Combined in vitro transcription and reverse transcription to amplify and label complex synthetic oligonucleotide probe libraries

Yusuf Murgha¹, Brian Beliveau², Kassandra Semrau¹, Donald Schwartz¹, Chao-ting Wu², Erdogan Gulari¹,³, and Jean-Marie Rouillard¹

¹ MYcroarray, Ann Arbor, MI, ² Department of Genetics, Harvard Medical School, Boston, MA, and ³ Department of Chemical Engineering, University of Michigan, Ann Arbor, MI

Keywords: in vitro transcription (IVT); reverse transcription (RT); single-stranded DNA (ssDNA) libraries; fluorescent in situ hybridization (FISH); oligonucleotide library pools; microarrays; Oligopaint FISH probes

Supplementary material for this article is available at www.BioTechniques.com/article/114298.

Oligonucleotide microarrays allow the production of complex custom oligonucleotide libraries for nucleic acid detection–based applications such as fluorescence in situ hybridization (FISH). We have developed a PCR-free method to make single-stranded DNA (ssDNA) fluorescent probes through an intermediate RNA library. A double-stranded oligonucleotide library is amplified by transcription to create an RNA library. Next, dye- or hapten-conjugate primers are used to reverse transcribe the RNA to produce a dye-labeled cDNA library. Finally the RNA is hydrolyzed under alkaline conditions to obtain the single-stranded fluorescent probes library. Starting from unique oligonucleotide library constructs, we present two methods to produce single-stranded probe libraries. The two methods differ in the type of reverse transcription (RT) primer, the incorporation of fluorescent dye, and the purification of fluorescent probes. The first method employs dye-labeled reverse transcription primers to produce multiple differentially single-labeled probe subsets from one microarray library. The fluorescent probes are purified from excess primers by oligonucleotide-bead capture. The second method uses an RNA:DNA chimeric primer and amino-modified nucleotides to produce amino-allyl probes. The excess primers and RNA are hydrolyzed under alkaline conditions, followed by probe purification and labeling with amino-reactive dyes. The fluorescent probes created by the combination of transcription and reverse transcription can be used for FISH and to detect any RNA and DNA targets via hybridization.

METHOD SUMMARY
Here, in vitro transcription and reverse transcription are combined to generate large quantities of single-stranded DNA probes. The individual methods utilized herein can be modified with no sacrifice in product quality, quantity, and manufacturing time, enabling researchers to alter the type and density of dye molecules on the single-stranded probes as well as simplifying intermediary and final product purification.
enables fine-resolution (a few kilobases) genome analysis (10–12).

However, the amount of oligonucleotide library generated from a single microarray is limiting. This necessitates pre-amplification steps followed by the introduction of fluorescent dyes to obtain probes. Currently, PCR is the preferred method to amplify and label oligonucleotide libraries. The fluorescent label is added to the oligonucleotides either using dye-conjugated primers or modified nucleotides. In the latter case, the modified nucleotides are incorporated by nick translation into the PCR amplicons followed by coupling with fluorescent dyes (13). PCR requires flanking primer binding sequences for amplification, which unless blocked may increase the probability of unwanted hybridization. Furthermore, the double-stranded amplicons may self- or cross-hybridize to other oligonucleotide library sequences, especially at primer binding sites, requiring increased probe concentration to obtain sufficient signal. Beliveau et al. proposed a PCR-based process that uses modified primers, nicking endonucleases, and polyacrylamide purification to obtain single-stranded DNA (ssDNA) fluorescent probes (Oligo-paint FISH probes) (11). Furthermore, they presented multiple protocols for sample preparation, interphase and metaphase FISH, and quick 3-D FISH adapted from Holmberg et al. (18).

Here we present a PCR-independent process to produce labeled single-stranded fluorescent probes starting from oligonucleotide libraries using a combination of in vitro transcription (IVT) and reverse transcription (RT). The probes are used for secondary detection of RNA on microarrays and in FISH. We describe multiple variations of the method to show versatility in the probe generation process. For example, the process is modified to generate multiple fluorescent sub-libraries for FISH from a single de novo synthesized library. Finally, the process can be scaled up to produce several hundred micrograms of labeled probe libraries (17).

Materials and methods

In vitro transcription

Custom double-stranded oligonucleotide libraries (MYlib) from MYcroarray (Ann Arbor, MI) were used as the templates for IVT using MEGAscript T7 Kit according to manufacturer’s recommendation (Life Technologies, Carlsbad, CA). Step-wise protocols for production of fluorescent probes are detailed in the Supplementary Material. Briefly, 2 pmol of MYlib were transcribed at 37°C for 4 h in a thermocycler, followed by hydrolysis of template DNA with 2 U TURBO DNase enzyme at 37°C for 15 min. EDTA (40 mM, pH 8.0) was added to stop DNase activity, followed by heat inactivation at 75°C for 10 min. The RNA was purified on RNeasy spin columns (Qiagen Inc., Valencia, CA) and RNA concentration was measured with a NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, DE).

Table 1. Primer sequences

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’–3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>TAATACGACTCACTATAGGGAGA</td>
</tr>
<tr>
<td>P2</td>
<td>AGTCATGAAGATCATACCGATATC</td>
</tr>
<tr>
<td>P3</td>
<td>AGTCATGAAGATCATACCGATATC</td>
</tr>
<tr>
<td>F1</td>
<td>/56-FAM/CTATCGTGCAAGGGTGAATGC</td>
</tr>
<tr>
<td>F2</td>
<td>/5Cy3/CTATCGTGCAAGGGTGAATGC</td>
</tr>
<tr>
<td>F3</td>
<td>/5Cy5/AGTCATGAAGATCATACCGATATC</td>
</tr>
</tbody>
</table>

Preparation of multiple fluorescent probe sets from a single oligonucleotide library for FISH

Reverse transcription. cDNA was synthesized using Superscript II Reverse Transcriptase (Life Technologies). The manufacturer’s standard reaction setup was modified to process a larger amount of RNA template per unit reaction volume, where the RNA input concentration was 10 pmol/μl, and the deoxynucleotide concentration was increased from 0.5 mM to 1.5 mM. Briefly, 1000 pmol of RNA library, 400 pmol each of 3 5’-end dye-labeled primers (F1, F2, and F3; Table 1) (Integrated DNA Technologies, Coralville, IA), and 40 U SUPERase-In (Life Technologies) were added to the mix and incubated at 42°C for 2 h, followed by a second addition of reverse transcriptase and a further 2 h incubation.

Preparation of RNA-DNA complex capture oligo-beads. MyOne Streptavidin C1 magnetic beads were prepared according to the manufacturer’s recommendations (Life Technologies). Here, 200 μl beads (10 μg/μl) were used per 1000 pmol input RNA template. The beads were washed (200 μl) once with 0.1 M NaOH, 0.05 M NaCl solution and twice with 2× Binding & Washing (B&W) buffer (2 M NaCl, 1 mM EDTA, 10 mM Tris-HCl, pH 7.5). Biotinylated oligonucleotide (1000 pmol 5’-biotinylated P2; Table 1) was immobilized on the beads (10 μg/μl) in 1× B&W buffer for 60 min at room temperature under constant shaking. The beads were collected on a magnetic stand for 2 min, and the supernatant was aspirated and discarded. The oligonucleotide-bound beads were washed once with 200 μl 2× B&W buffer and resuspended in 100 μl 2× B&W buffer.

Capture of RNA-DNA complexes. The RNA-DNA complex after reverse transcription was hybridized to P2-oligo magnetic beads at room temperature for 30 min with constant shaking. The beads were collected on a magnetic stand for 2 min, and the supernatant containing excess primers and uncaptured RNA-DNA hybrids was aspirated and saved for a second round of capture. The beads were washed once with 100 μl 1× B&W buffer followed by release of RNA-DNA hybrids from beads. The beads were reused for a second round of RNA-DNA capture.

Release of RNA-DNA complexes from magnetic beads. The protocol was adapted from Holmberg et al. (18). The RNA-DNA hybridized beads were suspended in 10 mM MgCl2 solution and incubated at 25°C for 5 s, followed by a slow ramp-up (+0.5°C/s) to 70°C, 5 s at 70°C, and immediate transfer to ice for 2 min to release the RNA-DNA hybrids from the magnetic beads. Following bead collection using a magnetic stand, the supernatant was aspirated out and saved for RNA hydrolysis. The beads were washed once with 100 μl 1× B&W buffer and the resuspended in 50 μl 2× B&W buffer for reuse.

Alkaline hydrolysis of RNA. The 100 μl solution obtained after 2 rounds of RNA-DNA complex release was
subjected to strand denaturation at 95°C for 5 min, followed by hydrolysis of RNA using 0.3 N NaOH for 15 min at 65°C. The solution was neutralized with equimolar HCl and buffered using 40 mM Tris-HCl, pH 8.0 and stored at -20°C for at least 30 min prior to ethanol precipitation. The fluorescent ssDNA pellet was resuspended in 20 µl nuclease-free water, and its concentration was measured with a Nanodrop spectrophotometer.

Preparation of fluorescent probes with multiple labels

RNA-DNA chimeric primer for cDNA synthesis. Amino-allyl incorporated cDNA (aa-cDNA) was synthesized with two modifications of the protocol outlined in the Reverse transcription section. First, we used 1.5-fold excess RNA-DNA chimeric primer (P3; Table 1), and second, the dNTP mix was changed as follows: 1 mM each dATP, dCTP, and dGTP; 0.3 mM dTTP; and 0.7 mM amino-allyl dUTP. Upon completion of the reaction, the RNA was degraded using the Alkaline hydrolysis of RNA protocol, and the buffered solution containing aa-cDNA was ethanol precipitated using 20 µg linear acrylamide, 0.3 M sodium acetate and 2.5 volumes ethanol.

Microarray analysis

Specific sample hybridization and washing protocols are described in the Supplementary Material. Briefly, the FISH libraries and the biotinylated RNA—either with or without dye-labeled ssDNA libraries—were hybridized overnight to Agilent gasket slides incubated in an oven at the set hybridization temperature. Following hybridization, the signal was removed from the array slide under 1× SSPE buffer (20x SSPE buffer: 0.02 M EDTA and 2.98 M NaCl in 0.2 M phosphate buffer, pH 7.4), washed twice for 3 min with 1× SSPE, and once with 0.25x SSPE solution under gentle stirring. The microarray was dried and scanned with an Axon GenePix 4000B scanner (Molecular Devices LLC, Sunnyvale, CA). To visualize biotinylated RNA, after the 0.25x SSPE wash the microarrays were immersed for 2 h at 16°C in 30 mL of a solution containing streptavidin Alexa Fluor 647 conjugate and then re-scanned.

For data analysis, the median spot signal intensity of probe spots corresponding to individual oligonucleotides in the oligonucleotide library were extracted using GenePix 4.1 software. Next, probe-target replicates were grouped together, and those that passed the quality filter threshold were used to calculate the trimmed median intensity of every oligonucleotide library sequence. Finally, all sequences with an intensity greater than 3 times the 95th percentile intensity of background probes were denoted as present, and the rest were denoted as absent.

DNA FISH with fluorescent probe sets

DNA FISH with fluorescent probe sets was performed as previously described (11,14). Briefly, 100 µL of a 1×10^6 cells/mL suspension of human diploid WI-38

---

**Table A**

<table>
<thead>
<tr>
<th>Step</th>
<th>Oligo sub-library</th>
<th>Chr X Start</th>
<th>Chr X Stop</th>
<th>Complexity</th>
<th>Label</th>
<th>Detection Array</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MYLib</td>
<td>73,633,445</td>
<td>73,948,034</td>
<td>3,020</td>
<td>6-FAM</td>
<td>NA</td>
</tr>
<tr>
<td>2</td>
<td>Qx13.2</td>
<td>73,948,034</td>
<td>74,892,014</td>
<td>7,000</td>
<td>Cy5</td>
<td>6963</td>
</tr>
<tr>
<td>3</td>
<td>Qx13.3</td>
<td>74,892,158</td>
<td>76,484,267</td>
<td>10,000</td>
<td>Cy3</td>
<td>9962</td>
</tr>
<tr>
<td>4</td>
<td>Qx13.3-q21.1</td>
<td>73,948,034</td>
<td>74,892,014</td>
<td>7,000</td>
<td>Cy5</td>
<td>6963</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(ATCC CCL-75) cells was added to 0.01% (v/v) poly-L-lysine treated glass microscope slides and allowed to adhere for 1 h at 37°C in a mammalian tissue culture incubator. The samples were then washed in 1× PBS and fixed in 1× PBS + 4% (w/v) paraformaldehyde for 5 min. The samples were transferred to 2× SSCT [0.3 M NaCl, 0.03 M sodium citrate, pH 7.0, and 0.1 % (v/v) Tween-20] and then 2× SSCT + 50% (v/v) formamide. Prior to the hybridization, the samples were incubated for 2.5 min in 2× SSCT + 50% formamide at 92°C, then for 10 min each in 2× SSCT and 0.2× SSC. Finally, the slides were mounted with SlowFade Gold + DAPI (Invitrogen, Waltham, MA) under a 22 × 30 cm #1.5 coverslip and sealed using nail polish. The DNA FISH was then imaged using a Zeiss (Thornwood, NY) LSM-780 laser scanning confocal microscope with a 63× oil NA 1.40 objective. Images were processed using Zeiss ZEN software and Adobe Photoshop.

**Results and discussion**

**Preparation of fluorescent probe sub-libraries from a master oligonucleotide library**

**Master library design.** The library design and process to prepare differentially labeled fluorescent sub-libraries (or probe sets) from one microarray library is depicted in Figure 1A. The 20,020 oligonucleotide dsDNA master library consists of 3 sub-libraries composed of fluorescent probes that target 3 adjacent regions of the human X chromosome, which together span a 2.8 Mb genomic region previously targeted by Oligopaint FISH probe sets (11). To assess sub-library isolation efficiency and identify sub-library coverage, the 3 ssDNA sub-libraries were labeled with 6-FAM (3020 oligonucleotides), Cy3 (7000 oligonucleotides), and Cy5 (10,000 oligonucleotides) dyes and co-hybridized to complementary detection microarrays. The Cy3 and Cy5 labeled ssDNA sub-libraries had greater than 99.5% coverage (Figure 1B), and images were processed using Zeiss ZEN software and Adobe Photoshop.
Preparation of fluorescent probes with multiple labels

**Reporter DNA and RNA library design.** A 4161 oligonucleotide dsDNA reporter library (MYlib) was used to demonstrate secondary detection of an RNA library using complementary fluorescent probes. The reporter library design and process to make aa-cDNA is shown in Figure 3. In parallel, we constructed an RNA library consisting of 4161 oligonucleotide sequences. Figure 4A shows the library structure and hybridization scheme for secondary detection. The 90-mer RNA sequence is split into two regions of equal length, one that is complementary to sequence L_N of the reporter DNA library and a second sequence (V_N) for hybridization to the detection array (microarray A, Figure 4A). Microarray B is used to determine the composition of the reporter library. The 45-mer sequences unique to the RNA were designed using OligoArray software (19).

**Library composition of RNA and reporter DNA libraries.** The reporter DNA and RNA libraries were independently hybridized to the respective detection microarrays to determine their composition and specificity. The reporter library detection microarray B has probes complementary to the RNA binding sequence L_N (Figure 4A). All 4161 oligonucleotides were qualified as present in the Cy3-labeled reporter library. There were 34 (0.82%) oligonucleotides that cross-hybridized to probes used to calculate background signal (Supplementary Table S1).

The composition of the Alexa Fluor 555 streptavidin-labeled biotinylated RNA library was determined for three independent IVT-RT amplification reactions. In all, 3769 (90.58%) RNA library sequences were qualified as present on 2 or more arrays, while the remaining 392 (9.42%) were absent (Supplementary Table S2). The distribution of RNA sequences per detection array is shown in Figure 4B. Finally, there were no false positive signals when the reporter library was hybridized to the RNA detection array.

**Detection of unlabeled RNA with reporter DNA.** A biotinylated RNA library and a Cy3-labeled reporter library were co-hybridized to the RNA detection array (microarray A, Figure 4A). The reporter library detected 3136 (83.21%) of the 3769 RNA library sequences previously qualified as present. It did not detect 645 previously reported sequences, while it reported the presence of 12 previously absent RNA library sequences (Supplementary Table S3). This result was
further validated by secondary labeling of the biotinylated RNA library. Post-labeling with Alexa Fluor 647–conjugated streptavidin, 3798 RNA library sequences were qualified as present. Of these, 29 previously undetected RNA library sequences were qualified as present because they passed the detection threshold criterion (Supplementary Table S4).

Across both detection methods (i.e., RNA detection by reporter DNA library and direct RNA labeling with Alexa Fluor dye), 3798 RNA library sequences were qualified as present and 631 (17%) RNA library sequences previously called present were not detected (Supplementary Table S5). Finally, 288 (6.92%) of the 4161 RNA molecules were not detected across the 9 arrays for the 3 detection experiments, namely: RNA library composition, unlabeled RNA library by reporter DNA library, and post-labeling of unlabeled RNA to confirm reporter DNA calls. The missing oligonucleotides in each of the respective custom libraries may be attributed to any of the following: synthesis failure, drop-out during in vitro transcription or reverse transcription, hybridization failure (probe synthesis failure, secondary structure, steric hindrance), or a concentration of individual oligonucleotides that is below the detection threshold (17).

We illustrate three variations of the IVT-RT combined protocol to make fluorescent probes. First, the fluorescent moiety can be added either directly using dye-conjugated RT primer, wherein the position and number of dye(s) can be controlled, or indirectly through conjugation of amine reactive moieties to amino-modified DNA. The adjacent dyes should be spaced greater than twice their Förster radius to prevent self-quenching. Also, the NHS-ester dyes (e.g., ATTO, Alexa Fluor) are released from the labeled probes during the alkaline hydrolysis step. Thus, an RNase enzyme cocktail is used to digest the RNA strand of RNA:DNA hybrids containing NHS-ester dyes. Second, the low efficiency of the RT reaction (30%–40%) leads to leftover primers. The primers are removed either by capture of RNA:DNA hybrids or using an RNA:DNA chimeric primer that is hydrolyzed into small fragments (3–5 nucleotides) and subsequently removed on filter spin-columns. Third, we describe differential labeling of sub-libraries from a single oligonucleotide library at the reverse transcription step. This reduces oligonucleotide library synthesis costs by packing many sub-libraries (100s to 1000s of probes per library) on a high-density microarray. As the differently labeled sub-libraries are obtained in a pool, they cannot be used independently unless the sub-libraries have unique internal primer sites. Here, individual sub-libraries are retrieved by nested PCR prior to amplification and labeling (20).
Normalized over reaction volume, the IVT-RT method gives a higher yield of ssDNA than the PCR-based method. A 20 μl IVT reaction yields ~40 μg RNA, resulting in ~10 μg of labeled ssDNA library, while the yield of a 20 μl PCR reaction of the same library is ~0.8 μg. In the PCR-based method, the fluorescent ssDNA probes are purified from intermediate products using PAGE (11), where the recovery of ssDNA probe varies from 30% to 90% (21,22). Instead, the use of lambda exonuclease to selectively hydrolyze the 5’ end-phosphorylated DNA strand gives consistent ssDNA yields (~70%) (17). Scaling both reactions to 96-well plates, where each well contains 100 μl PCR or 20 μl IVT reaction components, yields 4 μg PCR amplicons and 40 μg RNA, resulting in 3.6 μg and 10 μg ssDNA, respectively. The total amount of fluorescent probes is 345 μg from the PCR-lambda exonuclease method and 960 μg from the IVT-RT method. Finally, reverse transcriptases incorporate modified nucleotides more efficiently than the thermostable DNA polymerases used in PCR. Cox and Singer reported optimal amino-allyl dUTP and dTTP ratios to achieve a label density of 8 dyes per 100 bases during reverse transcription, compared with 2 to 4 dyes per 100 bases observed for PCR. The degree of labeling affects the FISH fluorescent signal, with the brightest signals obtained for 6–8 dyes per 100 bases (23).

Author contributions
Y.M. conceived, designed, developed, and performed the IVT-RT method and wrote the paper. B.B. conceived, designed, performed, and wrote the section on FISH experiments. K.S. assisted with conducting the IVT-RT protocol and final editing of the manuscript. D.S. supervised DNA microarray synthesis. C.-I.W. supervised the conception of experiments and the final editing of the manuscript. E.G. supervised the conception of experiments. J.-M.R. supervised the conception of experiments and the final editing of the manuscript.

Acknowledgments
This work was supported in part (oligonucleotide library synthesis) by a grant from NIH (1R43GM093579). This paper is subject to the NIH Public Access Policy.

Competing interests
Y. Murgha, K. Semrau, D. Schwartz, and J.-M. Rouillard are employees of MYCroarray. E. Gulari and J.-M. Rouillard are co-founders of MYCroarray. MYCroarray sells MYlib oligonucleotide libraries used in this project.

References

Received 06 November 2014; accepted 26 March 2015.

Address correspondence to Yusuf Murgha, MYcroarray, Ann Arbor, MI. E-mail: ymurgha@mycroarray.com

To purchase reprints of this article, contact: biotechniques@fosterprinting.com