Published online 1 September 2006 Nucleic Acids Research, 2006, Vol. 34, No. 15 e103
doi:10.1093/nar/gkl592

EASI—enrichment of alternatively spliced isoforms

Julian P. Venables* and John Burn

Institute of Human Genetics, International Centre for Life, Central Parkway, University of Newcastle-upon-Tyne, Newcastle-upon-Tyne, UK

Received May 25, 2006; Revised July 27, 2006; Accepted August 1, 2006

ABSTRACT

Alternative splicing produces more than one protein from the majority of genes and the rarer forms can have dominant functions. Instability of alternative transcripts can also hinder the study of regulation of gene expression by alternative splicing. To investigate the true extent of alternative splicing we have developed a simple method of enriching alternatively spliced isoforms (EASI) from PCRs using beads charged with *Thermus aquaticus* single-stranded DNA-binding protein (*T.Aq* ssb). This directly purifies the single-stranded regions of heteroduplexes between alternative splices formed in the PCR, enabling direct sequencing of all the rare alternative splice forms of any gene. As a proof of principle the alternative transcripts of three tumour suppressor genes, TP53, MLH1 and MSH2, were isolated from testis cDNA. These contain missing exons, cryptic splice sites or include completely novel exons. EASI beads are stable for months in the fridge and can be easily combined with standard protocols to speed the cloning of novel transcripts.

INTRODUCTION

The definition of alternatively spliced isoforms is 'transcripts from the same gene that have identical ends with at least one tract of internal difference'. The principle of enrichment of alternatively spliced isoforms from cDNA relies on these alternative transcripts forming heteroduplexes, in which the missing region of one strand causes the 'extra' region in the other strand to form a single-stranded bulge, which can bind to a single-stranded DNA trap. Two groups have so far reported use of this method; in both cases to enrich alternatively spliced isoforms from complex cDNA libraries (1,2). Because of the ambitious nature of their task both protocols contain complex denaturation and renaturation steps. Watahiki *et al*. (2) take full length cDNA libraries and transcribe them *in vitro* and then reverse-transcribe them again, just for their starting material. Thill *et al*. (1) perform three cycles of a procedure consisting of four different enzymatic steps per cycle to achieve a final 10-fold enrichment of alternatively spliced isoforms. Although both groups demonstrate modest global enrichment of alternatively spliced isoforms, the vast scale of cellular alternative splicing means that they are currently unable to approach saturating coverage at the individual gene level. Here we report a simple protocol (EASI) that enriches all the significant alternatively spliced isoforms of individual target genes from 'invisibility' on agarose gels. As this enrichment procedure adds just one hour's work to current cloning methods it should be widely applicable in the study of alternative splicing.

MATERIALS AND METHODS

Making EASI beads

Genomic DNA was isolated from *Thermus aquaticus* (YT1 strain from the centre for applied microbiology and research, Porton Down, Salisbury) and the single-stranded DNA-binding protein reading frame was amplified with the primers AAAAAAAACATATGGCTCGAGGCCTGAACCAA and AAAAAAAAGGATCCTCAAAACGGCAAATCCTCCTCC-GGCGGGAA and cloned into the NdeI and BamHI sites of PET15b (Novagen) to make the plasmid pEASI.

pEASI was then transformed into *Escherichia coli* BL21. 0.5 mM IPTG was added to a 400 ml exponential culture growing in the presence of 100 μM ampicillin. After three further hours at 37°C, cells were harvested by centrifugation and resuspended in 5 ml of phosphate-buffered saline (PBS) and repeatedly sonicated (3 × 20 s) without cooling and then heated to 70°C (to precipitate *E.coli* proteins) for 5 min, then micro-centrifuged at top speed for 10 min. The supernatant was frozen in 400 μl aliquots and stored at −80°C.

When needed, aliquots were thawed and further cleared by micro-centrifugation for 3 min and 80 μl (for each batch of lysate the saturating amount was determined by releasing bound protein from beads with Bradford reagent) Dynabeads TALON slurry (Dynal, Oslo) was washed in PBS and added to the supernatant and left on the bench for 5 min. The beads were magnetically separated and then incubated in (1 M NaCl and 20 mM Tris, pH 7.5) for 5 min to dissociate sheared *E.coli* DNA, separated again and then resuspended in 800 μl of PBS and stored in the fridge for use. The charged ssb-beads are stable indefinitely at 4°C.

*To whom correspondence should be addressed. Tel: +44 191 241 8636; Fax: +44 191 241 8666; Email j.venables@ncl.ac.uk

© 2006 The Author(s).
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 1. Nested primer pairs used

		Forward	Reverse
P53 Section A 5′-UTR-AA249	1st	GTCTAGAGCCACCGTCCAGGGAGCAGGT	GCCGGTCTCTCCCAGGACAGGCACA
	2nd	CGGGCTGGGAGCGTGCTTTCCACGA	CTCCGGTTCATGCCGCCCATGCAGGA
P53 Section B AA156-3′-UTR	1st	GACCTGCCCTGTGCAGCTGTGGGTTGA	GTCCTACTCCCCATCCTCCTCCCCACAA
	2nd	CGCGTCCGCGCCATGGCCATCTACA	GGGTGGGAGGCTGTCAGTGGGGAACA
MLH1 section A 5′-UTR-AA239	1st	CTGAAGGCAC TTCCGTTGAG CAT	GCACTTCTTC ACTGAGTAGT TTGC
	2nd	CCTTGGCTCT TCTGGCGCCA AA	AGGCTAGGGT TTTATCCTCA CATCCA
MLH1 section B AA206–AA496	1st	AACAAGGAGA GACAGTAGCT GA	CCTGGAGACT CAAAACACTA GTGA
	2nd	CTACCCAATG CCTCAACCGT GGA	GGGGGGTACA AGCTGCAGTC A
MLH1 section C AA459–AA718	1st	GGGATACAAC AAAGGGGACT TCAGA	GTGTGAGCGC AAGGCTTTAT AGACA
	2nd	GAGAAGAGAG GACCTACTTC CAGCA	GTGTTCCACA GTCCACTTCC AGGA
MSH2 section A AA6–AA311	1st	AGGTTTCGAC ATGGCGGTGC A	ACTTATTCAG CAAGGCAGCC AGA
	2nd	CGAAGGAGAC GCTGCAGTTG GA	GTTAAGGGCT CTGACTGCTG CAA
MSH2 section B AA265–AA512	1st	CCTCAACCGG TTGTTGAAAG GCAA	ATCCAAACTG TGCACTGGAA TCCA
	2nd	GTGCTGTATT GCCAGAAATG GAGA	GTTTAATCTG TTTGCCAGGG TCCAA
MSH2 section C AA441–AA742	1st	CAGGCTCTGG AAAAACATGA AGGAA	GTAGGTAGAA GTTCCTCTTC CCAA
	2nd	CTGATCTTCG TTCTGACTTC TCCA	GAATCTTTGG TTGCAGACCT GA
MSH2 section D AA680-3′-UTR	1st	GTTCCACATC ATTACTGGCC CCAA	AACAGTCCTC AGTTACAGCT CTCA
	2nd	GACAAACTGG GGTGATAGTA CTCA	AAGTTGATAG CCCATGGGCA CTGA

The forward and reverse, first and second round primers are shown for the gene sections indicated in the first column. The names of each section and the amino acid residue numbers covered are indicated. The nested PCR step was necessary as enrichment did not work on a PCR with a single pair of primers (data not shown). The reason is because although the simple PCR produced a similar amount of product, it is contaminated with a complex mixture of non-specific products that can cross hybridize with the main splice form of the target gene, thus attaching single-stranded DNA regions to it.

RT–PCR

Testis RNA was reversed transcribed with 'Transcriptor' reverse transcriptase (Roche) to avoid 'false splicing' artefacts caused by some non thermostable polymerases [data not shown and (3)]. First round hot-start nested PCR followed by [(94°C 30 s, annealing temperature 30 s, 72°C 60 s) × 20 cycles, 72°C 5 min] was performed on testis cDNA with Expand High Fidelity *T.Aq* polymerase (Roche) in PCR buffer (Promega, 1× contains 50 mM KCl, 10 mM Tris–HCl pH 9, 0.1% Triton) also containing 0.25 mM dNTPs and 2 mM $MgCl_2$. (Annealing temperatures were 69°C for TP53 and 59°C for MLH1 and MSH2, PCR volume 100 µl). A total of 0.1 µl of the first reaction was used as template for the second reaction with the same conditions as before but with 35 cycles. The first and second round primer pairs for amplification of different sections of TP53, MLH1 and MSH2 are shown in Table 1.

Splice enrichment protocol (1 h)

PCRs were treated by addition of 100 µl of 1× L buffer [Roche 1× L buffer contains 10 mM Tris–HCl pH 7.5, 10 mM $MgCl_2$, 1 mM dithioerythritol] containing 1 µl Exonuclease VII (USB) for 30 min at 37°C. No heat inactivation step is necessary and reactions can be stored at −20°C overnight either before or after this step.

Fifty microlitres of the ssb-beads (which corresponds to 4 µl of the original slurry) were magnetically separated and liquid was removed, then 100 µl of exonuclease-treated PCR was added and left on the bench for 5 min. The beads were washed three times in EASI buffer (100 mM NaCl and 20 mM Tris, pH 7.5) then bound heteroduplexes were eluted for 5 min in 10 µl of EASI buffer containing 200 mM imidazole.

Splice identification

One microlitre of the eluate was used in a 25-cycle PCR to amplify the enriched material. The final PCR (10 µl) was visualized on a 1.25% agarose TBE gel (Figure 1). Extra bands not annotated correspond to heteroduplexes of the different isoforms. For example in Figure 1b there are three strong un-annotated bands. These are the three possible heteroduplexes between the strong main band and isoforms 2 and 3. Each heteroduplex migrates about a quarter of the distance from the upper to the lower band that it is composed of, under our gel conditions.

The remaining 90 µl of the enriched material was treated with *T.Aq* DNA polymerase (Promega) and 0.25 mM dATP at 72°C for 15 min, cleaned (PCR cleanup, Qiagen) and then cloned into pCR 4 TOPO (Invitrogen) TA cloning vector, and transformed into *E.coli*. To screen for alternative splices a 'touch' of bacterial colonies was then used as template in 6–24 PCRs for each enrichment [95°C 3 min, (94°C 30 s, 59°C 30 s, 72°C 60 s) × 35, 72°C 3 min with *T.Aq* DNA polymerase (Promega)] with the M13uni(−43) AGGGTTTTCCCAGTCACGACGTT and M13rev(−49) GAGCGGATAACAATTTCACACAGG primers. After a quick centrifugation to remove cell debris, PCR products were imaged by agarose electrophoresis and reactions of each size not conforming to the wild-type length were purified (Qiagen PCR cleanup) and sequenced with the M13 primers.

RESULTS

We first developed the EASI technique using a known pair of alternative splice forms (from the HipK3 gene). A nested PCR from testis cDNA which contained both alternative splices in a >10:1 ratio (Figure 1a input lane) was bound (without any further treatment) to *E.coli* single-stranded DNA-binding protein (ssb) beads, *T.Aq* ssb beads or beads alone (Figure 1a). Bound material was eluted with imidazole and further amplified by PCR and visualized by agarose gel electrophoresis. Both ssbs caused enrichment of the rarer larger splice form but the *T.Aq* ssb beads enriched the isoform

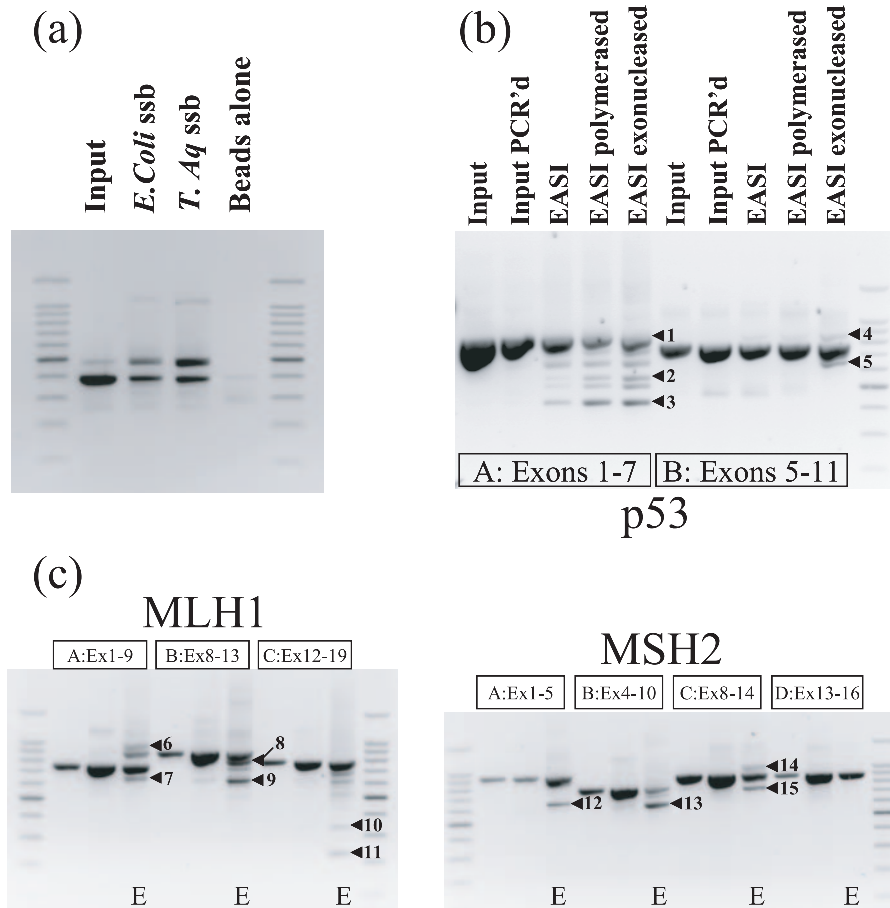

Figure 1. (**a**) EASI proof of principle. A known pair of alternative splices was bound to beads charged with *E.coli* or *T.Aq* ssb, then eluted, amplified by PCR and visualized by agarose electrophoresis. The alternatively spliced isoform of HipK3 (upper band) was enriched by the ssb beads. (**b**) Two overlapping sections of p53 were enriched after exonuclease treatment. Agarose gel showing two PCR product inputs to the EASI column (first lane on each side). The second lanes are a 10 000:1 dilution of the input subjected to a further 25 rounds of PCR (to show that enrichment is not a PCR artefact). Lanes 3, 4 and 5 of each side show the enriched output after applying three conditions: no treatment, addition of extra *T.Aq* polymerase for 10 min, or addition of Exonuclease VII. (**c**) Further confirmation. Three sections of MLH1 and four sections of MSH2 are shown. Each section shows input, diluted input PCR'd and EASI material in the third lane. The exons amplified in each section are indicated. (b and c). Sequences of the alternative splice forms found by EASI have been deposited in GenBank accession no. DQ648883–DQ648897. Arrows show alternative splice forms: (**1**) TP53+ins exon 1a, (**2**) TP53-Ex4, (**3**) TP53-Ex(2–4) (**4**) TP53ins9a (p53β), (**5**) TP53-Ex10, (**6**) MLH1ins ds 5′ss 1a, (**7**) MLH1-Ex6, (**8**) MLH1-Ex10, (**9**) MLH1-Ex(9–10), (**10**) MLH1-Ex(15–18), (**11**) MLH1-Ex(14–18), (**12**) MSH2-Ex3, (**13**) MSH2-Ex5, (**14**) MSH2+ins exon 9a, (**15**) MSH2-Ex10.

better (to the theoretical maximum of 50%) so *T.Aq* ssb was used for all subsequent work.

Our procedure was then used to discover new splice forms, from overlapping sections of the TP53 gene. Initially, two alternative splices were enriched in the 5′ section but no enrichment was observed in the 3′ end which was perplexing as there is a known alternative splice [p53insEx9a (4)] in the 3′ end. Therefore we used two treatments to reduce single-stranded DNA contamination from the input PCRs in the hope of improving enrichment (Figure 1b). Adding fresh polymerase to the PCR, for 10 min at 72°C, before enrichment, slightly improved enrichment of the 5′ end but had no effect on the 3′ end. However digesting single-stranded DNA (in both directions) from the PCR with exonuclease VII improved the 5′ result yet further, and significantly it caused the appearance of the p53beta isoform in the 3′ end. Exonuclease VII was therefore incorporated into the subsequent protocol (Materials and Methods).

To consolidate these results, three overlapping sections of MLH1 and four sections of MSH2 were enriched (Figure 1c). In all, 15 extra bands appeared in the EASI fractions for the three tumour suppressor genes (deposited in GenBank accession nos DQ648883-97). Of these 15, only four were known previously; the TP53ins9a is also known as the beta isoform (4) and three others have been deposited as ESTs very recently: TP53-Ex(2–4) accession DA901020, MLH1+ins1a accession DA499490 and MLH1-Ex6 accession DA343206 (5). The other 11 of these 15 sequences are unknown previously, validating the use of EASI for the discovery of new splice forms.

DISCUSSION

The EASI protocol described here contains several short-cuts from previously described methods that help to increase

its efficiency and usefulness. The *T.Aq* ssb column described in this work can be easily prepared from recombinant bacteria using heat to precipitate all host *E.coli* proteins without fear of denaturing the thermostable ssb. No complex protein purification is needed, as by mixing beads with the heat-cleared lysate the His-tagged protein can thus be stored bound to beads and ready for use for >3 months with no loss of activity (data not shown). As Talon-Dynabeads have a far smaller size than other nickel-agarose beads they have the advantage that no complex chromatography steps are needed in the enrichment as the beads stay suspended and effectively 'in solution' for the course of the binding reaction. Binding to the ssb-bead suspension is quick and extremely specific, as single-stranded DNA binding was unaffected by a 100-fold excess of double stranded DNA (data not shown). The recently discovered *T.Aq* ssb may perform better than *E.coli* ssb because it is a tandem dimer and does not need to tetramerize to become active (6). Similarly, tandem dimer red fluorescent protein has been shown to have enhanced kinetics of activation compared to the normal tetrameric form (7). The final cloning procedure also contains several time saving steps compared to the previous methods. To avoid restriction digestion, EASI material is directly ligated into a TA cloning vector and transformed into *E.coli*. Colony-PCR then allows all the alternative splice forms in the sample to give high quality sequence. This avoids the process of gel purification of bands that include heteroduplex bands that are difficult to clone and could lead to redundancy (8).

The development of a simple protocol for enrichment of alternatively spliced isoforms from PCRs represents a significant step forward from previous methods; in one method 'invisible' isoforms were rescued by PCR amplification of all the sections of agarose gel lanes that did not contain any visible bands (9). Our simple and robust approach takes the labour and chance out of alternative splicing discovery and will allow the prevalence and biological significance of splicing events to be fully explored. For example in testis there are several cell types: Sertoli, Leydig and germ cells. The germ cell population consists of cells in various stages of differentiation from stem cell to cells before and during meiosis to condensing haploid cells. It is thus difficult to find alternative splices in the stem cells as they constitute <10% of the tissue mass. The same problem applies to cancer stem cells in a tumour mass. Therefore with EASI it should be possible to define a limited set of all the alternative splices in a healthy or diseased tissue in order to study cell type specific splicing regulation. In addition we are proposing to develop EASI for high-throughput clinical diagnostics in the hope of identifying mutations that affect splicing responsible for genetic disease, especially in cases where the resulting isoforms are rendered 'invisible' in agarose gels by nonsense mediated decay.

ACKNOWLEDGEMENTS

This work was funded by the Newcastle Healthcare Charity. Funding to pay the Open Access publication charges for this article was provided by Newcastle Healthcare Charity.

Conflict of interest statement. None declared.

REFERENCES

1. Thill,G., Castelli,V., Pallud,S., Salanoubat,M., Wincker,P., de la Grange,P., Auboeuf,D., Schachter,V. and Weissenbach,J. (2006) ASEtrap: a biological method for speeding up the exploration of spliceomes. *Genome Res*., **16**, 776–786.
2. Watahiki,A., Waki,K., Hayatsu,N., Shiraki,T., Kondo,S., Nakamura,M., Sasaki,D., Arakawa,T., Kawai,J., Harbers,M. *et al*. (2004) Libraries enriched for alternatively spliced exons reveal splicing patterns in melanocytes and melanomas. *Nature Methods*, **1**, 233–239.
3. Cocquet,J., Chong,A., Zhang,G. and Veitia,R.A. (2006) Reverse transcriptase template switching and false alternative transcripts. *Genomics*, **88**, 127–131.
4. Bourdon,J.C., Fernandes,K., Murray-Zmijewski,F., Liu,G., Diot,A., Xirodimas,D.P., Saville,M.K. and Lane,D.P. (2005) p53 isoforms can regulate p53 transcriptional activity. *Genes Dev*., **19**, 2122–2137.
5. Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H. *et al*. (2006) Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res*., **16**, 55–65.
6. Dabrowski,S., Olszewski,M., Piatek,R., Brillowska-Dabrowska,A., Konopa,G. and Kur,J. (2002) Identification and characterization of single-stranded-DNA-binding proteins from *Thermus thermophilus* and *Thermus aquaticus*—new arrangement of binding domains. *Microbiology*, **148**, 3307–3315.
7. Campbell,R.E., Tour,O., Palmer,A.E., Steinbach,P.A., Baird,G.S., Zacharias,D.A. and Tsien,R.Y. (2002) A monomeric red fluorescent protein. *Proc. Natl Acad. Sci. USA*, **99**, 7877–7882.
8. Zacharias,D.A., Garamszegi,N. and Strehler,E.E. (1994) Characterization of persistent artifacts resulting from RT–PCR of alternatively spliced mRNAs. *Biotechniques*, **17**, 652–655.
9. Gorlov,I.P. and Saunders,G.F. (2002) A method for isolating alternatively spliced isoforms: isolation of murine Pax6 isoforms. *Anal. Biochem*., **308**, 401–404.