

# A sensitive procedure to detect alternatively spliced mRNA in pooled-tissue samples

Germán Gastón Leparc<sup>1,2</sup> and Robi David Mitra<sup>1,\*</sup>

<sup>1</sup>Department of Genetics and Center for Genome Sciences, Washington University in St Louis, 4444 Forest Park Parkway, Campus Box 8510, St Louis, MO 63108, USA and <sup>2</sup>WWTF Chair of Bioinformatics, Boku University Vienna, 1190 Muthgasse 18, Vienna, Austria

Received May 22, 2007; Revised and Accepted October 22, 2007

## ABSTRACT

One important goal of genomics is to explore the extent of alternative splicing in the transcriptome and generate a comprehensive catalog of splice forms. New computational and experimental approaches have led to an increase in the number of predicted alternatively spliced transcripts; however, validation of these predictions has not kept pace. In this work, we systematically explore different methods for the validation of cassette exons predicted by computational methods or tiling microarrays. Our goal was to find a procedure that is cost effective, sensitive and specific. We examined three ways of priming the reverse transcription (RT) reaction—poly-dT priming, random priming and pooled exon-specific priming. We also examined two strategies for PCR amplification—flanking PCR, which uses primers that hybridize to the constitutive exons flanking the predicted exon, and a semi-nested PCR with a primer that targets the predicted exon. We found that the combination of RT using a pool of gene-specific primers followed by semi-nested PCR resulted in a significant increase in sensitivity over the most commonly used methodology (97% of the test set was detected versus 14%). Our method was also highly specific—no false positives were detected using a test set of true negatives. Finally, we demonstrate that this method is able to detect alternative exons with a high sensitivity from whole-organism RNA, allowing all tissues to be sampled in a single experiment. The protocol developed here is an accurate and cost-effective way to validate predictions of alternative splicing.

## INTRODUCTION

Alternative splicing is an important cellular phenomenon that has been connected with a number of physiological and pathophysiological processes (1,2). It is estimated that ~74% of human protein coding genes are alternatively spliced (3). However, the detection of alternative splice variants is often challenging because, unlike constitutive splicing events, alternatively spliced exons often exhibit environmental, temporal or cell-type specific expression patterns. Most of the known alternative exons were discovered by sequencing large expressed sequence tag (EST) libraries. However, despite millions of available ESTs, the coverage of many organisms' transcriptome remains incomplete due to issues such as transcript end bias, library coverage limitations and sampling differences (4). To address these issues, new technologies to detect alternative splicing of mRNA are being developed. For example, splice junction micro-arrays have been designed to detect changes at the exon splice junctions of a large number of known genes (3). Whole genome tiling array (WGTA) analysis has been able to provide even more detail by using overlapping oligonucleotide probes to detect exons from whole sequenced genomes (5). Complementing these new experimental approaches, are computational algorithms designed to detect novel alternative exons from genomic sequences without relying on EST evidence (6–8). These new approaches often yield a large number of putative alternative splicing events. However, the specificity of these approaches is still relatively low compared with EST sequencing (9), and so these candidates must be validated experimentally. This crucial experimental step has become the bottleneck in discovering new alternatively spliced isoforms, so we sought to optimize this procedure, focusing on methods to validate cassette exons.

Validation of cassette exons is typically achieved in three steps: (i) perform reverse transcription (RT) of the RNA samples, (ii) perform the polymerase chain reaction

\*To whom the correspondences should be addressed. Tel: +(314)362-2751; Fax: (314)362-2156; Email: rmitra@genetics.wustl.edu  
Present address:

Germán Gastón Leparc, WWTF Chair of Bioinformatics, Boku University Vienna, 1190 Muthgasse 18, Austria

(PCR) on the resulting cDNA templates and (iii) clone and sequence the products to distinguish true splice variants from false positives. Each step can be carried out in different ways. For example, there are three ways to prime the RT reaction: poly-T priming, random priming and gene-specific priming. Alternatively, the PCR can be performed using primers that target the constitutive exons that flank the predicted splicing event or, in a semi-nested format, using a primer that targets the predicted exon. Currently there is no clear consensus to the optimal approach to use in splice variant validation. Most studies use poly-T priming for the RT reactions followed by flanking PCR (3,6–8), despite the fact that neither the sensitivity nor the specificity of this procedure has been characterized. Our goal was to develop a cost-effective approach that is sensitive, specific and allows for a large number of tissues to be analyzed. Therefore, we carried out a systematic survey to assess three different priming methods for the RT reaction and two different PCR methods. We measured their ability to detect a set of known alternative exons in a pooled sample containing RNA from 18 tissues. We also assessed the false positive rate of these methods using a set of pseudo-exons—intronic sequences that are not spliced but have donor and acceptor sites (10). Our results showed that there are significant differences in sensitivity between the different types of RT priming. In addition, we demonstrated that a semi-nested PCR approach is significantly more sensitive than the conventional flanking PCR approach utilized in most studies.

## MATERIALS AND METHODS

### Positive and negative control test sets

A set of 48 EST annotated alternatively spliced cassette exons were selected from the Alternative Splicing Database, splicing event annotations file (11) based on varying frequencies of exon inclusion (6–97%) into the canonical transcript (Supplementary Table A). Another set of 24 pseudo exons were used as negative controls for the semi-nested RT-PCR experiments. These pseudo exons were selected by randomly picking pairs of AG (Adenine-Guanine) acceptor and GT (Guanine-Thymine) donor dinucleotide sites from the introns of randomly selected RefSeq genes. We required these exons to be read through, thus preserving the reading frame and lacking any non-sense codons. We also compared these pseudo exons with the EST data to make sure that there was no evidence of this sequence being incorporated as part of a transcript. Specifically, we required that the pseudo-exon sequence was not present in any EST transcripts, while the flanking constitutive exon sequences were present in at least 100 independent ESTs. All alternative or pseudo-exon sequences and the adjacent constitutive exon sequences were used as templates for the primer design. For the selection of adjacent constitutive exons, we used transcript alignment information from the UC Santa Cruz genome browser Refseq and spliced-ESTs information tracks (<http://genome.ucsc.edu>) to select for exons that

were observed to be constantly incorporated into all documented spliced transcripts of the gene.

### Primer design of selected candidates

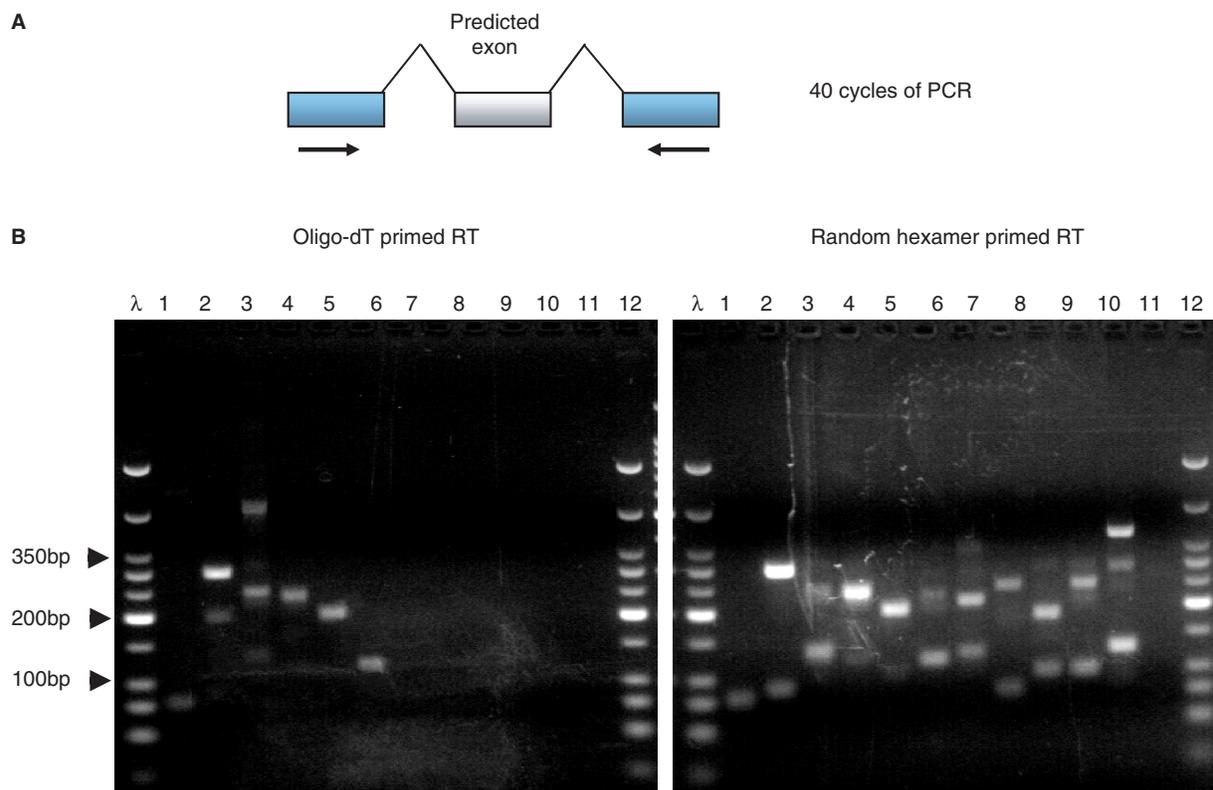
Flanking PCRs were designed with forward and reverse primers targeted to the constitutive exons, flanking the alternative cassette exons (Figure 1A). Semi-nested PCRs were designed specifically for two rounds of PCRs. In the first round of PCR, an ‘external’ forward primer is targeted to a 5′ upstream canonical exonic sequence and is used with a reverse primer targeted to the alternative cassette exon. The second round PCR then uses an ‘internal’ forward primer targeted to an exonic region between the ‘external’ forward primer and the previously used reverse exon primer. Figure 2A illustrates the semi-nested primer design. Primer design for all candidate exons was done with PRIMER3 software (12), using the following specific parameter settings: primer length minimum, 19 nt, desired, 25 nt and maximum, 32 nt; melting temperature minimum, 64°C, desired, 70°C and maximum, 73°C; minimum GC content of 45 and maximum 80; product length, 150–700 nt; and prefiltering of potentially mispriming sequences with the provided library of human repeats (See Supplementary Data for list of all primers used). Primer sequences were ordered from Integrated DNA Technologies (Coralville, Indiana).

### RNA samples

Total RNA samples of Universal Whole Mouse, 11-day, 15-day and 17-day mouse embryo, were obtained from Biochain (Cat No. R4334566, R1334XI-10, R1334XV-10 and R1334XVII-10, respectively). Total RNA samples of human liver, colon, stomach, thymus, lung, trachea, placenta, brain, retina, skeletal muscle, testis and kidney total RNA were obtained from Clontech (Cat No. 636 506, 636 553, 636 522, 636 512, 636 524, 636 541, 636 527, 636 530, 636 579, 636 534, 636 533 and 636 529, respectively), while skin, breast, ovary, pancreas, uterus and prostate total RNA samples were obtained from BioChain (Cat No. R1234218-50, R1234086-50, R1234086-50, R1234188-50, R1234274-50 and R1234201-50, respectively). All total RNA samples were collected by the vendors using isolation by acid guanidinium thiocyanate-phenol-chloroform extraction and DNase I treated. The majority of these tissue RNA samples match a ranking of tissues exhibiting the most alternative splicing events per tissue according to a study of ESTs (13) and were also matched with any tissues known to express the exons in our test set.

### Reverse transcription

Superscript III RT kit (Invitrogen catalog 18080-044) was used to create cDNA, using three different types of primers: oligo-dT, random hexamer and exon-specific reverse primers. The oligo-dT primer is a string of 20 deoxythymidylic acid residues. One RT reaction of each type was performed using 5 µg of RNA for each tissue analyzed. The RT reaction was done following manufacturer’s protocols with the following changes. The concentration of the primers for the initial 20 µl RT reaction



**Figure 1.** Flanking PCR is more sensitive at detecting exons using random hexamer primed cDNA templates than with oligo-dT primed templates. (A) Flanking PCRs were designed with forward and reverse primers targeted to the constitutive exons flanking the alternative cassette exons. (B) A subset of 12 alternative exons tested with the flanking PCR approach and the two different methods of priming RT. The oligo-dT primed experiment detected four alternative exons (lanes 2,3,4 and 5) in comparison to the random hexamer primed experiment that detected eight (lanes 2,4,5,7,8,9,10 and 11). DNA ladder  $\lambda$  shows 25 bp, 50 bp, 75 bp 100 bp, then 50 bp increments to 350 bp, then 500 bp and 766 bp. The strongest band represents 200 bp.

mixtures were as follows: oligo-dT primers added to a concentration of  $2\mu\text{M}$ , gene-specific primers added to a concentration of  $2\mu\text{M}$  for each, and for random hexamers, 50 ng were added as provided by the manufacturer. The reaction conditions were as follows: Oligo-dT priming, 1 h and 15 min at  $50^\circ\text{C}$ . Random hexamer priming, 10 min at  $25^\circ\text{C}$ , then 1 h and 15 min at  $50^\circ\text{C}$ . Gene-specific RT was done with 1 h and 15 min at  $50^\circ\text{C}$ . After RT, the RNA hydrolyzed with NaOH and then neutralized with Tris pH8.0. The intact cDNA was then purified using the Qiagen PCR clean up kit with a final elution of  $30\mu\text{l}$ . All cDNA stocks were tested for the presence of RNA polymerase II transcript as a quality control. For experiments in which a pooling strategy was utilized, 10-fold dilutions from cDNA stocks were pooled together to use as templates for PCR.

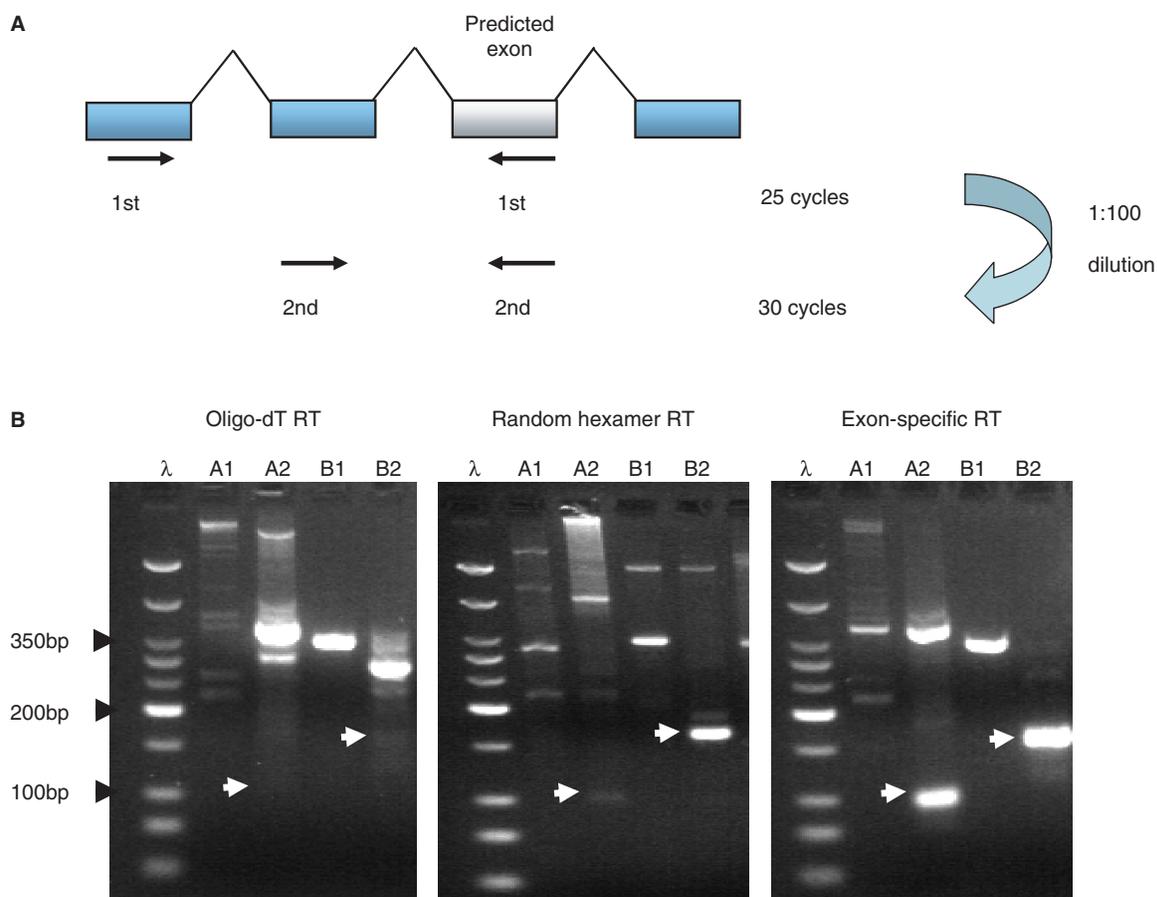
#### Flanking PCR reactions

Total of  $1\mu\text{l}$  of pooled cDNA template was used per  $25\mu\text{l}$  reaction tube. All PCRs were carried out with the Sigma Jumpstart Taq DNA polymerase kit (Cat D9307) on an MJ Research PTC-200 (Bio-Rad Laboratories). For the detection of exons by flanking PCR, a program of 40 cycles of melting ( $45\text{ s}$  at  $94^\circ\text{C}$ ), annealing ( $30\text{ s}$  at  $56^\circ\text{C}$ ), and extension ( $1\text{ min}$  at  $72^\circ\text{C}$ ) was used. PCR products were separated in 2% agarose gels supplemented with

ethidium bromide, and the DNA was visualized under a UV light. To evaluate the sensitivity of flanking PCR, titration experiments were performed on a subset of alternative exons (see 'Results'). In each instance, the alternative isoform was titrated against the constitutive form beginning with a one-to-one molar ratio (at 50 000 copies each), followed ratios of 1:100, 1:1000 and 1:10 000 of alternative to constitutive template. Total of 40 cycles of PCR were used for these experiments, however, three of the eight were further tested with 55 and 70 cycles (see 'Supplementary Data'). Analysis using Phoretix 1D (Nonlinear Dynamics) facilitated precise multiple-lane band size determinations for all PCR experiments. PCR reactions that contained a band within 30 base pairs of the expected size were selected for cloning and sequencing reactions.

#### Semi-nested PCR reactions

For the semi-nested PCR primer approach, a first round program of 25 cycles of melting ( $45\text{ s}$  at  $94^\circ\text{C}$ ), annealing ( $30\text{ s}$  at  $56^\circ\text{C}$ ) and extension ( $1\text{ min}$  at  $72^\circ\text{C}$ ) was used. A second round was performed for 30 cycles using the same program with a 1:100 dilution of first round reaction as the template. PCR products were separated in 2% agarose gels supplemented with ethidium bromide, DNA was visualized under a UV light. Analysis using Phoretix



**Figure 2.** Semi-nested PCR is more sensitive at detecting exons using exon-specific primed cDNA templates than with either random hexamer or oligo-dT primed templates. **(A)** Semi-nested PCRs were designed specifically for two rounds of PCRs. Thirty cycles of first round of PCR with an 'external' forward primer targeted to a 5' upstream canonical exonic sequence and used with a reverse primer targeted to the alternative cassette exon. Twenty-five cycles of the second round PCR then uses an 'internal' forward primer targeted to an exonic region between the 'external' forward primer and the previously used reverse exon primer. A 1:100 dilution of the first round reaction is used as a template for the second round reaction. **(B)** Two alternative exons tested with the flanking PCR approach and the two different methods RT. A1 and A2 are the first and second PCR reactions for the alternative exon in the *ABI1* gene. B1 and B2 are the first and second PCR reactions for the alternative exon in the *NCOA2* gene (See supplementary table A for the full list of genes). The expected PCR product sizes for the second round reactions are shown with white arrowheads. The oligo-dT primed experiment did not detect either of the alternative exons (lanes A2 and B2) in comparison to the random hexamer primed experiment that detected the second exon (lane B2) and the exon-specific primed experiment that detected both exons (lanes A2 and B2). DNA ladder  $\lambda$  shows 50 bp, 75 bp, 100 bp, then 50 bp increments to 350 bp, then 500 bp and 766 bp. The strongest band represents 200 bp.

1D (Nonlinear Dynamics) facilitated precise multiple-lane band size determinations for all PCR experiments. PCR reactions that contained a band within 30 base pairs of the expected size were selected for cloning and sequencing reactions.

#### Cloning and sequencing

Flanking PCR products and semi-nested second round PCR products of the expected predicted size were ligated into pGEM-T easy vectors (Promega catalog A1360) and transformed into High Efficiency GC10 chemically competent cells (GeneChoice catalog D-4). Bacterial clones were plated on LB/X-gal/IPTG agar plates and grown overnight at 37°C. A maximum of 12 colonies were picked from each plated transformation and used for colony PCR with standard M13 primers. Total of 2  $\mu$ l of this colony PCR product was then used as the template for cycle sequencing using Applied Biotech Inc. BigDye Terminator

Cycle Sequencing Mix (cat 4336917) and then run on an ABI 3700.

#### RESULTS

Our objective was to systematically evaluate the ability of different RT and PCR methods to detect alternative exons from a pooled RNA sample. We designed primers for a test set of 48 known alternatively spliced cassette exons from the Alternative Splicing Database (11). We also designed primers for a 'negative' set—a set of 24 'pseudo-exons' (10) that are unlikely to be spliced. These pseudo-exons were selected by randomly picking pairs of AG acceptor and GT donor sites from intronic DNA. We required that these pseudo-exons preserved the protein-reading frame, contained no stop codons, and were not present in the human EST database (see 'Methods'). We used these test sets to measure the sensitivity

**Table 1.** Detection of 48 alternative cassette exons from a pool of cDNAs by different RT-PCR methods

Method of PCR	Flanking PCR		Semi-nested PCR		
	Oligo-dT	Random hexamer	Oligo-dT	Random hexamer	Exon specific
Method of RT priming					
Number of alternatives detected	7/48 (14.6%)	20/48 (41.7%)	29/48 (60.4%)	42/48 (87.5%)	47/48 (97.9%)

Note: The sensitivity for each method is shown in parentheses.

[true positives/(true positives + false negatives)] and specificity [true negatives/(false negatives + true negatives)] for different validation methodologies.

### Comparison of PCR amplification strategies: flanking PCR versus semi-nested PCR

Flanking PCR is the approach most commonly used to detect alternative cassette exons from individual cDNA samples as well as pools of samples (3,6–8). It is usually designed with forward and reverse primers targeted to the constitutive exons flanking the alternative cassette exon (Figure 1A). Another approach that is potentially more sensitive is a semi-nested PCR where the predicted exon is targeted by one of the PCR primers (Figure 2A). Semi-nested PCR involves two rounds of PCR. In the first round of semi-nested PCR, an ‘external’ forward primer is targeted to a 5′ upstream canonical exonic sequence and is used with a reverse primer targeted to the alternative cassette exon. The second round of PCR is performed to remove mispriming events that occur in the first round and uses an ‘internal’ forward primer targeted to an exonic region between the ‘external’ forward primer and the previously used reverse exon primer.

Semi-nested PCR and flanking PCR was performed with primers designed to amplify the cassette exons in our test set. We pooled total RNA from 18 human tissues (see ‘Methods’) and used two different types of RT priming (oligo-dT and random hexamer) followed by PCR. The reaction products were analyzed by gel electrophoresis, and bands of the correct size ( $\pm 30$  bases) were verified by cloning and sequencing. If a reaction did not produce a band, produced a band of the wrong size, or produced a band of the correct size that was not validated by sequencing, then it was considered a negative. The results are shown in Table 1. Semi-nested PCR was able to detect more alternative exons than flanking PCR, regardless of which RT priming method was used. For example, flanking PCR was only able to detect 7/48 (sensitivity = 14.6%) exons in our test set with oligo-dT priming, but semi-nested PCR was able to detect 29/48 (sensitivity = 60.4%) under the same conditions. Thus, we conclude that semi-nested PCR is significantly more sensitive than flanking PCR.

Because the semi-nested PCR approach is so sensitive, a valid concern is that this approach might detect extremely rare splicing events in which pseudo-exons are spliced into transcripts due to errors by the splicing machinery (splicing ‘noise’). This would lead to the detection of splicing events that may not be biologically meaningful. To test this hypothesis, we performed semi-nested PCR on

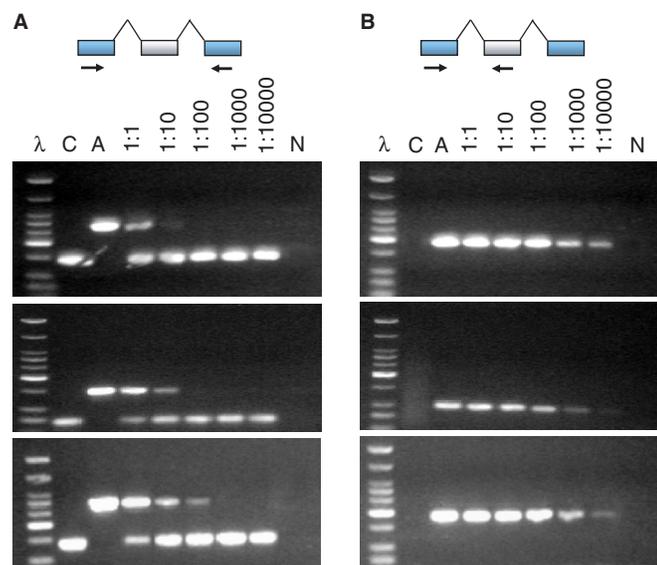
our negative test set of 24 pseudo-exons. We detected no alternative splicing events. In fact, we observed no PCR products of the expected size. Five of our reactions yielded products that were significantly larger than the expected size and thus did not pass our criteria to be advanced to the cloning and sequencing reactions (even so, we decided to sequence these bands and determined that they were result of mispriming in the PCR reaction) Thus, the semi-nested PCR approach did not pick up any splicing noise, indicating that this approach has high specificity (zero false positives in our test set).

### Comparison of priming strategies for reverse transcription

The results shown in Table 1 and Figure 1B demonstrate that random hexamer priming of the RT reaction performs significantly better than oligo-dT priming (although random priming was not superior for every alternative exon that we tested—see, e.g., lane 3 in Figure 1B). We hypothesized that this difference in sensitivity was the result of incomplete first strand synthesis by the reverse transcriptase. Therefore, to further improve the sensitivity of our method, we used a pool of gene-specific primers in the RT reaction followed by semi-nested PCR. A total of 72 primers (48 from the positive test set and 24 from the negative test set) that hybridize to the constitutive exon immediately 3′ of the cassette exons were used to prime the RT reaction. The combination of gene-specific priming of the RT reaction followed by semi-nested PCR detected 47/48 (98%) of the cassette exons, an improvement over random-hexamer priming (42/48) and oligo-dT priming (29/48). The specificity of the method was unchanged; zero exons were amplified from our negative test set (0/24). We conclude that the optimal way to prime RT reactions is to use a pool of gene-specific primers that are proximal to the predicted exon.

### Flanking PCR is limited by the molecular ratio of constitutive vs alternative splice form

Flanking PCR is widely used for the validation of predicted cassette exons, and so we sought to better understand why this method displays such low sensitivity in our experiments. Therefore, we designed an experiment where we could precisely control the relative concentrations of the constitutive and alternative splice forms (see ‘Supplementary Data’). We cloned 16 different cDNA—the constitutive and alternative form for eight members of our test set. We then mixed different ratios of the alternative and constitutive splice forms and performed flanking PCR to determine the point at which the



**Figure 3.** Flanking PCR is limited to detecting molar ratios of 1:10 to 1:100 of alternative to constitutive templates. Abbreviations: C and A lanes represent constitutive and alternative template controls. A titration was done of different ratios of the alternative and constitutive splice forms beginning with a one-to-one molar ratio of 50 000 copies each, followed by serial dilutions of 1:100, 1:1000 and 1:10 000 of alternative to constitutive templates. 40 cycles of PCR were used for both experiments. (A) Most of the alternative exons could not be detected beyond a 1:10 dilution of alternative to constitutive when using flanking PCR. Here, only one case showed that the flanking PCR was able to detect the alternative in at most a 1:100 molar ratio (bottom gel). (B) The exon-specific PCR approach demonstrated the greatest sensitivity in all three cases, detecting the alternative isoform even when the constitutive form was present at a 10 000-fold excess. DNA ladder  $\lambda$  shows 50 bp, 75 bp, 100 bp, then 50 bp increments to 350 bp, then 500 bp and 766 bp. The strongest band represents 200 bp.

alternative splice form could no longer be detected. We began with a one-to-one molar ratio, and then titrated the concentration of alternative form relative to the constitutive form. We used ratios of 1:1, 1:100, 1:1000 and 1:10 000 of alternative to constitutive template. Three of the experiments are shown in Figure 3A. Of the eight alternative exons we tested, four could not be detected beyond a 1:1 dilution of alternative to constitutive by flanking PCR. In another three instances, the alternative isoform could not be detected beyond a 1:10 ratio with the constitutive form. Only one instance demonstrated detection of the alternative at a 1:100 ratio with the constitutive form. We were unable to improve the sensitivity of this method by increasing the number of PCR cycles or by performing a two-step amplification (see 'Supplementary Data').

These results suggest that flanking PCR does not have sufficient sensitivity to detect alternative splicing in pooled samples, or for that matter, in single tissue samples in which the alternative form is expressed at lower levels than the constitutive form. This explains the low sensitivity of the flanking PCR approach seen previously. In those experiments, samples from 18 tissues were pooled, so if the alternative isoform is expressed in a tissue-specific fashion, the signal is likely to be undetectable by flanking PCR

due to the levels of constitutive isoform present in the other tissues.

We next repeated three of the titrations describe above but performed the PCR with a primer targeted to the alternative exon. For all three cases, we were able to detect the alternative isoform at much lower concentrations. In most cases, we detected the alternative isoform even when the constitutive form was present at a 10 000-fold excess (Figure 3B). These results suggest that the competition between the alternative and the constitutive isoforms severely limits the sensitivity of flanking PCR and that this limitation can be overcome by designing primers that target the predicted exon directly.

### Semi-nested PCR is sensitive enough to detect alternative Exons in whole mouse total RNA

The sensitivity of the semi-nested PCR combined with gene-specific priming led us to consider the possibility that it could be used to detect cassette exons using total RNA isolated from a whole organism. This would be desirable since there are approximately 200 different cell types present in mammals (14), so pooling RNA prepared separately from each of these cell types is expensive and inefficient. We selected 26 alternative exons present in the mouse EST database, many of which were expressed in tissues that are not commercially available (Supplementary Table B). We attempted to detect these cassette exons by performing a RT with gene-specific primers followed by semi-nested PCR. Our input was total RNA isolated from whole mice at 4 developmental stages (Adult whole mouse, 11-day, 15-day and 17-day mouse embryo). Alternative exons were efficiently detected in both adult and pooled-embryo developmental stages with a total of 25 out of 26 (96%) detected. Interestingly, almost all of the exons were observed in RNA from both adult and embryo samples, however there was one exon captured exclusively from the embryo samples. These results demonstrate that the combination of RT with gene-specific primers followed by semi-nested PCR is sensitive enough to validate predicted cassette exons in whole organisms.

Using RNA isolated from a whole organism instead of individual tissues may allow for the more efficient discovery of alternative exons. For example, in work published elsewhere, we tested a set of 384 computational predictions and found 26 novel human exons using our most sensitive semi-nested PCR approach (15). These predictions were tested using pooled-RNA stocks isolated from 18 human tissues. To see if we could improve upon this by using RNA isolated directly from a whole organism, we re-analyzed a subset of these predictions that were conserved in mouse. This subset contained 93 predictions, 9 of which were detected previously in our human experiments. By using RNA isolated from a whole mouse adult and embryos, we were able to detect 37 cassette exons (including all 9 detected in human), a significantly higher validation rate (40% versus 6.7%). These results suggest that the detection of alternative splicing events using whole-organism RNA is feasible and

that doing so may result in a higher rate of validation due to the inclusion of more tissues.

## DISCUSSION

The large-scale validation of genome-wide predictions has been limited by the large number of samples and conditions that must be tested. To date, there has not been a systematic analysis of the strengths and weaknesses of different approaches to isoform validation. In this work, we have quantified the sensitivity and specificity of several protocols for the validation of cassette exons. We examined three ways of priming the RT reaction—poly-dT priming, random priming and pooled exon-specific priming. We also examined two strategies for PCR amplification—flanking PCR, which uses primers that hybridize to the constitutive exons flanking the predicted exon, and a semi-nested PCR with a primer that targets the predicted exon. While all methods tested were highly specific (no method produced a false positive), we observed large differences in sensitivity. Surprisingly, the most commonly used method of validation (oligo-dT primed RT followed by flanking PCR) displayed the lowest sensitivity—only 14% of our test set was detected. This is significantly worse than the 97% sensitivity achieved by the best method tested, which uses a pool of gene-specific primers to prime the RT reaction followed by a semi-nested PCR. The high sensitivity of this method means that it can be used to detect tissue-specific cassette exons in an RNA sample isolated from a whole mouse, a strategy that allows for all tissues to be sampled in a cost-effective manner.

The results presented in Table 1 demonstrate that the method of priming the RT reaction is an important factor in the quality of the resulting cDNA. We found the best results were obtained when pooled gene-specific primer were used, compared with RT by random hexamer priming, and then oligo-dT priming. These results indicate that the processivity of the reverse transcriptase limits the yield of this reaction and has a large impact on the ability to detect splice forms. While the low efficiency of reverse transcriptase relative to DNA polymerase has been well documented (16), it was somewhat surprising that the high sensitivity of the subsequent PCR step was unable to fully compensate for differences in cDNA yield between the three priming methods.

Our results demonstrate that flanking PCR is not able to detect the alternative splice form when high levels of the constitutive form are present—often, a 10-fold molar excess of the constitutive form prevents the detection of the alternative form. When pooling tissues, this could be problematic when an alternative splice form is highly expressed in a single tissue, and yet the ratio of alternative to constitutive molecules is less than 1:10 in the pooled sample. Also, the analysis of a single tissue could be affected by the low sensitivity of flanking PCR, as most tissues are comprised of multiple cell types, only one of which may express a particular isoform. The semi-nested PCR presented here resolves this issue, as it can detect an alternative cassette exon even when there is a large

(>10 000-fold) excess of the constitutive form. One caveat of our semi-nested approach is that it only amplifies the 5' end of alternative exons. However, we believe that the benefits in detection due to this sensitive approach far outweigh this caveat, as follow up experiments with the primers designed around the 3' end can be easily done to completely characterize the exon (15).

There are other tradeoffs that should be considered when choosing between flanking PCR and an exon-specific approach. For example, by choosing flanking primers that are targeted to distant constitutive exons (e.g. exons 1 and 4), one has the potential to detect a larger number of splicing events, such as exon skipping events or mutually exclusive splicing, than would be possible with the semi-nested approach. Furthermore, the flanking design provides a nice internal control, as the constitutive isoform should always be amplified if the reaction is working correctly. Finally, the flanking approach can give a semi-quantitative measurement of the relative abundances of the constitutive to alternative isoforms. However, if one is using a pool of cDNA from different tissues, as we did in these experiments, then one cannot expect quantitative results from either approach. Also, the ability to detect multiple types of splicing events is less important if one is testing specific predictions. Therefore, although the semi-nested approach is not as flexible as the flanking approach, the greatly improved sensitivity makes it the method of choice.

We have shown that by combining gene-specific priming in the RT with a semi-nested PCR, alternative exons can be detected with a high sensitivity (96%) from RNA extracted from a whole mouse. This is a significant improvement over previous methods, since every tissue is sampled with this approach. With small modifications, we believe this approach will also be useful for the detection of micro-RNAs and other non-coding RNAs since these are often expressed in cell-type dependent manner.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We would like to thank Justin Sonnenburg and Jeff Sabina for reviewing and proofreading our manuscript. We would also like to acknowledge Tim Schedl for providing excellent suggestions concerning our experimental approach. We thank KT Varley, Yun Yue, Michael Brooks and Lee Tessler for helpful discussions. In addition, we thank Andreas Sommer and David Kreil for offering their help and lab space to complete this work. This work and the Open Access publication charges were funded by GATP training grant no. T32-HG000045, Whitaker Foundation (St. Louis, MO) and NIH grant no. 5P50HG003170-03.

*Conflict of interest statement.* None declared.

## REFERENCES

1. Stamm,S., Ben-Ari,S., Rafalska,I., Tang,Y., Zhang,Z., Toiber,D., Thanaraj,T.A. and Soreq,H. (2005) Function of alternative splicing. *Gene*, **344**, 1–20.
2. Venables,J.P. (2004) Aberrant and alternative splicing in cancer. *Cancer Res.*, **64**, 7647–7654.
3. Johnson,J.M., Castle,J., Garrett-Engle,P., Kan,Z., Loerch,P.M., Armour,C.D., Santos,R., Schadt,E.E., Stoughton,R. *et al.* (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*, **302**, 2141–2144.
4. Modrek,B. and Lee,C. (2002) A genomic view of alternative splicing. *Nat. Genet.*, **30**, 13–19.
5. Mockler,T.C., Chan,S., Sundaresan,A., Chen,H., Jacobsen,S.E. and Ecker,J.R. (2005) Applications of DNA tiling arrays for whole-genome analysis. *Genomics*, **85**, 1–15.
6. Ohler,U., Shomron,N. and Burge,C.B. (2005) Recognition of unknown conserved alternatively spliced exons. *PLoS Comput. Biol.*, **1**, 113–122.
7. Hiller,M., Huse,K., Platzer,M. and Backofen,R. (2005) Non-EST based prediction of exon skipping and intron retention events using Pfam information. *Nucleic Acids Res.*, **33**, 5611–5621.
8. Sorek,R., Shemesh,R., Cohen,Y., Basechess,O., Ast,G. and Shamir,R. (2004) A non-EST-based method for exon-skipping prediction. *Genome Res.*, **14**, 1617–1623.
9. Florea,L. (2006) Bioinformatics of alternative splicing and its regulation. *Brief Bioinform.*, **7**, 55–69.
10. Sun,H. and Chasin,L.A. (2000) Multiple splicing defects in an intronic false exon. *Mol. Cell Biol.*, **20**, 6414–6425.
11. Stamm,S., Riethoven,J.J., Le Texier,V., Gopalakrishnan,C., Kumanduri,V., Tang,Y., Barbosa-Morais,N.L. and Thanaraj,T.A. (2006) ASD: a bioinformatics resource on alternative splicing. *Nucleic Acids Res.*, **34**, D46–D55.
12. Rozen,S. and Skaletsky,H. (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.*, **132**, 365–386.
13. Yeo,G.W., Van Nostrand,E., Holste,D., Poggio,T. and Burge,C.B. (2005) Identification and analysis of alternative splicing events conserved in human and mouse. *Proc. Natl Acad. Sci. USA*, **102**, 2850–2855.
14. Freitas,R.A. (1999) *Appendix C: Catalog of Distinct Cell types in the Adult Human Body* Landes Bioscience, Georgetown, TX.
15. Leparc,G.G. and Mitra,R.D. (2007) Non-EST-based prediction of novel alternatively spliced cassette exons with cell signaling function in *Caenorhabditis elegans* and human. *Nucleic Acids Res.*, **35**, 3192–3202.
16. Malboeuf,C.M., Isaacs,S.J., Tran,N.H. and Kim,B. (2001) Thermal effects on reverse transcription: improvement of accuracy and processivity in cDNA synthesis. *Biotechniques*, **30**, 1074–1078.