Sequence-independent Selective Amplification of mRNAs over rRNAs

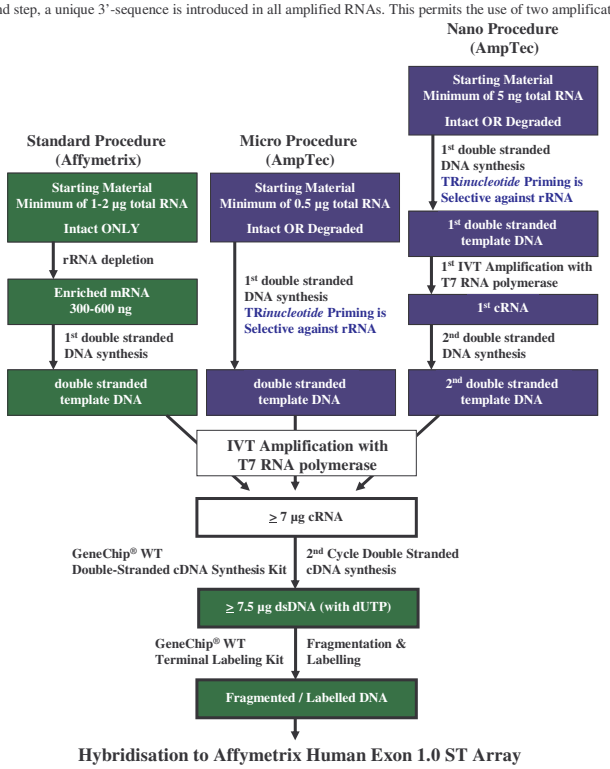
Extended Possibilities for Exon Arrays and for Bacterial Gene Expression Studies

John Arrand[1], Sim Sihota[1], Wenbin Wei[1] and Guido Krupp[2]
[1]Institute for Cancer Studies, Birmingham, U.K. and [2]AmpTec GmbH, Germany (krupp@amp-tec.com)

Standard mRNA amplifications for "All-Exon" microarrays and for bacterial RNAs are impossible with small samples (< 100 ng total RNA) and with degraded RNAs, because removal of rRNAs must precede universal, non-selective RNA amplification (Affymetrix and Ambion kits, respectively). This pre-treatment with magnetic beads is cumbersome, requires high amounts of starting material (>0.5 µg total RNA), is not universal for all species and degraded RNAs are not suitable. Omission of this step results in lost sensitivity [1].
However, AmpTec's unique TRinucleotide primer strategy renders reverse transcriptions selective for mRNA (fragments) and it provides sequence-independent, preferential 3'-proximal priming. This eliminates the rRNA removal step, thus high mRNA amplifications (input of <1 ng total RNA yields >10 µg amplified RNAs), and the use of degraded RNAs are possible.
A first Exon Array study with saliva RNA (<4ng of degraded RNA) is now in press [2].
[1] Pepper et al., 2007 "A Core Lab Case Study: Exon Array Challenges and Opportunities" Affymetrix Application Note.
[2] Hu et al., 2008 (Wong lab at UCLA) "Exon-level expression profiling: a comprehensive transcriptome analysis for oral fluids." Clinical Chemistry 54:5 (E-publication ahead of print).

Exon Arrays

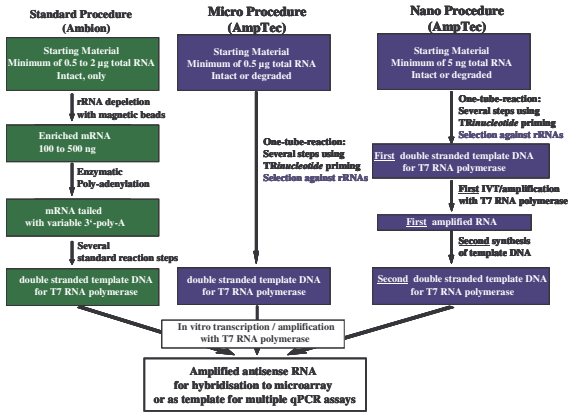
Comparison of Flow Diagrams for Whole Transcript (WT) Amplified Double-Stranded Target Assays
AmpTec's TR*inucleotide* priming technology results in full-length reverse transcription of mRNA sequences, independent of a 3'-polyA sequence. This permits the use of degraded RNA samples.
In a second step, a unique 3'-sequence is introduced in all amplified RNAs. This permits the use of two amplification rounds.



Hybridisation to Affymetrix Human Exon 1.0 ST Array

Bacterial mRNAs

Comparison of Flow Diagrams for Amplification of Bacterial mRNAs
AmpTec's TR*inucleotide* priming technology results in amplification of all mRNAs, independent of a universal 3'-sequence, and in effective exclusion of rRNA sequences.
In a second step, a universal 3'-sequence is introduced in all amplified mRNAs. This permits two amplification rounds.



Principle and Application of TR*inucleotide* Priming

Function of TR*inucleotide* Primers

Non-random distribution of priming sites resuls in
- essentially sequence-independent - priming near 3‘-ends of mRNAs and mRNA fragments

TR*inucleotide* primers have a unique sequence (BOX) at the 5‘-end, followed by a stretch of random and one chosen trinucleotide sequence at the 3‘-end. Transient primer binding for subsequent elongation requires the matching trinucleotide in the template sequence.
Polymerases can transiently bind to fork-like structures and to free 3‘-ends, which results in a high local polymerase concentration, as indicated by the blue area. The sterical proximity of both transient events results in prefential elongation of the primer close to the 3’-end of the template.

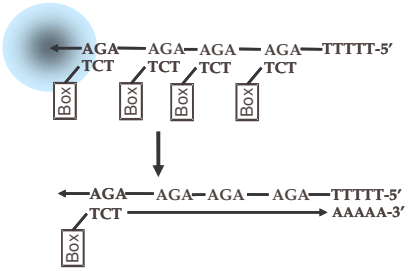
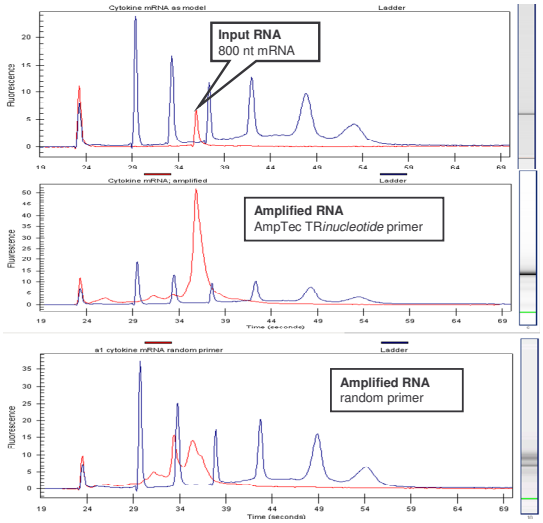
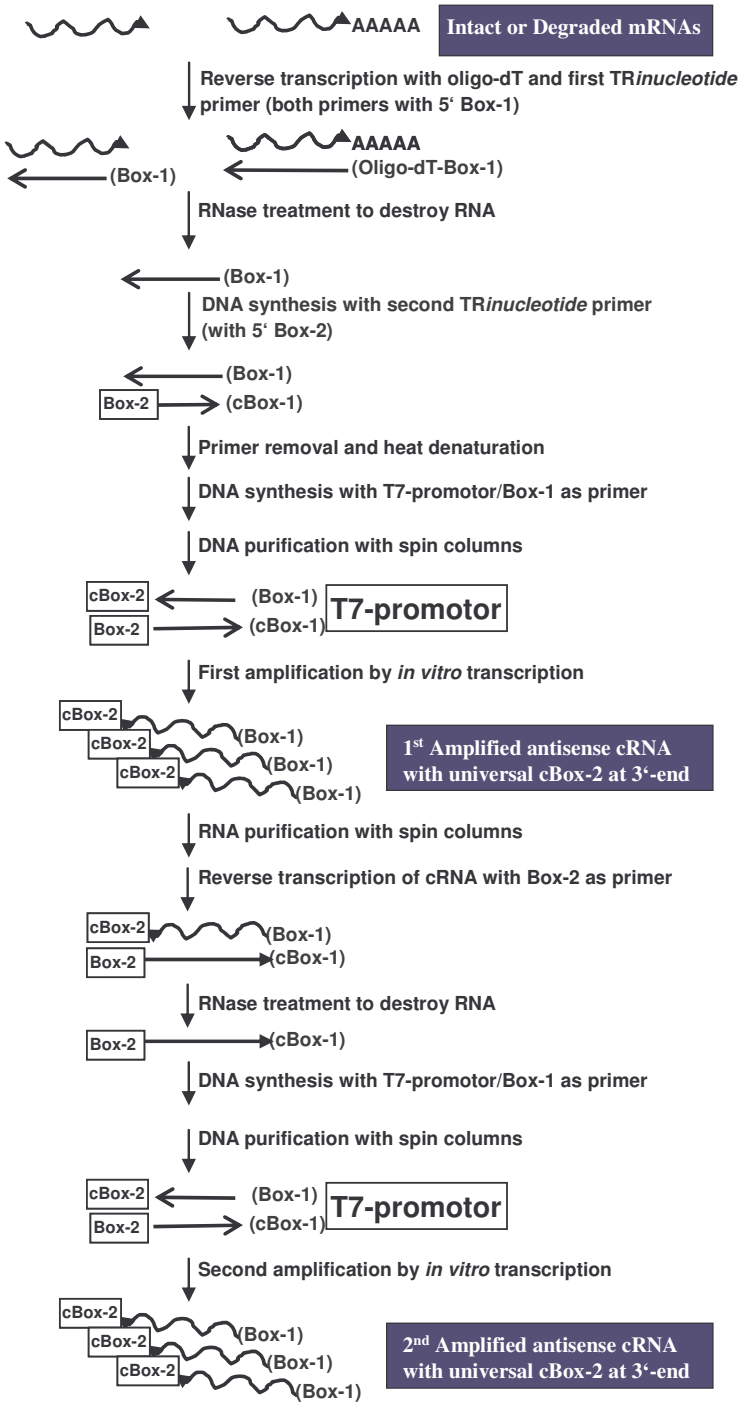


Illustration of preferential 3‘-priming by „full-length“ amplification of a model mRNA transcript

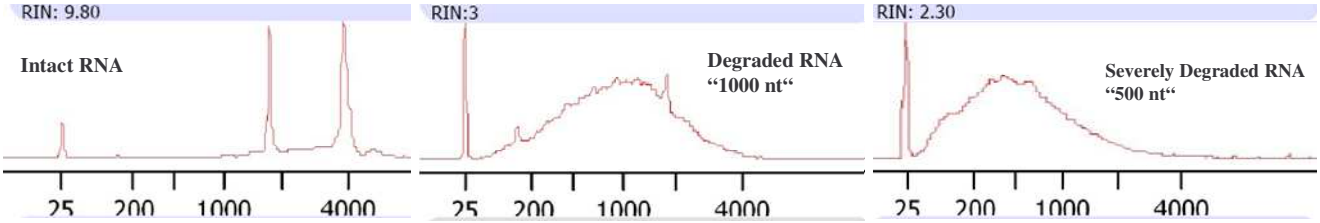


Scheme for TR*inucleotide*-based mRNA Amplification



EXON ARRAYS
Quality of Input RNA Samples and Overview of Hybridisation Results

Model experiment. Intact RNA samples were chemically degraded, to maintain full sequence complexity in degraded RNA samples

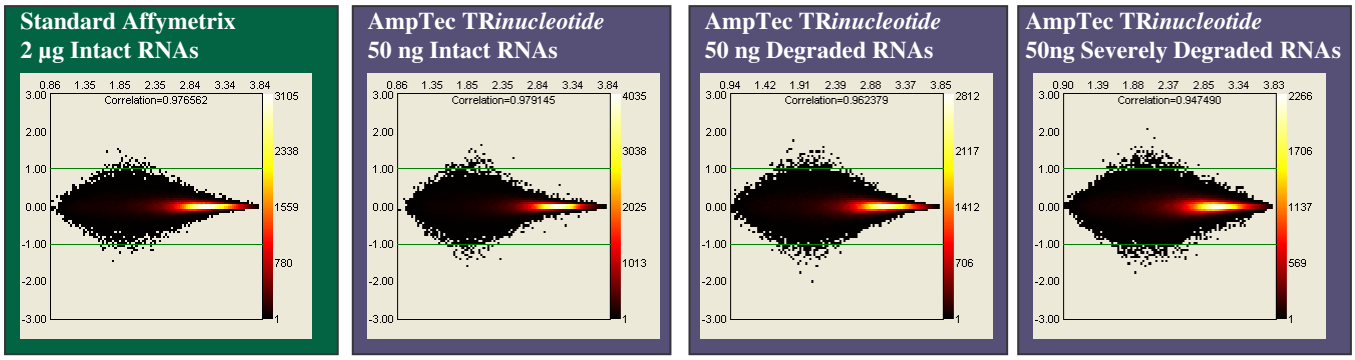


Overview of QC data. Three independently obtained biological samples (RNAs from human ell cultures) were processed and hybridised to Human Exon 1.0 ST Arrays. In all samples, the low rRNA amounts result in high sensitivity and in acceptable or good signal vs background ratios, as well as good ROC values (a measure of the false positive rate).

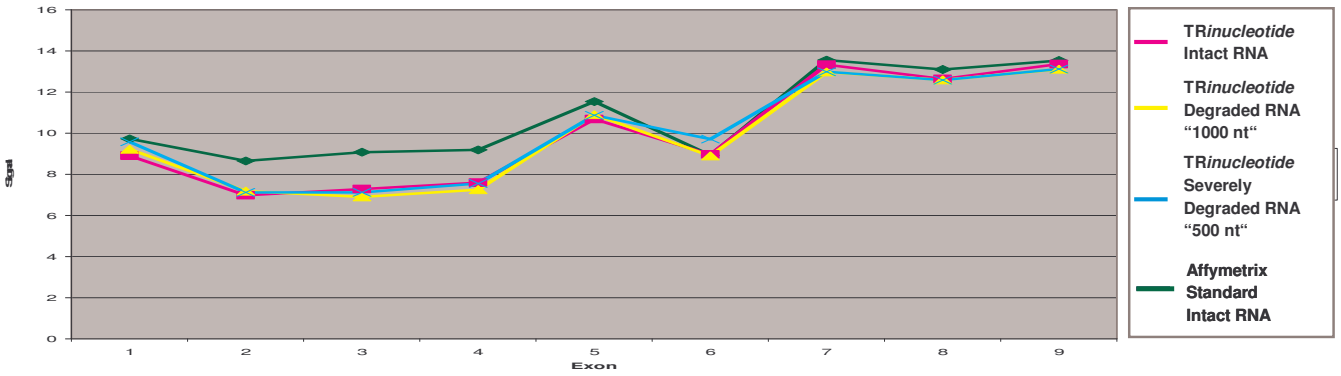
*Only 50% of the amplified RNAs from the first round were used for the second round.

Method for generating cRNAs	Starting material RNA amount / RNA quality	cRNA yields [µg]	Sensitivity [% P]	Mean Signal vs Background (Ratio)	Biological replicates [Pearson values]	ROC pos-neg cont
Standard Affymetrix	**2.0 µg / intact**	**21 ± 5**	**51 ± 1**	**280 vs 310 (0.9)**	**0.98**	**0.881 ± 0.003**
AmpTec TR*inucleotide*	**50 ng / intact**	**2 rounds* 62 ± 10**	**64 ± 2**	**360 vs 210 (1.7)**	**0.98**	**0.901 ± 0.004**
AmpTec TR*inucleotide*	**50 ng / degraded** ("1000nt")	**2 rounds* 58 ± 10**	**53 ± 2**	**280 vs 200 (1.5)**	**0.96**	**0.892 ± 0.004**
AmpTec TR*inucleotide*	**50 ng / severely degraded** ("500nt")	**2 rounds* 52 ± 5**	**47 ± 3**	**265 vs 250 (1.1)**	**0.95**	**0.875 ± 0.003**

Scatter plots (exon level) to compare pairs of biological replicates.



Signal Intensities for Exon Probes over the Complete Length of GAP-DH mRNA. Some variability (presumably structure related) was observed for all samples. Very high consistency for all AmpTec TR*inucleotide* amplified RNAs, irrespective of the divergent RNA qualities in the starting material.



BACTERIAL mRNA
Nano-samples for genome-wide bacterial gene expression studies
Selective amplification of prokaryotic mRNAs – without any mRNA enrichment step

Methods After performing a typical heat shock experiment with 7 min incubations at either 37°C or at 50°C (Richmond et al., 1999), total RNAs from *E.coli* K12 were extracted with Trizol, and genomic DNA was removed by DNase I treatment.
Conventional: 50 µg total bacterial RNA were used to generate dye-labelled cDNAs by reverse transcription, using random primers and Cy-3- or Cy-5-labelled dCTP's.
Nano-samples: 50 ng total bacterial RNA were amplified with the AmpTec ExpressArt Bacterial mRNA Amplification Kit to generate amplified antisense RNAs.
The ExpressArt TR*inucleotide* technology is based on the following unique steps. An especially designed TR*inucleotide* primer (Box1-random-trinucleotide primer; without T7-promoter) is used for reverse transcription. This results in preferential priming near the 3'-ends of all nucleic acids, while preventing reverse transcription of rRNAs. Then, a second "TR*inucleotide* primer" (Box2-random-trinucleotide primer) is used to generate almost full-length double stranded cDNAs. After denaturation, the second cDNA strand is primed with a T7-promoter/Box1 primer. This leads to double stranded cDNA with a functional T7-promotor at one end and the Box2 sequence at the other end. This dsDNA template results in amplified, antisense oriented RNA with defined sequences at both ends, Box-1 and Box-2, respectively.
This is a major advantage for the second amplification round, where size reductions of amplified RNAs are avoided. A second amplification round is straightforward and is initiated by reverse transcription of the amplified RNA with the Box-2 primer, followed by dsDNA synthesis with the T7-promoter/Box1 primer.
Fluorescent labelling: AminoAllyl-modified UTP`s were incorporated during in vitro transcription, followed by coupling with Cy-dyes. After fragmentation, cRNAs were hybridised with the Ocimum *E.coli* K12 Array.
Biotin labelling: Biotin-modified UTP`s and CTP's were directly incorporated during the in vitro transcription step. After fragmentation, cRNAs were hybridised with the Affymetrix *E.coli* Genome 2.0 GeneChips.
Richmond et al., 1999 "Genome-wide expression profiling in *Escherichia coli* K-12" Nucleic Acids Res. 27:3821-3835.

Fig. 1: Overlayed Electropherograms of TR*inucleotide* amplified RNAs. 50 ng of high quality *E.coli* total RNAs were used and after the first amplification round, appr. 1 µg of amplified RNAs were obtained; 0.5 µg aliquots in the second amplification round yielded > 50 µg amplified RNAs.
RNA samples were analyzed by capillary electrophoresis on the Agilent bioanalyzer 2100. RNA profiles are shown in blue (*E.coli* at 37°C) and in green (at 50°C) and the RNA ladder in red, with size markers of 0.2, 0.5, 1, 2, 4 and 6 kb; median size of amplified RNAs was 0.5 kb. Sizes are indicated by the "s" scale, this means seconds of run time. Please note the absence of amplified RNAs in the size range of 16S and 23S rRNAs (> 1.6 kb or appr. 40 sec run time).

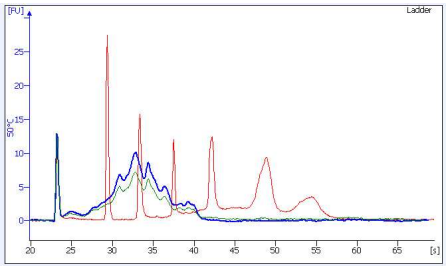


Figure 2: Reproducibility of mRNA amplification reactions. Four slides were hybridised with amplified cRNAs in these dye-swap experiments. The expression ratios of two slides with reverse labelling were averaged and compared to the replication of the experiment. High reproducibility was observed with a correlation coefficient (R) of 0.922.

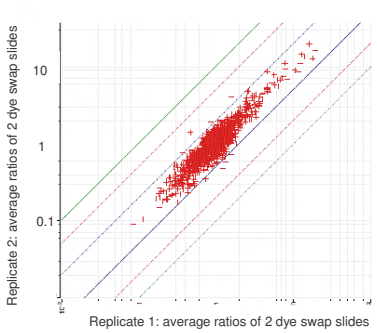


Table 1: Examples for genes induced during heat-shock in *E.coli*. Comparison of published data with results obtained with conventional cDNA labelling and with TR*inucleotide* amplification.
Quality control data with Affymetrix GeneChips. Presence Calls: 38.6% (37°C) or 46.3% (50°C); signal/background ratios: 45 – 50; scale factors: 10 – 11; average signals (P) >4,000.
Suppression of rRNA amplification: less than 2% rRNA signal intensities in amplified RNAs.

Gene	Gene description	published: Richmond et al., 1999	labelled cDNA & Ocimum Array	Amplified RNA & Ocimum Array	Amplified RNA & Affymetrix GeneChip
dnaj	chaperone with dnak, heat shock protein	high	medium	medium	medium
dnak	chaperone hsp70, dna biosynthesis, heat shock protein	medium	medium	medium	low
grpe	phage lambda replication, host dna synthesis, heat shock protein	medium	low	medium	medium
hslj	heat shock protein	medium	low	low	medium
hslu	heat shock protein hslvu, atpase subunit	low	low	low	low
hslv	heat shock protein hslvu, proteasome-related peptidase subunit	medium	low	medium	medium
htpx	integral membrane protein, heat shock protein	medium	low	low	medium
ibpa	heat shock protein	high	high	high	high
ibpb	heat shock protein	high	high	high	high
lon	dna-binding, atp-dependent protease la, heat shock k-protein	medium	low	medium	low

Figure 3: Differential gene expression. With conventional cDNA labelling and with TR*inucleotide* amplification, the same genes were identified as differentially expressed, although a slight reduction of the dynamic range was observed. Red lines connect selected genes which are known to be induced or repressed in heat-shock experiments (Richmond et al., 1999; see also Table 1). Each experiment was performed twice (rep 1 and rep 2).
Please note: Complete concordance of gene assignments as induced or repressed; no cross-over of the red lines.