Expression profiling of antisense transcripts on DNA arrays

Andreas Werner,1 Gabriele Schmutzler,1 Mark Carlile,1 Colin G. Miles,2 and Heiko Peters2

1Epithelial Research Group, Institute for Cell and Molecular Biosciences, and 2International Centre for Life, Newcastle University, Newcastle upon Tyne, United Kingdom

Submitted 20 June 2006; accepted in final form 6 November 2006

Werner A, Schmutzler G, Carlile M, Miles CG, Peters H. Expression profiling of antisense transcripts on DNA arrays. Physiol Genomics 28: 294–300, 2007. First published November 14, 2006; doi:10.1152/physiolgenomics.00127.2006.—The majority of mouse genes are estimated to undergo bidirectional transcription; however, their tissue-specific distribution patterns and physiological significance are largely unknown. This is in part due to the lack of methodology to routinely assess the expression of natural antisense transcripts (NATs) on a large scale. Here we tested whether commercial DNA arrays can be used to monitor antisense transcription in mouse kidney and brain. We took advantage of the reversely annotated oligonucleotides on the U74 mouse genome array from Affymetrix that hybridize to NATs overlapping with the sense transcript in the area of the probe set. In RNA samples from mouse kidney and brain, 11.9% and 10.1%, respectively, of 5,652 potential NATs returned positive and about half of the antisense RNAs were detected in both tissues, which was similar to the fraction of sense transcripts expressed in both tissues. Notably, we found that the majority of NATs are related to the sense transcriptome since corresponding sense transcripts were detected for 92.5% (kidney) and 74.5% (brain) of the detected antisense RNAs. Antisense RNA transcription was confirmed by real-time PCR and included additional RNA samples from heart, thymus, and liver. The randomly selected transcripts showed tissue specific expression patterns and varying sense/antisense ratios.

The results indicate that antisense transcriptomes are tissue specific, and although pairing of sense/antisense transcripts are known to result in rapid degradation, our data provide proof of principle that the sensitivity of commercial DNA arrays is sufficient to assess NATs in total RNA of whole organs.

natural antisense transcript; DNA array; RNA; antisense transcriptome

THE NUMBER OF PREDICTED ANTISENSE transcripts in mammalian genomes has risen considerably in the last five years from a few hundred to >50% of all genes in the latest report (15, 26). These estimates are based on large-scale cDNA sequencing and expressed sequence tag (EST) database mining with a strong bias toward fully processed transcripts. Antisense RNA synthesis involves RNA polymerase II and general processing of the primary transcript by capping, polyadenylation, and splicing. The biological functions of these transcripts as well as many mechanistic aspects of antisense mediated gene regulation remain speculative. A recently suggested concept that integrates functional and genomic aspects predicts at least two different categories of antisense transcripts (32, 33). A group of long noncoding transcripts are related to monoallelic expression as observed in imprinting, X chromosome inactivation, or allelic exclusion in leukocytes. Examples of such noncoding transcripts include the antisense RNA Air, Tsix/Xist, or, to a certain extent, the intergenic transcripts in V(D)J recombination (19, 28). These transcripts are likely to recruit silencing factors [or in the case of V(D)J allow recruitment of a recombination machinery] but do not require an overlap with the cognate sense transcripts (2, 13, 27). The events are restricted to the allele from which the noncoding (antisense) RNA is transcribed, thus acting in cis. The second category of natural antisense transcripts (NATs) is far more abundant than the first one but less understood. These NATs are significantly under represented on mammalian X chromosomes indicating that biallelic expression of overlapping transcripts may represent some form of evolutionary advantage (6, 17). A putative interaction of sense and antisense transcripts (most likely via RNA-RNA hybridization) is supported by the fact that overlapping RNAs are likely to be expressed in the same tissues and share responsiveness to specific hormones (5, 9, 14, 15, 30).

The current lack of suitable tools to assess the expression of sense and antisense RNAs on a routine basis limits progress in antisense research (18). The methodologies that currently provide this information such as tiling arrays or large-scale sequencing strategies are not suitable to routinely address physiological questions. At present commercial arrays are designed to hybridize mainly to (protein encoding) sense transcripts and neglect most antisense transcripts. An experimental method to label both strands of an RNA preparation independently is conceivable; however, this strategy may prove unsuccessful due to the low level of antisense expression and the often limited coverage of the DNA arrays. Here we provide the proof of principle that both coverage and sensitivity of commercial DNA oligo arrays are sufficient for monitoring antisense RNA expression in total RNA on a genome-wide scale. In addition, we show that a significant proportion of randomly selected sense/antisense pairs show tissue-specific expression patterns, strongly suggesting that antisense transcripts may play a physiological role.

MATERIALS AND METHODS

RNA extraction. RNA was extracted from 2-wk-old C57BL/6J mice or, for one DNA array experiment, 5-wk-old male 129SvJ mouse brain. The animals were killed by cervical dislocation according to the UK Home Office guidelines and the relevant personal and project licences. Tissues (kidney, brain, thymus, heart, and liver) were dissected, rinsed in ice-cold PBS, and either directly used for RNA extraction or frozen in liquid N2 and stored at −70°C. RNA was extracted with TRIzol (Invitrogen) for array experiments or TriReagent (Sigma) for RT-PCR according to the supplier’s protocol. Concentration of the RNA samples was determined spectrophotometrically.

DNA arrays. Two generations of Affymetrix mouse genome arrays were used, the U74A and U74B chips that each contained ~25% probes annotated in reverse orientation and the 430 array as a positive reference. The U74A array carries 12,588 probes, of
which 2,611 were in reverse complement orientation; on the U74B chips of 12,570 probes, 3,041 were wrongly annotated. The U74C chip that contains 60% wrongly annotated probe sets was not available for this study. U74A and parts of U74B contained probes representing fully annotated genes and ESTs, whereas the remaining part of U74B and U74C contained only ESTs (likely the major source of the wrongly annotated sequence information). The mouse 430 array contains 45,102 probe sets. [The 430 array is supplied in two versions, one that contains the probe set on one chip (430.2, used for kidney samples); the other has the identical probe set spotted on two chips (430A and 430B, used for brain samples). Over all, the probe sets are identical and will be referred to as “mouse genome array 430”].

Isolation of total RNA, reverse transcription, second-strand DNA synthesis, and synthesis of biotin-labeled cRNA were performed with kits and reagents according to the manufacturer’s protocol (Affymetrix). In brief, 15 μg of total RNA were primed with an oligo dT adaptor primer carrying a T7 polymerase promoter. Reverse transcription using SuperScriptII was performed at 42°C for 1 h. After second-strand cDNA synthesis and cDNA clean-up, biotin-labeled cRNA was synthesized using T7 RNA polymerase (note that cDNA that may be synthesized from contaminating traces of genomic DNA during second-strand cDNA synthesis will not be transcribed into cRNA due to the absence of the T7 promoter). The cRNA was then purified by removal of unincorporated NTPs and fragmented to 35–200 base fragments by metal-induced hydrolysis. Hybridization and washing steps were performed according to the manufacturer’s protocol (Affymetrix). The microarray analysis suite 5.0 (Affymetrix) was used to quantify the fluorescent signals and to determine probability values for expression [as: (P)resent, (M)arginal, (A)bsent]. A mask provided by Affymetrix identified the wrongly annotated probe sets on the U74A and U74B chips. The data from the U74 and 430 arrays were compared and annotated using the Affymetrix web portal NetAffx. The array data are available at GEO accession number GSE6106.

Real-time RT-PCR. We reverse transcribed 0.5 μg of total RNA in a 20-μl reaction using oligo dT primers and reagents from Qiagen. Omniscipt reverse transcriptase was used in the presence of 2 U/μl RNase inhibitor (RNasin, Promega). The reaction was performed at 37°C for 1 h. A reaction without enzyme was run as a negative control. After denaturation for 5 min at 95°C, aliquots of the reaction were either directly used for real-time PCR or stored at −20°C. We used 0.6 μl of cDNA for PCR in a 10-μl reaction containing 250 nM of forward and reverse primer, 4 mM MgCl₂, and 1 × Absolute QPCR capillary SYBR green mix (without MgCl₂) (Abgene). Reactions were performed in a Roche Light Cycler. Cycling initiated with a polymerase activation step of 15 min at 95°C followed by 50 cycles denaturation at 95°C (10 s), annealing at 55°C (10 s), and elongation 72°C (15 s). The reaction was terminated with the recording of the melting curve. Selected end point amplicons were analyzed by 2.5% Agarose gel electrophoresis. All PCR experiments were repeated at least once under identical conditions and at least four times with different batches of RNA or cDNA.

Primer design. Primers were designed using the net platform http://biotools.umassmed.edu/bioapps/primer3_web.cgi. The melting temperature of the primers was 60°C. The generated amplicons of ~120 base pairs were required to span an intron. The sequences of all primers as well as the cognate sequence accession numbers are given in Supplementary Table 1. The online version of this article contains supplementary material.

RESULTS

The first version of the mouse genome U74 arrays contained a considerable number of oligonucleotides in reverse complement orientation (20). These probes will hybridize to antisense RNAs that overlap with the cognate sense RNA in these exons. A schematic description of probe hybridization to correctly or wrongly annotated oligonucleotides is given in Fig. 1. These “faulty” DNA arrays enabled us to assess whether the coverage and the sensitivity of the arrays were sufficient to detect a significant number of antisense transcripts in a routine RNA preparation. Two probes were generated from total mouse kidney and brain RNA. Aliquots were hybridized to the mouse U74A and U74B chips as well as to the mouse genome array 430. The raw data from the U74A and B array contained signals derived from correctly and wrongly annotated probe sets. With the help of a spreadsheet-based mask from Affymetrix, the reversely oriented probes were identified and separated. This resulted in two U74-related probe sets, one representing the sense transcripts, the other representing the antisense transcripts. The U74 sense dataset was qualitatively comparable to the 430 dataset (not shown), thus validating our approach. Comparison of antisense U74 and 430 datasets revealed

![Fig. 1. Schematic representation of a bidirectionally transcribed exon with adjacent intronic sequences. The spliced sense and antisense transcripts are hybridized to correctly annotated probe (left) and to the faulty probe (right).](image-url)
expression levels of overlapping sense/antisense pairs. The strategy applied for data analysis is outlined in Fig. 2.

In a first step, \( P \) values and signal intensity of selected sense transcripts were assessed on both the U74 and the 430 arrays. A comparable expression pattern validated the data obtained with the U74 chips (not shown). Next, we focused on the reversely annotated probe sets. The mask from Affymetrix identified a total of 5,652 reverse-complement probes on the readout from the experiments performed with both kidney and brain probes. In the kidney sample, 672 (11.9\%) positive calls, indicative for antisense transcripts were obtained; in brain 573 (10.1\%). We found 309 transcripts in both kidney and brain (46.0 and 53.9\%, respectively). Comparable values for sense transcripts were obtained with the 430 arrays with 51.8\% (23,240) positive calls for kidney and 42.4\% (19,138) for brain. We found 10,910 transcripts to be expressed in kidney (46.9\% of 23,240) and brain (57.0\% of 19,138). Therefore, the relative numbers of coexpressed transcripts are similar for sense and antisense RNAs. A summary of the data is given in Table 1, top.

The expressed antisense transcripts were related to the cognate sense transcripts as monitored on the 430 arrays. The comparison of the two datasets was hampered by the fact that the probes on the 430 array derive from an updated genome version, thus carrying new identities, and were designed with a new algorithm. As a consequence, the comparison was done manually and revealed that only 268 of the 672 positive calls with kidney RNA on U74 could be assigned to probe sets on the 430 array. Of the 268 sense/antisense pairs 248 (92.5\%) showed coexpression, 19 (7.1\%) of the sense RNAs were not expressed. From the 573 positive calls in the brain, 212 could be related to probes on the 430 array. Of these 212 pairs, 158 (74.5\%) were coexpressed and 49 (23.1\%) sense RNAs were absent. We detected 129 of coexpressed transcript pairs in both kidney and brain. Most interestingly bidirectionally active genes appear to be preferentially coexpressed in kidney and brain (Table 1, bottom).

To corroborate the array results we selected 10 bidirectionally transcribed genes from the 129 coexpressed transcript pairs for validation by real-time RT-PCR. The selection depended on the accessibility of information concerning the genomic structure of potential sense/antisense pairs. We preferentially used the sense/antisense database of the PHANTOM3 consortium (http://fantom31p.gsc.riken.jp/s_as/) and compared the results to Ensembl annotations (http://www.ensembl.org/Mus_musculus/index.html). This process was hampered by the fact that we identified numerous antisense transcripts that were not represented in the database. We often found evidence for bidirectional transcription of the cognate loci in the form of close genes in tail-to-tail arrangement but no documented overlap that would enable primer design. Figure 3 gives an example of an antisense transcript identified on the U74 array, how the locus is represented in the FANTOM and the Ensembl database, and where the primers are localized. To exclude genomic DNA from contaminating the RT-PCR we designed primer pairs to locate on different exons. RNA was isolated from kidney, brain, heart, thymus, and liver and used for real-time RT-PCR following standard procedures. The 10 primer pairs for sense and antisense transcripts were used in parallel reactions. \( \beta \)-Actin primers served as positive controls; reactions without reverse transcriptase and reactions without template served as negative controls. MgCl\(_2\) concentration and primer concentrations were adjusted to optimize the results. Three parameters were taken into account to evaluate the data: The cycle threshold \((cT)\) values, the melting curves, and agarose gel electrophoresis of the end phase amplicon. Based on these criteria, five of the sense specific primer pairs and seven of the antisense specific primers produced specific amplicons (Table 2). The high \( cT \) values observed for a number of primer pairs (6 and 10 for sense, 2 for antisense) may reflect inefficient amplification rather than low abundance of the template, or both. A summary of the results is given in Table 2. The values were compared with actin and a cut-off of \((cT\text{ actin} + 10\text{ cycles})\) was set to exclude the very inefficiently amplified fragments (sense, 6 and 10; antisense, 2 and 6). As shown in Fig. 4, sense and antisense transcripts were expressed in a tissue-specific manner. Deviations of up to 39-fold for both sense and antisense, within the assessed tissues, were observed. However, it has to be taken into account that the level of \( \beta \)-actin may vary between the different tissues; therefore, the RT-PCR results are qualitative. Nevertheless, the results confirm that the expression of antisense transcripts is regulated in

Table 1. Summary of the mouse genome array experiment

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Coexpressed in Kidney/Brain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Antisense probe sets</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kidney RNA</td>
<td>5,652</td>
</tr>
<tr>
<td></td>
<td>Brain RNA</td>
<td>5,652</td>
</tr>
<tr>
<td></td>
<td>Sense probe sets</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kidney RNA</td>
<td>45,102</td>
</tr>
<tr>
<td></td>
<td>Brain RNA</td>
<td>45,102</td>
</tr>
<tr>
<td></td>
<td>NATs Detected on U74</td>
<td></td>
</tr>
<tr>
<td>Kidney RNA</td>
<td>672</td>
<td>268</td>
</tr>
<tr>
<td>Brain RNA</td>
<td>573</td>
<td>212</td>
</tr>
</tbody>
</table>

\( cT \): readout from the hybridization experiment. The data for both U74 and 430 arrays are given. The signals derived from the U74 chip were filtered with the mask from Affymetrix to detect reverse complement probes. The present calls are indicated as well as the proportion of probes that showed expression in both kidney and brain RNA. Bottom: comparison of sense and antisense transcripts in kidney and brain RNA. The transcripts with positive calls were related to the correctly spotted 430 arrays to obtain the cognate sense RNA derived signals (on 430).
a tissue-specific manner. The widespread coexpression of sense and antisense transcripts suggests that the expression patterns may be related. Previous reports on the tissue-specific expression of sense and antisense transcripts corroborate our findings thus validating the DNA array approach (5, 15, 16, 31).

DISCUSSION

A major obstacle in deciphering the physiological roles of bidirectional transcription is the deficit in available tools that allow the routine coassessments of sense/antisense transcription. We report the proof of principle that commercial (Affymetrix) DNA chips provide sufficient coverage and sensitivity to detect a representative proportion of natural antisense transcripts. We found that antisense transcripts show tissue-specific expression similar to protein encoding sense transcripts. Interestingly, antisense transcriptomes in different tissues overlap and show significant coexpression with the cognate sense transcripts.

Antisense transcripts are increasingly recognized to have gene-regulating potential. Their expression pattern is time and tissue dependent, a finding that is corroborated by few specific examples of antisense transcripts implicated in embryonic development and cancer progression (25). To elucidate the largely unknown impact of NATs on biological processes would require monitoring the antisense transcriptome under physiological and pathological circumstances. The reversely annotated probes on the U74 chip series offer the opportunity to deliver the proof of principle for this kind of experiment. The probe sets only cover ~25% of each sequence. This raised the question of whether a significant proportion of polyadenylated antisense transcripts overlap with and bind to these probes. In addition, antisense RNAs are usually expressed at a low level, which again may be critical for detection. We
Expression of sense and antisense RNA pairs by quantitative RT-PCR. Amplicons were chosen to have a length of 120 bp with primer melting temperatures of 60°C. Standard protocols were followed for reverse transcription and SYBR green qPCR. At least 2 independent cDNA synthesis and 4 PCR were done with every primer pair. The RNA was DNase treated; “minus reverse transcriptase” samples and “no DNA” or “water” PCRs were run as negative controls. Only primer pairs that resulted in specific amplicons of the correct size are included. The numbers represent the cycle threshold (cT) values of a representative experiment (nd, not detectable). The relatively high cT values are discussed in the text. The accession numbers of genes related to actin (cT sample in combination with exon arrays will bring routine optimization of the cognate reactions. These values were not included into the assessment of tissue-specific expression of sense/antisense pairs (Fig. 4). Nevertheless, we could confirm a significant number of transcripts identified on the chip. To conclude, we show that commercial DNA arrays offer a tool to monitor antisense expression. A protocol that allows the independent labeling of sense and antisense strand of an RNA sample in combination with exon arrays will bring routine examinations of the antisense transcriptome within reach.

obtained ~10% positive calls in both kidney and brain, which is lower than recent predictions (15, 17). The most relevant figure for this study comes from the FANTOM consortium, estimating that 72% of all transcriptional units in mouse show evidence of overlapping RNA (15). To explain this discrepancy we have to take the limited coverage of the array into account and to acknowledge that the reversely annotated oligonucleotides of the U74 chips may not provide a fully representative selection. This bias will become irrelevant with the use of exon arrays (10).

Recent reports by Kiyosawa et al. (16) and Cheng et al. (7) found a substantial number of nonpolyadenylated transcripts compared with polyadenylated (overlapping) RNAs. The intimate relation between transcription and RNA processing, in this case polyadenylation, implies divergent synthesis pathways for the two groups of RNA and possibly different cellular fates (24). Common labeling procedures, including our own, select for polyadenylated transcripts. This bias may prove advantageous to distinguish between transcripts of different biological function. Fully processed RNAs are more stable and thus less likely to be byproducts of transcription (transcriptional noise) (8, 29).

The validation of the array results was performed by real-time PCR. Primer design was hampered by two major constraints. First, records of both sense and antisense transcripts (ESTs, cDNAs) were required. Our findings comply with others that antisense transcripts are often unannotated. Second, RNA preparations contain varying traces of genomic DNA that may interfere with the amplification. To avoid this pitfall we designed all primers to contain the upstream intron adjacent to the probe set. This requirement may have had impact on the quality of the primers, and as a consequence high cT values were observed in a number of samples. The input from reverse transcription is limited to 10% of the PCR volume; therefore, we ran 50 cycles to monitor expression of the transcript pairs. This resulted in a number of borderline values (cT >40, Table 2) that would require either redesign of the primers or further optimization of the cognate reactions. These values were not included into the assessment of tissue-specific expression of sense/antisense pairs (Fig. 4).

Table 2. Summary of cT values from RT-PCRs

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sense</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kidney</td>
<td>20.8</td>
<td>27.2</td>
<td>25.9</td>
<td>44.1</td>
<td>20.6</td>
<td>41.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brain</td>
<td>21.6</td>
<td>29.3</td>
<td>26</td>
<td>35.2</td>
<td>25.1</td>
<td>38.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liver</td>
<td>23</td>
<td>29.9</td>
<td>30.4</td>
<td>nd</td>
<td>25.6</td>
<td>nd</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thymus</td>
<td>23.3</td>
<td>34.3</td>
<td>29.4</td>
<td>nd</td>
<td>25.7</td>
<td>nd</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heart</td>
<td>23.7</td>
<td>30</td>
<td>30.8</td>
<td>44.5</td>
<td>26.8</td>
<td>43.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Antisense</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kidney</td>
<td>20.8</td>
<td>24</td>
<td>36.7</td>
<td>26.7</td>
<td>25.4</td>
<td>30.8</td>
<td>26.1</td>
<td>28.5</td>
</tr>
<tr>
<td>Brain</td>
<td>21.6</td>
<td>24.3</td>
<td>41.2</td>
<td>22.8</td>
<td>26.5</td>
<td>34.9</td>
<td>27.5</td>
<td>28.2</td>
</tr>
<tr>
<td>Liver</td>
<td>23</td>
<td>27.8</td>
<td>nd</td>
<td>30.2</td>
<td>31.2</td>
<td>37.6</td>
<td>29.5</td>
<td>32.6</td>
</tr>
<tr>
<td>Thymus</td>
<td>23.3</td>
<td>24.7</td>
<td>nd</td>
<td>27.9</td>
<td>30</td>
<td>38.3</td>
<td>29.8</td>
<td>30</td>
</tr>
<tr>
<td>Heart</td>
<td>23.7</td>
<td>30</td>
<td>44.3</td>
<td>30.9</td>
<td>32.9</td>
<td>40</td>
<td>31.8</td>
<td>35.1</td>
</tr>
</tbody>
</table>

Fig. 4. A: expression levels of overlapping transcripts in different tissues. The bars represent cycle threshold (cT) values obtained with the indicated primer pair and RNA related to actin (cT sample – cT actin). At least 4 qualitative equivalent sets of data were generated. The error bars represent SD.

The validation of the array results was performed by real-time PCR. Primer design was hampered by two major constraints. First, records of both sense and antisense transcripts (ESTs, cDNAs) were required. Our findings comply with others that antisense transcripts are often unannotated. Second, RNA preparations contain varying traces of genomic DNA that may interfere with the amplification. To avoid this pitfall we designed all primers to contain the upstream intron adjacent to
We have found extensive coexpression of sense and antisense transcripts in the RNA preparations from the kidney and brain. In general terms, similar observations have been made when the expression pattern of antisense transcripts in mammalian systems was evaluated by tiling arrays and large-scale sequencing approaches (1, 5, 15, 16, 23). Also our experimental validation of a few antisense transcripts showed tissue-specific expression patterns that reflected a developmental and hormonal state (15, 23). Note that our findings confirm that expression of sense and antisense does not follow a general pattern of co- or reciprocal regulation (15); interestingly, however, the data suggest that this pattern may vary for a specific transcript pair between different organs. Another striking feature represents the strong bias of coexpressed sense/antisense pairs toward expression in both kidney and brain. This could indicate that coexpression of sense and antisense transcripts provides a very general positive impact on mRNA and/or protein synthesis. The two transcripts could theoretically derive from different cell populations; however, the extent of coexpression argues against a strict local separation of sense and antisense transcription. We therefore assume that a significant number of NATs are coexpressed with their cognate sense RNAs in the same cell population. Together with the preferred conservation of NAT pairs with their cognate sense RNAs in the same cell population, the coexpression would suggest a putative regulatory mechanism that involves hybrid formation of fully processed sense and antisense transcripts.

Our results are based on the reversely annotated probe set on the Affymetrix U74A and B arrays. This resulted in a serendipitous probe “selection” that excludes any bias toward the genes screened for antisense transcription. We found more antisense transcripts in the kidney compared with the brain and sense/antisense coexpression also seemed higher in the kidney than in the brain. The lower general transcriptional activity in brain might account for this observation. Sun et al. (31) reported higher antisense expression in mammalian brain compared with other tissues. It is tempting to speculate about biological explanations for these findings especially since regulatory RNAs are potential molecular mediators of higher eukaryote brain complexity (11, 21). We suggest that the reason for the discrepancies between our data and published reports is the biological background of the different datasets. Our data derive from mice with defined genetic background and age, whereas in silico compiled datasets usually do not discriminate strains and developmental stages. Considering the regulated and potentially dynamic expression of the antisense transcriptome it is conceivable that the snapshot presented here may differ from the bioinformatic average. In addition, experimental variations may add to the observed differences. For example, the dataset is relatively small, limited by unique chip design, and thus more susceptible to experimental variations. Several reports have highlighted the fact that cDNA preparation gives rise to a small proportion of reverse complement sequences (7). These traces, however, should not account for the variations. The probe synthesized for chip hybridization strictly derives from primed RNA since contaminations from genomic DNA will not include the T7 RNA polymerase promoter. The RT-PCR results are again unlikely to derive from unspecific priming because the all amplicons span an intron.

Natural antisense RNAs help to organize fundamental biological processes such as embryonic development and cancer formation (4, 12, 22). Consequently, the expression of NATs is tightly regulated with respect to tissue and age (3). We provide proof of principle that commercial DNA arrays are suitable tools to map the antisense transcriptome. This will eventually help to unravel the putative role of NATs in eukaryotic gene regulation.

ACKNOWLEDGMENTS

We thank Affymetrix for generously providing the U74 arrays, Michael Jackson and Maria Lastowska for sharing data, and the Biotechnology and Biological Sciences Research Council for funding (studentship to M. Carlie).


