Anais da Academia Brasileira de Ciências (2010) 82(4): 933-939
(Annals of the Brazilian Academy of Sciences)
ISSN 0001-3765
www.scielo.br/aabc

The central role of RNA in the genetic programming of complex organisms

JOHN S. MATTICK

Institute for Molecular Bioscience, University of Queensland, Brisbane 4072, Australia

Manuscript received on September 17, 2010; accepted for publication on October 15, 2010

ABSTRACT

Notwithstanding lineage-specific variations, the number and type of protein-coding genes remain relatively static across the animal kingdom. By contrast there has been a massive expansion in the extent of genomic non-protein-coding sequences with increasing developmental complexity. These non-coding sequences are, in fact, transcribed in a regulated manner to produce large numbers of large and small non-protein-coding RNAs that control gene expression at many levels including chromatin architecture, post-transcriptional processing and translation. Moreover, many RNAs are edited, especially in the nervous system, which may be the basis of epigenome-environment interactions and the function of the brain.

Key words: development, noncoding RNA, epigenome, gene regulation, RNA editing, brain.

INTRODUCTION

It appears that the genetic programming of complex organisms has been misunderstood for the past 50 years, because of the assumption – largely true for the unicellular prokaryotes, but apparently not for multicellular eukaryotes – that most genetic information is transacted by proteins. This assumption is based upon the central dogma which holds that 'DNA makes RNA makes protein', implying that RNA functions primarily as an intermediate between a gene and its encoded protein, which in turn are responsible for the core functions of the cell, including regulatory functions. Reciprocally it has been assumed that the vast tracts of non-protein-coding sequences that are present in animal and plant genomes are largely non-functional. However, this assumption may be incorrect (Mattick 1994), and the emerging evidence suggests that these non-coding sequences actually specify a vast and hitherto hidden layer of regulatory information that is transacted by RNAs, in conjunction with generic protein complexes that interact with them (Mattick 2001, 2003, 2004, 2007, Mattick and Gagen 2001).

E-mail: j.mattick@imb.uq.edu.au

INFORMATION SCALING IN COMPLEX ORGANISMS

The human genome specifies an anatomically complex and cognitively advanced organism comprised of $\sim 10^{14}$ cells, with exquisitely precise architecture of its different muscles, bones, many organs and the brain, which itself contains $\sim 10^{10}$ neurons each with an estimated 10^{14} synaptic connections in the neocortex alone (Andersen et al. 2003). Surprisingly the human genome contains only ~20,000 conventional protein-coding genes (Goodstadt and Ponting 2006, Clamp et al. 2007), which are similar in number and largely share orthologous functions with those in nematodes that have only ~1,000 somatic cells. Indeed, not withstanding clade-specific variations and innovations (such as RNA editing proteins, see below), the core proteome and extent of protein-coding sequences has not changed greatly since the origin of the metazoa, despite enormous increases in their developmental and cognitive complexity (Taft et al. 2007).

On the other hand, the extent of non-protein-coding DNA in the genome increases with increasing complexity, reaching 98.8% in humans (Taft et al. 2007), suggesting that much of the information required to program our development resides in these sequences, and is presum-

ably regulatory in nature. Indeed theoretical considerations suggest that regulatory information scales quadratically with organizational complexity in all functionally integrated systems (Mattick and Gagen 2005), which is supported by empirical data showing that the number of regulatory genes increases as a square function of the number of total genes in bacteria (Croft et al. 2003), contrary to the expectations of combinatoric control. Thus regulatory architecture increasingly dominates genomic information content as organismal complexity increases.

GLOBAL TRANSCRIPTION OF THE GENOME

Moreover, irrespective of the extent of non-protein-coding sequences, it is now evident that the majority of all genomes is transcribed, mainly into non-protein-coding RNAs (ncRNAs), of which there are tens if not hundreds of thousands in mammals (Carninci et al. 2005, Katayama et al. 2005), arranged in complex interlacing and overlapping patterns (for reviews see Frith et al. 2005, Mattick and Makunin 2006, Kapranov et al. 2007). These ncRNAs generally fall into two size classes: (i) small RNAs that are less than 200 nt, including infrastructural RNAs like tRNAs, rRNAs and small nuclear / spliceosomal RNAs (snRNAs), as well as various types of regulatory RNAs, including microRNAs (miRNAs), small interfering RNAs (siRNAs), piwi-interacting RNAs (piRNAs) and small nucleolar RNAs (snoRNAs) (Mattick and Makunin 2005); and (ii) long noncoding RNAs (lncRNAs) that can range from a few hundred bases up to well over 100 kilobases in length (Furuno et al. 2006, Pang et al. 2007, Mercer et al. 2009).

REGULATED EXPRESSION OF NONCODING RNA

These lncRNAs show tissue-specific and physiologically-responsive expression (Ravasi et al. 2006), as well as dynamic expression profiles in differentiating embryonal stem cells (Dinger et al. 2008b), neuronal cells (Mercer et al. 2010), T-cells (Pang et al. 2009), muscle cells (Sunwoo et al. 2009), and other developmental contexts in animals and plants (Amaral and Mattick 2008, Ben Amor et al. 2009, Dinger et al. 2009). They also show many other signatures of functionality, with increasing numbers of validated examples, as well as altered expression in cancer and other diseases (for reviews see (Mattick 2009b, Mercer et al. 2009, Taft et al. 2010a).

Approximately half of all lncRNAs show highly specific expression patterns in different regions of the brain, and many are trafficked to specific sub-cellular locations (Mercer et al. 2008). Moreover particular ncRNAs are associated with known and novel sub-nuclear domains (Sone et al. 2007, Sunwoo et al. 2009), suggesting a key role for lncRNAs in cell biology that has yet to be explored. While ncRNAs exhibit a wide range of conservation (Pang et al. 2006), this is to be expected given that their sequences are subject to different structure-function constraints (i.e., may be more plastic) than proteins, and that regulatory innovation underpins much if not most of phenotypic variation (Pheasant and Mattick 2007). There is also an underexplored subterranean strata of differentially expressed repeat-derived RNAs (Lunyak et al. 2007, Faulkner et al. 2009), which may also play an important role in developmental regulation (Faulkner and Carninci 2009, Mattick et al. 2010).

RNA REGULATION OF EPIGENETIC PROCESSES

A major function of ncRNAs appears to be the regulation of the epigenetic processes that underpin differentiation and development (Amaral and Mattick 2008), by guiding relatively generic chromatin-modifying complexes to their sites of action (Mattick et al. 2009). Many chromatin-modifying proteins contain RNA binding domains, as indeed do major classes of transcription factors (Shi and Berg 1995, Mattick 2003, 2007, Bernstein and Allis 2005). An increasing number of lncRNAs have been shown to be associated with chromatin-modifying complexes and different forms of modified histones (Rinn et al. 2007, Dinger et al. 2008b, Nagano et al. 2008, Pandey et al. 2008, Terranova et al. 2008, Zhao et al. 2008, Khalil et al. 2009, Swiezewski et al. 2009). Indeed, ncRNA-directed regulatory circuits underpin most, if not all, complex epigenetic phenomena in eukaryotes, including RNA interference-related processes such as transcriptional and post-transcriptional gene silencing, position effect variegation, hybrid dysgenesis, chromosome dosage compensation, parental im-

printing and allelic exclusion, paramutation (see below), and possibly transvection and transinduction (see Mattick and Gagen 2001, Mattick 2009b). In addition exons are preferentially associated with nucleosomes in somatic and sperm cells in vertebrates (Nahkuri et al. 2009), indicating that epigenetic regulation acts not just the level of the gene, but at the level of individual exons, which potentially explains the basis of the long-standing mystery of how alternative splicing is regulated, a prediction that has recently gained experimental support (Luco et al. 2010).

MULTIPLE CLASSES OF SMALL RNA

Small RNAs of the miRNA, piRNA and siRNA families play important roles in a wide range of developmental and physiological processes in animals and plants (Bartel 2004, Jones-Rhoades et al. 2006, Stefani and Slack 2008, Ghildiyal and Zamore 2009), and many are dysregulated in diseases such as cancer (Esquela-Kerscher and Slack 2006, Medina and Slack 2008). Recently, we have discovered a number of new classes of small RNAs, including tiny RNAs associated with transcription initiation sites (tiRNAs) (Taft et al. 2009c) that appear to be related to nucleosome positioning (Taft et al. 2009a), similarly-sized RNAs associated with splice junctions (spliRNAs) (Taft et al. 2010b), and a range of small RNAs derived from snoRNAs (sdRNAs) (Taft et al. 2009b), some of which appear to function as miRNAs (Ender et al. 2008), indicating an interplay between the snoRNA- and miRNA-mediated regulatory systems (Politz et al. 2009, Taft et al. 2009b).

RNA COMMUNICATION AND PLASTICITY

Finally, it appears that RNA is trafficked between cells (Dinger et al. 2008a). It also appears to be the substrate for the transmission of environmental information into endogenous epigenetic networks, via RNA editing (Mattick 2010). RNA editing occurs via two classes of enzymes, the ADARs (one of which, ADAR3, is brain-specific) that catalyze adenosine deamination to inosine (Bass 2002, Valente and Nishikura 2005) and the APOBECs (two of which, APOBEC1 and APOBEC3 are specific to mammals, with the latter having expanded under positive selection in the primate lineage) that act variously on RNA or DNA to catalyze cytosine or 5-methylcytosine deamination to uracil or thymine (Morgan et al. 2004, Sawyer et al. 2004, Zhang and Webb 2004, Mikl et al. 2005, Navaratnam and Sarwar 2006). RNA editing occurs in most if not all tissues, appears to play an important role in development (Bhutani et al. 2010, Sato et al. 2010), and is particularly active in the brain (Bass 2002, Valente and Nishikura 2005). Intriguingly, there is ~30 times more RNA editing observed in human than in mouse, the vast majority of which occurs in Alu primate-specific elements (Athanasiadis et al. 2004, Blow et al. 2004, Kim et al. 2004, Levanon et al. 2004). The amount of editing has also increased during primate evolution associated with new human-specific Alu insertions in genes of neuronal function (Paz-Yaacov et al. 2010). Alu sequences also appear to have been subject to positive selection (Lander et al. 2001), possibly associated with the evolution of advanced brain function, which also involves processes that are similar to those in the immune system (Mattick and Mehler 2008, Mattick 2010). Finally it appears that RNA is the mediator of transgenerational epigenetic inheritance, referred to as 'paramutation' (Chandler 2007), a process that is also influenced by editing (Nadeau 2009).

CONCLUSION

The emerging evidence suggests that, rather than oases of protein-coding sequences in a desert of junk, the genomes of humans and other complex organisms should be viewed as islands of protein-coding sequences in a sea of regulation (Mattick 2004, Ovcharenko et al. 2005), most of which is transacted by RNA (Amaral et al. 2008, Mattick 2010). Moreover it appears that RNA, rather than simply being an ephemeral intermediate between gene and protein, actually comprises the computational engine of the cell (Mattick 1994, Mattick and Gagen 2001) and the substrate for epigenome-environment interactions (Mattick 2010). It is a remarkably versatile molecule (Leontis and Westhof 2003, Lescoute and Westhof 2006, Cruz and Westhof 2009), with capacity to form sophisticated structures, possess catalytic functions and engage in sequence-specific interactions, which may be allosterically controlled and interact with various sorts of effector proteins, thereby coupling ana-

log and digital functions (Mattick 2007, St Laurent and Wahlestedt 2007). What was dismissed as junk because it was not understood may hold the key to understanding human evolution, development and cognition, as well as our individual differences and susceptibilities to complex diseases (Mattick 2009a).

ACKNOWLEDGMENTS

This article is an extended summary of plenary lectures presented at the IRT 2010 (XIX International Round Table on Nucleosides, Nucleotides and Nucleic Acids), Lyon, France, and at the 56th Brazilian Congress of Genetics, Guarujá, Brazil. The author gratefully acknowledges the support of Australian Research Council Federation Fellowship FF0561986.

RESUMO

Apesar das variações linhagem-específicas, o número e tipo de genes codificadores de proteínas permanecem relativamente estáticos no reino animal. Em contraste, houve uma expansão maciça da quantidade de sequências genômicas não-codificadoras de proteínas com o aumento da complexidade do desenvolvimento. Essas sequências não codificadoras são, de fato, transcritas de maneira regulada para produzirem numerosos RNAs grandes e pequenos não-codificadores de proteínas que controlam a expressão de genes em vários níveis, incluindo a arquitetura da cromatina, o processamento pós-transcricional e a tradução. Além disso, muitos RNAs são editados, especialmente no sistema nervoso, o que pode ser a base de interações epigenoma-ambiente e a função do cérebro.

Palavras-chave: desenvolvimento, RNA não-codificador, epigenoma, regulação gênica, edição de RNA, cérebro.

REFERENCES

AMARAL PP AND MATTICK JS. 2008. Noncoding RNA in development. Mamm Genome 19: 454–492.

AMARAL PP, DINGER ME, MERCER TR AND MATTICK JS. 2008. The eukaryotic genome as an RNA machine. Science 319: 1787–1789.

ANDERSEN BB, GUNDERSEN HJ AND PAKKENBERG B. 2003. Aging of the human cerebellum: a stereological study. J Comp Neurol 466: 356–365.

ATHANASIADIS A, RICH A AND MAAS S. 2004. Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. PLoS Biol 2: e391.

BARTEL DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116: 281–297.

BASS BL. 2002. RNA editing by adenosine deaminases that act on RNA. Annu Rev Biochem 71: 817–846.

BEN AMOR B ET AL. 2009. Novel long non-protein coding RNAs involved in Arabidopsis differentiation and stress responses. Genome Res 19: 57–69.

BERNSTEIN E AND ALLIS CD. 2005. RNA meets chromatin. Genes Dev 19: 1635–1655.

BHUTANI N, BRADY JJ, DAMIAN M, SACCO A, CORBEL SY AND BLAU HM. 2010. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature 463: 1042–1047.

BLOW M, FUTREAL PA, WOOSTER R AND STRATTON MR. 2004. A survey of RNA editing in human brain. Genome Res 14: 2379–2387.

CARNINCI P ET AL. 2005. The transcriptional landscape of the mammalian genome. Science 309: 1559–1563.

CHANDLER VL. 2007. Paramutation: from maize to mice. Cell 128: 641–645.

CLAMP M, FRY B, KAMAL M, XIE X, CUFF J, LIN MF, KELLIS M, LINDBLAD-TOH K AND LANDER ES. 2007. Distinguishing protein-coding and noncoding genes in the human genome. Proc Natl Acad Sci USA 104: 19428–19433.

CROFT LJ, LERCHER MJ, GAGEN MJ AND MATTICK JS. 2003. Is prokaryotic complexity limited by accelerated growth in regulatory overhead? Genome Biology Preprint Depository http://genomebiology.com/qc/2003/5/1/p2.

CRUZ JA AND WESTHOF E. 2009. The dynamic landscapes of RNA architecture. Cell 136: 604–609.

DINGER ME, MERCER TR AND MATTICK JS. 2008a. RNAs as extracellular signaling molecules. J mol endocrinol 40: 151–159.

DINGER ME ET AL. 2008b. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res 18: 1433–1445.

DINGER ME, PANG KC, MERCER TR, CROWE ML, GRIMMOND SM AND MATTICK JS. 2009. NRED: a database of long noncoding RNA expression. Nucleic Acids Res 37: D122–126.

ENDER C, KREK A, FRIEDLANDER MR, BEITZINGER M, WEINMANN L, CHEN W, PFEFFER S, RAJEWSKY N AND MEISTER G. 2008. A human snoRNA with microRNA-like functions. Mol Cell 32: 519–528.

ESQUELA-KERSCHER A AND SLACK FJ. 2006. Oncomirs – microRNAs with a role in cancer. Nat Rev Cancer 6: 259–269.

FAULKNER GJ AND CARNINCI P. 2009. Altruistic functions for selfish DNA. Cell Cycle 8: 2895–2900.

FAULKNER GJ ET AL. 2009. The regulated retrotransposon transcriptome of mammalian cells. Nat Genet 41: 563–571.

FRITH MC, PHEASANT M AND MATTICK JS. 2005. The amazing complexity of the human transcriptome. Eur J Hum Genet 13: 894–897.

FURUNO M ET AL. 2006. Clusters of internally primed transcripts reveal novel long noncoding RNAs. PLoS Genet 2: e37.

GHILDIYAL M AND ZAMORE PD. 2009. Small silencing RNAs: an expanding universe. Nat Rev Genet 10: 94–108.

GOODSTADT L AND PONTING CP. 2006. Phylogenetic reconstruction of orthology, paralogy, and conserved synteny for dog and human. PLoS Comput Biol 2: e133.

JONES-RHOADES MW, BARTEL DP AND BARTEL B. 2006. MicroRNAs and their regulatory roles in plants. Annu Rev Plant Biol 57: 19–53.

KAPRANOV P, WILLINGHAM AT AND GINGERAS TR. 2007. Genome-wide transcription and the implications for genomic organization. Nat Rev Genet 8: 413–423.

KATAYAMA S ET AL. 2005. Antisense transcription in the mammalian transcriptome. Science 309: 1564–1566.

KHALIL AM ET AL. 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci USA 106: 11667–11672.

KIM DD, KIM TT, WALSH T, KOBAYASHI Y, MATISE TC, BUYSKE S AND GABRIEL A. 2004. Widespread RNA editing of embedded alu elements in the human transcriptome. Genome Res 14: 1719–1725.

LANDER ES ET AL. 2001. Initial sequencing and analysis of the human genome. Nature 409: 860–921.

LEONTIS NB AND WESTHOF E. 2003. Analysis of RNA motifs. Curr Opin Struct Biol 13: 300–308.

LESCOUTE A AND WESTHOF E. 2006. The interaction networks of structured RNAs. Nucleic Acids Res 34: 6587–6604.

LEVANON EY ET AL. 2004. Systematic identification of abundant A-to-I editing sites in the human transcriptome. Nat Biotechnol 22: 1001–1005.

LUCO RF, PAN Q, TOMINAGA K, BLENCOWE BJ, PEREIRA-SMITH OM AND MISTELI T. 2010. Regulation of alternative splicing by histone modifications. Science 327: 996–1000.

LUNYAK VV ET AL. 2007. Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. Science 317: 248–251.

MATTICK JS. 1994. Introns: evolution and function. Curr Opin Genet Dev 4: 823–831.

MATTICK JS. 2001. Non-coding RNAs: the architects of eukaryotic complexity. EMBO Reports 2: 986–991.

MATTICK JS. 2003. Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. Bioessays 25: 930–939.

MATTICK JS. 2004. RNA regulation: a new genetics? Nat Rev Genet 5: 316–323.

MATTICK JS. 2007. A new paradigm for developmental biology. J Exp Biol 210: 1526–1547.

MATTICK JS. 2009a. Deconstructing the dogma: a new view of the evolution and genetic programming of complex organisms. Ann NY Acad Sci 1178: 29–46.

MATTICK JS. 2009b. The genetic signatures of noncoding RNAs. PLoS Genet 5: e1000459.

MATTICK JS. 2010. RNA as the substrate for epigenome-environment interactions: rNA guidance of epigenetic processes and the expansion of RNA editing in animals underpins development, phenotypic plasticity, learning, and cognition. Bioessays 32: 548–552.

MATTICK JS AND GAGEN MJ. 2001. The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. Mol Biol Evol 18: 1611–1630.

MATTICK JS AND GAGEN MJ. 2005. Accelerating networks. Science 307: 856–858.

MATTICK JS AND MAKUNIN IV. 2005. Small regulatory RNAs in mammals. Hum Mol Genet 14: R121–R132.

MATTICK JS AND MAKUNIN IV. 2006. Non-coding RNA. Hum Mol Genet 15: R17–29.

MATTICK JS AND MEHLER MF. 2008. RNA editing, DNA recoding and the evolution of human cognition. Trends in neurosciences 31: 227–233.

MATTICK JS, AMARAL PP, DINGER ME, MERCER TR AND MEHLER MF. 2009. RNA regulation of epigenetic processes. Bioessays 31: 51–59.

MATTICK JS, TAFT RJ AND FAULKNER GJ. 2010. A global view of genomic information–moving beyond the gene and the master regulator. Trends Genet 26: 21–28.

MEDINA PP AND SLACK FJ. 2008. microRNAs and cancer: an overview. Cell Cycle 7: 2485–2492.

MERCER TR, DINGER ME, SUNKIN SM, MEHLER MF AND MATTICK JS. 2008. Specific expression of long

noncoding RNAs in the mouse brain. Proc Natl Acad Sci USA 105: 716–721.

MERCER TR, DINGER ME AND MATTICK JS. 2009. Long non-coding RNAs: insights into functions. Nat Rev Genet 10: 155–159.

MERCER TR, QURESHI IA, GOKHAN S, DINGER ME, LI G, MATTICK JS AND MEHLER MF. 2010. Long noncoding RNAs in neuronal-glial fate specification and oligodendrocyte lineage maturation. BMC Neurosci 11: 14.

MIKL MC, WATT IN, LU M, REIK W, DAVIES SL, NEUBERGER MS AND RADA C. 2005. Mice deficient in APOBEC2 and APOBEC3. Mol Cell Biol 25: 7270–7277.

MORGAN HD, DEAN W, COKER HA, REIK W AND PETERSEN-MAHRT SK. 2004. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. J Biol Chem 279: 52353–52360.

NADEAU JH. 2009. Transgenerational genetic effects on phenotypic variation and disease risk. Hum Mol Genet 18: R202–210.

NAGANO T, MITCHELL JA, SANZ LA, PAULER FM, FERGUSON-SMITH AC, FEIL R AND FRASER P. 2008. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science 322: 1717–1720.

NAHKURI S, TAFT RJ AND MATTICK JS. 2009. Nucleosomes are preferentially positioned at exons in somatic and sperm cells. Cell Cycle 8: 3420–3424.

NAVARATNAM N AND SARWAR R. 2006. An overview of cytidine deaminases. Int J Hematol 83: 195–200.

OVCHARENKO I, LOOTS GG, NOBREGA MA, HARDISON RC, MILLER W AND STUBBS L. 2005. Evolution and functional classification of vertebrate gene deserts. Genome Res 15: 137–145.

PANDEY RR, MONDAL T, MOHAMMAD F, ENROTH S, REDRUP L, KOMOROWSKI J, NAGANO T, MANCINI-DINARDO D AND KANDURI C. 2008. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol Cell 32: 232–246.

PANG KC, FRITH MC AND MATTICK JS. 2006. Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. Trends Genet 22: 1–5.

PANG KC, STEPHEN S, DINGER ME, ENGSTROM PG, LENHARD B AND MATTICK JS. 2007. RNAdb 2.0–an expanded database of mammalian non-coding RNAs. Nucleic Acids Res 35: D178–182.

PANG KC, DINGER ME, MERCER TR, MALQUORI L, GRIMMOND SM, CHEN W AND MATTICK JS. 2009. Genome-wide identification of long noncoding RNAs in CD8+ T cells. J Immunol 182: 7738–7748.

PAZ-YAACOV N, LEVANON EY, NEVO E, KINAR Y, HARMELIN A, JACOB-HIRSCH J, AMARIGLIO N, EISENBERG E AND RECHAVI G. 2010. Adenosine-to-inosine RNA editing shapes transcriptome diversity in primates. P Nat Acad Sci USA 107: 12174–12179.

PHEASANT M AND MATTICK JS. 2007. Raising the estimate of functional human sequences. Genome Res 17: 1245–1253.

POLITZ JC, HOGAN EM AND PEDERSON T. 2009. MicroRNAs with a nucleolar location. RNA 15: 1705–1715.

RAVASI T ET AL. 2006. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. Genome Res 16: 11–19.

RINN JL ET AL. 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129: 1311–1323.

SATO Y, PROBST HC, TATSUMI R, IKEUCHI Y, NEUBERGER MS AND RADA C. 2010. Deficiency in APOBEC2 leads to a shift in muscle fiber type, diminished body mass, and myopathy. J Biol Chem 285: 7111–7118.

SAWYER SL, EMERMAN M AND MALIK HS. 2004. Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. PLoS Biol 2: E275.

SHI Y AND BERG JM. 1995. Specific DNA-RNA hybrid binding by zinc finger proteins. Science 268: 282–284.

SONE M, HAYASHI T, TARUI H, AGATA K, TAKEICHI M AND NAKAGAWA S. 2007. The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. Journal of cell science 120: 2498–2506.

ST LAURENT G 3RD AND WAHLESTEDT C. 2007. Noncoding RNAs: couplers of analog and digital information in nervous system function? Trends neurosci 30: 612–621.

STEFANI G AND SLACK FJ. 2008. Small non-coding RNAs in animal development. Nat Rev Mol Cell Biol 9: 219–230.

SUNWOO H, DINGER ME, WILUSZ JE, AMARAL PP, MATTICK JS AND SPECTOR DL. 2009. MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. Genome Res 19: 347–359.

SWIEZEWSKI S, LIU F, MAGUSIN A AND DEAN C. 2009. Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. Nature 462: 799–802.

TAFT RJ, PHEASANT M AND MATTICK JS. 2007. The relationship between non-protein-coding DNA and eukaryotic complexity. Bioessays 29: 288–299.

TAFT RJ, KAPLAN CD, SIMONS C AND MATTICK JS. 2009a. Evolution, biogenesis and function of promoter-associated RNAs. Cell Cycle 8: 2332–2338.

TAFT RJ, GLAZOV EA, LASSMANN T, HAYASHIZAKI Y, CARNINCI P AND MATTICK JS. 2009b. Small RNAs derived from snoRNAs. RNA 15: 1233–1240.

TAFT RJ ET AL. 2009c. Tiny RNAs associated with transcription start sites in animals. Nat Genet 41: 572–578.

TAFT RJ, PANG KC, MERCER TR, DINGER M AND MATTICK JS. 2010a. Non-coding RNAs: regulators of disease. J Pathol 220: 126–139.

TAFT RJ ET AL. 2010b. Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans. Nat Struct Mol Biol 17: 1030–1034.

TERRANOVA R, YOKOBAYASHI S, STADLER MB, OTTE AP, VAN LOHUIZEN M, ORKIN SH AND PETERS AH. 2008. Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. Dev Cell 15: 1–12.

VALENTE L AND NISHIKURA K. 2005. ADAR gene family and A-to-I RNA editing: diverse roles in posttranscriptional gene regulation. Prog Nucleic Acid Res Mol Biol 79: 299–338.

ZHANG J AND WEBB DM. 2004. Rapid evolution of primate antiviral enzyme APOBEC3G. Hum Mol Genet 13: 1785–1791.

ZHAO J, SUN BK, ERWIN JA, SONG JJ AND LEE JT. 2008. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 322: 750–756.