Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Mol Cell. Author manuscript; available in PMC 2017 June 27.

Published in final edited form as:

Mol Cell. 2015 November 05; 60(3): 346–347. doi:10.1016/j.molcel.2015.10.015.

Human gene promoters are intrinsically bidirectional

Robin Andersson[1,#,*], **Yun Chen**[1,#], **Leighton Core**[2,*], **John Lis**[3,*], **Albin Sandelin**[1,*], and **Torben Heick Jensen**[4,*]

[1]The Bioinformatics Centre, Section for Computational and RNA Biology, Department of Biology, University of Copenhagen, Ole Maaloes Vej 5, DK-2200 Copenhagen, Denmark

[2]Department of Molecular & Cell Biology, Institute for Systems Genomics, University of Connecticut, Storrs, CT 06269, USA

[3]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA

[4]Centre for mRNP Biogenesis and Metabolism, Department of Molecular Biology and Genetics, C.F. Møllers Alle 3, Bldg. 1130, DK-8000 Aarhus, Denmark

A recent Molecular Cell paper by Duttke et al. concluded that ‘human promoters are intrinsically directional' (Duttke et al., 2015). This, at least on the surface, contradicts earlier observations (Almada et al., 2013; Core et al., 2008; Ntini et al., 2013; Seila et al., 2008). As part of the disagreement could be based on a different usage of scientific terms, we find it important to contribute our perspective hoping to reach common definitions. In addition, we present a reconsideration of an important question asked by Duttke et al., namely whether divergent transcription is a generic feature of human promoters.

A core promoter has been classically defined by biochemical studies as the minimal region of DNA required to support assembly of a general transcription complex and initiate transcription. This activity is often encompassed by the +/- 50bp surrounding the transcription start site (TSS) (Kadonaga, 2012). Core promoters rarely act alone *in vivo*, requiring nearby transcription factor binding to regulate their activity. Therefore, the term ‘promoter’ often includes more extended upstream regions that contribute to regulation of initiation at the core promoter. The terms ‘core promoter’, ‘TSS’ and ‘promoter’ have at times been used interchangeably in the literature, which in the light of current genomic data is likely to cause confusion.

Although well described *in vitro*, core promoters can be difficult to define *in vivo*, making high throughput mapping of TSSs paramount for promoter identification. Systematic genomic studies have demonstrated that mammalian mRNA TSSs are typically accompanied by upstream TSSs firing in the opposite direction to produce so-called PROMoter uPstream Transcripts (PROMPTs) or Upstream Antisense RNAs (uaRNAs), which are most often unstable and non-coding (Almada et al., 2013; Andersson et al., 2014; Core et al., 2008; Core et al., 2014; Flynn et al., 2011; Ntini et al., 2013; Preker et al., 2008; Seila et al. 2008).

[*]Corresponding authors (email: robin@binf.ku.dk; leighton.core@uconn.edu; johnlis@cornell.edu; albin@binf.ku.dk; thj@mbg.au.dk).
[#]Shared first authors

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Transcription initiation of mRNA and PROMPT/uaRNA is characteristically tightly spaced (110-250bp between TSSs), centered around a high density of transcription factor binding, and encompassed by positioned nucleosomes that define a nucleosome depleted region (NDR). Based on this, and on data discussed below, we therefore suggest that a mammalian gene promoter is generally demarcated by an NDR that contains two oppositely oriented core promoters (Figure S1A). Because of this property, such promoters were called divergent or bidirectional. Although language used in some of the early reports on divergent promoters could be interpreted to support a model where a single core promoter supports both mRNA and PROMPT/uaRNA transcription, this was to our knowledge not the intention. Nor is such a notion supported by high resolution TSS mapping in mammals (Almada et al., 2013; Andersson et al., 2014; Core et al., 2014; Ntini et al., 2013) or ChIP-exo data from yeast (Rhee and Pugh, 2012). The model depicted in Figure S1A is also used by Duttke et al. and by others (e.g. Scruggs et al. Mol Cell 2015).

An important question arising from these observations is whether divergent transcription is an inherent feature of human gene promoters. Duttke et al. set out to answer this by using a 5′-GRO-seq protocol to obtain nascent RNA 5′ends from HeLa-S3 cell nuclei. These 5′ends were used to classify NDR-defined promoters (detected by DNase I hypersensitivity) to UCSC gene TSSs as either unidirectional or divergent, and they found 56% (2237/3978) and 44% (1741/3978) respectively in each class (Duttke et al., 2015). It was further argued that unidirectional transcription was caused by a lack of core promoter signals in the reverse direction of the respective gene TSSs.

Distinguishing whether a promoter is divergently transcribed will critically depend on the method used and its sensitivity. To visualize this, we sorted unidirectionally- and divergently-classified promoters by the widths of their DNase I hypersensitive sites (DHSs) and plotted around DHS midpoints the frequencies of mapped reads of previously published data sets, including CAGE and RNA-seq data from control and exosome (hRRP40)-depleted HeLa-S3 cells as well as GRO-seq (HeLa-S3) and GRO-cap (similar to 5′-GRO-seq) data from K562 and GM12878 cells. It is evident that the clear majority of claimed unidirectional promoters also support transcription initiation from the reverse DNA strand (Figure S1B). The 5′-GRO-seq assay by Duttke et al. does not detect these transcription events. Consistently, a substantially lower fraction of 5′-GRO-seq reads map within 500bp of uniform ENCODE DHS midpoints compared to the CAGE and GRO-cap data (Figure S1C). We suggest that this decreased ability to detect promoter-proximal RNAs is due to differences in the published GRO-seq protocols (see Methods), rather than because of differently used thresholds.

What then is the degree of divergent transcription initiation from human gene promoters? To address this, we assessed which fraction of transcribed gene promoters analyzed by Duttke et al. also display detectable reverse-strand transcription as a function of thresholds on the number of mapped 5′ends per million library reads (TPMs). We estimated library-specific noise level distributions and used these to derive thresholds (dashed vertical lines in Figure S1D, see Methods) to call promoter-associated transcription initiation events distinguishable from noise. CAGE-hRRP40 data found 69% of unidirectionally classified transcribed gene promoters (1433/2089) and 82% of all classified transcribed gene promoters (3400/4157) to

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

be divergently transcribed (Figure S1D). GRO-cap data yielded comparable numbers; 78% (1491/1919) and 88% (3472/3966) for GM12878 and 83% (1594/1929) and 90% (3591/3971) for K562 cells, respectively. The union of calls suggests 93% (3974/4285) of all classified transcribed promoters, as defined by Duttke et al., to be divergently transcribed. These estimates are lower boundaries, since the sensitivity of any of the employed methods will improve with greater sequencing depth. Although the fraction of promoter reads in either orientation varies considerably for different genes, reverse-strand reads constituted on average 46.5% and 25.5% of total GRO-cap reads associated with classified divergent and unidirectional promoters, respectively (see Methods and Fig. S1B, last two panels). Thus, we conclude that the vast majority of human gene promoters are intrinsically bidirectional, and that their amount of reverse transcription is not negligible.

The identification of reverse-strand TSSs at classified unidirectional promoters enabled improved core promoter analysis at well-defined focus points. Consequently, we detected TATA- and Inr-sequences ~30bp upstream of and at CAGE-defined reverse-strand TSSs, respectively (Figure S1E). In support of PIC formation at these sites, TBP and TFIIB recruitment was observed by ChIP-exo in K562 cells, and RNAPII was found downstream at the pause site (Figure S1F, data from Short Read Archive SRP019209). It thus appears that PICs form at NDR edges for most, if not all, transcriptionally active loci, and that PIC positioning is influenced by core promoter elements at either edge.

Although the mechanistic significance of reverse-oriented transcription remains to be delineated, a near-universal existence of gene promoter-associated divergent transcription initiation is important to note. We are presently in a flux in the definitions of transcription regulatory elements in general, driven by insights from new technologies. Here, our understanding of any roles and causes of divergent promoter transcription may be vital for comprehending basic properties of promoters and core promoters and how they interact with the surrounding transcription factors and chromatin architecture. Such knowledge might be crucial to appreciate how these genetic elements have evolved to maintain their activity and receptiveness to regulation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Almada AE, Wu X, Kriz AJ, Burge CB, Sharp PA. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. Nature. 2013; 499:360–363. [PubMed: 23792564]

Andersson R, Refsing Andersen P, Valen E, Core LJ, Bornholdt J, Boyd M, Heick Jensen T, Sandelin A. Nuclear stability and transcriptional directionality separate functionally distinct RNA species. Nat Comms. 2014; 5:5336.

Core LJ, Martins AL, Danko CG, Waters CT, Siepel A, Lis JT. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. Nat Genet. 2014; 46:1311–1320. [PubMed: 25383968]

Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science. 2008; 322:1845–1848. [PubMed: 19056941]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Duttke SHC, Lacadie SA, Ibrahim MM, Glass CK, Corcoran DL, Benner C, Heinz S, Kadonaga JT, Ohler U. Human promoters are intrinsically directional. Mol Cell. 2015; 57:674–684. [PubMed: 25639469]

Flynn RA, Almada AE, Zamudio JR, Sharp PA. Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. Proc Natl Acad Sci USA. 2011; 108:10460–10465. [PubMed: 21670248]

Kadonaga JT. Perspectives on the RNA polymerase II core promoter. Wiley Interdiscip Rev Dev Biol. 2012; 1:40–51. [PubMed: 23801666]

Ntini E, Järvelin AI, Bornholdt J, Chen Y, Boyd M, Jørgensen M, Andersson R, Hoof I, Schein A, Andersen PR, et al. Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. Nat Struct Mol Biol. 2013; 20:923–928. [PubMed: 23851456]

Preker P, Nielsen J, Kammler S, Lykke-Andersen S, Christensen MS, Mapendano CK, Schierup MH, Jensen TH. RNA exosome depletion reveals transcription upstream of active human promoters. Science. 2008; 322:1851–1854. [PubMed: 19056938]

Rhee HS, Pugh BF. Genome-wide structure and organization of eukaryotic pre-initiation complexes. Nature. 2012; 483:295–301. [PubMed: 22258509]

Scruggs BS, Gilchrist DA, Nechaev S, Muse GW, Burkholder A, Fargo DC, Adelman K. Bidirectional Transcription Arises from Two Distinct Hubs of Transcription Factor Binding and Active Chromatin. Molecular Cell. 2015; 58:1101–1112. [PubMed: 26028540]

Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, Young RA, Sharp PA. Divergent transcription from active promoters. Science. 2008; 322:1849–1851. [PubMed: 19056940]