COBISS: 1.01
Agris category code: /

NEW PRIMER COMBINATIONS WITH COMPARABLE MELTING TEMPERATURES DETECTING HIGHEST NUMBERS OF *nosZ* SEQUENCES FROM SEQUENCE DATABASES

Blaž STRES [1], Boštjan MUROVEC [2]

Delo je prispelo 20. novembra 2008, sprejeto 27. februarja 2009.
Received November 20, 2008; accepted February 27, 2009.

New primer combinations with comparable melting temperatures detecting highest numbers of nosZ sequences from sequence databases

We explored existing primer sequences targeting nitrous oxide reductase (*nosZ*) gene in order to explore their capability to recognize variant *nosZ* sequences. Published *nosZ* sequences longer than 380 AA residues were obtained from FunctionalGene Database /Repository (http://flyingcloud.cme.msu.edu/fungene/) and used for explorations with PrimerChart program. The numbers of sequences recovered using all possible forward and reverse primer combinations were determined and the stringency of primer site recognition was further varied by allowing 1, 2, or 3 primer mismatches to DNA binding site. We identified novel primer combinations resulting in satisfactory amplicon length (> 500 bp) and increased sequence recognition capabilities at comparable forward and reverse primer melting temperatures. Overall, this study indicates that current state of the art molecular methods can be and should frequently be further refined by the use of targeted bioinformatic approaches.

Key words: microbiology / molecular biology / denitrification / nitrous oxide reductase / melting temperature / detection

Nove kombinacije začetnih oligonukleotidov s primerljivimi temperaturami taljenja zaznavajo najvišje število sekvenc nosZ v podatkovnih bazah

V tej študiji sva raziskala obstoječe sekvence začetnih oligonukleotidov, s katerimi se pomnožujejo fragmenti gena za reduktazo N_2O (*nosZ*), da bi proučila njihovo zmožnost prepoznavanja variant sekvenc *nosZ*. Objavljene sekvence gena *nosZ* daljše od 380 aminokislninskih ostankov sva pridobila od FunctionalGene Database /Repository (http://flyingcloud.cme.msu.edu/fungene/) in jih analizirala s programom PrimerChart. Raziskala sva število, ki ga prepoznajo posamične možne kombinacije začetnih oligonukleotidov. V nadaljevanju sva spreminjala natančnost prileganja začetnih oligonukleotidov na tarčno DNK tako, da sva dovolila 1, 2, or 3 napačna parjenja med začetnim oligonukleotidom in DNK. Tako sva identificirala nove kombinacije začetnih oligonukleotidov, ki ustvarijo ustrezno dolge fragmente (> 500 bp), s povišano sposobnostjo prepoznavanja sekvenc pri primerljivi temperaturi taljenja začetnih oligonukleotidov. Prav tako so se nakazale nove možnosti za izboljšanje začetnih oligonukleotidov z vnosom novih degeneriranih mest. Ta študija nakazuje, da je novejše molekularne metode možno in tudi potrebno pogosto nadgrajevati s ciljanimi bioinformatskimi pristopi.

Ključne besede: mikrobiologija / molekularna biologija / denitrifikacija / dušikov oksid / reduktaza / temperatura taljenja / zaznavanje

1 INTRODUCTION

Knowledge of abundances and kinds of organisms in an ecosystem is widely recognized as an important step towards understanding the ecology of the system (Prosser *et al.*, 2007). As most prokaryotic species cannot be readily studied by cultivation dependent approaches, culture-independent methods have been widely employed to explore microbial diversity and understand community dynamics. These methods are believed to provide a snapshot of the relative abundances of underlying microbial populations. However, as the number of

1 Univ. of Ljubljana, Biotechnical Fac., Dept. of Animal Science, Groblje 3, SI-1230 Domžale, Slovenia, Ph.D., M.Sc., e-mail: blaz.stres@bfro.uni-lj.si
2 Univ. of Ljubljana, Fac. of Electrical Engineering, Tržaška 25, SI-1000 Ljubljana, Slovenia, Assist.Prof., Ph.D., M.Sc.

sequences deposited to public databases increases and basic molecular approaches to studying microbial communities continue to differ, it is ever increasingly hard to grasp and contemplate the outcomes of numerous studies. A closer inspection of published literature (http://www.ncbi.nlm.nih.gov/PubMed or http://www.sciencdirect.com) reveals that a number of studies deployed unbalanced tools, the unbalance spanning from sampling, over DNA extraction to molecular tool development, their use and interpretation.

NosZ is the crucial enzyme in denitrification that is responsible for conversion of potent greenhouse gas N_2O to molecular dinitrogen. In an accompanying paper (Stres and Murovec, 2007) we explored the differences in predicted melting temperatures of available forward and reverse primers used in amplification of *nosZ* sequences. In this work we explored the capability of these primers to recognize *nosZ* sequences using 0, 1, 2, 3 primer mismatch thresholds in order to identify novel primer combinations resulting in increased sequence recognition capabilities at comparable melting temperatures. To achieve this in controlled manner, the aligned published sequences of nitrous oxide reductase (*nosZ*) gene of sufficient length were obtained from FunctionalGene Database /Repository (http://flyingcloud.cme.msu.edu/fungene/) and used as a model community.

2 MATERIALS AND METHODS

2.1 DATA SELECTION

Literature on the molecular methods used for amplification of target denitrification genes from environmental samples was explored as described in accompanying paper (Stres and Murovec, 2007). The primer sequences were extracted and organized into two dictionaries, containing forward and reverse primer sequences. Each primer sequence was characterized with a primer published name, first base binding location and its DNA binding sequence.

A selection of HMM aligned *nosZ* database currently containing 2025 sequences was downloaded from FunGene Repository / Pipeline (http://flyingcloud.cme.msu.edu/fungene/). The sequences were selected according to primary criterion length (L > 380 AA) and HMM score (s > 20) thus resulting in a dataset containing1985 sequences.

2.2 DATA ANALYSIS

Newly developed software PrimerChart (Murovec and Stres, unpublished) was used for analysis. The numbers of sequences recovered using all possible forward and reverse primer combinations were explored. The stringency of primer site recognition was varied by allowing 1, 2, or 3 primer mismatches to DNA binding site. The combinations of forward and reverse primer pairs were ranked according to the number of sequences detected.

3 RESULTS AND DISCUSSION

In the present study we used previously described primer sets to sample a model microbial community comprised of the longest available and aligned *nosZ* sequences. Figure 1 shows schematic distribution of primer binding sites to the gene of *nosZ* according to *Pseudomonas aeruginosa* 2192 complete genome full nitrous oxide reductase sequence under accession number NZ_AAKW01000028. As it can be seen, the primer binding sites are mainly designed and distributed in the center region of *nosZ* gene spanning between ~1000 bp and ~2000 bp covering roughly 1000 kb region between two conserved *CuA* and *CuZ* sites (Hoeren *et al.*, 1993). In this respect, this region contains the highest sequence coverage.

Figure 2 shows the distribution of hits as detected by different forward and reverse primer combinations used in this study. Each primer combination was assigned a number key after they were sorted according to the number of recognized sequences. As it can be seen, a small number of primer sets could be identified as potential candidates for primers with highest recognition capabilities of sequences from model *nosZ* community.

Figure 1: *The schematic representaion of primer-binding sites of primers used in this study. For more details on primers please see Stres and Murovec (2007).*
Slika 1: *Shematičen prikaz mest naleganja začetnih oligonukleotidov, uporabljenih v tej študiji. Za podrobnosti glede začetnih oligonukleotudov glej Stres and Murovec (2007).*

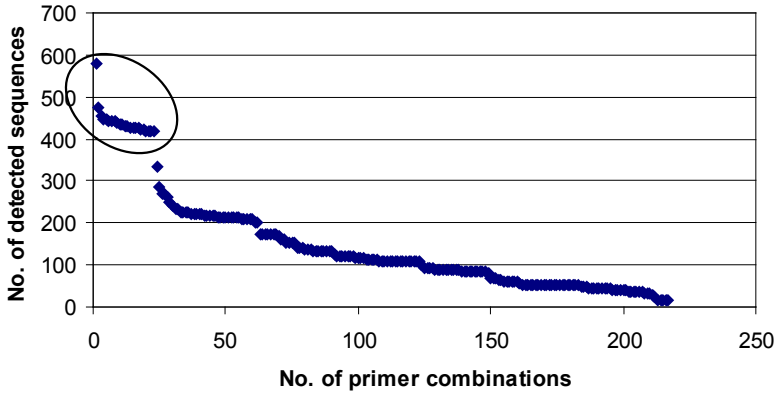

Figure 2: *The schematic representaion of primer-binding sites of primers used in this study. For more details on primers please see Stres and Murovec (2007).*
Slika 2: *Shematičen prikaz mest naleganja začetnih oligonukleotidov, uporabljenih v tej študiji. Za podrobnosti glede začetnih oligonukleotudov glej Stres and Murovec (2007).*

As primers were covering comparable gene region we explored the effect of additional primer mismatches to DNA primer binding site. Increasing the number of recognition sites from 0 to 3 resulted in roughly 50% more detected sequence (dat not shown). This indicates that the degeneracy of already degenerated primers should and could be further increased to incorporate unaccounted variations in primer binding sites. However, this was not the scope of our current research.

In theory, the melting temperature of forward and reverse primers used in a pair should be kept as comparable as possible (Ausubel *et al.*, 1999). Therefore the *in-silico* melting temperatures (Stres and Murovec, 2007) were taken as a measure of comparability of primers melting temperatures and a measure of their compatibility in order to be used as a primer pair in amplification. In this respect the following combination of forward and reverse primers can be suggested for further use in molecular studies (Table 1). However, differences in average melting temperatures of novel primer combinations should be

Table 1: *The combination of forward and reverse primers suggested for further use in molecular studies exploring nosZ diversity in complex samples. Designations ?, +, ++, +++ indicate > 10 °C, < 6 °C,< 4 °C and < 2 °C difference in average melting temperatures of paired oligonucleotides, respectively.*
Preglednica 1: *Predlagane kombinacije začetnih oligonukleotidov za uporabo v molekularnih študijah raznolikosti nosZ in complex samples. Oznake ?, +, ++, +++ kažejo > 10 °C, < 6 °C,< 4 °C in < 2 °C razlike v povprečnih temperaturah taljenja začetnih oligonukleotidov v paru.*

nameF	average FTm	sd	nameR	average RTm	sd	DNA Matches	Δ Tm	Δ Tm suitability	Fragment lenght
24nosZf436	48.52	2.59	1nos1319R	61.94	2.34	580	−13.42	?	883
24nosZf436	48.52	2.59	4nos1527R	59.33	1.97	474	−10.81	?	1 091
35PsNosZ175F	60.07	0.00	1nos1319R	61.94	2.34	454	−1.87	+++	1 144
2Nos1527F	64.65	1.87	17nosZ1773b	61.07	2.84	442	3.58	++	246
27nosZ-F-1181	66.01	1.83	19nosZ1R1421	65.42	2.31	441	0.58	+++	240
27nosZ-F-1181	66.01	1.83	17nosZ1773b	61.07	2.84	439	4.94	+	592
25nosZ-F1211	66.18	1.61	19nosZ1R1421	65.42	2.31	436	0.76	+++	210
25nosZ-F1211	66.18	1.61	17nosZ1773b	61.07	2.84	432	5.11	+	562
2Nos1527F	64.65	1.87	1773R	61.18	2.19	426	3.47	++	246
27nosZ-F-1181	66.01	1.83	1773R	61.18	2.19	423	4.83	+	592
25nosZ-F1211	66.18	1.61	1773R	61.18	2.19	418	5.00	+	562

noted before adopting these combinations for research. Two of the most promising primer combinations resulted in more than 10 °C difference in melting temperatures rendering them least suitable. Both have numerous degenerated sites as can be seen from accompanying standard deviations of their average melting temperatures. The most suitable primer pair satisfying (i) the need for sufficient amplicon length, (ii) comparable average melting temperatures, and (iii) sequence recognition capabilities appears to be NosZ175F and nosZ1319R.

However, the following problems still remain: (i) relatively low number of sequences deposited to public databases, (ii) low number of sequences of sufficient quality (containing only characters A, T, G, C), (iii) unequal melting temperatures of most suitable suggested primer pairs and (iv) relatively low resolution resulting from short sequences amplified by some of the suggested primers. Future studies, especially metagenomic studies and direct reconstructions of genomes from environment, are going to provide valuable data on the uncovered *nosZ* gene variants in environment.

4 CONCLUSIONS

The analysis of primer combinations revealed that existing primers sequence could be further modified to accommodate novel degenerated sites and thus be able to detect a broader sequence diversity. Further, some previously untested primer combinations were explored resulting in higher number of recognized sequences, sufficient length of amplicon (> 500 bp) and comparable melting temperatures, thus indicating their potential for future use in molecular studies. Future work is going to be directed towards detailed analysis of primer binding sites in order to generate combinations of primers targetting widest array of availabe sequences.

Overall, this study indicates that current state of the art molecular methods can be and should frequently be further refined by the use of targeted bioinformatic approaches.

5 REFERENCES

Ausubel F.M., Brent R., Kingston R.E., Moore D.D. Seidman J.G., Smith J.A., Struhl K. 1999. Current protocols in molecular biology. New York, John Wiley and Sons, N. Y.

Hoeren F.U., Berks B.C., Ferguson S.J., McCarthy J.E.G. 1993. Sequence and expression of the gene coding the respiratory nitrous-oxide reductase from *Paracoccus denitrificans*: new and conserved structural and regulatory motifs. Eur. J. Biochem., 218: 49–57

Prosser J.I., Bohanan B.J.M., Curtis T.P., Ellis R.J., Firestone M.K., Freckleton R.P., Green J.L., Green L.E., Killham K., Lennon J.J., Osborn A.M., Solan M., van der Gast C.J., Young J.P.W. 2007. The role of ecological theory in microbial ecology. Nature Rev. Microbiol., 5: 384–392

Stres B., Murovec B. Melting temperatures of degenerated oligonucleotides targetting nitrous oxide reductase (*nosZ*) genes. *Acta Agric. Slov*. (submitted)