VKM Ac-1883 strains, respectively. The molecular weight of specific PCR product indicated above the track.

Some additives and company-specific reagents can significantly increase the success of amplification of a GC-rich DNA (Strien et al., 2013; Naz, Fatima, 2013). We repeat the reactions with TaqSE, Herculase II, and PfuUltra II Fusion HS DNA polymerases in presence of *High GC enhancer* and 3 % DMSO. The number of amplified specific products increases up to 10-fold. From eight for TaqSE polymerase to fifteen for Herculase II with 3 % DMSO (Table 1). The pattern of amplified products on agarose gels was slightly different for reactions supplemented with *High GC-enhancer* and 3 % DMSO for same DNA-polymerase (Fig. 3).

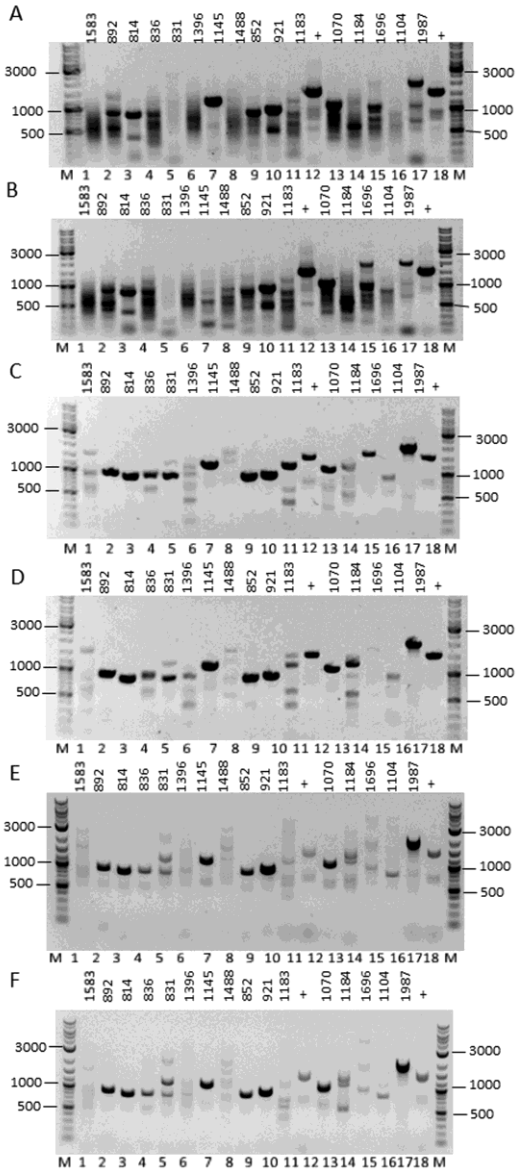

Fig. 3. Results of testing TaqSE (A,B); Herculase II (C,D); PfuUltra II Fusion HS (E,F) in presence of *High GC-enhancer* (A,C,E) and 3 % DMSO (B,D,F). The reactions have same positions as described for Fig. 1.

Because Q5 DNA polymerase amplify large number of the target genes and has higher fidelity in comparison with Herculase II, we test Q5 enzyme for the ability to amplify 187 target genes. The PCR amplifications considered as a successful when at least 150 ng of specific DNA recovered from 100 µL of the PCR reaction mixture purified from the specific DNA band of agarose gel by a QIAgen gel extraction kit (cat. no. 28706). In our conditions, 161 (87 %) target genes successfully amplified.

In eight of nine strains, the number of amplified genes was higher than 79 % (Fig. 4A). For two of nine strains, 100 % of target products successfully amplified. In the case of the *Thermomonospora curvata* VKM Ac-1241 strain, the number of amplified genes was 60 %, because only six of ten reactions gave the sufficient yield of the target DNA. However, in three of four negative reactions, the specific PCR products detected on the agarose gel. This observation indicates that the amount of successfully amplified genes can be increased via the optimisation of PCR conditions.

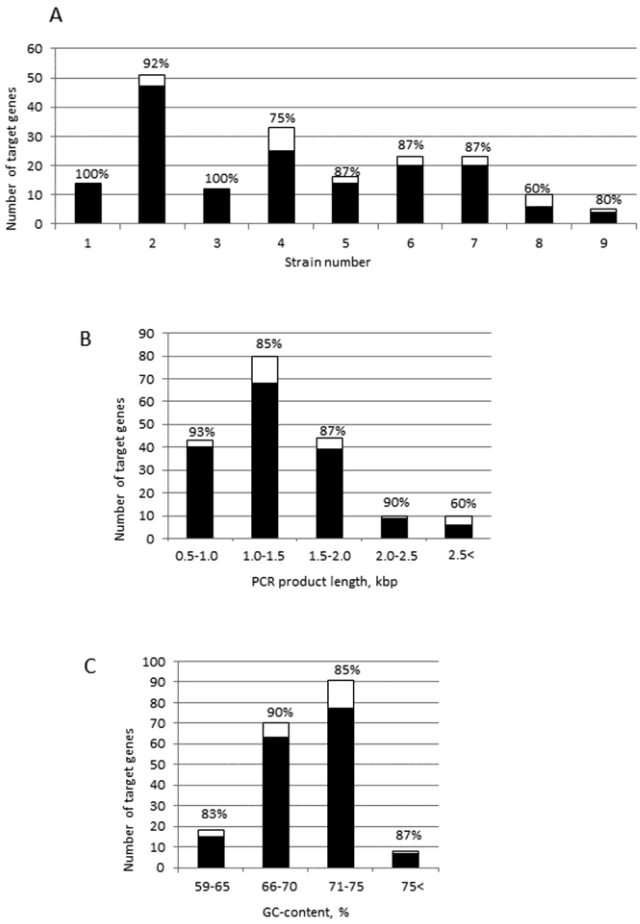

Fig. 4. Correlation between the number of successfully amplified genes and the parameters of DNA amplification. The number of successfully amplified genes shown as black fraction of the columns. The ratio between successfully amplified genes and total number of genes in the set shown above the columns. (A) Dependence on the specificity of the strains: 1 – *Nocardiopsis synnemataformans* VKM Ac-2518 (GC content 72.3 %); 2 – *Streptomyces avermitilis* VKM Ac-1301 (GC content 70.7 %); 3 – *Streptomyces mobaraensis* VKM Ac-928 (GC content 73.3 %); 4 – *Saccharopolyspora erythraea* VKM Ac-1189 (GC content 71.1 %); 5 – *Saccharopolyspora rectivirgula* VKM Ac-810 (GC content 68.9 %); 6 – *Saccharothrix espanaensis* VKM Ac-1969 (GC content 72.2 %); 7 – *Meiothermus ruber* VKM B-1258 (GC content 63.4 %); 8 – *Thermomonospora curvata* VKM Ac-1241 (GC content 71.6 %); 9 – *Nocardiopsis alba* VKM Ac-1883 (GC content 69.6 %). The average GC content in the amplified target genes for each strain is given in brackets. (B) Dependence on the length of target genes. (C) – Dependence on the GC content in the target genes.

The lengths of amplified target genes were in the range from 592 to 3834 bp. The target genes up to 2500 bp long successfully amplified. For products longer than 2500 bp, the portion of amplified genes decreased from 90±3 % to 60 % (Fig. 4B). Presumably, the amplification of long target genes can be improved by increasing the elongation time. However, for high-throughput amplification procedures where the majority of target genes (95 % in our set) are shorter than 2500 bp, we recommend to use a 45-s PCR elongation step to prevent the synthesis of nonspecific products.

The GC content in the whole set of amplified target sequences varied from 57 to 77 %, and the majority (86 %) of target sequences contained from 66 to 75 % GC residues (Fig. 4C). In our set of PCR procedures, we did not identify any significant correlation between the amplification efficiency and the GC content of target genes.

4. Conclusion

Summing up, our results show that Herculase II and Q5 DNA polymerases are most effective high-fidelity DNA polymerases for high-throughput amplification of large number of GC-rich target genes. Presence of additives such us *High GC-enhancer* or DMSO is necessary for successful amplifications of GC-rich DNA. Together with the primer design procedure, our approach affords 87 % of successful amplifications of target genes in amounts sufficient for downstream applications such as cloning, restriction analysis, sequencing, *etc*. The absence of optimisation allows one to apply this protocol for the high-throughput cloning of *Actinomycetales* genetic material and other GC-rich genes.

5. Acknowledgments

This work was supported by the Ministry of Education and Science of the Russian Federation (project no. RFMEFI60714X0013) [Agreement No. 14.607.21.0013]. A.S.A. was supported by the Foundation for Assistance to Small Innovative Enterprises in Science and Technology (project no. 5843 GU 2015 of 10.06.2015).

References

Desper, Gascuel, 2004 – *Desper, R., Gascuel, O.* (2004). Theoretical foundation of the balanced minimum evolution method of phylogenetic inference and its relationship to weighted least-squares tree fitting. *Mol. Biol. Evol.* Vol. 21, № 3. pp. 587-598.

Horakova et al., 2011 – *Horakova, H., Polakovicova, I., Shaik, G.M., Eitler, J., Bugajev, V., Draberova, L., Draber, P.* (2011). 1,2-propanediol-trehalose mixture as a potent quantitative real-time PCR enhancer. *BMC Biotechnology*. Vol. 11, № 4. pp. 1-6.

Jirasripongpun, 2002 – *Jirasripongpun, K.* (2002). The characterization of oil-degrading microorganisms from lubricating oil contaminated (scale) soil. *Lett. Appl. Microbiol.* Vol. 35, № 4. pp. 296-300.

Koressaar, Remm, 2007 – *Koressaar, T., Remm, M.* (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*. Vol. 23, № 10. pp. 1289-1291.

Li et al., 2011 – *Li, L.Y., Li, Q., Yu, H., Zhong, M., Yang, L., Wu, Q.H., Qiu, Y.R., Luo, S.Q.* (2011). A primer design strategy for PCR amplification of GC-rich DNA sequences. *Clinical Biochemistry*. Vol. 44, № 8-9. pp. 692–698.

Naz, Fatima, 2013 – *Naz, S., Fatima, A.* (2013). Amplification of GC-rich DNA for high-throughput family based genetic studies. *Mol. Biotechnol.* Vol. 53, № 3. pp. 345–350.

Orpana et al., 2012 – *Orpana, A.K., Ho, T.H., Stenman, J.* (2012). Multiple Heat Pulses during PCR Extension Enabling Amplification of GC-Rich Sequences and Reducing Amplification Bias Anal. Chem. Vol. 84, № 4. pp. 2081–2087.

Sahdev et al., 2007 – *Sahdev, S., Saini, S., Tiwari, P., Saxena, S., Saini, K.S.* (2007). Amplification of GC-rich genes by following a combination strategy of primer design, enhancers and modified PCR cycle conditions. *Molecular and Cellular Probes*. Vol. 21, № 4. pp. 303–307.

Schrempf, 2001 – *Schrempf, H.* (2001). Recognition and degradation of chitin by streptomycetes. *Antonie Van Leeuwenhoek*. Vol. 79, № 3. pp. 285-289.

Strien et al., 2013 – *Strien, J., Sanft, J., Mall, G.* (2013). Enhancement of PCR amplification of moderate GC-containing and highly GC-rich DNA sequences. *Mol. Biotechnol.* Vol. 54, № 3. pp. 1048-1054.

Untergasser et al., 2012 – *Untergasser, A., Cutcutache, I., Koressaar, T., Faircloth, J.Ye, B.C., Remm, M., Rozen, S.G.* (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Res.* Vol. 40, № 15. pp. 1-12.

Wilson, 1992 – *Wilson, D.B.* (1992). Biochemistry and genetics of actinomycete cellulases. *Crit. Rev. Biotechnol.* Vol. 12, № 1-2. pp. 45-63.