Bioscience Methods
BM 2011, Vo.2, No.6 http://bm.sophiapublisher.com

Research Article Open Access

Transcriptome Subtractive Hybridization and Its Reliability Validated by an *E. coli* Model

Yijing Ma, Zhiping Hong, Qin Wu, Xinling Yao
School of Life Science, Ningxia University, Yinchuan, 750021, P.R., China
Corresponding author email: chinanoahl@163.com; Author
Bioscience Methods 2011, Vol.2 No.6 DOI:10.5376/bm.2011.02.0006
Received: 22 Nov., 2011
Accepted: 28 Dec., 2011
Published: 30 Dec., 2011

Preferred citation for this article as:
Ma et al., 2011, Transcriptome Subtractive Hybridization and its Reliability Validated by an *E. coli* Model, Bioscience Methods, Vol.2 No.6 (doi:10.5376/bm.2011.02.0006)

Abstract In the application of subtractive hybridization methods to identify differentially expressed genes, It is no doubt that would increase the efficiency of subtractive hybridization method by simplifying experimental procedures and reducing experimental time-consuming. In this study, we reported a modified method, called transcriptome subtractive hybridization (abbr. TSH). In order to cover different types of RNA as much as possible and reduce the experimental procedure, TSH in the protocol neither uses restriction endonuclease for adopter ligation, nor use oligo-d (T) for the reverse transcription reaction. Instead, tester RNA directly hybridizes with driver single strand cDNA, then the hybrid of RNA/cDNA was digested by *Hae*Ⅲ, residual targeting tester RNA was enriched by reverse transcription PCR amplification. As a result, the experimental protocol was simplified to reduce experimental time-consuming.

To validate use reliability of the TSH, TSH was employed to identify cell-specific expressing ESTs in both *E. coli* JM109 with and without a recombinant plasmid. The results showed that the non-target RNA finally eliminated through direct hybridization and digestion of RNA/cDNA hybrids. Six reference genes within the recombinant plasmid were detected as expected in the tester library constructed from the JM109 RNA with the recombinant plasmid.

Obviously, the verified results showed that the TSH, a modified protocol, would be a reliable and effective method, which can be applied to identify differential gene expression between the two cell types.
Keywords Transcriptome subtractive hybridization; Differential expression; Reliability; An *E. coli* model

Background

Several methods are available to identify gene differential expression between two types of different cells, such as developed and undeveloped cells, or treated and untreated cells. The approach called suppression subtractive hybridization (SSH) is well known used in the fields of Biosciences.

Duplex-specific nuclease (DSN), firstly isolated from the *Kamchatka crab* (Veronika et al., 2008) can be able to cleave double-stranded DNA or DNA in DNA-RNA hybrids, but not to cleave single-stranded DNA or single or double-stranded RNA. Due to specific functions of the DSN, the SSH technique can enrich over 1 000–fold DNAs from rare sequences in a single round of subtractive hybridization (Rebrikov et al., 2004).

Recently, by using SSH technology, many target genes were identified from specific cells in different species, such as melanoma (Landreville et al., 2011), amaranth (Aguilar-Hernández et al., 2011), beetle (Vogel et al., 2011) and mycorrhiza (Murat et al., 2011). Generally, there are 10 000~20 000 rare expressed mRNAs in a typical cell accounting for only less than 20% of the mass of mRNA (Carninci et al., 2000). Although SSH works well in all of these reported studies, it is not yet clear how deep the SSH could dig in the transcriptome of a cell.

To reach more than 0.01% of the fractional concentration of a target gene in a cell by effective enrichment in SSH, the concentration ratio (R) of a target gene requires more than 5 folds between two cDNA prepa-

rations (Wan et al., 2002). However, there is no report to reach this level of the concentration ratio so far.

Furthermore, endonucleases are employed in the procedures of SSH for adaptor ligation, usually, it might be uncertain whether or not the used endonucleases could recognize all kinds of cDNA sequences, which indicated that target genes should be missed in the first step of SSH operation.

SSH normally uses mRNA with poly-adenine (poly(A)) tails as template, therefore RNA without poly(A) tails will be excluded in experimental steps that might lose target genes. It is obvious that SSH might not work well as exact as we expected when transcriptome profiling is taken into account in the research.

In order to break above mentioned limitations, we developed a modified method, named transcriptome suppression subtractive hybridization (TSH) in this research. The new method starts with the total RNA from cells rather than mRNA only and no endonuclease is employed at all in the THS.

To validate use reliability of the TSH, TSH was employed to identify cell-specific expressing ESTs in both *E. coli* JM109 with and without a recombinant plasmid. The results showed that the non-target RNA finally eliminated through direct hybridization and digestion of RNA/cDNA hybrids. Six reference genes within the recombinant plasmid were detected as expected in the tester library constructed from the JM109 RNA with the recombinant plasmid.

1 Results and Analysis

1.1 Screening cDNA libraries

Two cDNA libraries were constructed based on TSH protocol. One library was generated by using cDNA from JM109 that contained pC-SG after having done subtractive hybridization with RNA from JM109 without pC-SG. This library was named JM109$^+$ cDNA library containing approximately 300 colonies. While, the other library, named JM109$^-$ cDNA library containing approximately 200 colonies was cons-tructed by using cDNA from JM109 without pC-SG, after having done subtractive hybridization with RNA from JM109 with pC-SG.

Sixty-eight colonies were randomly selected from the JM109$^+$ cDNA library. Each selected colony was identified by colony PCR by using TSH primers. 37 positive colonies were identified out of the 68 colonies. The results showed in figure 1 exhibited that the sizes of the majority of the inserts varied from 200 bp to x bp. Eighteen inserts from 37 positive colonies were selected for sequencing.

The sequences were aligned by using the BLAST online software to identify non-redundant sequences. Thirty seven identified sequences represented fifteen different transcripts. As shown in Table 1, 12 ESTs

Table 1 BLAST hits of ESTs from JM109$^+$ and JM109$^-$ cDNA library to *E. coli* K12 MG1655

	Gene	Accession number	Annotation
ESTs from JM109$^+$	*wcaK*	G7096	Pedicted colanic acid biosynthesis pyruvyl transferase
	ygjK	G7599	Glycoside hydrolase
	uvrC	EG11063	Excinuclease ABC subunit C
	clpB	EG10157	Potein disaggregation chaperone
	puuC	EG10036	Gma-aminobutyraldehyde dehydrogenase
	mepA	EG10580	Pnicillin-insensitive murein endopeptidase
	yraP	G7644	Lpoprotein
	STSSU	DQ297414	Potato ADP-glc PPase small subunit
	*NPT*Ⅱ	AF354046	Nomycin phosphotransferaseⅢ
	SCR	AF354046	Pasmid DNA replicase
	*HPT*ⅢⅢ	AF354046	Hgromycin phospho-transferase
	35S promoter	AF354046	CaMV 35S promoter
ESTs from JM109$^-$	*dmsB*	EG10233	Dmethyl sulfoxide reductase, chain B
	hslU	EG11881	ATPase component of HslUV protease
	ISIT	CU928160	ISI transposase

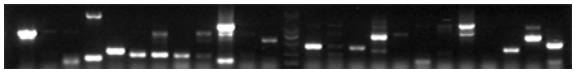

Figure 1 JM109$^+$ cDNA library screening by colony PCR
Note: 1~24: PCR products from different colonies; M: DL 2000 mlecular mrker

were identified from the JM109$^+$ cDNA library, of which 7 come from JM109 genome and 5 from the plasmid pC-SG, while 2 ESTs from JM109$^-$ cDNA library. The annotated function of each non-redundant EST was indicated in the table.

1.2 Targeting genes known

The *STSSU* gene in the pC-SG was employed to be an internal reference because of no *STSSU* gene in the JM109 genome, whereas the *glgC* gene existing in both the JM109 and plasmid genomes was set up as a negative control. As expected, *STSSU* was found only in the JM109$^+$ cDNA library and *glgC* was not detected in either library.

To figure out indirectly the efficiency of the TSH procedure, four of the known genes in the plasmid, *STSSU, SCR, HPT*Ⅱ and *NPT*Ⅱ except *glgC* and *GUS*, were set up as targeting genes. The detection frequency of the targeting genes might represent the efficiency of the TSH procedure, which higher detection rate would correspond to a higher efficiency. The results showed that the four targeting genes were hit in the JM109$^+$ cDNA library.

It was reasonable that *GUS* was undetectable in JM109$^+$ because *GUS* only existed in JM109 genome, which was subtracted in the procedures of THS. As unexpected 35S promoter was detectable, the reason was clear that 35S promoter existing in the JM109$^+$ was no homologous sequence in JM109 genome to be subtracted. Meanwhile, 72 colonies were selected randomly from the JM109$^-$ cDNA library. Sixteen positive colonies were obtained from the tested 72 colonies. Subsequently, 10 colonies were selected to be sequenced. The BLAST results indicated that this selection contained two non-redundant ESTs shown in Table 2.

Table 2 Genes and functions known in plasmid pC-SG

Genes	Functions	Detected expectation
glgC	*E. Coli* ADP-glc PPase	No
STSSU	Potato ADP-glc PPase small subunit	Yes
SCR	Plasmid DNA replicase	Yes
GUS	Beta-glucuronidase	No
*HPT*Ⅱ	Hygromycin phospho transferase	Yes
*NPT*Ⅱ	Neomycin phosphotransferase`	Yes

2 Discussions

As yet, it might be difficult to analyze the expression of genes that are dependent on the expression of another genome. However, bacteria such as *E. coli* can provide a suitable system for such gene analysis due to independent on plasmid DNA for cell growth. In other words, plasmids could constitute an alternative optional "mini-genome" to *E. coli*.

Nowadays, *in silico* tools perform strong abilities to deal with a huge amount of sequencing information, which makes it possible to work with non-subtractive library or full transcriptionist profile and even RNA-Seq. However, more and more studies focused on the effect of a single independent variable with small gradual changes. In this case, subtraction-based library still has its advantage on low-cost and simplicity. To better use this technique, it would be certain that reliability and efficiency of subtractive library has to be improved. The SSH has a low coverage (less than 60%) of targeting mRNAs since it is not all cDNA reverse transcripted from mRNAs to be recognized by endonucleases used in SSH procedures. In addition, mRNA without poly-adenine tails would be losing during reaction of reverse transcription. An alternative to overcome those shortages might be to work with all RNA with or without the polyadenine tails and to work with cDNA without endonuclease digestion.

To make suppression subtractive hybridization working better, evaluating the efficiency of the approach should be necessary. In this study, we employed total RNA from the cells and eliminated endonuclease digestion to evaluate the reliability of TSH with a precise system. Further study for comparison of TSH and SSH on reliability will be done in the future. Although the genome sequences of many *E. coli* strains are available, genes involving in interactions between the chromosomal genome and plasmid DNA have yet to be annotated. TSH approach might be reliable and precise tool to identify differential expressing genes.

3 Methods

3.1 Reactions of reverse transcription

The experiment was carried out using *E. coli* strain JM109 and JM109 that contained the plasmid pC-SG. The pC-SG contains 6 genes as shown in table 2. All the strains constructed in this work were verified by PCR using appropriate primers. The LB medium was supplemented with 30 μg/ml kanamycin. Cells were grown in LB at 37°C with stirring shaker (200 rpm) overnight.

Total RNA was extracted from samples collected during the exponential growth phase using Trizol Reagent (Invitrogen, USA) following the manufacturer's instructions. The RNA yield was determined by measuring the absorbance at point of 260 nm.

To set up the ligation reaction of RNA and adapter, 1 μL 0.8 μg/μL RNA, 1 μL 1.2 pM/L TSH adapter and 6μL ddH_2O were added into a tube. After keeping it at RT for 5 min, 1 μL 10× T_4–DNA ligase buffer, 0.5 μL 350 U/μL T_4–DNA ligase, and 2.5 μL ddH_2O were added in the ligation reaction. The reaction was incubated at 16℃ for 2 h, then terminated at 75℃ water bath for 10 min and kept on ice bath at least 2 min.

To set up the reaction of reverse transcription (RT reaction), 0.5 μL 10 mM/L dNTP, 0.25 μL 40 U/μL RNase inhibitor and 1μL 200 U/μL reverse transcriptase (superscriptII from Promega) were added in the ligation reaction. The RT reaction was incubated at 42℃ for 1h, stopped at 75℃ water bath for 10 min, kept it on icebath for 2 min. 1 μL 60 U/μL RNase H was added into the RT reaction to set up the RNase H digestion reaction. The digestion reaction was incubated at 30℃ for 1 h, stopped at 75℃ water bath for 10 min, kept it on icebath for 2 min.

3.2 Hybridization reactions

To set up the hybridization reaction, 1 μL 1 μg/μL driver RNA was add to the digestion reaction. The hybridization reaction was performed at 94℃ for 2 min, 60℃ for 5 min, 50℃ for 15 min in a thermocycler, then kept on icebath. To digest excessing driver RNA, 1 μL 10 mg/mL RNase A was added in the hybridization reaction and incubated at 37℃ for 1 h.

To collect mixture of cDNA/RNA and cDNA in the reactions, 1 volume fresh mixture of Tris-phenol: chloroform: isoamylalcohol (25:24:1) was added in the RNase A digestion reaction and mixed well. Centrifuged at 8 000 rpm for 10 minutes at 4℃, and transferred supernatant into a new tube, then repeated above collecting step twice, added 2.5 volumes ice-cold ethanol and 1/10 volume 3 M NaAc in collecting solution. After kept the solution at -20℃ for 30 min, centrifuged the solution at 10 000 rpm for 5 minutes at 4℃, having pellet air-dried for 30 min.

3.3 cDNA/RNA digestion

To set cDNA/RNA digestion reaction, dissolving the dried pellet in 8.5 μL ddH_2O, incubating at 75℃ water bath for 10 min and keep on icebath at least 2 min. Add 1 μL universal 10× buffer M, 0.5 μL 10 U/μL *Have*III, incubating the cDNA/RNA digestion reaction at 37℃ for 1.5 h, Adding 1 vol chloroform:isoamyl alcohol and mixing well. Centrifuged at 8 000 rpm for 10 minutes at 4℃. Transferred supernatant into a new tube, adding 2.5 volumes ice-cold ethanol and 1/10 volume 3 M NaAc, keep for 30 min or more at -20℃. Centrifuged at 10 000 rpm for 5 minutes at 4℃, had pellet air-dried for 30 min.

To set the ligation reaction of the second adapter, re-dissolving the dried pellet above in 5 μL ddH_2O, 1 μL 1.2 pM/L TSH adapter. After kept the ligation reaction at RT for 5 min, add 1 μL 10× T_4-DNA ligase buffer, 0.5 μL T_4–DNA ligase, 10 μL ddH_2O. Incubated the reaction at 16℃ for 2 h, terminated the reaction at 75℃ water bath for 10 min and kept on icebath at least 2 min.

3.4 Amplification of tester specific cDNA

To amplify tester specific cDNA with PCR, a reaction was set by adding 2 μL mixture of the ligation reaction above, 2.5 μL 10× PCR buffer, 2 μL 2.5 mM/L dNTP, 1 μL 10 μM/L TSH primers for each of forward and reverse, 16.25 μL ddH_2O and 0.25 μL 5 U/μL *Taq* DNA polymerase in a PCR tube. The reaction was carried out as following denature at 95℃ for 30 sec，annealing at 55℃ for 45 sec, annealing temperature rising 0.5℃ for each cycle, extension at 72℃ for 2.5 min, 10 cycles for total, extension at 72℃ for 10 min and kept at 4℃ for the last cycle. PCR products were recovered with a gel recovery kit (Promega).

3.5 Construction and screen of subtractive library

To set up the ligation reaction of PCR products and vector, a reaction was done with 3 μL recovered PCR products and 1 μL pGEM-T easy following instruction of the pGEM-T easy kit. 2 μL ligation reaction products were transformed into *E. coli* super competent cells by heat shock, transformed cells were plated on a LB Patri dish. Cells grew overnight. Once white-blue clone visible, clones were screened by strain PCR with TSH primers. Positive clones were selected by gel electrophoresis and confirmed by sequencing (ABI 3730XL DNA Sequencer at the Shenggong Gene Center, Shanghai, China).

3.6 List of abbreviations

PCR: polymerase chain reaction; RT: reverse transcripttion; TSH: transcriptome suppression subtractive hybridization; SSH: suppression subtractive hybridization; pC-SG: a recombinant plasmid contained 6 genes as shown in the table 1; JM109: an *E. coli* strain; $JM109^{-}$: JM109 without any plasmid; $JM109^{+}$:JM109 contained plasmid of pC-SG: Other abbreviations as shown in the table 1 and table 2.

Authors' contributions

XLY designed this experiment, primers and adaptor, and wrote and revised the manuscript; YJM performed subtractive hybridization, amplification of tester specific cDNA and establishment of tester specific cDNA library; YJM, ZPH and QW contributed to literature retrieve, BLAST search and ESTs sequence analysis. All authors read and approved the final manuscript.

Acknowledgements

This work was jointly supported by National Sci-Tech Support Program in China (2009BADC5B01, to XY) and National High Tech-Dev Program (863) in China (2006AA100107, to XY).

References

Aguilar-Hernández H.S., Santos L., León-Galván F., Barrera-Pacheco A., Espitia-Rangel E., De León-Rodríguez A., Guevara-González R.G., and Barba de la Rosa A.P., 2011, Identification of calcium stress induced genes in amaranth leaves through suppression subtractive hybridization. J. Plant Physiol., 168(17): 2102-2109 http://dx.doi.org/10.1016/j.jplph.2011.06.006 PMid:21794947

Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., and Hayashizaki Y., 2000, Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes, Genome Res., 10(10): 1617-1630 http://dx.doi.org/10.1101/gr.145100 PMid:11042159 PMCid:310980

Landreville S., Lupien C.B., Vigneault F., Gaudreault M., Mathieu M., Rousseau A.P., Guérin S.L., and Salesse C., 2011, Identification of differentially expressed genes in uveal melanoma using suppressive subtractive hybridization, Mol.Vis., 17: 1324-1333 PMid:21647268 PMCid:3107995

Murat C., Zampieri E., Vallino M., Daghino S., Perotto S., and Bonfante P., 2011, Genomic suppression subtractive hybridization as a tool to identify differences in mycorrhizal fungal genomes, FEMS Microbiol. Lett., 318(2): 115-122 http://dx.doi.org/10.1111/j.1574-6968.2011.02248.x PMid:21362020

Rebrikov D.V., Desai S.M., Siebert P.D., and Lukyanov S.A., 2004, Suppression subtractive hybridization, Methods Mol. Biol., 258: 107-134 PMid:14970460

Veronika E.A., Denis V.R., Dmitry A.S., Valery B.K., Natalia I.M., Dmitry B.S., Rustam Z., Laura L.V., Valery A.R., Sergey A.L., and Alex S.S., 2008, Isolation, characterization and molecular cloning of duplex-specific nuclease from the hepatopancreas of the kamchatka crab, BMC Biochem., 9: 14 http://dx.doi.org/10.1186/1471-2091-9-14 PMid:18495036 PMCid:2413221

Vogel H., Badapanda C., and Vilcinskas A., 2011, Identification of immunity-related genes in the burying beetle *Nicrophorus vespilloides* by suppression subtractive hybridization, Insect Mol. Biol., 20(6): 787-800 http://dx.doi.org/10.1111/j.1365-2583.2011.01109.x PMid:21929718

Wan J., Matthew B.W., Li C., Angel F., and Klaus L., 2002, Efficacy of SSH PCR in isolating differentially expressed genes, BMC Genomics, 3: 12 http://dx.doi.org/10.1186/1471-2164-3-12 PMid:12033988 PMCid:115870