TaqMan Noncoding RNA Assays help characterize hidden RNA regulatory networks

Long noncoding RNAs represent a large, new class of gene expression regulators

Advances in genomic analysis technologies have uncovered new classes of regulatory noncoding RNAs (ncRNAs) in mammals and many other complex organisms [1-3]. Unlike messenger RNAs (mRNAs), which may participate in regulation of gene expression through protein intermediaries, ncRNAs are postulated to represent a vast regulatory layer that functions more directly. This network includes both microRNAs (miRNAs), a large group of small regulatory RNAs that have been characterized extensively, and long ncRNAs (>200 nt), which have emerged as a focus of biological research.

Long ncRNAs were first discovered by large-scale sequencing of full-length cDNA libraries and expression profiling using high-resolution genome tiling arrays [4,5]. They vary from several hundred to thousands of nucleotides in length, have little or no protein-coding capacity, and are abundantly expressed in a developmentally regulated manner. In mice, large numbers of long ncRNAs are specifically expressed during embryonic stem cell differentiation [6], and they exhibit precise subcellular localization in the brain [7], suggesting specific regulatory roles in development. Early discoveries pertaining to certain groups of long ncRNAs imply that they are involved in a surprisingly wide array of cellular functions, including epigenetic silencing, transcriptional regulation, and RNA processing and modification [8]. In addition, long ncRNAs have been associated with human diseases such as cancer [9], Alzheimer's disease [10], and heart disease [11]. Better understanding of long ncRNA functional roles has tremendous potential to advance our understanding of cell regulatory and disease mechanisms.

Quantification of long ncRNA using TaqMan Assay technology

To meet the specific needs of researchers investigating long ncRNAs, we have developed a comprehensive set of predesigned Applied Biosystems™ TaqMan® Noncoding RNA Assays for reliable and accurate quantification of long ncRNA expression levels. Based on proven TaqMan probe–based technology and developed with a state-of-the-art assay design pipeline, the TaqMan Noncoding RNA Assays provide the confidence and familiar workflow of Applied Biosystems™ TaqMan® Assay technology to accelerate discoveries in the ncRNA field.

ThermoFisher SCIENTIFIC

TaqMan Noncoding RNA Assay design

TaqMan Noncoding RNA Assays are designed using proprietary and functionally verified design algorithms similar to those developed for Applied Biosystems™ TaqMan® Gene Expression Assays (read "The design process for a new generation of quantitative gene expression analysis tools", Pub. No. CO011189, for details). This sophisticated assay design pipeline integrates an extensive array of bioinformatics tools, including comprehensive target sequence analysis, computational and experimentally tested assay design rules, assay quality control (QC) strategies, and robust assay selection criteria (Figure 1). Briefly, ncRNA target sequences are retrieved from NCBI and other ncRNA databases, and mapped to the genome. Single-nucleotide polymorphisms (SNPs), repeats, and areas of sequence discrepancy are masked, and suitable locations for assay design (i.e., exon–exon junctions) are identified. Next, the most robust primer/probe sets are designed based on both thermodynamic and chemical properties, including optimal T_m requirements, GC content, secondary structure, optimal amplicon size, and primer-dimer minimization.

Predesigned TaqMan Assays for >43,000 ncRNAs

Currently, there is a comprehensive set of over 43,000 predesigned TaqMan Noncoding RNA Assays available for ncRNAs from human, mouse, rat, and other species (Table 1). Examples of assay coverage for known disease-associated ncRNA targets are shown in Table 2.

Table 1. Available numbers of TaqMan Noncoding RNA Assays.

Species	No. of assays
Human	23,067
Mouse	10,214
Rat	6,570
Goat	1,235
Cow	478
Horse	437
Marmoset	361
Rhesus monkey	361
Dog	253
Pig	242
Mole rat	81
Sheep	71

Total: 43,370 as of February 2019.

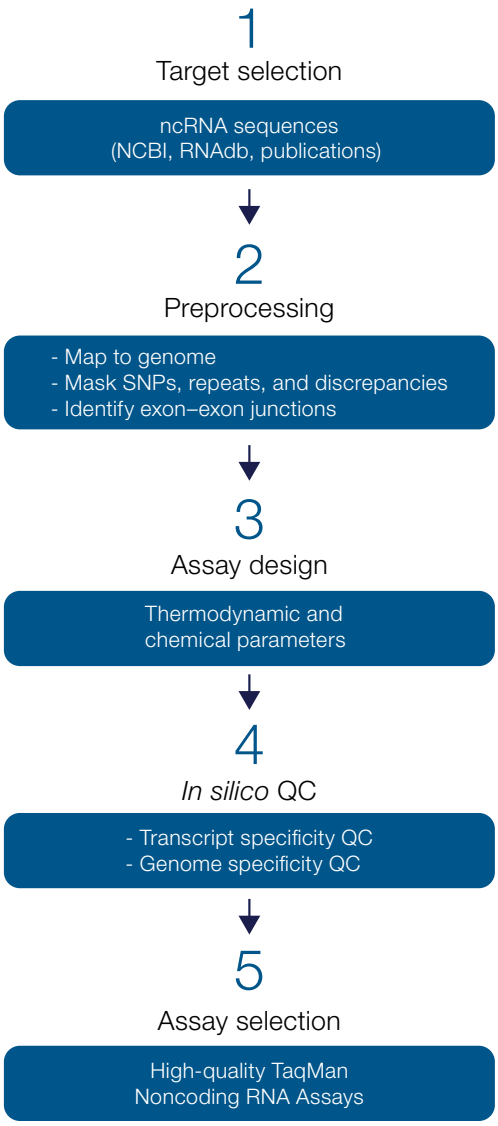

Figure 1. TaqMan Noncoding RNA Assay design pipeline.

Because ncRNAs are expressed pervasively throughout the genome, the specificity of TaqMan Noncoding RNA Assays at both transcript and genomic levels is critically important. The *in silico* QC process scores potential ncRNA assays for transcript specificity using a process similar to that used for TaqMan Gene Expression Assays. For genome specificity, however, the *in silico* QC includes a position-specific, genome-wide alignment matrix to identify assays with minimal potential cross-reactivity with nontarget transcripts or genomic sequences. This extensively verified assay design and QC pipeline ensures that the TaqMan Noncoding RNA Assays are highly specific to their target and discriminate between highly homologous genes and gene families.

In addition, all TaqMan Noncoding RNA Assays are carefully screened to detect only the targeted noncoding transcript, and not coding NCBI-annotated transcripts—even when they colocate within the same genomic locus (Figure 2).

Table 2. Examples of ncRNAs associated with human diseases.

ncRNA	Accession number	TaqMan Noncoding RNA Assay ID	Function	Disease	Reference
PCA3	NR_015342.2	Hs03309852_g1	Epigenetically silencing p15 (*CCKN2B*) tumor suppressor gene	Cancer	Yu W, Gius D, Onyango P et al. (2008) *Nature* 451(7175):202–206.
FMR1-AS1	NR_024499.3	Hs03680976_g1	Anti-apoptotic function	Fragile X syndrome	Khalil AM, Faghihi MA, Modarresi F et al. (2008) *PLoS One* 3(1):e1486.
ATXN8OS	NR_002717.3	Hs01382089_m1	Hypothesized to deregulate pre-mRNA splicing	Spinocerebellar ataxia	Chen WL, Lin JW, Huang HJ et al. (2008) *Brain Res* 1233:176–184.
MALAT1	NR_002819.2	Hs00273907_s1	Predicts metastasis and survival in early-stage non–small cell lung cancer	Cancer	Ji P, Diederichs S, Wang W et al. (2003) *Oncogene* 22(39):8031–8041.
HIF1A-AS2	NR_045406.3	Hs03454328_s1	Biomarker for breast cancer prognosis	Cancer	Cayre A, Rossignol F, Clottes E et al. (2003) *Breast Cancer Res* 5(6):R223–230.
MIAT	NR_003491.1	Hs00402814_m1	SNP in the gene increases risk of the disease	Cardiovascular disease	Ishii N, Ozaki K, Sato H et al. (2006) *J Hum Genet* 51(12):1087–1099.
PCA3	NR_015342	Hs03462121_m1	Overexpressed in prostate cancers	Prostate cancer	de Kok JB, Verhaegh GW, Roelofs RW (2002) *Cancer Res* 62(9):2695–2698.
HOTAIR	NR_015342.3	Hs03296680_s1	Epigenetically represses *HOXD* genes, a known predictor of breast cancer metastasis	Breast cancer	Rinn JL, Kertesz M, Wang JK et al. (2007) *Cell* 129(7):1311–1323.
KCNQ1OT1	NR_002728.3	Hs03665990_s1	Imprinting	Beckwith-Wiedemann syndrome	Arima T, Kamikihara T, Hayashida T et al. (2005) *Nucleic Acids Res* 33(8):2650–2660.
H19	NR_002196.2	Hs00262142_g1	Imprinting	Cancer	Yoshimizu T, Miroglio A, Ripoche MA et al. (2008) *Proc Natl Acad Sci U S A* 105(34):12417–12422.

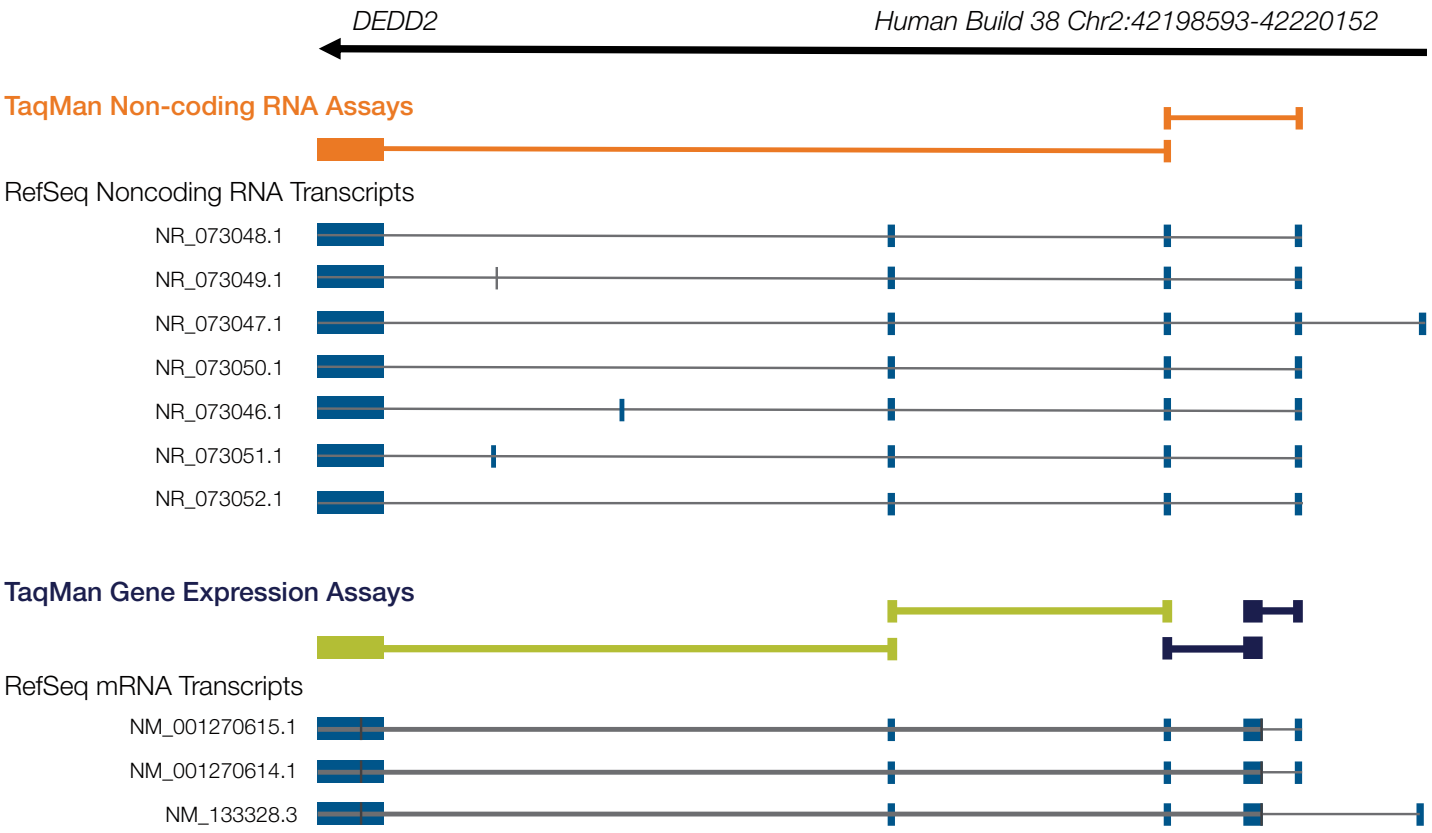

Figure 2. TaqMan Noncoding RNA Assays are specific to ncRNA targets. An example of TaqMan Noncoding RNA Assay specificity is shown on the human *DEDD2* gene locus, which is associated with 3 coding transcripts (prefix NM_) and 7 noncoding transcript variants (prefix NR_). Whereas some TaqMan Gene Expression Assays can detect both coding and noncoding transcripts (colored in green), all TaqMan Noncoding RNA Assays are carefully screened to specifically detect only ncRNA transcripts. Transcript exons are shown as blue boxes, introns as gray bars. For TaqMan Gene Expression Assays, blue and green boxes indicate the exons spanned by the probe; for TaqMan Noncoding RNA Assays, orange boxes indicate the exons spanned by the probe.

Comprehensive mapping and annotation information facilitates assay selection

The TaqMan Assay search tool (**thermofisher.com/taqman**) includes an alignment map and comprehensive annotation with direct links to the NCBI GenBank database for each TaqMan Noncoding RNA Assay. Annotation on adjacent or overlapping coding genes for each TaqMan Noncoding RNA Assay is also provided to illustrate the genomic context of each target ncRNA so that researchers can develop initial hypotheses for functional studies (Table 3).

TaqMan Noncoding RNA Assay workflow and performance

Like current TaqMan Gene Expression Assays, all TaqMan Noncoding RNA Assays are designed for real-time PCR using standardized qPCR conditions with universal cycling and buffer concentrations. In addition, they are also compatible with preamplification, which expands small samples to enable analysis with hundreds of TaqMan Noncoding RNA Assays using only nanogram amounts of input RNA.

We experimentally verified the performance of TaqMan Noncoding RNA Assays extensively. To evaluate sensitivity and dynamic range of input nucleic acid, 10-fold serial dilutions of genomic DNA (gDNA) were amplified using a panel of 96 TaqMan Noncoding RNA Assays targeting sequences within single exons (thus, assays can detect the targets from either gDNA or cDNA). Our results showed that the TaqMan Noncoding RNA Assays are highly sensitive: 72% of the assays tested with gDNA showed detectable signals (C_t value <35) with an estimated 15 copies of target, while 97% of tested assays detected 150 copies of target. The TaqMan Noncoding RNA Assays also showed a wide dynamic range: greater than 5 orders of magnitude, from an estimated 10 to 10^6 copies of input target (Figure 3A), with a PCR amplification efficiency of 100% ±10% (Figure 3B).

Table 3. An example of annotation of genomic context (adjacent and/or overlapping coding genes) for the TaqMan Noncoding RNA Assays.

	Gene name	Gene symbol	Alias	Distance	Relative orientation	NCBI chromosome location
Upstream	Homeobox C12	*HOXC12*	*HOC3F, HOX3, HOX3F*	–10.7 kb	Antisense	Chr.12:52634981–52636617
Downstream	Homeobox C11	*HOXC11*	*HOX3H, MGC4906*	4.4 kb	Antisense	Chr.12:52653177–52656470
Overlapping	HOX transcript antisense RNA (non–protein coding)	*HOTAIR*	*FLJ41747, HOXAS, NCRNA00072*	NA	Sense	Chr.12:52642496–52648727

Case study: characterization of ncRNAs in human *HOX* loci

The *HOX* gene family includes 39 HOX transcription factors clustered on 4 genomic loci (*HOXA* through *HOXD*); it is essential for defining the positional identities of cells during embryonic development. The embryonic patterns of *HOX* expression are faithfully retained by adult human fibroblast cells to maintain regional identities via complex epigenetic regulation. Previous transcriptional surveys conducted using high-resolution tiling arrays of the 4 *HOX* loci across 11 anatomic sites identified 231 HOX ncRNAs [12]. These HOX ncRNAs appeared to be spatially expressed along developmental axes and were hypothesized to play an important role in epigenetic regulation of *HOX* genes [12]. In collaboration with Dr. Howard Chang's group at Stanford, we used the *HOX* gene family as a model system and investigated the role of HOX ncRNAs in gene regulation during development.

A panel of 96 TaqMan Noncoding RNA Assays was developed targeting 43 HOX ncRNAs and 39 HOX transcription factors across the 4 *HOX* loci. Human fibroblast cells from lung and foot, representing two distinct anatomic origins, were profiled with this HOX assay panel.

The resulting transcription profile of the HOX ncRNAs showed positional expression patterns consistent with published results generated using genomic tiling arrays [12].

These experiments highlighted some of the advantages of TaqMan Noncoding RNA Assays over tiling arrays; the real-time PCR approach provided better sensitivity and dynamic range, and the data were both reproducible and less ambiguous (Figure 4A). More interestingly, the position-specific expression patterns of HOX-coding and -noncoding RNAs were also highly correlated with the diametric chromatin domains previously defined by ChIP-chip experiments [12]. For example, in the *HOXA* locus, a switch of expression patterns between foot and lung fibroblast cells occurred between *HOXA7* and *HOXA9* (Figure 4B, upper panel). This expression switch precisely correlated with the previously defined chromatin domains that undergo diametrically opposing modification (H3K27 trimethylation) in an anatomically specific manner (Figure 4B, lower panel). These results provided further evidence that transcription from *HOX* loci in adult fibroblasts is regulated by opposing epigenetic modifications. The coordinated expression pattern of HOX ncRNAs also suggests that they may play an important role in establishing this epigenetic regulation.

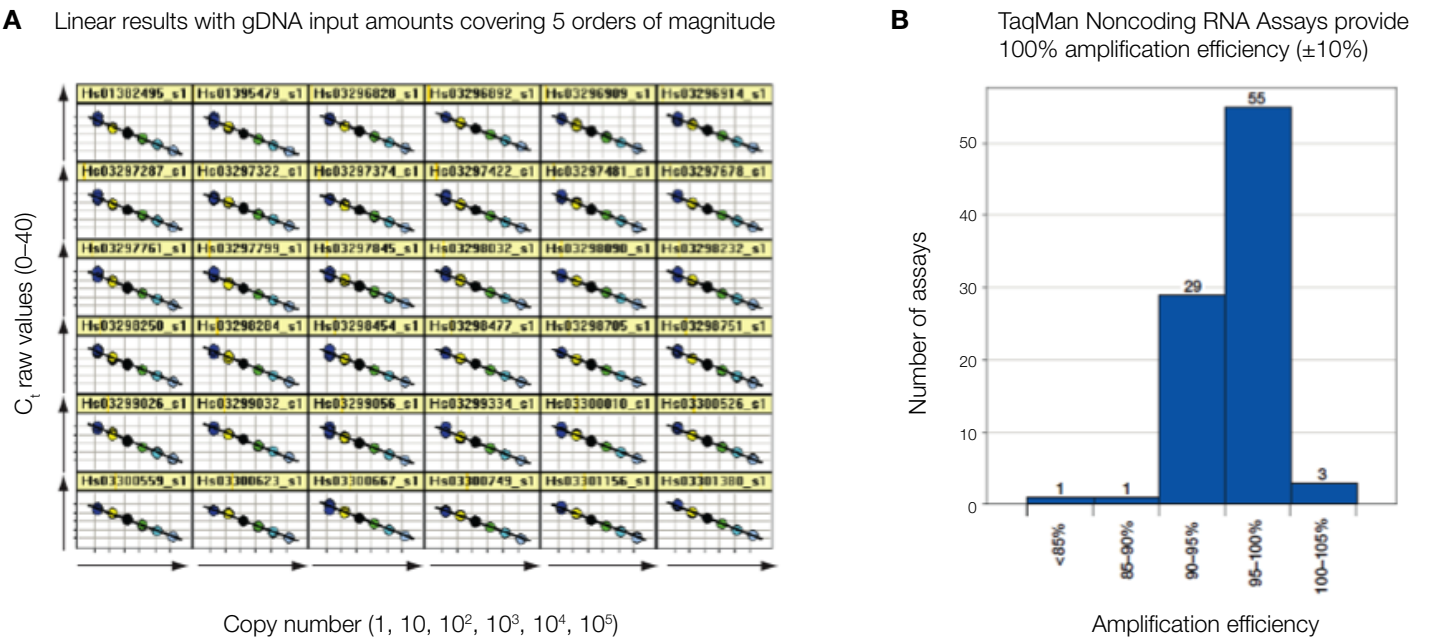

Figure 3. TaqMan Noncoding RNA Assays provide >5 orders of magnitude for dynamic range and 100% amplification efficiency. (A) 10-fold dilution series using gDNA. Copy number of gDNA was estimated using spectrophotometry, and 10-fold serial dilutions were prepared for the estimated copy number shown (covering 5 logarithmic units). Each dilution was combined with the TaqMan Noncoding RNA Assay and TaqMan Gene Expression Master Mix in 4 replicate 10 µL reactions and run on an Applied Biosystems™ 7900HT Fast Real-Time PCR System using universal cycling conditions. **(B)** Distribution of amplification efficiencies measured across at least 5 logarithmic units for 89 TaqMan Noncoding RNA Assays. Amplification efficiency was determined from the slope of the gDNA dilution series using the equation, amplification efficiency (E) = $10^{(-1/\text{slope})} - 1$, multiplied by 100. Data show that the TaqMan Noncoding RNA Assays provide 100% amplification efficiency (±10%).

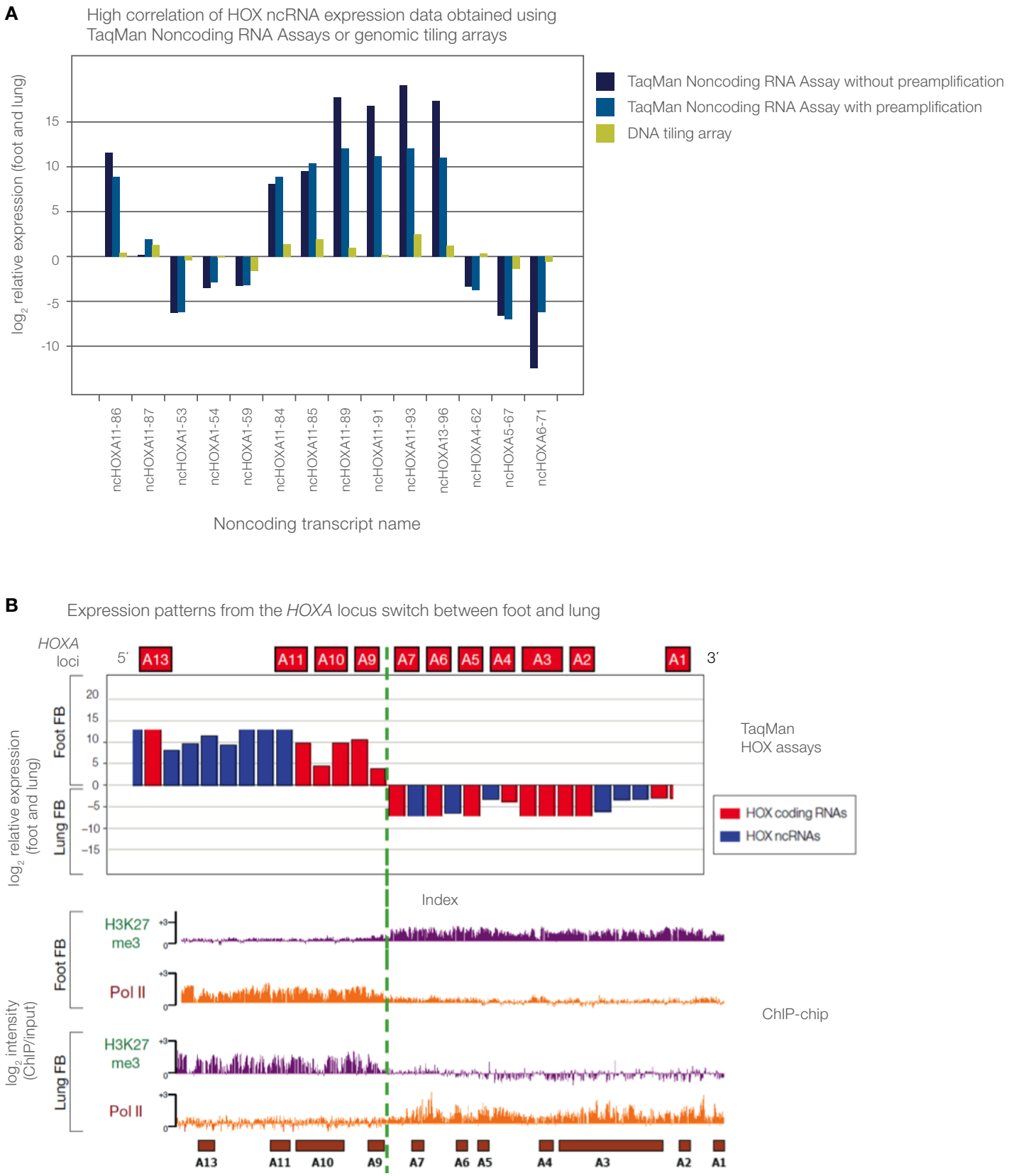

Figure 4. HOX ncRNA showed position-specific expression patterns colinear with coding *HOX* genes. (A) In general, results from real-time PCR using HOX TaqMan Noncoding RNA Assays correlated well with published tiling array results. The TaqMan Assay approach, however, showed significantly higher relative expression changes (therefore less ambiguous) than tiling arrays, presumably due to their wider dynamic range and higher signal-to-noise ratio. **(B)** Upper panel: TaqMan Assay expression profiling showed a position-specific, colinear expression pattern of coding and noncoding HOX transcripts within the *HOXA* locus, with a switch of expression patterns between foot and lung fibroblasts (FB) and *HOXA7* and *HOXA9* loci. Lower panel: This expression switch precisely correlated with the previously defined boundary of diametrically opposed chromatin modifications and transcriptional accessibility in the *HOXA* locus based on a tiling array study [12].

applied
biosystems

TaqMan Noncoding RNA Assays put you at the forefront of ncRNA studies

Noncoding RNAs represent a critical area of biological research and molecular medicine research. Our comprehensive set of predesigned TaqMan Noncoding RNA Assays make quantification of the expression of long ncRNA transcripts reliable and accurate. With sophisticated assay design and reliable TaqMan Assay performance, TaqMan Noncoding RNA Assays provide a powerful tool to help uncover the hidden RNA regulatory network in gene regulation.

Acknowledgments

We would like to thank Yulei Wang, Benjamin Kong, Pius Brzoska, Fangqi Hu, Caifu Chen, Madhu Augustine, Kathy Lee, and Allen Nguyen, Thermo Fisher Scientific Inc., for their contributions to this application note.

The HOX ncRNA study was conducted in collaboration with Dr. Rajnish Gupta and Dr. Howard Chang from Stanford University.

References

1. Birney E, Stamatoyannopoulos JA, Dutta A et al. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447(7146):799–816.
2. Carninci P, Kasukawa T, Katayama S et al. (2005) The transcriptional landscape of the mammalian genome. *Science* 309(5740):1559–1563.
3. Kapranov P, Willingham AT, Gingeras TR (2007) Genome-wide transcription and the implications for genomic organization. *Nat Rev Genet* 8(6):413–423.
4. Kapranov P, Drenkow J, Cheng J et al. (2005) Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays. *Genome Res* 15(7):987–997.
5. Okazaki Y, Furuno M, Kasukawa T et al. (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* 420(6915):563–573.
6. Dinger ME, Amaral PP, Mercer TR et al. (2008) Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res* 18(9):1433–1445.
7. Mercer TR, Dinger ME, Sunkin SM et al. (2008) Specific expression of long noncoding RNAs in the mouse brain. *Proc Natl Acad Sci U S A* 105(2):716–721.
8. Mercer TR, Dinger ME, Mattick JS (2009) Long non-coding RNAs: insights into functions. *Nat Rev Genet* 10(3):155–159.
9. Yu W, Gius D, Onyango P et al. (2008) Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature* 451(7175):202–206.
10. Faghihi MA, Modarresi F, Khalil AM et al. (2008) Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med* 14(7):723–730.
11. Broadbent HM, Peden JF, Lorkowski S et al. (2008) Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. *Hum Mol Genet* 17(6):806–814.
12. Rinn JL, Kertesz M, Wang JK et al. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129(7):1311–1323.
13. Shuman S (2007) Transcriptional networking captures the 7SK RNA 5'-gamma-methyltransferase. *Mol Cell* 27(4):517–519.

Find out more at **thermofisher.com/ncrna**

ThermoFisher SCIENTIFIC