reflections of the underlying cognitive manifold and recognize the fictive aspect of spatial and temporal expressions in language on the other.

1. Clark, H. H., In *Cognitive Development and the Acquisition of Language* (ed. Moore, T. E.), Academic Press, New York, 1973, pp. 28–64.
2. Haspelmath, M., *From Space to Time: Temporal Adverbials in the World's Languages*, Lincom, München, 1997.
3. Lakoff, G. and Johnson, M., *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*, University of Chicago Press, Chicago, 1999.
4. Gentner, D., *Spatial Metaphors in Temporal Reasoning*, The MIT Press, Massachusetts, 2001.
5. Casasanto, D., Fotakopoulou, O. and Boroditsky, L., *Congnit Sci*., 2010, **34**, 387–405.
6. Casasanto, D. and Boroditsky, L., *Cognition*, 2008, **106**, 579–593.
7. Heine, B., *Linguist. Africaine*, 1989, **2**, 77–127.
8. Heine, B., Ulrike, C. and Hunnemeyer, F., *Grammaticalization: A Conceptual Framework*, University of Chicago Press, Chicago, 1991.
9. Levinson, S. C., *Space in Language and Cognition*, Cambridge University Press, Cambridge, 2003.
10. Gibbs Jr, R. W., *Embodiment and Cognitive Science*, Cambridge University Press, Cambridge, 2005.
11. Levinson, S. C. and Wilkins, D. P. (eds), In *Grammars of Space*, Cambridge University Press, Cambridge, 2006.
12. Traugott, E. C., In *Universals of Human Language: Word Structure* (ed. Greenberg, J. H.), Stanford University Press, Stanford, 1978, vol. 3, pp. 369–400.
13. Svorou, S., *The Grammar of Space*, John Benjamin Publishing Company, Amsterdam, 1994.
14. Alverson, H., *Semantics and Experience: Universal Metaphors of Time in English, Mandarin, Hindi and Sesotho*, John Hopkins University Press, Baltimore, 1994.
15. Gentner, D., Imai, M. and Boroditsky, L., *Lang. Cognit. Process*., 2000, **17**(5), 537–565.
16. Evans, V., *The Structure of Time: Language, Meaning and Temporal Cognition*, John Benjamins, Amsterdam, 2003.
17. Boroditsky, L., *Cognition*, 2000, **75**, 1–28.
18. Talmy, L., *Toward a Cognitive Semantics: Concept Structuring Systems*, The MIT Press, Massachusetts, 2000, vol. I.
19. Wilson, M., *Psychon. Bull. Rev*., 2002, **9**, 625–636.
20. Yu, C., Ballard, D. H. and Aslin, R. N., *Cognit. Sci*., 2005, **29**, 961–1005.
21. Mouene, J., Hidaka, S. and Smith, L. B., *Cognit. Sci*., 2008, **32**, 1200–1216.
22. Feist, M. I., *Cognit. Sci*., 2008, **32**, 1177–1199.
23. Edelman, G. M., *Second Nature: Brain Science and Human Knowledge*, Yale University Press, New Haven, 2006.
24. Jackendoff, R., *Semantics and Cognition*, The MIT Press, Cambridge, 1983.
25. Habel, C. and Eschenbach, C., In *Foundations of Computer Science. Potential – Theory – Cognition* (eds Freksa, C., Jantzen, M. and Valk, R.), Springer-Verlag, Berlin, 1997, pp. 369–378.

Samir Karmakar and Rajesh Kasturirangan are in the National Institute of Advanced Studies, Indian Institute of Science Campus, Bangalore 560 012, India.*
**e-mail: samirkrmkr@yahoo.co.in*

Pervasive transcription

A. Abhinaya

Pervasive transcription or interleaved transcription is the transcription of the interspersed genes which are embedded within the normal coding sequence. The quintessential factor is that, it is believed that the entire stretch of the genome is transcribed, whether it is coding for a particular protein or not. The other underlying factor is that not all coding sequences lie juxtaposed; they may also overlap one another. In other words, they may lie interspersed. The unconventional fact is that these overlapping sequences may be a non-coding region or an interspersed coding region which is still transcribed. The question is, how often are these interspersed regions getting transcribed and does all of the transcribed sequence get translated into protein coding messages?

Transcription is a term associated with the genic region. However, recent findings help extend this concept to include intergenic region, many non-functional elements, pseudogenes, etc.[1–8]. The shear size of the mRNA transcripts produced does not correspond to the mRNA translated; neither do the data relating to the number of coding regions correspond to the multitude of transcripts obtained[9,10]. The genomic framework and transcription process have to be reviewed for better understanding of transcription of interleaved and intragenic regions. Non-coding, stably unannotated transcripts (SUTs) are also produced, which may have a vital function as a regulatory molecule[11]. Some of these might translate to peptides and hence are called transcripts of unknown function (TUFs)[12–14]. Projects such as 'ENCODE' are aimed at exploring what is transcribed to what is translated to what is expressed. The results of the ENCODE project revealed new sectors of RNA and new layers of the transcription machinery. However, the precision of the positioning of the CUTs and SUTs is still to be studied, because of the diverse mechanism of generation of transcripts and also that the number of transcripts generated does not correlate with the findings of Xu *et al.*[1]. Some interesting cases that involve promoter-associated pervasiveness are discussed below, as they are the major source of SUTs[1].

Multiple transcription start site (TSS) points, which control transcription of single annotated genes, are found by mapping the 5′-end of capped RNAs to their corresponding TSSs[8,14–16]. TSS shifts result in the production of snRNAs and snoRNAs, as seen in the shift between upstream and downstream TSS of *IMD2* gene in yeasts[17]. Additionally, pervasive

transcription is also associated with the coding region[18]. Putatitve promoters found within retrotransposons (154 in mice and 579 in humans) are a source of novel isoforms of protein-coding genes[15,19]. Internal promoters found within intron-7 in mouse *c-erbAα* gene or long interspersed nuclear-like retro-transposon called I factor produced a new transcript[20]. These are activated by the class-III group of RNA polymerase II (ref. 21). They help duplicate downstream genes which may not lie within the control of upstream promoters in operons or other genes, thereby transcribing CUTs which are exon-originated during the process[1,22,23].

In case of bidirectional promoters, the interleaved sequence is associated with the nucleosome-free region and is often found associated with a promoter of another/adjacent protein-coding sequence, thereby promoting the activation of such regions[24]. PROMPTs are produced in this way[1,18]. A cryptic promoter also promotes transcription of a comparatively silenced region, promotion of open chromatin region by modulating the de-acetylases which modify the histone proteins[25].

Apart from these varied promoters there are other factors which might influence transcription. In promoter occulsion of *SER3* gene, the RNA polymerase II fails to terminate at the termination site of an upstream gene due to the presence of SRG1, and causes it to interfere with the promoter of *SER3*, but transcribes SRG1 (refs 1, 18). Stalled polymerases are also responsible for producing tiRNAs generated from TSS of several protein-coding genes[1,18,26]. Genome-wide analysis reveals that nucleosome occupancy within the exon is higher compared to pseudoexons[27]. Therefore, modulating the nucleosomes will aid in exposing the exons for transcription. Additionally, pseudoexons, overlapping transcripts and intergenic regions also have a fair chance of getting transcribed. The curvature and bendability of the DNA may also aid in transcription initiation in this region[28].

About 68% of these unannotated transcripts are associated with the 5′-nucleosome-free region aiding the transcription of sense PARs, but most of them are in the antisense direction[1,8]. Usually 70% of the human genes have at least one antisense transcript that overlaps the coding transcript. This leads to the formation of RNA–RNA hybrids, which induce RNAi-mediated silencing as in XIST[29]. RNA polymerase V imparts silencing of overlapping/adjacent genes by transcribing the intragenic region and the transcripts act as silencers[30]. Modification of the histones is also involved in repressing the overlapping genes[29]. The intragenic transcripts at the 3′-end are eliminated by the histone modification factors and histone chaperones (Spt6, Spt16 and HET set-2)[1]. But the most common form of eviction of these cryptic transcripts is via the exosome complex with the aid of the TRAMP complex[31,32]. Here too the complexity lies in which binding proteins (KRSP or TTP) bind to the ARE elements, based on which the transcript is either degraded or accumulated[33].

These transcription procedures still account for only a meagre level. Future experiments may prove that the entire genome is being transcribed and that there might be more such layered mechanisms that meets the eye. With the onset of the minimal genome concept, it is more certain that these unannotated transcripts are involved in the regulatory affairs of the cell just like ncRNAs or SUTs, rather than being translated. It is important to study this aspect and explore the role of the non-coding transcripts as the expression of some ncRNAs is related to cancer. Possibilities of harbouring ncRNAs/SUTs that aid in avoiding illness may be a reality.

1. Berretta, J. and Morillion, A., *EMBO Rep.*, 2009, **10**(9), 973–982.
2. Gerstein, M. B. *et al.*, *Genome Res.*, 2007, **17**, 669–681.
3. Mortazavi, A., Williams, B. A., CcCue, K., Schaeffer, L. and Wold, B., *Nature Methods*, 2008, **5**(7), 621–628.
4. Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M. and Snyder, M., *Science*, 2008, **320**(5881), 1344–1349.
5. Wang, Z., Gersteing, M. and Snyder, M., *Nature Rev. Genet.*, 2009, **10**(1), 57–63.
6. Dinger, M. E., Armaral, P. P., Mercer, T. R. and Mattick, J. S., *Brief. Funct. Genomics*, 2009, **8**(6), 407–423.
7. Jacquier, A., *Nature Rev. Genet.*, 2009, **10**, 833–844.
8. ENCODE Project Consortium, *Nature*, 2007, **447**, 800–816.
9. Kapranov, P., *Genome Med.*, 2009, **1**(5), 50.1–50.4.
10. Le Roch, K. G. *et al.*, *Genome Res.*, 2004, **14**, 2308–2318.
11. Kapranov, P. *et al.*, *Science*, 2007, **316**, 1484–1488.
12. Willingham, A. T. and Gingeras, T. R., *Cell*, 2006, **125**, 1215–1220.
13. Kapranov, P., Willingham, A. T. and Gingeras, T. R., *Nature Rev.*, 2007, **8**, 413–423.
14. He, H. *et al.*, *Nucleic Acids Res.*, 2006, **34**(10), 2976–2983.
15. Kapranov, P., *Genome Biol.*, 2009, **10**(4), 217.1–217.4.
16. Callahan, K. P. and Butler, J. S., *Genome Biol.*, 2008, **9**(4), 218.1–218.4.
17. Harley Jenks, M., O'Rourke, T. W. and Reines, D., *Mol. Cell. Biol.*, 2008, **28**(12), 3883–3893.
18. Taft, R. J. *et al.*, *Nature Genet.*, 2009, **41**, 572–578.
19. Faulkner, G. J. *et al.*, *Nature Genet.*, 2009, **41**, 563–571.
20. Chassande, O. *et al.*, *Mol. Endocrinol.*, 1997, **11**(9), 1278–1290.
21. McLean, C., Bucheton, A. and Finnegan, D. J., *Mol. Cell. Biol.*, 1993, **13**(2), 1042–1050.
22. Huang, P. *et al.*, *Genome Res.*, 2007, **17**, 1478–1485.
23. Mankin, A. S., Skripkin, E. A. and Kangramanova, V. K., *FEBS Lett.*, 1987, **219**(2), 269–273.
24. Xul, Z. *et al.*, *Nature Lett.*, 2009, **457**, 1033–1038.
25. Kuehner, J. N. and Brow, D. A., *Mol. Cell*, 2008, **31**, 201–211.
26. Tamkun, J. W., *Nature Genet.*, 2007, **39**, 1421–1422.
27. Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcárcel, J. and Guigó, R., *Nature Struct. Mol. Biol.*, 2009, **16**(9), 996–1002.
28. Kanhere, A. and Bansal, M., *Nucleic Acids Res.*, 2005, **35**(10), 3165–3175.
29. Bernstein, E. and David Allis, C., *Genes Dev.*, 2009, **19**, 1635–1655.
30. Wierzbicki, A. T., Haag, J. R. and Pikaard, C. S., *Cell*, 2008, **135**, 635–648.
31. LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A. and Tollervey, D., *Cell*, 2005, **121**, 713–724.
32. Saira Mian, I., *Nucleic Acids Res.*, 1997, **25**(16), 3187–3195.
33. Schilders, G., van Dijk, E., Raijmakers, R. and Pruijin, G. J. M., *Int. Rev. Cytol.*, 2006, **251**, 159–208.

A. Abhinaya is in the Department of Biotechnology, St Peter's Engineering College, Anna University, Chennai 600 050, India. e-mail: a_abhinaya@ymail.com