

# UC San Diego

## UC San Diego Previously Published Works

### Title

Perspectives on Unidirectional versus Divergent Transcription.

### Permalink

<https://escholarship.org/uc/item/9z31433m>

### Journal

Molecular cell, 60(3)

### ISSN

1097-2765

### Authors

Duttke, Sascha HC  
Lacadie, Scott A  
Ibrahim, Mahmoud M  
et al.

### Publication Date

2015-11-01

### DOI

10.1016/j.molcel.2015.10.014

Peer reviewed

# Perspectives on Unidirectional versus Divergent Transcription

Sascha H.C. Duttke,<sup>1,8</sup> Scott A. Lacadie,<sup>4,8</sup> Mahmoud M. Ibrahim,<sup>4,5</sup> Christopher K. Glass,<sup>2,3</sup> David L. Corcoran,<sup>7</sup> Christopher Benner,<sup>6</sup> Sven Heinz,<sup>2,6</sup> James T. Kadonaga,<sup>1,\*</sup> and Uwe Ohler<sup>4,5,\*</sup>

<sup>1</sup>Section of Molecular Biology

<sup>2</sup>Department of Cellular and Molecular Medicine

<sup>3</sup>Department of Medicine

University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

<sup>4</sup>Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, 13125 Berlin, Germany

<sup>5</sup>Department of Biology, Humboldt University, 10115 Berlin, Germany

<sup>6</sup>Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA

<sup>7</sup>Center for Genomic and Computational Biology, Duke University, Durham, NC 27708, USA

<sup>8</sup>Co-first author

\*Correspondence: [jkadonaga@ucsd.edu](mailto:jkadonaga@ucsd.edu) (J.T.K.), [uwe.ohler@mdc-berlin.de](mailto:uwe.ohler@mdc-berlin.de) (U.O.)

<http://dx.doi.org/10.1016/j.molcel.2015.10.014>

In their Letter to the Editor, [Andersson et al. \(2015\)](#) examined divergent and unidirectional human promoter regions reported in our February 2015 *Molecular Cell* paper ([Duttke et al., 2015](#)) and compared our results with data from their concurrent studies ([Andersson et al., 2014](#); [Core et al., 2014](#)). They concluded that the nature of our transcription start site (TSS) data generated by the 5'-GRO-seq method ([Lam et al., 2013](#)) led to an inflation of the percentage of unidirectional promoter regions, and further suggested that unidirectional promoter regions may not exist.

We appreciate the opportunity to address these issues. To place this discussion in the context of our findings ([Duttke et al., 2015](#)), it is first useful to restate the four summary highlights of our paper verbatim:

- (1) Basal RNA polymerase II machinery and core promoters are inherently directional.
- (2) Divergent transcripts arise from their own core promoters at the edge of open chromatin.
- (3) Unidirectional promoters are frequent and depleted of reverse core promoter sequences.
- (4) Reverse-directed core promoters are associated with a unique chromatin signature.

Thus, our findings indicated that transcription from human promoters does not occur spontaneously in both directions, but is instead intrinsically directional because both unidirectional and

divergent/bidirectional promoter regions contain unidirectional core promoters. In the Letter, [Andersson et al. \(2015\)](#) are in agreement with conclusions #1 and #2 of the highlights and incorporate these findings into the model shown in [Figure S1A](#) of their Letter. They disagree with conclusion #3 and do not address conclusion #4.

Hence, the key question relates to conclusion #3: are there human unidirectional promoter regions that contain only one forward-directed core promoter? In this regard, it should be noted that in [Table S1](#) of [Duttke et al. \(2015\)](#), we had analyzed the exosome knockdown (hRRP40) CAGE data available at that time ([Ntini et al., 2013](#)) and found approximately 47% (3,188/6,828) unidirectional promoters. These results indicate that substantial unidirectional transcription can be seen with TSS data generated by exosome knockdown CAGE as well as by 5'-GRO-seq methods.

It therefore appears that the basis for the different conclusions is largely in the analysis of the data. It is particularly important to consider what is defined to be a meaningful level of (reverse-directed) transcriptional activity. In their work, [Andersson et al. \(2015\)](#) calculate the sum of all reads anywhere within a window from -100 to +50 relative to the DNase I hypersensitive site (DHS) edge and derive a library-specific cutoff based on read frequencies in equally sized windows in negative reference regions. For their negative reference regions, the authors chose inactive chromatin by excluding any genic regions as well as any annotated

DHSs or enhancer regions. This strategy yields a positive reverse-direction signal in their new hRRP40 knockdown CAGE data ([Andersson et al., 2014](#)) with three reads across all three pooled replicates totaling ~45 million mapped reads. With these criteria, about 18% of promoters are found to be unidirectional. The lower number of unidirectional promoters reported in their Letter is based on pooling not only replicates but also separate results obtained from different cell types with different protocols.

We felt that it would be useful to address the concerns of [Andersson et al. \(2015\)](#) by analyzing their data ([Andersson et al., 2014](#); with results from the same cell type) and by using their approach, with two differences. First, we calculated TSS activity by the maximal read count in a 10-nt window from -100 to +1 relative to the DHS edge. (This is in contrast to the sum of all reads in the window from -100 to +50 relative to the DHS edge, as in [Andersson et al. \[2015\]](#).) Based on our previous observations ([Duttke et al., 2015](#)), this criterion reflects the properties of authentic TSSs. Second, instead of using inactive chromatin as the background reference (as in [Andersson et al., 2015](#)), we used 101-nt central segments of open chromatin from intergenic DHS regions (excluding promoters and genes) that do not overlap with the DHS edges. The analysis of the new hRRP40 knockdown CAGE data by this approach yielded "alternatively classified" promoter regions. For comparison, we refer to the promoter regions described in [Duttke et al. \(2015\)](#) as the "original" promoter regions.

With this approach, we found that 42% of the alternatively classified promoter regions (including annotated bidirectional promoters) are unidirectional. This percentage is similar to the 47% unidirectional observed with exosome knockdown CAGE data and the 51% unidirectional observed with 5'-GRO-seq data by using our original method of analysis (Duttke et al., 2015). To determine whether the alternatively classified unidirectional promoters are distinct from the alternatively classified divergent promoters, we compared their properties. First, promoter sequence models (Frith et al., 2008) show nearly the same high peak scores for forward-directed transcription in unidirectional and divergent promoter regions, a lower but distinct increase in the score for reverse-directed transcription in divergent promoter regions, and a nearly negligible increase in the score in reverse-directed unidirectional promoter regions (Figure S1A). Second, examination of TFIIB ChIP-exo data from HeLa cells (Venters and Pugh, 2013) reveals a strong bimodal pattern for divergent promoter regions and a clear unimodal pattern for unidirectional promoter regions (Figure S1B). Third, the chromatin signature, as seen in Promoter State 2, is different in unidirectional versus divergent promoter regions (Figure S1C). This latter point was the fourth major conclusion of our previous study (Duttke et al., 2015) and was not described by Andersson et al. (2014, 2015) or Core et al. (2014). These analyses show many significant differences between unidirectional and divergent promoter regions and thus provide some validation of the methods that we used to classify them.

Thus, our further analysis of the data has revealed distinct unidirectional and divergent promoter regions. Depending on the particular method of promoter classification, the percentage of unidirectional

promoters in a single cell type has been observed to vary from 18% to 51%. In this light, it is relevant to note that two recent studies have reported 24.6% unidirectional transcription in mouse macrophages (Scruggs et al., 2015) and 23% unidirectional transcription in human cells (HeLa or HEK293T) (Mayer et al., 2015). As discussed above, the percentages of unidirectional transcription will vary with the methods of detection and analysis. Nevertheless, it can be seen that there is a substantial fraction of unidirectional transcription in all of these estimates.

From a more biological perspective, it is important to consider whether there are distinct functions for unidirectional versus divergent promoters. In other words, rather than debate whether a handful of reverse-direction reads relative to thousands of forward-direction reads is “unidirectional” or “divergent,” it may be more useful to determine if there is an underlying basis for a promoter to be unidirectional or divergent. It can be seen that the features of the promoters that we classify as unidirectional are distinct from those of the promoters that we classify as divergent (Figure S1; Duttke et al., 2015). These differences in DNA sequences, transcription factor occupancy, and chromatin modifications may reflect differences in biological function. Hence, in the future, the classification of promoters may not be as “unidirectional” or “divergent,” but perhaps rather as promoter class A, B, C, and so on, where each promoter class has its own distinct functional characteristics, one of them being its degree of unidirectionality.

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes one figure and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2015.10.014>.

#### AUTHOR CONTRIBUTIONS

S.A.L., M.M.I., and U.O. designed the overall analysis framework. M.M.I. performed all peak calling and generated the HMM and its associated analyses. S.A.L. performed all other analyses. S.A.L., S.H.C.D., M.M.I., C.K.G., D.L.C., C.B., S.H., J.T.K., and U.O. wrote the manuscript.

#### ACKNOWLEDGMENTS

J.T.K. is the Amylin Chair in the Life Sciences. This work was supported by a grant from the National Institutes of Health to J.T.K. (R01 GM041249). M.M.I. is supported by the MDC-NYU graduate program.

#### REFERENCES

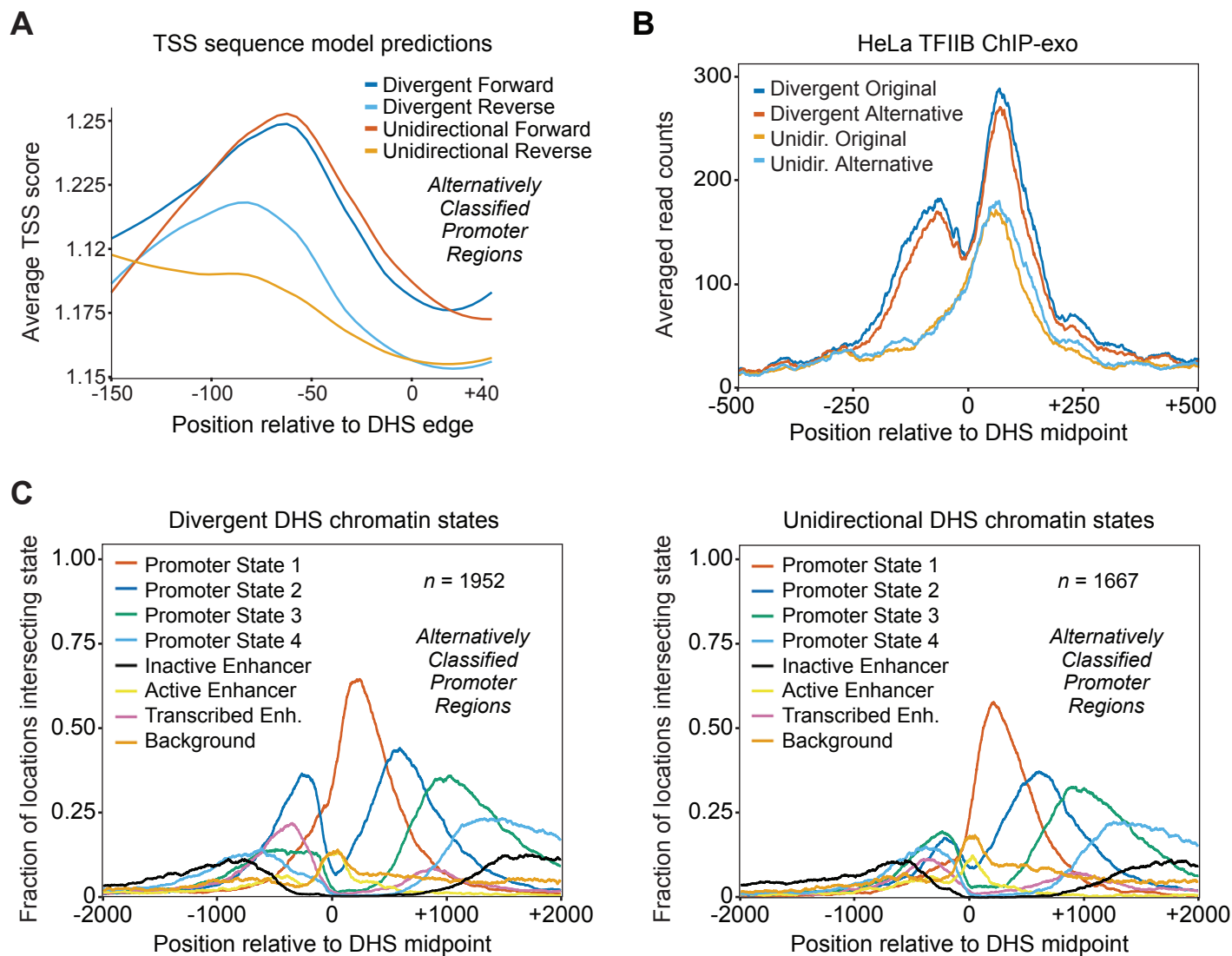
- Andersson, R., Refsing Andersen, P., Valen, E., Core, L.J., Bornholdt, J., Boyd, M., Heick Jensen, T., and Sandelin, A. (2014). *Nat. Commun.* 5, 5336.
- Andersson, R., Chen, Y., Core, L., Lis, J., Sandelin, A., and Heick Jensen, T. (2015). *Mol. Cell* 60, this issue, 346–347.
- Core, L.J., Martins, A.L., Danko, C.G., Waters, C.T., Siepel, A., and Lis, J.T. (2014). *Nat. Genet.* 46, 1311–1320.
- Duttke, S.H., Lacadie, S.A., Ibrahim, M.M., Glass, C.K., Corcoran, D.L., Benner, C., Heinz, S., Kadonaga, J.T., and Ohler, U. (2015). *Mol. Cell* 57, 674–684.
- Frith, M.C., Valen, E., Krogh, A., Hayashizaki, Y., Carninci, P., and Sandelin, A. (2008). *Genome Res.* 18, 1–12.
- Lam, M.T., Cho, H., Lesch, H.P., Gosselin, D., Heinz, S., Tanaka-Oishi, Y., Benner, C., Kaikkonen, M.U., Kim, A.S., Kosaka, M., et al. (2013). *Nature* 498, 511–515.
- Mayer, A., di Iulio, J., Maleri, S., Eser, U., Vierstra, J., Reynolds, A., Sandstrom, R., Stamatoyannopoulos, J.A., and Churchman, L.S. (2015). *Cell* 161, 541–554.
- Ntini, E., Järvelin, A.I., Bornholdt, J., Chen, Y., Boyd, M., Jørgensen, M., Andersson, R., Hoof, I., Schein, A., Andersen, P.R., et al. (2013). *Nat. Struct. Mol. Biol.* 20, 923–928.
- Scruggs, B.S., Gilchrist, D.A., Nechaev, S., Muse, G.W., Burkholder, A., Fargo, D.C., and Adelman, K. (2015). *Mol. Cell* 58, 1101–1112.
- Venters, B.J., and Pugh, B.F. (2013). *Nature* 502, 53–58.

**Molecular Cell**

**Supplemental Information**

# **Perspectives on Unidirectional versus Divergent Transcription**

**Sascha H.C. Duttke, Scott A. Lacadie, Mahmoud M. Ibrahim, Christopher K. Glass,  
David L. Corcoran, Christopher Benner, Sven Heinz, James T. Kadonaga, and Uwe  
Ohler**



**Figure S1. Classification and analysis of unidirectional and divergent promoter regions with exosome knockdown CAGE data from Andersson et al. (2014)**

DHSs were classified based on exosome knockdown CAGE data (Andersson et al., 2014) and a thresholding approach similar to that used by Andersson et al. (2015). Our method requires that the maximum read count in a 10 nt window within 101 bp upstream of and including the DHS edge must be over the 95th percentile of that same value measured on a negative control set, defined as both strands of equally-sized windows in the centers of 7379 intergenic-annotating DHSs (14758 strand-specific regions) large enough to prevent overlaps with the DHS edge region (i.e., the background intergenic DHSs must be  $\geq 303$  bp). This value corresponds to 12 tags. We used the original promoter regions reported in Duttke et al. (2015) and classified those in which both 101 bp edge windows (forward and reverse) meet the threshold as divergent and those in which only the forward window meets the threshold as unidirectional. **A.** Position specific Markov chain sequence model (Frith et al., 2008) assessment of alternatively classified promoter regions. Average sequence model scores in 10 bp windows are shown anchored on respective DHS edges and calculated as described in Duttke et al. (2015). Divergent promoter regions ( $n = 1952$ ). Unidirectional promoter regions ( $n = 1667$ ). **B.** Comparison of HeLa cell TFIIB ChIP-exo (Venters and Pugh, 2013) signal for original 5'-GRO-seq classified and exosome knockdown CAGE alternatively classified promoter regions. ChIP-exo reads were extended by the fragment length of 53 nucleotides as determined by peakzilla (Bardet et al., 2013), and average counts were plotted (single base positions in windows 500 bp upstream and downstream of DHS centers). Divergent original ( $n = 1741$ ). Divergent alternatively classified ( $n = 1952$ ). Unidirectional original ( $n = 2237$ ). Unidirectional alternatively classified ( $n = 1667$ ). **C.** Hidden Markov model-based chromatin state intersections as described in Duttke et al. (2015) for alternatively classified promoter regions. The plots show the fraction of each position intersecting a given state 2000 bp upstream and downstream of DHS centers. Exosome knockdown CAGE reads were trimmed with fastx\_trimmer from the fastx toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/commandline.html](http://hannonlab.cshl.edu/fastx_toolkit/commandline.html)). Trimmed CAGE reads and those from the HeLa TFIIB ChIP-exo were mapped with the default settings of Bowtie2 (Langmead and Salzberg, 2012), disregarding all non-uniquely mapped reads.

## Supplemental References

Andersson, R., Refsing Andersen, P., Valen, E., Core, L.J., Bornholdt, J., Boyd, M., Heick Jensen, T., and Sandelin, A. (2014). *Nat. Commun.* *5*, 5336.

Andersson, R., Chen, Y., Core, L., Lis, J., Sandelin, A., and Heick Jensen, T. (2015). *Mol. Cell* *60*, this issue.

Bardet, A.F., Steinmann, J., Bafna, S., Knoblich, J.A., Zeitlinger, J., and Stark, A. (2013). *Bioinformatics* *29*, 2705–2713.

Duttke, S.H., Lacadie, S.A., Ibrahim, M.M., Glass, C.K., Corcoran, D.L., Benner, C., Heinz, S., Kadonaga, J.T., and Ohler, U. (2015). *Mol. Cell* *57*, 674–684.

Frith, M.C., Valen, E., Krogh, A., Hayashizaki, Y., Carninci, P., and Sandelin, A. (2008). *Genome Res.* *18*, 1–12.

Langmead, B., and Salzberg, S.L. (2012). *Nat. Methods* *9*, 357–359.

Venters, B.J., and Pugh, B.F. (2013). *Nature* *502*, 53–58.