

TranscriptionPath™ ChIP-Seq and ChIP-on-chip

# Comprehensive discovery of transcribed regions across whole genomes

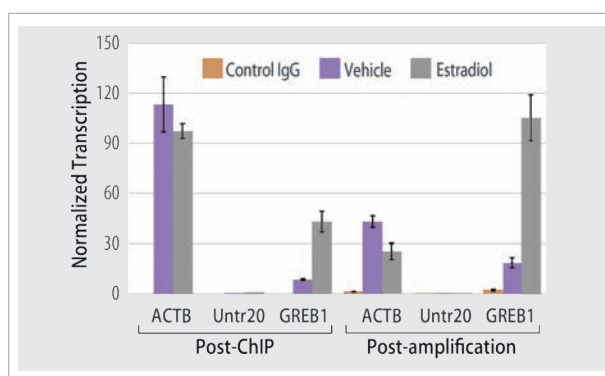
TranscriptionPath™ is a powerful assay that identifies and quantifies genomic DNA sequences undergoing active transcription. This novel assay was developed at Active Motif using the technique of chromatin immunoprecipitation (ChIP) in concert with antibodies that recognize the phosphorylated carboxy terminal domain (CTD) of RNA Polymerase II. As RNA Polymerase II is released from gene promoters and moves into gene bodies, serine 2 of the CTD becomes phosphorylated. Chromatin immunoprecipitation reactions that target this specific phosphorylation event enrich regions of the genome that are undergoing active transcription. The enriched DNA population can be used for ChIP-Seq or ChIP-on-chip. Using either of these two platforms allows transcribed and differentially transcribed genes to be identified on a whole-genome basis by measuring the occupancy of RNA Polymerase II phosphoserine 2 across the length of each gene.

This Application Note presents data generated using TranscriptionPath in combination with whole-genome tiling arrays. In this assay, the ChIP DNA is amplified, labeled and hybridized to one or more Affymetrix GeneChip® Tiling Arrays. These arrays are typically Human or Mouse Tiling 2.0R Arrays, which contain non-repetitive tiling probes of 25 nucleotides separated by 10-base pair gaps across sets of 7 arrays/genome. Active Motif's internal research and development has shown that whole-genome TranscriptionPath data generated from ChIP-on-chip and ChIP-Seq platforms are largely overlapping. Although the data provided in this document was generated using ChIP-on-chip, ChIP-Seq data would be presented in the same manner and be expected to give similar results.

### TranscriptionPath ChIP-on-chip in Induced vs. Uninduced MCF-7 Cells

MCF-7 breast cancer cells were treated with estradiol for 1 hour and the chromatin was immunoprecipitated using an antibody specific for RNA Polymerase II phosphoserine 2. To confirm the quality of the ChIP DNA before hybridization to the arrays, a TranscriptionPath qPCR assay was carried out using primers that amplify a genomic region in the 1<sup>st</sup> introns of the uninduced ACTB (Actin B) gene and the estradiol-induced GREB1 gene. Figure 1 shows similar enrichment of the ACTB gene in uninduced and induced cells and shows induced enrichment of the GREB1 gene in response to estradiol treatment (5-fold). qPCR was performed again following amplification of the ChIP DNAs, which was required to generate adequate DNA for hybridization; the relative enrichment levels were maintained.

The amplified DNA was labeled using the Affymetrix GeneChip® WT Double-Stranded DNA Terminal Labeling Kit and hybridized to the human tiling arrays. Arrays were washed and scanned according to Affymetrix standard procedures. Raw data from the scans was analyzed using Affymetrix® Tiling Analysis Software (TAS) and the results were visualized using the Inte-

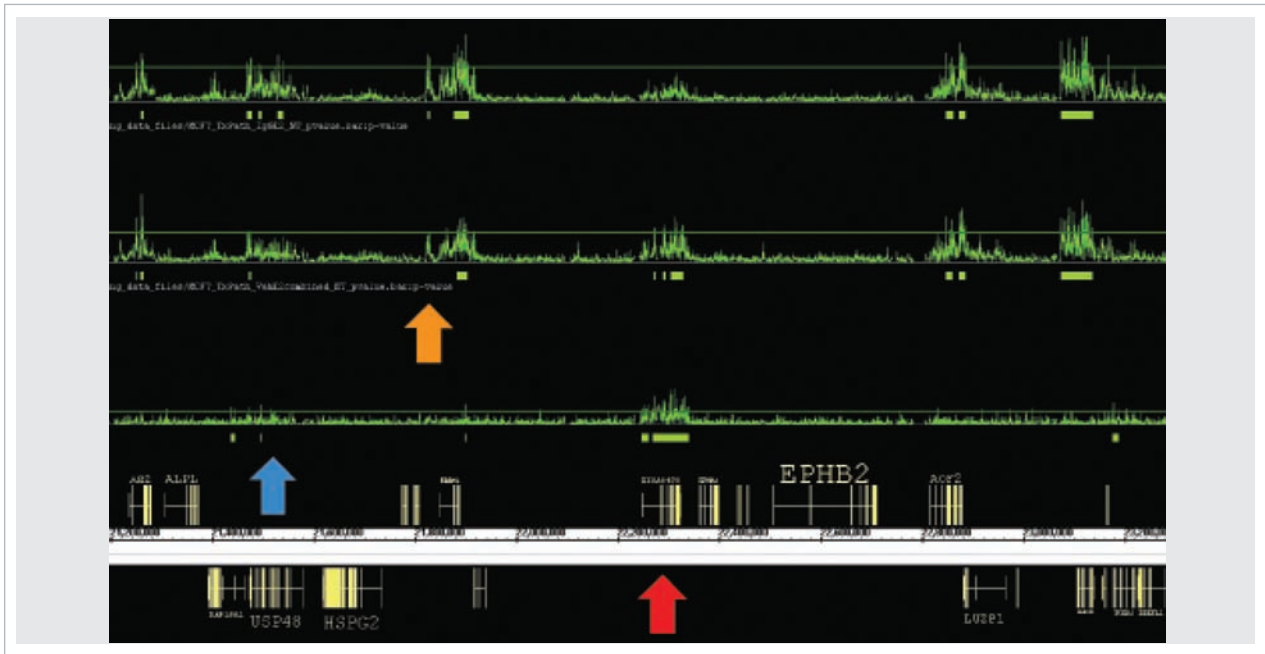


**Figure 1.** TranscriptionPath qPCR showing enrichment of RNA Pol II-associated ChIP DNA in the estradiol-treated MCF-7 cells vs. untreated cells. IgG ChIP was included as a negative control. qPCR was performed both after ChIP and after amplification of ChIP DNA.

grated Genome Browser (IGB). Active Motif's proprietary ChIP Analysis Software was used to compile the data into tables that contain information about binding site locations, binding intensities, and associated gene and genomic annotations.

### Identification of Transcribed and Differentially Transcribed Regions Throughout the Genome

Analysis of the arrays hybridized with RNA Polymerase II ChIP DNA showed that genomic regions corresponding to transcribed genes are evident as areas of elevated signal. In most cases, the transcribed regions correspond to annotated genes, as demonstrated in Figures 2 and 3. Gene annotations are shown at the bottom of each figure. In Figure 2, 12 transcribed genes are apparent. By zooming in on a single transcribed region, as illustrated in Figure 3, individual transcribed regions can easily be assigned to specific gene annotations. In addition, it is not unusual to identify alternate transcriptional start sites by comparing the start of the elevated signal to the annotated transcriptional start site.

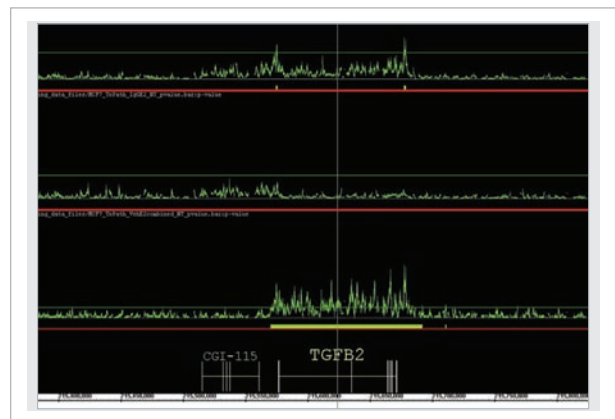


**Figure 2.** Transcribed genes are indicated by elevated array signals as visualized using the Integrated Genome Browser. The KIAA0478 gene (red arrow) is transcribed at a higher level in estradiol-treated cells (middle green line) relative to untreated cells (upper green line). The p-values for results normalized between the two samples (lower green line) are also elevated at this gene, supporting differential transcription. Similarly, the USP48 gene (blue arrow) is transcriptionally downregulated. The orange arrow points to a transcribed region that is unannotated.

Table 1 lists summary information for more than 1,300 genes on chromosomes 1 and 6 found to be undergoing transcription in this assay. Of value is the fact that TranscriptionPath ChIP-Seq and ChIP-on-chip not only detect transcription of annotated genes, but it also detects transcribed sequences in unannotated genomic regions. In fact, approximately 10% of the genes analyzed were located in unannotated regions of the genome and represent novel genes. An example of such a region is shown in Figure 2 (orange arrow). These genes are relatively easy to define because RNA Pol II is bound along the entire region undergoing transcription (exons plus introns).

Another important advantage of TranscriptionPath ChIP-Seq and ChIP-on-chip is the ability to detect genes undergoing differential transcription. Overall, approximately 10% of the genes analyzed exhibited different transcription levels between the induced vs. uninduced MCF-7 cells (Table 1). This is demonstrated in Figures 2 and 3 by comparing the hybridization signals for the untreated and treated cells, which are shown on the top and middle rows, respectively, each of which was normalized to the control IgG array signals. The third row shows the p-values from normalization of the treated vs. untreated signals, and highlights further those genomic areas that are differentially transcribed. Figure 2 shows that the USP48 gene is downregulated (blue arrow) and the KIAA0478 gene is upregulated (red

arrow) in the estradiol-treated cells compared to untreated cells. The other genes depicted in this figure are transcribed non-differentially. In Figure 3, it is clear that the TGFB2 gene is downregulated as a result of estradiol treatment, while the CGI-115 gene immediately upstream is transcribed equally in the two cell samples.



**Figure 3.** Differential transcription of TGFB2 as visualized using the Integrated Genome Browser. Green lines represent TranscriptionPath results for untreated (upper) and estradiol-treated cells (middle), and p-values for results normalized between the two samples (lower). The CGI-115 gene is non-differentially transcribed.

**Verification of Elevated Tiling Array Signals as Transcribed Genomic Regions**

To validate that the regions with elevated signal on the tiling arrays were indeed transcribed and/or differentially transcribed genes, TranscriptionPath qPCR was performed on 17 regions that appeared to be differentially transcribed. Controls included two housekeeping genes (ACTB, and PPIB) and the negative control untranscribed region, Untr12. As shown in Figure 4, 100% of the putative regions tested were verified as undergoing transcription when compared to Untr12. Further, 11/11 of the genes that appeared to be upregulated on the array were also shown to be upregulated when measured by qPCR (genes CAPON to RBM24 in Figure 4), and 6/6 of the genes that appeared to be downregulated on the array were verified as downregulated by qPCR (genes TGFB2 to GPR126 in Figure 4).

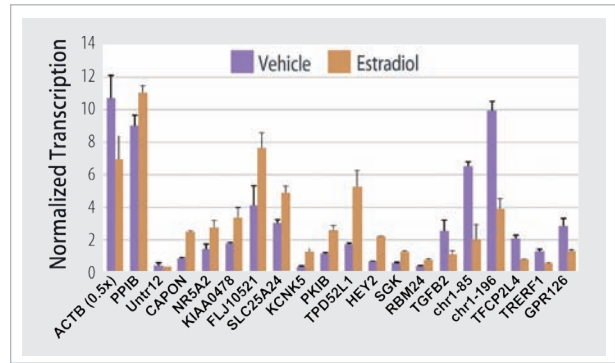
**TranscriptionPath ChIP-Seq and ChIP-on-chip: A Novel Assay to Measure Transcription**

TranscriptionPath ChIP-Seq and ChIP-on-chip provide the first assays for comprehensive identification of sequences undergoing active transcription across the genome. Detection of transcribed regions includes both known genes and previously unknown genes. Because signals are generated across the entire transcribed regions (exons plus introns), novel genes are discovered with greater confidence than when cDNA (from RNA) is hybridized to tiling arrays. Unlike the RNA approach, which can yield variable signals from relatively short exons, there is no need to group together exons that can be spread over significant distances.

At the same time, the ability of this assay to detect differential transcription of both known and novel genes across the genome and at the DNA level, which offers greater reproducibility than RNA, adds even more to the value of this novel assay. TranscriptionPath ChIP-on-chip is highly informative in a wide variety of applications, including disease and drug mechanisms, detection of biomarkers, and target validation, in areas ranging from basic research to drug discovery to preclinical and clinical testing.

**Table 1.** Characteristics of genomic regions undergoing transcription

	Vehicle-treated cells	Estradiol-treated cells	Differential
Number of regions	1360	1301	136
On chromosome 1	890	863	76
On chromosome 6	470	438	60
Annotated genes	1239	1160	98
Unannotated regions	121	141	38



**Figure 4.** Validation by TranscriptionPath qPCR that elevated RNA Pol II signals on the tiling arrays represent transcribed genes. Total (100%) concordance was observed between differential array signals and differential transcription levels by qPCR. Chr1-85 and chr1-96 are located in unannotated regions of the genome.

**Overview of Active Motif Services**

Active Motif services have wide utility in the discovery and development of targeted drugs, novel diagnostic biomarkers, and the identification of causal disease and pharmacological pathways.

We offer a complete suite of epigenetic services for understanding gene regulation and genetic pathways:

- **FactorPath™** for discovery of transcription factor binding sites
- **TranscriptionPath™** for discovery of genomic DNA sequences undergoing active transcription in cells
- **MethylPath™** and **HistonePath™** for discovery and analysis of epigenetic events such as DNA methylation and histone modification
- **Bisulfite Sequencing** for determining the methylation status of cytosines at specific CpGs
- **Antibody Qualification** to validate those antibodies that are suitable for use in ChIP, ChIP-Seq and ChIP-on-chip

All of our services include expert consultation on experimental design, quality controls at multiple steps and comprehensive data analysis. Please contact us with any questions you might have about initiating a project in your cell system.