REGULATION OF GENE 
EXPRESSION BY NON-
PROTEIN-CODING RNAs

Mohammad Ali Faghihi M.D.
TO:

My wife, Farzaneh, and my daughters, Shamim and Zahra
ABSTRACT

The human genome sequencing projects revealed that the human genome contains over 3 billion DNA base pairs, but only 20,000–25,000 protein-coding genes. In fact, only about 1.2% of the genome codes for proteins. Surprisingly, the number of human genes seems to be almost equal to lower mammals like rodents, and less than a factor of two greater than that of many much simpler organisms, such as the roundworm and the fruit fly. On the other hand, recent studies have revealed that eukaryotic genomes are almost entirely transcribed, generating an enormous number of non-protein-coding RNAs (ncRNAs). Thus there may be a vast reservoir of biologically meaningful ncRNAs that greatly exceed the ~1.2% of the genome that corresponds to conventional protein coding genes. Several classes of functional ncRNAs have been identified in recent years. One prominent and complex class of ncRNAs is natural antisense transcripts (NATs). NATs are RNA molecules transcribed from the opposite strand of conventional genes often overlapping in part with mature sense mRNA. Indeed a large fraction of NATs is expressed in specific regions of the brain, supporting involvement of these ncRNAs in sophisticated regulatory brain functions as well as in complex neurological disorders. Recent research on NATs, including several large-scale expression-profiling studies, has conclusively established the existence of NATs in eukaryotic genomes. In fact, the consensus opinion is that natural antisense transcripts, most of which represent ncRNAs, occur abundantly in the mammalian genome. However, there are many unanswered questions that still exist concerning NATs biological functions and their heterogeneous mode of actions in various cells. For instance, what fraction of NATs may have functional significance, and how many different regulatory mechanisms may exist for these RNA molecules? NATs appear to be utilizing various cellular pathways, but it is still not clear which intrinsic properties of natural antisense RNA molecules or extrinsic features, such as protein interactions, cellular and developmental context are decisive for any given pathway. How is the expression of these ncRNAs regulated in various cells, and what are the extrinsic factors that affect the regulatory output of antisense RNA transcripts? Based on what we know about the broad expression of NATs in different tissues and cell types, and their varied proposed functions, NATs appear to be a heterogeneous group of regulatory RNAs with a wide variety of biological roles.

During the course of my studies, I initially tried to uncover some general aspects of NAT-mediated regulation of gene expression. Thereafter I have investigated, in further detail, the functional significance of a number of these regulatory RNA elements. I have also reviewed all the reported cases of NATs and summarized them in the introduction section of my thesis. In conclusion, I found that there are widespread occurrences of NATs in mammalian genomes and that many of these regulatory elements are indeed functionally relevant in controlling conventional (sense) gene expression. Considering tissue- and cell type-specific expression patterns of NATs and their heterogeneous proposed functions, it seems that we have, so far, only touched parts of an elephant in the dark. The big picture, in the light of future studies, probably will include these parts, but it could be dissimilar to our current understanding. My work, like any other scientific project, has generated many more questions than answers. Several other Ph.D. assignments are needed to address these questions and to generate more questions for future projects and this is the nature of growing sciences.

This thesis is based on the following articles, which will be referred to by their roman numerals in the text.

I. A noncoding RNA is elevated in Alzheimer’s disease and drives rapid feed-forward regulation of β-secretase expression
   Mohammad Ali Faghihi, Farzaneh Modarresi, Ahmad Khalil, Douglas E. Wood, Barbara G. Sahagan, Todd E. Morgan, Caleb E. Finch, Georges St Laurent III, Paul J. Kenny, and Claes Wahlestedt.
   Nature Medicine 2008 14 (7), 723-730

II. A Novel RNA Transcript with Antiapoptotic Function Is Silenced in Fragile X Syndrome
    Mohammad Ali Faghihi, Ahmad M. Khalil, Farzaneh Modarresi, Shaun P. Brothers and Claes Wahlestedt (MAF and AK are first co-authors with equal contribution to this publication)

III. The human PINK1 locus is regulated in vivo by a non-coding natural antisense RNA during modulation of mitochondrial function.
    BMC Genomics. 2007 Mar 15;8:74

IV. RNA interference is not involved in natural antisense mediated regulation of gene expression in mammals
    Mohammad Ali Faghihi and Claes Wahlestedt
    Genome Biology Genome Biol. 2006;7(5):R38

V. Antisense transcription in the mammalian transcriptome
    Science. 2005 Sep 2;309(5740):1564-6

VI. A High-Throughput RNAi Screen Reveals Widespread Biological Function For Mammalian Noncoding Antisense Transcripts
    Mohammad Ali Faghihi, Jannet Kocerha, Farzaneh Modarresi, Par G. Engstrom, Alistair Chalk, Georges St. Laurent II and Claes Wahlestedt
    Manuscript
OTHER PUBLICATIONS

I have also contributed to the following publications, which are not included in this thesis:

1. A small molecule enhances RNA interference and promotes microRNAs processing

2. Identification of functional SNPs in the 5-prime flanking sequences of human genes.

   Mohammad Ali Faghihi, Mottagui-Tabar S, Wahlestedt C.

4. MicroRNA-219 modulates NMDA receptor mediated neurobehavioral dysfunction
   Jannet Kocerha, Mohammad Ali Faghihi, Miguel A. Lopez-Toledano, Jia Huang, Amy J. Ramsey, Marc G. Caron, Nicole Sales, David Willoughby, Joacim Elmen, Henrik F. Hansen, Henrik Orum, Sakari Kauppinen, Paul J. Kenny and Claes Wahlestedt
   Proceedings of the National Academy of Sciences (PNAS), In press
INTRODUCTION

1.1 Natural antisense transcript definition ............................................... 1

1.1.1 Widespread expression of natural antisense transcripts ....... 1

1.2 Table 1: Reported Sense-Antisense (S-AS) pairs ....................... 2

1.3 Historical overview ............................................................. 5

1.3.1 Prokaryotic and mammalian natural antisense transcripts ... 5
1.3.2 Human natural antisense transcripts .............................. 5
1.3.3 Large-scale detection of natural antisense transcripts .......... 5

1.4 Reported ncRNA transcript related to NATs ............................ 5

1.4.1 Small natural antisense transcripts (small-NATs) ............. 5
1.4.2 Promoter- and termini-associated small RNA ................. 5
1.4.3 Promoter-directed antigen RNA .................................... 6
1.4.4 Natural antisense miRNA and siRNA .......................... 6
1.4.5 Other non-protein-coding RNAs ................................. 6
1.4.6 MicroRNA: Synthesis and function.............................. 9

1.5 Natural antisense transcript databases ................................. 9

1.5.1 Non-protein-coding RNA database, RNAdb .................. 9
1.5.2 cis-natural antisense transcript database ....................... 9
1.5.3 Trans-SAMap ............................................................ 9

1.6 Classification of natural antisense transcripts ....................... 9

1.6.1 Classification based on cis vs. trans NATs: ..................... 10
1.6.2 Classification based on coding potential ......................... 10
1.6.3 Classification based on S-AS format: ............................ 11
1.6.4 Classification based on NAT-mediated regulation: .......... 12

1.7 Natural antisense transcripts evolutionary Conservation .......... 13

1.7.1 Motif conservation vs. sequence conservation ................. 13
1.7.2 Conservation and function ......................................... 14
1.7.3 Low abundance of ncRNA ....................................... 14
1.7.4 Reported conserved NATs ....................................... 14
1.7.5 Role of NATs in organismic and organismic complexity:... 14

1.8 The regulatory mechanisms by which NATs act are diverse ......... 15

1.8.1 Transcriptional interference ...................................... 15
1.8.2 Epigenetic changes, imprinting, methylation and demethylation.. 16
1.8.3 Involvement of NATs in genomic imprinting: ................. 16
1.8.4 Involvement of NATs in DNA/chromatin modifications... 17
1.8.5 Involvement of NATs in DNA recombination .................. 18
1.8.6 Involvement of NATs in X chromosome inactivation ....... 19
1.8.7 Alternative initiation, splicing and polyadenylation........... 20
1.8.8 Editing, nuclear retention and transport: ...................... 21
1.8.9 Changes in mRNA stability and translation: .................. 21
1.8.10 Cytoplasmic RNA duplex formation ........................ 23
1.8.11 Formation of endogenous siRNA ............................ 23

1.9 Clinical significance ...................................................... 25

1.9.1 NATs associated with cancers .................................... 25
1.9.2 NAT involvement in certain forms of anemia ................. 27
1.9.3 NAT-related disorders linked to imprinting........................27
1.9.4 Involvement of NATs in neurological disorders..............27
1.10 Alzheimer’s disease......................................................................................28
  1.10.1 Alzheimer’s disease an enormous unmet medical need......28
  1.10.2 Amyloid hypothesis related to Alzheimer’s disease ........29
  1.10.3 APP cleavage and generation of Aβ peptides.....................30
  1.10.4 BACE1 expression is tightly regulated ...........................30
  1.10.5 Unexpected complexity in the BACE1 gene locus ..........30
  1.10.6 BACE1-AS as an Alzheimer’s disease biomarker............31
1.11 Conclusion ........................................................................................................32
2 Present investigation ........................................................................................33
  2.1 Scientific background .............................................................................33
  2.2 FANTOM project revealed high abundance of NATs..................33
  2.3 Mechanistic aspects of NATs .................................................................35
  2.4 Natural antisense transcript for PINK1 ..............................................36
  2.5 FMR4 Noncoding RNA involved in Fragile X Mental retardation 36
  2.6 BACE1-AS a NAT involved in Alzheimer’s disease ..................37
  2.7 Functional relevance of NATs ...............................................................38
3 Prospective ........................................................................................................39
  3.1 Future research goals ..........................................................39
4 Acknowledgements ...................................................................................40
5 References ....................................................................................................41
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NATs</td>
<td>Natural Antisense Transcript</td>
</tr>
<tr>
<td>ncRNAs</td>
<td>Non-protein-coding RNAs</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>EST</td>
<td>Expressed sequence tag</td>
</tr>
<tr>
<td>FANTOM</td>
<td>Functional annotation of the mammalian transcriptome</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>TU</td>
<td>Transcription unit</td>
</tr>
<tr>
<td>UTR</td>
<td>Un-translated region</td>
</tr>
<tr>
<td>ENCODE</td>
<td>Encyclopedia of functional DNA elements</td>
</tr>
<tr>
<td>BACE1</td>
<td>β-secretase-1</td>
</tr>
<tr>
<td>BACE1-AS</td>
<td>BACE1-antisense</td>
</tr>
<tr>
<td>Aβ 1-42</td>
<td>Amyloid-β peptide 1-42</td>
</tr>
<tr>
<td>APP</td>
<td>Amyloid precursor protein</td>
</tr>
<tr>
<td>S-AS</td>
<td>Sense-antisense RNA</td>
</tr>
<tr>
<td>nt</td>
<td>Nucleotide</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>ASSAGE</td>
<td>Asymmetric strand-specific analysis of gene expression</td>
</tr>
<tr>
<td>PASR</td>
<td>Promoter-associated small RNA</td>
</tr>
<tr>
<td>TASR</td>
<td>Termini-associated small RNA</td>
</tr>
<tr>
<td>agRNA</td>
<td>Promoter-directed antigen RNA</td>
</tr>
<tr>
<td>TSS</td>
<td>Transcription start site</td>
</tr>
<tr>
<td>miRNA</td>
<td>MicroRNA</td>
</tr>
<tr>
<td>nat-miRNA</td>
<td>Natural antisense miRNA</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering RNA</td>
</tr>
<tr>
<td>Nat-siRNA</td>
<td>Natural antisense siRNA</td>
</tr>
<tr>
<td>DCL1</td>
<td>Dicer-like-1 enzyme</td>
</tr>
<tr>
<td>Endo-siRNA</td>
<td>Endogenous siRNA</td>
</tr>
<tr>
<td>DCL2</td>
<td>Dicer-like-2 enzyme</td>
</tr>
<tr>
<td>CAGE</td>
<td>Cap analysis of gene expression</td>
</tr>
<tr>
<td>XIC</td>
<td>X chromosome inactivation center</td>
</tr>
<tr>
<td>XIST</td>
<td>X-inactive specific transcript</td>
</tr>
<tr>
<td>TSIX</td>
<td>X (inactive)-specific transcript, antisense</td>
</tr>
<tr>
<td>Xi</td>
<td>Inactive X chromosome</td>
</tr>
<tr>
<td>Xa</td>
<td>Active X chromosome</td>
</tr>
<tr>
<td>ARE</td>
<td>AU rich element</td>
</tr>
<tr>
<td>dsRNA</td>
<td>double stranded RNA</td>
</tr>
<tr>
<td>RISC</td>
<td>RNA-induced silencing complex</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

1.1 NATURAL ANTISENSE TRANSCRIPT DEFINITION

Natural antisense transcripts (NATs) are RNA molecules that are transcribed from the opposite strand of conventional (sense) genes and overlap in part with mature sense mRNA. The overlap can occur in \textit{cis} (sense and antisense transcripts reside in the same gene locus) or in \textit{trans} (distinct loci). It has become clear, over years of reports on NATs, that the genome of a mammalian organism in many cases encodes two distinct genes by using both strands of the same DNA \cite{1, 2}.

1.1.1 Widespread expression of natural antisense transcripts

Mammalian transcriptome analyses have unraveled the existence of a large number of NATs, which in many cases have been proposed to be involved in the regulation of sense gene expression. The largest transcriptome profiling effort, that one pursued by the FANTOM-3 consortium, identified NATs for more than 70\% of transcription units (TU)\footnote{A transcription unit (TU) is a group of ESTs/mRNAs, usually with alternative splice pattern, in which the ESTs/mRNAs share exonic overlap of at least one nucleotide and are in the same chromosomal orientation} within mouse and human genomes \cite{3}. Over 20\% of human and mouse mRNA had previously been predicted to form sense-antisense (S-AS) pairs \cite{4-9}.

NATs have been suggested to regulate gene expression by causing transcriptional and post transcriptional changes in sense mRNA levels \cite{10-13}. Evolutionary studies and numerous well documented experiments have ruled out the possibility that this genome-wide NAT transcription represents a leakage of the RNA transcription machinery \cite{14}. In fact, the high abundance of antisense transcripts observed in gene expression studies have been established as real overlap which has considerably affected vertebrate genome evolution \cite{14}, and suggested to display pivotal role in complexity of the higher organism nervous system \cite{15}.

Table-1 contains all reported functional NATs in mammalian genomes characterized to date. This growing list of validated S-AS transcripts includes many important developmental genes as well as genes known to be involved in complex human disorders.
1.2 **TABLE 1: REPORTED SENSE-ANTISENSE (S-AS) PAIRS IN THE MAMMALIAN GENOME AND THEIR PROPOSED FUNCTION / DISEASE RELEVANCE AND POSSIBLE MECHANISM OF ACTION.**

Species; H: Human, M: Mice, R: Rat or rodent, Ch: Chicken, P: Primate, C: Canine, Mm: mammals, B: Bovine, S: Snail, F: Fungus, Z: Zebrafish

<table>
<thead>
<tr>
<th>S-AS transcripts</th>
<th>Proposed function / disease relevance</th>
<th>Suggested mechanism</th>
<th>Species</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>BACE1 &amp; BACE1-AS</td>
<td>Alzheimer’s disease</td>
<td>Stability</td>
<td>H, M</td>
<td>[16]</td>
</tr>
<tr>
<td>APOE &amp; APOE AS1</td>
<td>Alzheimer’s disease</td>
<td></td>
<td>H, M</td>
<td>[17]</td>
</tr>
<tr>
<td>PU.1 and PU.1-AS</td>
<td>Hematopoiesis</td>
<td>Translational block</td>
<td>H, M</td>
<td>[18]</td>
</tr>
<tr>
<td>HOXD &amp; HOXAIR</td>
<td>Embryonic development</td>
<td>Transcriptional gene silencing</td>
<td>H, M</td>
<td>[19]</td>
</tr>
<tr>
<td>Δ5-desaturase &amp; reverse Δ5-desaturase</td>
<td>Fatty acid metabolism</td>
<td>Translational block, Transcriptional interference, mRNA Stability</td>
<td>H, R</td>
<td>[20]</td>
</tr>
<tr>
<td>P15 &amp; P15AS</td>
<td>Tumor suppressor</td>
<td>Chromatin modification</td>
<td>H</td>
<td>[21]</td>
</tr>
<tr>
<td>P21 &amp; P21-AS</td>
<td>Tumor Suppressor</td>
<td>Chromatin modification</td>
<td>H</td>
<td>[22]</td>
</tr>
<tr>
<td>NKx2.2 &amp; NKx2.2AS</td>
<td>Neuronal cell differentiation</td>
<td></td>
<td>H</td>
<td>[23]</td>
</tr>
<tr>
<td>Zfh-5 &amp; Zfh-5AS</td>
<td>Transcription factor</td>
<td></td>
<td>H, M</td>
<td>[24]</td>
</tr>
<tr>
<td>PRGesteron receptor &amp; PR-AS</td>
<td>PR activation/inhibition</td>
<td>Promoter activation/inhibition by heterochromatin protein 1</td>
<td>H</td>
<td>[25]</td>
</tr>
<tr>
<td>HAR1F &amp; HAR1R</td>
<td>Neuro-development</td>
<td></td>
<td>H</td>
<td>[26]</td>
</tr>
<tr>
<td>WT1 &amp; WT1-AS</td>
<td>Kidney development</td>
<td>Methylation</td>
<td>H, M</td>
<td>[27]</td>
</tr>
<tr>
<td>BDNF &amp; BDNFOS</td>
<td>Neurotrophic factor</td>
<td>RNA duplex formation</td>
<td>P</td>
<td>[28, 29]</td>
</tr>
<tr>
<td>PINK1 &amp; naPINK1</td>
<td>Mitochondrial function</td>
<td></td>
<td>H</td>
<td>[30]</td>
</tr>
<tr>
<td>FMR1 &amp; ASFMR1</td>
<td>Fragile X Mental Retardatin</td>
<td>Epigenetic changes</td>
<td>H, M</td>
<td>[31, 32]</td>
</tr>
<tr>
<td>EPO-R &amp; asEPO-R</td>
<td>Lung growth</td>
<td>Stability, translation</td>
<td>H, C</td>
<td>[33]</td>
</tr>
<tr>
<td>Ghrelin &amp; ghrelinOS</td>
<td>Anxiety, depression</td>
<td></td>
<td>H</td>
<td>[34]</td>
</tr>
<tr>
<td>Rad 18 &amp; NAT-Rad18</td>
<td>Apoptosis</td>
<td>Post transcriptional</td>
<td>H, R</td>
<td>[35]</td>
</tr>
<tr>
<td>HFE &amp; HFE antisense RNA</td>
<td>Iron storage disorder</td>
<td>Translation repression</td>
<td>H</td>
<td>[36]</td>
</tr>
<tr>
<td>Zeb2 &amp; Zeb2 NAT</td>
<td>Epithelial-mesenchymal transition</td>
<td>Splicing</td>
<td>H, M</td>
<td>[37, 38]</td>
</tr>
<tr>
<td>TSP1 &amp; TSP1-AS</td>
<td>Platelet aggregation</td>
<td></td>
<td>H</td>
<td>[39]</td>
</tr>
<tr>
<td>Urocortin, Ucn &amp; Ucn-AS</td>
<td>Neuro-transmission</td>
<td>Post transcriptional</td>
<td>R</td>
<td>[40]</td>
</tr>
<tr>
<td>Sphk1 &amp; Khip1</td>
<td>Calcium mobilization</td>
<td>Demethylation</td>
<td>H, R</td>
<td>[41]</td>
</tr>
<tr>
<td>Pdc2d &amp; Tbp</td>
<td>Apoptosis</td>
<td>Editing, Alternative splicing, polyadenylation</td>
<td>H, M, Ch</td>
<td>[42]</td>
</tr>
<tr>
<td>Msh4 &amp;Hipa5</td>
<td>Meiotic recombination</td>
<td>DNA degradation</td>
<td>M</td>
<td>[43]</td>
</tr>
<tr>
<td>Pax6,2 &amp; Pax6,2OS</td>
<td>Eye development</td>
<td>-</td>
<td>M, H</td>
<td>[44]</td>
</tr>
<tr>
<td>Gene/Protein</td>
<td>Function</td>
<td>Regulation</td>
<td>Cell Type</td>
<td>References</td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>------------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>Six3.6 &amp; Six3.6 OS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Otx2 &amp; Otx2OS, Crx &amp; CrxOS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rax &amp; RaxOS, Vax2 &amp; Vax2OS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyaluronan Synthase 2 &amp; HASNT</td>
<td>Hyaluronan biosynthesis</td>
<td>-</td>
<td>M, H</td>
<td>[45]</td>
</tr>
<tr>
<td>Mx1 &amp; Mx1 AS</td>
<td>Skeletal differentiation</td>
<td>Splicing, Imprinting</td>
<td>R, H</td>
<td>[46]</td>
</tr>
<tr>
<td>FGF-2 &amp; FGF-AS (bFGF &amp; bFGF-AS)</td>
<td>Hematological tumors, endometriosis</td>
<td>Polyadenylation, Translational Block, Editing, Stability</td>
<td>Mm</td>
<td>[47-50]</td>
</tr>
<tr>
<td>p53</td>
<td>Differentiation</td>
<td>Transport</td>
<td>M</td>
<td>[51]</td>
</tr>
<tr>
<td>N-myc</td>
<td>Oncogenesis</td>
<td>Splicing</td>
<td>M, H</td>
<td>[52]</td>
</tr>
<tr>
<td>Tsix &amp; Xist</td>
<td>X inactivation</td>
<td></td>
<td>Mm</td>
<td>[53]</td>
</tr>
<tr>
<td>HIF-1α &amp; aHIF</td>
<td>Poor prognosis marker in breast cancer, renal cancer</td>
<td>RNA destabilization, RNA Splicing</td>
<td>H, R</td>
<td>[54]</td>
</tr>
<tr>
<td>Survivin &amp; EPR-1</td>
<td>Colon cancer</td>
<td>Not known</td>
<td>H</td>
<td>[55]</td>
</tr>
<tr>
<td>α-globulin &amp; LUC7L</td>
<td>A- Thalassemia</td>
<td>Methylation</td>
<td>H</td>
<td>[56]</td>
</tr>
<tr>
<td>IGF2R &amp; Air</td>
<td></td>
<td></td>
<td>H, M</td>
<td>[57-58]</td>
</tr>
<tr>
<td>KvLQT1</td>
<td>Beckwith-Wiedemann</td>
<td>Imprinting</td>
<td>H</td>
<td>[59]</td>
</tr>
<tr>
<td>SNURF-SNRPN &amp; UBE3A</td>
<td>Prader-Willi, Angelman syndrome</td>
<td>Imprinting</td>
<td>H</td>
<td>[60]</td>
</tr>
<tr>
<td>GNAS</td>
<td>Signal transduction</td>
<td>Imprinting</td>
<td>H, M</td>
<td>[61]</td>
</tr>
<tr>
<td>BCMA &amp; Antisense BCMA RNA</td>
<td>B-cell maturation</td>
<td>Translation block, Editing</td>
<td>H</td>
<td>[62, 63]</td>
</tr>
<tr>
<td>Bcl2 &amp; IgH</td>
<td>Follicular B-cell lymphoma</td>
<td>RNA Stabilization</td>
<td>H</td>
<td>[64]</td>
</tr>
<tr>
<td>c-erbA &amp; Rev-ErbAα</td>
<td>Thyroid hormone receptor</td>
<td>Splicing</td>
<td>H, R</td>
<td>[65-67]</td>
</tr>
<tr>
<td>Thymidylate synthase &amp; rTSo</td>
<td>DNA replication and repair</td>
<td>Editing</td>
<td>H</td>
<td>[68]</td>
</tr>
<tr>
<td>CHRNA3 &amp; CHRNA5</td>
<td>Neuronal nicotinic receptor</td>
<td>Stabilization</td>
<td>H, B</td>
<td>[69]</td>
</tr>
<tr>
<td>Myelin Basic Protein (MBP &amp; MBP-AS)</td>
<td>Myelin formation</td>
<td>Transport</td>
<td>M</td>
<td>[70]</td>
</tr>
<tr>
<td>eNOS &amp; NOS3AS (sONE)</td>
<td>Vascular disease</td>
<td>Inverse S-AS correlation</td>
<td>H, M</td>
<td>[71]</td>
</tr>
<tr>
<td>Neuronal Nitric Oxide Synthase &amp; NOS</td>
<td>Nervous system signaling</td>
<td>Post transcriptional, Translation</td>
<td>S, H, R</td>
<td>[72]</td>
</tr>
<tr>
<td>Inducible nitric oxide synthase (iNOS &amp; iNOS AS)</td>
<td>Inflammatory diseases</td>
<td>Stability</td>
<td>R</td>
<td>[73]</td>
</tr>
<tr>
<td>NOS2A &amp; anti-NOS2A</td>
<td>Neuronal differentiation</td>
<td>Inverse S-AS correlation</td>
<td>P</td>
<td>[74]</td>
</tr>
<tr>
<td>SMAD5 &amp; DAMS</td>
<td>TGF-beta/BMP</td>
<td>Transcriptional interference, Translational</td>
<td>H, R</td>
<td>[75]</td>
</tr>
<tr>
<td>Gene/Expression</td>
<td>Function</td>
<td>Regulator</td>
<td>Chromosome</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
<td>----------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>eIF2α</td>
<td>Inhibitory signals</td>
<td>T cell mitogenesis</td>
<td>RNA degradation</td>
<td>H [76]</td>
</tr>
<tr>
<td>ERCC-1, RAF49 (ASE-1) &amp; α1 Collagen</td>
<td></td>
<td>DNA repair</td>
<td>Stability, localization</td>
<td>H [77]</td>
</tr>
<tr>
<td>α1 Collagen</td>
<td>Chondro-genesis</td>
<td>Competitive transcriptional interference</td>
<td>Ch [78]</td>
<td></td>
</tr>
<tr>
<td>MKRN2 &amp; RAF1 (ase-1)</td>
<td>Cancer</td>
<td>Polyadenylation</td>
<td>Mm [79]</td>
<td></td>
</tr>
<tr>
<td>Hoxa 11</td>
<td>Development</td>
<td>epigenetic</td>
<td>Mm [80]</td>
<td></td>
</tr>
<tr>
<td>Cardiac Troponin 1 &amp; pMCH &amp; pMCH antisense</td>
<td>Myocardial function</td>
<td>Translation</td>
<td>H, R [81]</td>
<td></td>
</tr>
<tr>
<td>CDY &amp; CDY-AS</td>
<td>Spermatogenesis</td>
<td>RNA degradation</td>
<td>B [84]</td>
<td></td>
</tr>
<tr>
<td>FGFR-3 &amp; psiFGFR-3</td>
<td>Bone and hematopoietic maturation</td>
<td>RNA degradation, translation inhibition</td>
<td>M [85]</td>
<td></td>
</tr>
<tr>
<td>TOP1 &amp; TOP1-AS</td>
<td>Cell cycle</td>
<td>Translational regulation</td>
<td>H [86]</td>
<td></td>
</tr>
<tr>
<td>EP1 prostanoid receptor &amp; PKN protein kinase</td>
<td>Intracellular signaling</td>
<td>-</td>
<td>M [87]</td>
<td></td>
</tr>
<tr>
<td>EMX2 &amp; EMX2OS</td>
<td>Development</td>
<td>Splicing, Polyadenylation</td>
<td>H, M [88]</td>
<td></td>
</tr>
<tr>
<td>Thymidine kinase &amp; TK-AS</td>
<td>Cell cycle</td>
<td>Inverse S-AS correlation</td>
<td>M [89]</td>
<td></td>
</tr>
<tr>
<td>DIPLA1 &amp; DIPAS</td>
<td>Placenta specific</td>
<td>-</td>
<td>H [90]</td>
<td></td>
</tr>
<tr>
<td>GnRH &amp; SH</td>
<td>Gonadotropin-releasing hormone (GnRH)</td>
<td>Alternative splicing</td>
<td>R [91]</td>
<td></td>
</tr>
<tr>
<td>HLA-J cluster &amp; HZFw &amp; HZFc &amp; HZFw &amp; HCGV &amp; HTEX6 &amp; HTEX4</td>
<td>MHC class I</td>
<td>Transcriptional interference and/or Promoter methylation</td>
<td>H, M [92]</td>
<td></td>
</tr>
<tr>
<td>MHC IIA, IIX, IIB &amp; Antisense aII, xII, bII</td>
<td>Skeletal muscle myosin heavy chain regulation</td>
<td>-</td>
<td>R [93]</td>
<td></td>
</tr>
<tr>
<td>Cardiac MHC &amp; AS-βMHC</td>
<td>Cardiac myosin heavy chain alpha-beta gene switching</td>
<td>Transcriptional regulation at promoter</td>
<td>H, R [94]</td>
<td></td>
</tr>
<tr>
<td>ABO &amp; ABOAS</td>
<td>Blood group, ABO gene expression</td>
<td>Post-transcriptional, methylation</td>
<td>H [95]</td>
<td></td>
</tr>
<tr>
<td>Frequency, frq &amp; antisense-frq</td>
<td>Circadian clock function</td>
<td>Inverse S-AS correlation</td>
<td>F [96]</td>
<td></td>
</tr>
<tr>
<td>ORC7 &amp; ORC7LS</td>
<td>Wilms tumor</td>
<td>Imprinting</td>
<td>H [97]</td>
<td></td>
</tr>
<tr>
<td>Tenasin-X &amp; P450c21B</td>
<td>Adrenal function</td>
<td>Post transcriptional</td>
<td>H [98]</td>
<td></td>
</tr>
<tr>
<td>NPT &amp; NPT-AS</td>
<td>Na/Pi cotransporter, Phosphate homeostasis</td>
<td>Translation interference</td>
<td>M, Z [99]</td>
<td></td>
</tr>
<tr>
<td>PKN &amp; EP1</td>
<td>Protein kinase</td>
<td>Alternative polyadenylation</td>
<td>M [87]</td>
<td></td>
</tr>
<tr>
<td>COX10 &amp; C17orf1</td>
<td>Charcot-Marie-Tooth</td>
<td>Post transcriptional</td>
<td>H [100]</td>
<td></td>
</tr>
</tbody>
</table>
|c-myc & c-myc-antisense| Oncogene| Pre-mRNA processing, Transcription interference| R, H [2, 101]
1.3 HISTORICAL OVERVIEW

1.3.1 Prokaryotic and mammalian natural antisense transcripts
Overlapping antisense transcripts were first identified in viruses e.g. polyoma virus and prokaryotes [102-106]. In 1986 Trevor Williams and Mike Fried reported, for the first time, antisense transcription from opposite strands of DNA in mammalian systems [1]. The authors identified a mouse genetic locus at which two processed poly (A)+ RNA species transcribed from opposite strands overlap by 133 nucleotides at their 3’ ends and suggested that endogenous RNA double strand formation from overlapping transcripts can prevent RNA processing and/or transport. Other natural antisense RNAs were reported from rodent [2, 91] and Drosophila [107-109] genomes in or around 1986.

1.3.2 Human natural antisense transcripts
Three years later Van Duin et. al. reported, for the first time, a human example of conserved overlapping antisense transcription for ERCC-1 (Excision repair cross-complementing rodent repair deficiency, complementation group 1), DNA repair gene regions [77]. The antisense transcript for ERCC-1 forms a tail-to-tail duplex with the sense ERCC-1 mRNA, and the duplex RNA suggested to mediate S-AS transport to a common cytoplasmic location, where it affect translation and/or stability of both transcripts [77].

1.3.3 Large-scale detection of natural antisense transcripts
These scattered reports were followed by numerous studies such as large scale sequencing of cDNA clones [3, 12, 110, 111], tiling arrays [112-115], analysis of RefSeq and EST databases [5-10, 116, 117], hybridization techniques [21, 118] SAGE libraries [119, 120], strand specific microarrays [121-123], and most recently with a technique called asymmetric strand-specific analysis of gene expression (ASSAGE) [124]. These studies demonstrated the widespread occurrence of antisense transcription in mammalian genomes.

1.4 REPORTED ncRNA TRANSCRIPT RELATED TO NATs

1.4.1 Small natural antisense transcripts (small-NATs)
Overlapping transcription of small RNA (<50 nt) in the sense and antisense direction has been documented by utilizing strand-specific genomic tiling arrays in the ENCODE\(^1\) region of the human genome [125]. These small sense-antisense transcripts do not correspond to annotated NATs. There is no evidence for double stranded RNA or hairpin RNA precursors that could represent intermediates in the biogenesis of such small natural antisense transcripts. These small NATs are thought to be involved in “housekeeping” functions ensuring the basic structural and metabolic needs of living cells [125].

1.4.2 Promoter- and termini-associated small RNA (PASR & TASR)

---
\(^{1}\) ENCODE, the ENCyclopedia Of DNA Elements is a publicly founded project that aims to find functional elements in human genome
Promoter-associated small RNA (PASR) and Termini-associated small RNA (TASR) are two classes of small RNA less than 200 nucleotide which have been identified by genomic tiling arrays and are enriched in the 5' UTR and 3' UTR of genes, respectively [125, 126]. The enrichment of NATs in both promoter and termini regions is confirmed by unbiased technique called asymmetric strand-specific analysis of gene expression (ASSAGE) [124].

1.4.3 Promoter-directed antigen RNA

There is a class of newly described synthetic antisense RNA, promoter-directed antigen RNA (agRNA), which can bind to the transcription start site (TSS) of genes and can activate or block transcription of the target gene dependent on the cellular context [127]. Natural antisense RNA has been shown to be essential for agRNA mediated gene activation/silencing, providing a scaffold for suppressor or activator proteins to bind to the promoter region [25]. The potency and generality of silencing with agRNA are consistent with the suggestion that RNA-mediated recognition of TSS may be a natural mechanism for regulation of gene expression. Given the fact that significant fraction of PSAR overlap with the transcription-initiation sites of genes [125], PASR could possibly act as an endogenous agRNA, interacting with NATs to regulate transcriptional output at the DNA level.

1.4.4 Natural antisense miRNA (nat-miRNA) and siRNA (nat-siRNA)

Natural antisense miRNA (nat-miRNA) [128] and natural antisense siRNA (nat-siRNA) [129] are classes of small noncoding RNA that originate from overlapping region of sense-antisense RNA pairs. Canonical miRNAs are generated from characteristic hairpin structure, in miRNA precursor. Similarly, nat-miRNAs derive from intra-molecular (stem-loop formation) interactions and their production in plant depends on Dicer-like-1 (DCL1) enzyme. Nat-miRNAs have been reported in the fly where the Hox miRNA locus generates miRNAs from both sense and antisense transcripts [130]. Nat-miRNAs have also been demonstrated in mice, originating from NATs to an imprinted locus retrotransposon-like gene (Rtl1) [131]. These nat-miRNA, unlike canonical miRNAs, are fully complementary to their target mRNA.

On the other hand, nat-siRNA’s, which are essentially the same as the recently reported endogenous siRNA (endo-siRNAs) [132], derive from inter-molecule (double strand-RNA formation) interactions between sense and antisense transcripts. Nat-siRNA production in plant, unlike nat-miRNA, depends on Dicer-like2 (DCL2) enzyme. Both RNA species can form a perfect match with their target mRNA in plants, inducing mRNA cleavage [128]. Importantly, presence of many cis-NAT in eukaryotic systems suggests a great potential for generation of nat-miRNA and nat-siRNA small RNA species.

1.4.5 Other non-protein-coding RNAs

Table-2 and text box-1 summarizes the various categories of non-protein-coding RNA (ncRNA) and their special features.
Table 2: Features of various ncRNA classes; double stranded RNA: dsRNA

<table>
<thead>
<tr>
<th>Non-coding RNA</th>
<th>Size</th>
<th>Proposed function</th>
<th>Origin/target</th>
<th>Features</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small natural antisense transcripts</td>
<td>&gt; 50 nt</td>
<td>Housekeeping</td>
<td>ENCODE</td>
<td>No dsRNA or hairpin RNA precursor</td>
<td>[125]</td>
</tr>
<tr>
<td>(small-NATs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Promoter-associated small RNA (PASR)</td>
<td>&gt; 200 nt</td>
<td>Transcription initiation</td>
<td>5’ UTR</td>
<td>Housekeeping</td>
<td>[125]</td>
</tr>
<tr>
<td>Terminally-associated small RNA (TASR)</td>
<td>&gt; 200 nt</td>
<td>Post-transcriptional</td>
<td>3’ UTR</td>
<td>Housekeeping</td>
<td>[125]</td>
</tr>
<tr>
<td>Promoter-directed antigen RNA (agRNA)</td>
<td>20 nt</td>
<td>Transcription activation/suppression</td>
<td>TSS</td>
<td>Synthetic</td>
<td>[127]</td>
</tr>
<tr>
<td>Natural antisense miRNA (nat-miRNA)</td>
<td>~20 nt</td>
<td>Translational suppression</td>
<td>Antisense RNA</td>
<td>Derive from intra-molecule hairpins</td>
<td>[128, 131]</td>
</tr>
<tr>
<td>Natural antisense siRNA (nat-siRNA)</td>
<td>~20 nt</td>
<td>RNAi</td>
<td>Overlapping region</td>
<td>Derive from inter-molecule dsRNA</td>
<td>[129]</td>
</tr>
<tr>
<td>Endogenous siRNA (endo-siRNA)</td>
<td>~20 nt</td>
<td>RNAi</td>
<td>Overlapping region, Pseudogenes, Repetitive elements</td>
<td>Derive from inter-molecule dsRNA</td>
<td>[133-137]</td>
</tr>
<tr>
<td>Transacting siRNA (ta-siRNA)</td>
<td>~20 nt</td>
<td>RNAi</td>
<td>Plant miRNA cleavage product</td>
<td>Need RdRP</td>
<td>[138]</td>
</tr>
<tr>
<td>Small temporally regulated RNAs (siRNA)</td>
<td>~20 nt</td>
<td>RNAi</td>
<td>C. elegans miRNA, Let-7 Lin-4</td>
<td>Translation repression</td>
<td>[139, 140]</td>
</tr>
<tr>
<td>MicroRNA (miRNA)</td>
<td>~20 nt</td>
<td>RNAi</td>
<td>Imperfect complementarity to target RNA</td>
<td>Translational repression, mRNA decay/degradation</td>
<td>Reviewed at [141]</td>
</tr>
<tr>
<td>MacroRNA</td>
<td>~1kb-1kb</td>
<td>Heterogeneous</td>
<td>miRNA regulation</td>
<td>Spliced, Cupped and polyadenylated</td>
<td>(See text)</td>
</tr>
<tr>
<td>Small nucleolar RNAs (snorRNAs)</td>
<td>200 nt</td>
<td>Methylation, Pseudouridylation</td>
<td>rRNA processing</td>
<td>Reside in nucleolus</td>
<td>[142]</td>
</tr>
<tr>
<td>Telomerase RNA</td>
<td>~1kb</td>
<td>Maintenance of telomere</td>
<td>Telomeres</td>
<td>Reside in nucleus</td>
<td>[143]</td>
</tr>
<tr>
<td>Small nuclear RNA (snRNA)</td>
<td>~150 nt</td>
<td>mRNA splicing</td>
<td>Splicosome</td>
<td>Reside in nucleus</td>
<td>[144, 145]</td>
</tr>
<tr>
<td>Small interfering RNA (siRNA)</td>
<td>~20 nt</td>
<td>RNAi</td>
<td>mRNA degradation</td>
<td>Synthetic</td>
<td>Reviewed at [146]</td>
</tr>
<tr>
<td>PIf-1-interacting RNA (piRNA)</td>
<td>~30 nt</td>
<td>Chromatin modification</td>
<td>Maintaining germline DNA integrity</td>
<td>Germline silencing of repeat transcripts</td>
<td>Reviewed at [147]</td>
</tr>
<tr>
<td>Repeat-associated small interfering RNA (rasiRNA)</td>
<td>~20 nt</td>
<td>Chromatin modification</td>
<td>Retroposon</td>
<td>Germline silencing of repeat transcripts</td>
<td>Reviewed at [146]</td>
</tr>
<tr>
<td>Transfer RNA (rRNA)</td>
<td>74-95 nt</td>
<td>mRNA translation</td>
<td>Protein-coding mRNA</td>
<td>Anticodon loop</td>
<td></td>
</tr>
<tr>
<td>Ribosomal RNA (rRNA)</td>
<td>18S, 28S</td>
<td>mRNA translation</td>
<td>Protein-coding miRNA</td>
<td>Tandem repeats</td>
<td></td>
</tr>
</tbody>
</table>
**Text Box 1. Some mammalian RNA species**

1. **Messenger RNA (mRNA)**, well known class of RNA with average size of 2 kb. It is transcribed from DNA and processed before leaving the nucleus. The processed mRNA, which is located in cytoplasm, contains polyA tail, cap structure, open reading frame and it is frequently spliced, in many cases alternatively.

2. **MicroRNA (miRNA)** is a small non-coding regulatory RNA. The miRNA precursor (pri-premiRNA) is transcribed into a single stranded RNA transcript of approximately 150-250 nucleotides in length. A ‘hairpin’ secondary structure is formed in pri-premiRNA which is then processed by the enzyme Drosha and exported to the cytoplasm. Pre-miRNA is further processed by the enzyme Dicer to create a stable, ~22 nucleotide single-stranded mature miRNA from one arm of the hairpin. The mature miRNA sequence tends to be highly conserved.

3. **Small nucleolar RNAs (snoRNAs)** are a class of small RNA molecules that guide chemical modifications (methylation or pseudouridylation) of ribosomal RNAs (rRNAs) and other RNA genes (tRNAs and other small nuclear RNAs (snRNAs)). snoRNAs are commonly referred to as guide RNAs but should not be confused with the guide RNAs (gRNA) that direct RNA editing in trypanosomes. The snoRNAs are less than 70 nucleotides in length including 10-20 nucleotides of antisense elements for base pairing.

4. **Small nuclear RNA (snRNA)** is a class of small RNA molecules that are found within the nucleus of eukaryotic cells. They are involved in a variety of processes such as RNA splicing, regulation of transcription factors (7SK RNA) or RNA polymerase II (B2 RNA), and maintaining the telomeres.

5. **Piwi-interacting RNA (piRNA)** is a class of small RNA molecules that is expressed in mammalian testes and forms RNA-protein complexes with Piwi proteins. These piRNA complexes (piRCs) have been linked to transcriptional gene silencing of retrotransposons and other genetic elements in germ line cells, particularly those in spermatogenesis. Purification of these complexes has revealed that these oligonucleotides are approximately 29-30 nucleotides long.

6. **Rapid associated RNA (RasiRNA):** presumable derives from long double stranded RNA (dsRNA) and match to repetitive sequence elements in antisense orientation. In the *Drosophila* germline, rasiRNAs ensure genomic stability by silencing endogenous selfish genetic elements such as retrotransposons and repetitive sequences.

7. **Natural antisense transcripts (NAT)** are single-stranded RNAs that are complementary to mRNAs. NAT regulate mRNAs in a concordant or discordant manner. The average length of NAT is 2 kb, but in some cases it is extremely long (over 100 kb). NAT in some cases is spliced and contains polyA, cap structure or even open reading frame.

8. **Other long non-coding RNA transcripts** (sometimes referred to as macroRNA) are diverse and not necessarily well conserved; they are often processed, containing polyA tail and/or cap structure. There is no significant open reading frame for macroRNAs and their functions are largely unknown.

9. **Ribosomal RNA (rRNA) and transfer RNA (tRNA)** are well studied components of the protein synthesis machinery.
1.4.6 MicroRNA: Synthesis and function

MiRNAs are a class of small ncRNAs that have recently generated much interest [148-150]. The enzymatic machinery and sequence of events, involved in the biogenesis of miRNAs are highly conserved across animals and plants. Specifically, miRNA precursor (immature miRNA) is transcribed into a single stranded RNA transcript of approximately 50-120 nucleotides in length, which forms a ‘hairpin’ secondary structure [151, 152]. This precursor miRNA hairpin is exported from the nucleus to the cytoplasm, where it is processed by Dicer, in combination with Argonaute proteins, and the RISC complex (RNA-induced silencing complex) to yield a stable, ~22 nucleotides single-stranded mature miRNA from one arm of the pre-miRNA hairpin [152]. This mature miRNA sequence is highly conserved across species [148, 149]. In plants, miRNAs often demonstrate complete or precise complementary base-pairing with target mRNA transcripts, resulting in the cleavage and degradation of target mRNA transcripts, via RNA interference (RNAi) machinery [153, 154]. In contrast to plant, animal miRNAs are generally thought to recognize and bind to the target mRNA transcripts by incomplete complementary base pairing. Such imperfect base pairing with target transcripts results in translational inhibition and down-regulation of associated proteins. Thus, miRNAs may represent ‘master regulators’ of gene expression that orchestrate the expression levels of clusters of associated proteins. Indeed, it has been estimated that more than 33% of human gene products may be regulated by miRNAs [155].

1.5 NATURAL ANTISENSE TRANSCRIPT DATABASES

1.5.1 Non-protein-coding RNA database, RNAdb

RNAdb is a comprehensive database of mammalian ncRNA, which provide nucleotide sequences and annotations for tens of thousands of ncRNAs, including a wide range of mammalian microRNAs, small nucleolar RNAs and larger mRNA-like ncRNAs [156].

1.5.2 cis-natural antisense transcript database

There are some more specialized databases for NATs like LEADS-Antisensor [6], SADB (http://fantom31p.gsc.riken.jp/s_as/), AntiHunter [157, 158], NATsDB [159] and antiCODE [160], among these NATsDB has the most coverage of various species and antiCODE is the most comprehensive dataset comprising most of currently detected NAT pairs and introduces a simple classification system to facilitate studies of natural antisense transcripts [160].

1.5.3 Trans-SAMap

Trans-SAMap [161], is a dataset of trans-NATs in human and nine more species freely available at http://trans.cbi.pku.edu.cn/.

1.6 CLASSIFICATION OF NATURAL ANTISENSE TRANSCRIPTS

There are different classifications for functional RNA molecules. NATs are classified by their expression pattern, their alignment with regard to the sense mRNA, coding potentials and the type of regulation exerted on the sense gene. Classifications of NATs are useful for defining various transcripts and more importantly for predicting their regulatory function. The more common and practical classifications of NATs are outlined below:
1.6.1 Classification based on cis vs. trans NATs:

Most NATs are cis-encoded antisense RNA [162, 163]. By definition, cis-NATs are complementary RNA with an overlapping transcriptional unit (TU) at the same chromosomal locus. Trans-NATs are complementary RNA transcribed from different chromosomal locations [118, 161, 164].

1.6.1.1 Abundance of Trans-NATs

A recent study on Trans-NAT showed that the abundance of these regulatory elements is much more than previously expected. Although the authors had applied very stringent criteria for selecting trans-NATs, eliminating all the NATs originating from repeat regions and pseudogenes, they reported presence of trans-NATs as high as 4.13% among transcriptional units of various species [161]. Particularly, trans-NATs have been reported for nearly 3,000 human TUs (or 2.89% of all human TUs), which involve ncRNA partner at least in one-fourth of the reported cases [161].

1.6.1.2 Pseudogenes and trans-NATs

Trans-NATs often originate from pseudogenes or repeat regions. Repetitive sequences in genome and pseudogenes have long been considered to be non-functional artifacts of transposition pathways. However, an increasing number of reports point to the functional role for repetitive elements in post-transcriptional events [165]. Anti-sense transcription of pseudogenes may constitute a mechanism for controlling their cognate (parental) genes.

1.6.1.3 Examples of pseudogenes-related trans-NATs

Such a regulatory role has been demonstrated for topoisomerase I, neural nitric oxide synthase, inducible nitric oxide synthase (NOS2A/anti-NOS2A) and fibroblast growth factor receptor-3 pseudogenes [72, 74, 85, 86]. Importantly, recent reports proposed a role for a subset of mammalian pseudogenes in the production of endogenous siRNAs (endo-siRNA) through formation of double stranded RNA [166-168].

1.6.1.4 Chimeric NATs

Chimeric NATs are RNAs with identity to more than one region of the genome and may have some function in controlling retrotransposons [169, 170]. Chimeric NATs and in some cases trans-NATs offer partial complementarity to more than one target transcript [161], therefore are capable of regulating many sense mRNA at the same time reminiscent of miRNA-target mRNA interactions.

1.6.2 Classification based on coding potential

NATs can be protein coding or non-protein-coding (Table-3). Noncoding RNA’s (ncRNA) appear to be the most abundant form of NATs in the mammalian genome, in which there is a ncRNA overlapping with the protein coding target mRNA [3].

Coding–coding partners are also common in the mammalian genome for example Thymidylate synthetase (TS) and Enolase superfamily member (rTSalpha) are two protein coding genes in S-AS orientation [68].
Table-3: *Cis*-NAT from FANTOM-3 dataset with exon overlap.

<table>
<thead>
<tr>
<th>TU pairing types</th>
<th>Cis-S/AS pairs</th>
<th>% in group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coding-coding</td>
<td>1,687</td>
<td>37%</td>
</tr>
<tr>
<td>Coding-noncoding</td>
<td>2,478</td>
<td>55%</td>
</tr>
<tr>
<td>Noncoding-noncoding</td>
<td>355</td>
<td>8%</td>
</tr>
<tr>
<td>Total</td>
<td>4,520</td>
<td>100%</td>
</tr>
</tbody>
</table>

1.6.3 Classification based on S-AS format:

There is enrichment for the presence of NATs in both 5’ and 3’ ends of the sense genes. Some reports indicate that over 70% of *cis*-encoded NATs have a tail-to-tail 3’ overlap, while 15% have a 5’ head-to-head format with the 3’UTR or 5’UTR of their target mRNA, respectively. The remaining NATs have full, intronic or coding overlap (Figure-1). These figures vary widely and some reports [3] revealed more prevalent head-to-head orientation among NATs.

Figure-1: Illustration of prominent transcriptomics patterns relating to complex loci in human and mouse genomes.

(A) *cis*-NATs, like *BACE1-AS* [16] in which two converging transcripts, from opposite strands of DNA, have overlapping exons shown as hashed parts.
(B) Bi-directional promoters as, for example, for FMR4 [32] drive transcription of two RNAs in opposite directions. The transcripts may share the same transcription start site (TSS), or even exhibit overlapping 5’UTRs.
(C) Full or intronic transcripts in which antisense RNA (blue) is inside the boundary of the sense transcript (red). Even if the fully processed RNAs do not contain overlapping sequences, RNA duplexes can still form between unprocessed transcripts. Alternatively, antisense RNA can bind to the DNA and exert its regulatory function.

1.6.3.1 Examples of 5’UTR S-AS overlapping pattern

NATs with 5’ overlapping pattern, e.g. the antisense transcript to Wilm’s tumor suppressor gene (*WT-AS*) [171] or tumor suppressor *p15 (P15AS)* [21] or antisense transcript to hemochromatosis protein isoform 1 precursor (HFE-AS) [36] can exert their regulatory effects through epigenetic mechanisms like methylation of the sense mRNA promoter region[172].
1.6.3.2 Examples of 3' UTR S-AS overlapping pattern

Tail-to-tail orientation is reported to be much more common in mammalian genomes and has a much higher probability (>5 times) of evolutionary conservation and discordant regulation (inverse expression pattern see below) [173]. Tail to tail overlapping S-AS pairs, like ERCC-1, RAF49 (ASE-1) [77] Makorin-2 and RAF1 [79], may change the polyadenylation of sense mRNA, cover the miRNA binding sites or modulate post transcriptional events by a variety of mechanisms.

1.6.3.3 Example of coding S-AS overlapping pattern

NATs interacting with the coding part of their target mRNAs, like ghrelinOS [34], change the stability or splicing of the sense mRNA. Thus, classification of NATs by their overlapping pattern is important for prediction of functional properties.

1.6.4 Classification based on NAT-mediated regulation:

There are broadly two types of regulation between sense and antisense transcripts, namely concordant and discordant regulation.

1.6.4.1 Examples of concordant regulation

In concordant regulation, like in the case of BACE1-AS [16], aHIF1 [174], asEPO-R [33] and Zeb2-AS [37], the NATs augment the level of the sense RNA, or corresponding protein levels.

BACE1-AS transcript: We have recently identified a conserved noncoding antisense transcript for β-secretase-1 (BACE1), a critical enzyme in Alzheimer’s disease pathophysiology. The BACE1-antisense transcript (BACE1-AS) concordantly up-regulates BACE1 mRNA and subsequently BACE1 protein expression in vitro and in vivo [16].

Zeb2-AS transcript: Concordant regulation is also reported for NATs to a transcriptional repressor of E-cadherin called zinc finger homeobox 1b (zeb2). The antisense RNA for Zeb2 inhibits 5'UTR splicing, which in turn makes an internal ribosome entry site (IRES) accessible and causes increase in Zeb2 protein level without changes in mRNA transcript. This NAT is an example of concordant regulation in which antisense RNA induces Zeb2 protein up-regulation and subsequently induces epithelial-mesenchymal transition [37].

asEPO-R transcript: Concordant regulation is reported in the case of as-EPO-R transcript. Erythropoietin receptor (EPO-R) mRNA and protein levels are shown to be positively controlled by a cis-NAT, called asEPO-R [33]. Both S-AS transcripts as well as EPO-R protein are increased in canine lung after pneumonectomy, suggesting a role for antisense regulation during lung growth. Although asEPO-R has two potential open reading frames (ORF), antisense-induced up-regulation of EPO-R seems to be protein independent [33].

1.6.4.2 Examples of discordant regulation

In discordant regulation, e.g. naPINK1 [30], RevΔ5-desaturase [20], zfh-5AS [24] or antisense-frq [96], the antisense transcripts have negative (opposing) effects on sense transcripts.
**RevΔ5-desaturase transcript:** The noncoding antisense RNA for Δ5-desaturase regulates fatty acid metabolism during the transition between fasting and refeeding by altering the expression of the sense gene. The regulation was shown to be discordant *i.e.* a diet enriched in fish oil produced a reciprocal increase in antisense and decrease in sense transcript [20].

**zfh-5AS transcript:** Consistent with discordant regulation, gene-targeted knock-out of zfh-5 NAT, a long spliced and polyadenylated RNA, caused up-regulation of zfh-5 mRNA, *in vivo*, in the brain of mice [24].

**frq-AS transcript:** Another documented discordant regulation is reported for a NAT of *frequency* (*frq*) transcripts, related to *Neurospora crassa* circadian clock function. This ncRNA transcript is important in synchronizing internal and external time by reducing *frq* RNA transcript as well as FRQ protein [96].

### 1.6.4.3 Features of concordant vs. discordant regulation

Interestingly, in most cases both types of regulation (concordant & discordant) are primarily unidirectional in that the antisense transcript regulates the sense RNA [175].

In most reported cases of concordant regulation, the NATs show a low degree of RNA processing, are not spliced or display short introns [176] that are stay in the nucleus and have a much shorter half-life time than their coding partner [16]. The rapid transcription and processing of this NAT class, similar to what suggested for (*nimble* genes) [176], implies that they are implicated in acute stress responses.

In contrast, discordant NATs are mainly fully processed RNA transcripts with multiple exons and more prominent in the cytoplasm. High cytoplasmic abundance of this group of NATs suggests a long lasting and housekeeping regulatory role through cytoplasmic RNA duplex formation (see proposed cytoplasmic mechanisms).

### 1.7 NATURAL ANTISENSE TRANSCRIPTS EVOLUTIONARY CONSERVATION

The presence of NATs is already shown in a spectrum of eukaryotic organisms including human [4-8], mouse [12, 111] cow [84], dog [161], frog, zebrafish, chicken [9] rat, nematodes [177] *Drosophila* [9, 178], rice [179], Arabidopsis [180, 181] and yeast [182, 183]. This widespread occurrence in various organisms indicates that the overall regulation of gene expression through NATs is a very well conserved phenomenon [184, 185].

#### 1.7.1 Motif conservation vs. sequence conservation

The lack of strong nucleotide sequence conservation among individual NATs despite their occurrence in different species may indicate that only very short and specific parts of the whole NAT sequence or a specific secondary structure is required for NAT-mediated gene regulation.

Unlike protein coding genes where complete nucleotide conservation is required to keep the amino-acid sequence functional, ncRNA may only need to maintain certain motifs to preserve their function. Therefore, selective pressure in ncRNA is likely only applied to specific motifs rather than to the entire nucleotide sequence. Indeed,
ncRNA transcripts in the HOX gene clusters demonstrate certain significantly enriched sequence motifs related to their expression pattern in the body [19].

1.7.2 Conservation and function
It is worth noting that lack of conservation does not necessarily indicate lack of function and that evolutionary conservation varies considerably among RNA classes [186]. The ENCODE project reported that as many as 50% of the experimentally identified functional elements, especially ncRNAs, do not show evidence of evolutionary constraint across mammals [187].

1.7.3 Low abundance of ncRNA
Lack of strong evolutionary conservation may indeed related to the low abundance of these ncRNA molecules, which make sequence detection more demanding than conventional protein-coding genes. This problem could be subsided, by applying deep sequencing approach, where the sequence detection limits will be pushed to very low copy number RNA molecules.

1.7.4 Reported conserved NATs
Nevertheless, many NATs are reported to be well conserved during evolution [9]. Up to 40% of trans-NATs [161] and at least 1,000 of cis-NATs from FANTOM-3 are well conserved between human and mice [188] and displayed identical expression pattern between the two species [189]. In a recent study, evolutionary conservation has been reported for 27% of the overlapping genes (NATs), in the sense and antisense direction, between human and mouse [190].

1.7.4.1 Primate specific NATs
There are reports of primate-specific NATs, such as BDNF-AS, FMR4, Anti-NOS2A and BMC transcripts [28, 29, 32, 74, 191, 192].

1.7.4.2 Human specific NATs
Human specific NATs have been also reported expressed from the human accelerated region 1 (HAR1). HAR1 is a genomic region that is conserved among mammals, but nevertheless has changed rapidly in the human lineage, and gives rise to multiple antisense-overlapping ncRNAs, one of which is specifically expressed in Cajal-Retzius neurons of the developing neocortex of humans [26].

The non-conserved sequences have been linked to the emergence of human-specific brain features [26]. Although evolutionary conservation may not be a reliable signature of functional NATs, it can be an effective resource given the various reports that correlate conservation with function.

1.7.5 Role of NATs in organismic and organistic complexity:
It has previously reported that the percentage of the genome transcribed into ncRNA increases with the complexity of the organism [195]. In spite of numerous large intergenic spaces in the genome of higher mammals, many genes are still overlapping, suggesting that such a genomic arrangement must be functionally beneficial.

Abundance of antisense transcription varies between multicellular animals; however, a correlation between antisense transcription and organismic complexity was not observed [177]. In fact, NATs are not more enriched in human brain than in mouse brain, undermining their role in organismic complexity.
However, antisense transcription was more prominent in nervous system compared to other tissues, which suggest a role for NATs in organ complexity [177, 194]. Long ncRNAs have a highly specific distribution pattern in mouse brain comparing to protein coding mRNAs [195]. Utilizing the Allen Brain Atlas [196], Mercer et al. identified expression of 849 ncRNA, including NATs in adult mouse brain, the majority of which were expressed in specific neuroanatomical regions, cell types or subcellular compartments [195]. The authors further predicted the expression of another 20,000 long ncRNAs in brain, supporting the idea that ncRNA underlie the functional complexity of the brain [195].

1.8 THE REGULATORY MECHANISMS BY WHICH NATS ACT ARE DIVERSE.

The NATs have been suggested to regulate gene expression by controlling various levels of gene expression including chromatin architecture/epigenetic memory, transcription, transcript localization, translation and turnover [10-13]. NATs have also been shown to be involved in methylation, demethylation [41], parental gene imprinting [197], chromosome X inactivation [198], RNA splicing [52, 65, 199], transport [51], polyadenylation [42, 47, 200], editing and stabilization [201, 202].

1.8.1 Transcriptional interference

The transcriptional collision model (Figure-2) is based on the assumption that during cis-NATs transcription, RNA polymerases bind to the promoters of both sense and antisense transcripts, and move toward the 3'-end of the genes. RNA polymerase complexes collide in the overlapping region blocking further transcription [203]. Transcription interference has been observed in Drosophila bithorax (bxd) ncRNAs [204] and Saccharomyces cervisiae [205].

![Figure-2 NATs molecular mechanisms](image)

Figure-2 NATs molecular mechanisms; Transcription interference model in which RNA synthesis from one DNA strands would possibly collide with transcription of the other strand. According to this model, transcription only occurs in one direction at any given time, and active antisense transcription would suppress sense RNA transcription.

This function might be true for a subset of mammalian cis-encoded NATs; however, transcription at the same locus may occur independently, with individual turn for each RNA or simply sense and antisense RNA may transcribed from different paternal or maternal alleles. Allelic-specific transcription may be supported by the fact that the X chromosome, present in only one active copy in mammals, shows a significantly lower degree of antisense transcription [5, 9, 12].
Epigenetic changes, imprinting, methylation and demethylation

The NATs can also be involved in transcriptional regulation [206], like methylation and monoallelic expression (Figure-3). Monoallelic expression includes X-chromosome inactivation, imprinting and allelic exclusion in B and T lymphocytes.

1.8.2 Involvement of NATs in genomic imprinting:

Imprinted genes are genes for which only one allele, maternal or paternal, is actively transcribed. Fifteen percent of imprinted genes were previously reported to have an associated antisense transcript [207]. In a more recent report, this fraction significantly increased to 24-47% of human and mouse imprinted loci [9] and up to 81% when antisense transcripts to introns were included [3].

There are more than 160 imprinted genes, identified so far in human and mice, which are generally organized into clusters (http://igc.otago.ac.nz/home.html). NATs have been shown to be involved in imprinting of maternal genes through chromatin modification or methylation of CpG islands.

1.8.2.1 Examples of NATs involved in genomic imprinting

There are several studies suggesting a pivotal role for NATs in imprinted genes, such as IGF2R [57], KCNQ1 [59], UBE3A [60], ATP10C, MKRN3, MAGEL2, NDN [208], Slc22a2, Slc22a3 [209], GNAS [61] and Gnas [210]. Although, some controversy still exists [211, 212], guided chromatin and DNA modification by antisense RNA and spreading to the neighboring genes was clearly shown for some imprinted genes, such as insulin-like growth-factor type-2 receptor (IGF2R) [209] and potassium voltage-gated channel, KQT-like (Kcnq1) imprinting control region [213].

Kcnq1ot1 transcript:

Antisense transcription and its functional role in bidirectional silencing of Kcnq1 imprinted locus is a well-studied case that provides a model for involvement of NATs in allelic exclusion [214]. The antisense RNA, Kcnq1ot1, is essential for the silencing activity of Kcnq1 imprinting control region, which in turn controls the imprinting of a cluster of neighboring genes on chromosome 11 [213].

The effect of antisense RNA is not mediated through RNA interference (RNAi). NATs rather appear to recruit repressor complexes, turning chromatin into an inactive state. Suppressive chromatin modifications are spreading in both direction to the neighboring genes, similar to the X chromosome inactivation but with a limited penetrance. Kcnq1ot1 antisense RNA was shown to be involved in both establishment of transcriptional silencing as well as in maintenance of silencing through subsequent cell divisions [215]. Furthermore, hypomethylation of this region was documented in 40% of sporadic cases of Beckwith-Wiedemann syndrome (BWS) patients [59, 216].

Air transcript:

A region that contains three imprinted, maternally expressed protein-coding genes (Igf2r/Slc22a2/Slc22a3) on chromosome 6 has been shown to be controlled by a paternally expressed noncoding Air RNA [209]. Expression of Air (named Air, for antisense Igf2r RNA) correlates with repression of all three genes [217]. Air is a 108-kb unspliced and repeat-rich transcript, overlaps with just one of these genes in an antisense orientation [218]. However, Air is required for silencing of all three genes, which is likely through a similar mechanism shown for Kcnq1 imprinted locus.
**UBE3A-ATS transcript:**
In the case of ubiquitin ligase E3A (UBE3A), natural antisense transcript, **UBE3A-ATS**, is a very long (460-kb), spliced noncoding RNA located on chromosome 15. Paternal expression of **UBE3A-ATS** is responsible for monoallelic (maternal) expression of the **UBE3A** gene in the brain [60, 219, 220]. **UBE3A-ATS** lies within a highly complex locus containing several other imprinted genes including **ATP10C** and **UBE3A**, both show genomic imprinting [211, 221].

Association studies have revealed an important link to **ATP10C** in autistic patients [222, 223]. Prader-Willi syndrome and Angelman syndrome also result from the disturbance of **UBE3A** loci imprinted gene expression [224, 225]. These human disorders emphasize the importance of studying the role of NATs in genomic imprinting.

![Figure-3](image)

Figure-3 **NATs molecular mechanisms;** RNA-DNA interaction model; in which newly-formed RNA transcript from Watson’s strands can bind Crick’s strand DNA and guide methylation, demethylation, acetylation of DNA or chromatin and thereby modulate the chromatin architecture/epigenetic memory. Similar models have been proposed for imprinted genes as well as for mammalian X-chromosome inactivation. RNA-DNA binding also can cause alternative initiation or termination of the sense mRNA.

### 1.8.3 Involvement of NATs in DNA and chromatin modifications
NATs have been proposed to cause DNA modification at non-imprinted autosomal loci. Several different NAT-mediated DNA and chromatin modifications have been documented, which indicates complexity of this proposed function. Modifications of chromatin structure induced by noncoding RNA are suggested by the observation that many histone methyl transferase complexes lack DNA-binding domains but possess RNA-binding motifs [226].

Both trimethylated K4 on Histone H3 (a mark of active transcription) and trimethylated K27 on histone H3 (a mark of repressed chromatin) were reported to associate with ncRNAs, suggesting a scaffold model for the ncRNA in guiding chromatin modification [227].

### 1.8.3.1 Examples of DNA and chromatin modifications

**NAT for α-globulin gene:**
NAT-mediated DNA methylation is documented from studies of the human haemoglobin gene [56] where an antisense RNA for the α-globulin genes can induce DNA methylation leading to silencing of α-globulin gene.
**P15AS transcript:**
Another example is the silencing of the tumor suppressor gene *p15* via DNA methylation and heterochromatin formation induced by the NAT, *p15AS*. Epigenetic silencing of *p15* was not caused by transcriptional interference and was not Dicer-dependent, excluding intermediate small RNAs originating from *p15AS* as mediating the silencing effect [21].

**P21-AS transcript:** Low-copy promoter directed NAT of tumor suppressor gene, *p21*, mediates epigenetic modification of the sense promoter region [22]. Suppression of sense mRNA is directed by antisense-mediated induction of H3K27me3 at the p21 sense promoter region. Knockdown of antisense transcript caused relief of p21 promoter suppression, by loss of the H3K27me3, repressed chromatin mark [22].

**HOTAIR transcript:**
Histone modifications are well studied for homeobox (*HOX*) genes that encode key regulators of embryonic development [19]. Specifically, *HOX* antisense intergenic RNA (*HOTAIR*) is shown to repress the transcription in *trans* across a 40 kb region of the *HOXD* gene cluster by recruiting a regulatory complex that produces H3K27 trimethylated histone [19]. *HOTAIR* is a conserved polyadenylated and spliced long noncoding NAT, differentially expressed throughout the body [19]. Spatial expression of *HOTAIR* in various tissues induces epigenetic changes in the *HOXD* sense promoter region and accounts for the difference in *HOXD* sense gene expression [19].

**PR-AT transcript:**
NATs for progesterone receptor, *PR-AT1* and *PR-AT2*, are fully processed, spliced and polyadenylated and can mediate agRNA-induced gene activation or suppression. *PR-AT1* and *PR-AT2* bind to the argonaute (Ago) protein and providing a scaffold for suppressor or activator proteins to assemble in the progesterone receptor promoter region [25].

Another case of RNA-mediated modification of the promoter region is the inhibition of dihydrofolate reductase (*DHFR*) expression by RNA transcript synthesized from an upstream promoter, involving DNA-RNA triple-helix formation [228]. Yet another example is the control of elongation factor 1α (*EF1α*) expression through interactions between promoter-associated RNA transcript and agRNA directed to the promoter region [229]. These reports point out to the fact that there should be more examples of RNA-mediated regulation of transcriptional output at the promoter region.

### 1.8.4 Involvement of NATs in DNA elimination, recombination and genomic rearrangement:
Immunoglobulin production in B-lymphocytes and receptor selection of the T-lymphocytes depend on transcription from hyper-variable regions. To generate variability, T- and B-lymphocytes need a sophisticated silencing/recombination process and hypermutation in the variable regions of immunoglobulin and T cell receptor genes.

Activation-induced cytidine deaminase (AID), which deaminitates deoxycytidine to deoxyuridine in single-stranded DNA, is required for hypermutation process. Antisense transcription in the variable region makes the single-stranded DNA accessible for the AID [230, 231]. NATs are frequently observed in these regions.
[232-235] and may be involved in remodeling of chromatin structure in order to make the DNA sequence accessible for recombination.

1.8.5 Involvement of NATs in X chromosome inactivation

X chromosome inactivation accounts for balancing expression of the genes on the X chromosome in female mammals. Silencing of one of the two copies of the X chromosome is a process that ensures that females, with two X chromosomes, do not have twice as many X chromosome gene products (dosage compensation).

Two long non-protein coding genes are transcribed from the X chromosome inactivation center (XIC), XIST (X-inactive specific transcript) and TSIX (X (inactive)-specific transcript, antisense), and control the silencing of the X chromosome. The XIC is necessary and sufficient for X chromosome inactivation. Expression of the XIST, a 17 kb RNA transcript that triggers X inactivation [236], is regulated in cis by an antisense gene, TSIX, transcribed along the entire XIST gene, and it is reported to be involved in X chromosome inactivation [198, 237].

1.8.5.1 Tsix silences Xist by chromatin modification at promoter region

The Tsix gene encodes a large non-protein coding RNA, transcribed antisense to Xist [238]. Tsix is a negative regulator of Xist through a mechanism requiring overlapping transcription and blocks inactivation on the future active X chromosome (Xa).

It has been shown that Tsix silences Xist through modification of the chromatin structure in the Xist promoter region. Premature termination of Tsix transcription, by introduction of polyadenylation signal, abolishes the repressive chromatin configuration at the Xist promoter on the mutated X [237].

1.8.5.2 Xist is required for X inactivation

The XIST gene is the only gene expressed from the inactive X chromosome (Xi) but not from the active X chromosome [236]. X chromosomes that lack the Xist gene cannot be inactivated.

Prior to inactivation, both X chromosomes weakly express Xist RNA. During the inactivation process, the future Xa ceases to express Xist, whereas the future Xi dramatically increases Xist RNA production. On the future Xi, Xist RNA progressively coats the chromosome, spreading out from the XIC. The silencing of genes along Xi occurs soon after coating by Xist RNA. Therefore, dosage compensation occurs through heterochromatin formation along the inactive X-chromosome, which is coated by Xist.

1.8.5.3 Tsix is only express from active X chromosome

Like Xist, prior to inactivation, both X chromosomes weakly express Tsix RNA. At the onset of X inactivation, the future Xi ceases to express Tsix RNA (thereby increasing Xist expression), whereas Xa continues to express Tsix for several days.

Alterations in Tsix expression lead to distorted inactivation patterns, with deletions of the Tsix promoter resulting in preferential silencing of the mutant chromosome [239-241] and upregulation of Tsix expression inhibiting X chromosome inactivation [242, 243].

All of these well-documented reports about X chromosome inactivation, genomic imprinting and methylation or chromatin modifications of autosomal loci suggest that
an ample fraction of NATs are involved in RNA-DNA binding and control transcription through RNA-directed epigenetic modifications.

1.8.6 Alternative initiation, splicing, polyadenylation and termination:

NATs can form a triplex with DNA strands and cause alternative initiation and termination of the sense RNA [199]. Antisense RNA may also bind to the sense RNA and mask the splice sites and thereby change the balance between splice variants [10] (Figure-4).

1.8.6.1 Examples of NATs role in alternative splicing

In the case of N-Myc antisense, N-cym, the NAT forms an RNA duplex with the donor site of the first exon of N-Myc gene and thereby modulates the splice variants [52]. Thyroid hormone receptor alpha gene (TRα) is another example where antisense transcript called RevErbAα influences splicing of TRα1 and TRα2 mRNAs [66, 67, 244].

![Figure-4 NATs molecular mechanisms](image)

Figure-4 NATs molecular mechanisms; Nuclear sense-antisense RNA pairing, right after transcription, may inhibit sense RNA processing. NATs in this scenario can cover donor and acceptor splice sites to change the alternative splicing patterns. Additionally, by altering polyadenylation of the sense transcript, subsequently alterations of 3’UTR of sense transcript, NATs are able to affect stability and transport of mRNA. Yet another consequence of nuclear RNA duplex formation would be A to I editing by the ADAR enzyme, which can cause a change in amino acids, nuclear retention and/or degradation of hyper-edited transcript by inosine specific nucleases.

1.8.6.2 Alternative termination generates S-AS pairs

If two oppositely oriented neighboring genes undergo alternative polyadenylation and termination, they might form overlapping S-AS pattern. For instance, alternative polyadenylation of two oppositely oriented genes, the programmed cell death-2 (PDCD2) and TATA box binding protein (Tbp), induces formation of a S-AS pair [42]. Alternative termination of transcripts originating from the TP53Bp1/TUBGCP4 and CCNE2/FLJ20530 loci were also reported to produce S-AS pairs. These alternate transcripts originate from two neighboring convergent genes could potentially affect the cognate mRNA levels through S-AS pairing [6]. Therefore, it has been suggested that NATs may play a role in controlling the balance of transcripts differing in their 3’ terminus [14].
1.8.6.3 Overlapping S-AS pairs are not leakage of transcription
Given that at least half of all human genes encode transcripts with alternative 3’ termini [245], it is hypothesized that antisense transcription is a “leakage” of RNA transcription machinery originating from un-terminated transcription of downstream genes. However, evolutionary studies concerning genomic organization of neighboring genes were not in favor of the transcriptional leakage idea [14]. Moreover, unambiguous sequencing of human cell transcriptomes [124] did not support this idea. Indeed, there were poor correlations between the density of antisense-originated sequence tags and the density of sequence tags originated from the closest downstream genes [124].

1.8.7 Editing, nuclear retention and transport:

1.8.7.1 Transport
NATs can modulate mRNA nuclear transport by a mechanism, which involves nuclear duplex formation between S-AS pre-RNAs (Figure-4). For instance, the antisense RNA for the non-spliced p53 RNA binds to intron-1, preventing transport of p53 RNA to the cytoplasm [51].

1.8.7.2 Nuclear retention
Nuclear retention of the antisense RNA is commonly observed for NATs and could account for some antisense RNA-mediated regulation [201, 246]. NATs can in turn also be the subject of nuclear retention. Some cellular stressors, such as hypoxia, serum starvation and hydrogen peroxide can change the nuclear retention pattern of NATs and thereby alter the levels of their sense partners [16]. Nuclear retention of NATs is likely caused by direct interactions with nuclear proteins or other nuclear RNAs. Elucidation of these RNA-protein interactions causing nuclear retention is likely important for developing therapeutic interventions.

1.8.7.3 RNA editing
NATs have also been linked to mRNA editing [63, 247]. Interaction between the Drosophila 4f-rnp gene and its cis-NATs, sas-10, is reported to induce A-to-G editing in the overlapping region of 4f-rnp mRNA. Developmentally regulated expression of sas-10 overlapping transcript causes hyper-edition and subsequent degradation of the sense mRNA [247].

A-to-G RNA editing is induced by double stranded RNA (dsRNA) formation, in turn recruiting an enzyme called ADAR (adenosine deaminases that act on RNA), leading to deamination of targeted adenosine to inosine [248]. There are few documented reports on involvement of NATs in A-to-G RNA editing [50, 68, 247] and the overlapping regions of many S-AS pairs contain A-G mismatches [16, 63], suggesting possible RNA editing. However, there is a bioinformatics report arguing that the overlapping regions of the S-AS transcripts are not extensively edited, which may not exclude the possibility of alternative forms or degrees of editing of S-AS transcripts [249]. Nevertheless, editing is sometimes followed by the degradation of sense RNA via inosine-specific RNAse activity [201, 250].

1.8.8 Changes in mRNA stability and translation:

1.8.8.1 RNA stability
Cytoplasmic S-AS duplex formation can alter sense mRNA stability and translation efficiency (Figure-5). We previously showed that cytoplasmic S-AS duplexes are not normally subject to Dicer cleavage in mammalian cells and thus do not generally form
a basis for endogenous RNA interference [251]. However, duplex formation is still biologically relevant should antisense transcripts modulate the stability of their sense partners.

1.8.8.2  BACE1-AS is an example of alteration in mRNA stability

The overlapping region might affect mRNA stability by reducing mRNA decay whereby mRNA undergoes endo/exonucleolytic degradation by various RNAses. Indeed, we have recently demonstrated that BACE1-AS increases the stability of BACE1 mRNA, through a mechanism involving RNA duplex formation. We hypothesize that transient RNA duplex formation may alter the secondary or tertiary structure of BACE1 and thereby increase its stability [16].

![Figure-5 NATs molecular mechanisms](image)

**Figure-5 NATs molecular mechanisms:** Cytoplasmic S-AS RNA duplex formation can possibly have its own effects on sense mRNA. Hiding or exposing AU rich elements in sense transcript can affect RNA stability. Changes in the RNA secondary structure, upon binding to antisense RNA, can alter translation, sub-cellular localization, and accessibility of the RNA degradation machinery. According to this model, NATs can potentially “mask” miRNA-binding sites and release the miRNA-induced block of translation.

1.8.8.3  Other examples of alteration in mRNA stability

Antisense transcripts for inducible nitric oxide synthase, iNOS, an important gene in inflammatory diseases, increases the stability of iNOS mRNA [73]. Increased stability of iNOS mRNA, by iNOS AS, is basis of the observed concordant regulation of these transcripts. Enhancement of iNOS mRNA stability is mediated through interactions of antisense RNA molecule with the AU-rich element-binding HuR protein. HuR protein in turn may suppress RNA degradation by inhibiting deadenylase or exonuclease enzymes [73].

The NAT for basic fibroblast growth factor (bFGF) has been also reported to alter the stability of bFGF mRNA [47, 48].

The antisense transcript for hypoxia inducible factor, HIF-1α (aHIF) is yet another example of decrease in RNA stability induced by NATs. The aHIF ncRNA destabilizes one isoform of HIF-α mRNA and shifts the balance in favor of the other variant [252, 253]. Two isoforms of HIF, HIF-1α and HIF-2α, are both responsive to hypoxia but HIF-1α has a more ubiquitous expression pattern than HIF-2α. Prolonged hypoxia cause induction of antisense transcript, aHIF, which in turn binds to sense mRNA, leading to reduction of HIF-1α transcript. Importantly, destabilization takes place by exposing the AU rich elements (ARE) in HIF-1α mRNA following antisense binding to its 3’UTR [202, 252, 254].

An AU rich element (ARE) is a region in RNA transcript with frequent “A” and “U” nucleotides, such as “AUUUA”, that targets the RNA for degradation. Alterations in
HIF-1α secondary structure expose the ARE and make this RNA prone for degradation. Stabilization of mRNA by an opposite mechanism, covering the AU rich element, has been suggested for antisense transcript of the Bcl-2/IgH hybrid gene [64].

1.8.8.4 Translation block

Translational inhibition is yet another proposed function for some NATs as reported in the case of B cell maturation antigen transcript (BCMA), where over-expression of the antisense transcript has been reported to reduce sense protein but not sense mRNA level [62, 63]. BCMA, belong to the tumor necrosis factor receptor (TNF-R) family, and it has been shown to be controlled at the translational level by an antisense RNA transcript. BCMA antisense transcript is a fully processed RNA with an ORF; however, the discordant regulation of the sense protein takes place independent of antisense coding potentials [63].

Another well-documented case of translational inhibition is the NAT for PU.1 mRNA. Transcription factor PU.1 is an important regulator of hematopoiesis and suppressor of leukemia transformation. PU.1 mRNA translation is inhibited by a noncoding NAT [18]. Both sense and antisense transcripts of the PU.1 are co-regulated by an upstream regulatory element (URE). PU.1 antisense RNA is a polyadenylated transcript with a lower concentration but a longer half-life time than the sense PU.1 transcript and is equally distributed between cytoplasm and nucleus [18]. Processed antisense RNA in the cytoplasm may bind to the sense transcript and stall translation between initiation and elongation steps [18].

1.8.9 Cytoplasmic RNA duplex formation and “masking” miRNA binding sites:

We propose and do have some preliminary data supporting the idea that NATs can conceivably cover miRNA binding sites upon cytosolic RNA duplex formation (Figure-5). In contrast with plant miRNA, most animal miRNAs are predicted to have their binding site in the 3’ UTR of target mRNA [255]. Although most web tools for miRNA binding sites are designed to screen 3’ UTR regions of transcripts, there is no evidence that miRNA does not bind to the coding region. Binding of miRNA to the coding region of the mRNA, or even 5’ UTR, has been shown in plants and recently in animals [256-258]. We propose that one of the regulatory functions of NATs could occur by an ability of the antisense transcript to “mask” the miRNA binding site on the sense mRNA.

1.8.10 Formation of endogenous siRNA from double stranded RNA

1.8.10.1 RNA interference

RNA interference (RNAi) is a cellular surveillance mechanism that responds to exogenous double stranded RNA (dsRNA) molecules by destroying mRNAs containing sequences homologous to the dsRNA [259]. Duplex RNAs first cleaved by an RNase III enzyme, Dicer, to generate short interfering RNAs (siRNAs). SiRNAs include two strands, guide and passenger strands, and it become incorporated into a multiprotein RNA-induced silencing complex (RISC) that unwinds the helical structure of the siRNA duplex. The guide strand is retained in RISC, which then guides the entire complex to a target mRNA. Based on the complementarity between the target mRNA sequence and the guide strand sequence, RISC initiates the endonucleolytic cleavage or translational arrest of the target mRNA [260]. Processing of short RNA (>200 nt) from long and largely un-annotated nuclear RNAs were suggested from high resolution tilling array studies of human cell lines [126].
Gene regulation by endogenous siRNAs has been frequently observed in organisms possessing RNA-dependent RNA polymerase (RdRP) [261-263].

1.8.10.2 Endogenous siRNA
The presence of endogenous processing machinery for exogenous siRNA, which is mediated sequence-specific knockdown of targeted genes, implies that endogenous siRNA should exist. Endogenous siRNAs derived from a NAT were observed in Arabidopsis, where they regulate salt tolerance. Two types of siRNAs were shown to be generated from the overlapping region of Pyrroline-5-carboxylate dehydrogenase (P5CDH), a stress-related gene, and SRO5, a gene of unknown function[129]. Plant endogenous siRNAs are documented, derived from sense-antisense RNA duplex formation of several genes. For instance, Sho gene NAT transcript [264], SRLK/AtRAP NAT [265], and 64% of protein-coding cis-NAT in Arabidopsis are reported to generate endogenous siRNAs [266].

1.8.10.3 Mammalian endogenous siRNA
In mammals, where no RNA-dependent RNA polymerase (RdRP) activity has yet been found, biogenesis and function of endogenous siRNAs remain largely unknown [132]. Endogenous siRNAs derived from transposable elements and pseudogenes have been identified in mouse oocytes and cultured human cells [136, 137, 166-168]. Endogenous siRNAs originating from mRNAs and their corresponding NATs were recently identified in mouse oocytes [167] and human HepG2, liver carcinoma cells [168].

Both 25-27-nucleotide Piwi-interacting RNAs (piRNAs, Dicer independent) [147, 267] and approximately 21-nucleotide siRNAs originating from messenger mRNAs were found in mouse oocytes [167]. Transposable elements, inverted repeat structures, bidirectional transcription of sense-antisense genes (cis-NATs) and antisense transcripts from remote loci (trans-NATs) have been recognized as sources of dsRNAs and subsequent, Dicer-dependent, endogenous siRNA production [167, 168].

1.8.10.4 Intra-molecular vs. intermolecular dsRNA formation
Inverted repeats with an intra-molecular dsRNA structure are more likely to act as precursors of endogenous siRNAs as they can form more stable dsRNA structures required for their cleavage by Dicer. However, endo-siRNA has also been reported from intermolecular S-AS double stranded RNA formation. For instance, Kinesin family member 4A (KIF4A) and PDZ domain containing 11 (Pdzd11) are two genes located on opposite strands of X chromosome, where the two genes are orientated in a head-to-head manner. The overlapping transcripts in the Pdzd11/Kif4A locus generating endogenous siRNA derived from cis-NAT. Importantly, almost all of the endo-siRNAs in that locus (117 unique sequences) were derived from the overlapping region of the sense and antisense transcript, suggesting that these endo-siRNAs were produced from an intermolecular dsRNA formed between the oppositely oriented transcripts.

In Dicer mutants, levels of the siRNAs derived from the Pdzd11/Kif4 locus were decreased and both Pdzd11 and Kif4 mRNA levels were increased, suggesting that Pdzd11 and Kif4 expression is regulated by an endogenous siRNA pathway [167]. Endo-siRNAs may regulate both sense and antisense transcript levels. There is further evidence indicating that endogenous RNAi is also used as a defense mechanism to silence selfish genetic elements [168, 268].
1.8.10.5 *Endo-siRNA is not a prime consequence of S-AS duplex RNA formation*

Although NATs are abundant in mammalian cells, high throughput small RNA sequencing did not yield many endo-siRNAs mapping to NATs suggesting that RNAi is not the predominant mode of action of NATs.

Co-expression of NATs with their sense counterpart [251] as well as frequently observed concordant regulation of sense and antisense RNAs in many tissues and cell lines argue against endogenous siRNA being a prime mechanism of NAT-mediated regulation of gene expression. In addition, most co-expressed *cis*-NATs in Drosophila S2 cells did not generate endo-siRNAs [269, 270].

It is unclear how the majority of co-expressed NATs escape the endo-siRNA formation pathway. It is also not clear if there is an active selection for entry into the RNAi pathway and endo-siRNA formation.

Sub-cellular compartmentalization and coating of transcripts by proteins might act as potential barriers for the formation of dsRNA and subsequent endogenous siRNA. Nevertheless, biogenesis of endogenous siRNAs from some NATs has been documented. The reports mentioned above have revealed a biological role for endogenous siRNAs in mammals and show that organisms lacking RdRP activity can produce functional endogenous siRNAs from naturally occurring dsRNAs [166-168].

1.8.10.6 *Endogenous siRNA production in oocytes*

The production of dsRNAs by interactions between sense and antisense transcripts has been most frequently reported in oocytes suggesting that dsRNA formation requires a unique environment found only in oocytes. Oocytes substantially lack a protein kinase R response (a dsRNA-induced general translational repression pathway).

1.9 **CLINICAL SIGNIFICANCE**

Several important physiological roles have been proposed for NATs, such as balancing gene expression during spermatogenesis [189] cardiac gene regulation [271], cardiac [94, 272] and skeletal [93] myosin gene organization, and regulation of circadian clock function [96].

NATs involvement in various disorders has also been reported [273, 274]. Significant association with complex human disorders has been demonstrated for a number of NATs (Table-4).

1.9.1 **NATs associated with cancers**

1.9.1.1 **NATs for tumor suppressor genes**

Tumor suppressor genes are frequently suppressed in cancer. An elegant study by Yu *et al.* has documented the presence of NATs for many tumor suppressor genes and they have identified NATs for each one of 21 well-known tumor suppressor genes [21].

Specifically, *p15*, a tumor suppressor gene involved in a wide variety of tumors including leukemia, melanoma, glioma, lung cancers and bladder carcinomas, is epigenetically controlled by its NAT, *p15*AS. The expression of *p15* S-AS transcripts has an inverse correlation (discordant regulation) in leukemic patients where 70% of
leukemic patients showed increased expression of p15AS and reduced expression of p15 mRNA [21].

Another tumor suppressor gene, p21, is also shown to negatively regulated by an antisense RNA molecule, p21-AS [22]. Imbalance expression of antisense p21, p21-AS, can potentially suppress p21 expression, leading to tumor growth.

Table 4: Human disease related natural antisense transcripts

<table>
<thead>
<tr>
<th>Sense-Antisense pairs</th>
<th>Human Disease</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>BACE1 &amp; BACE1-AS</td>
<td>Alzheimer’s Disease</td>
<td>[275]</td>
</tr>
<tr>
<td>FMR1, FMR4 &amp; ASFMR1</td>
<td>Fragile X mental retardation, Fragile X-associated tremor and ataxia syndrome (FXTAS)</td>
<td>[31, 32]</td>
</tr>
<tr>
<td>PINK1 &amp; naPINK1</td>
<td>Parkinson disease, Mitochondrial disorders</td>
<td>[30]</td>
</tr>
<tr>
<td>C6orf37 &amp; C6orf37OS</td>
<td>Diffuse panbronchiolitis</td>
<td>[276]</td>
</tr>
<tr>
<td>FGF-2 &amp; GFG</td>
<td>Endometriosis, Carcinogenic progression</td>
<td>[277-279]</td>
</tr>
<tr>
<td>HIF-1α &amp; aHIF</td>
<td>Poor prognosis marker in breast and renal cancer</td>
<td>[54]</td>
</tr>
<tr>
<td>Survivin &amp; EPR-1</td>
<td>Colon cancer</td>
<td>[55]</td>
</tr>
<tr>
<td>WT1 and WT1-AS</td>
<td>Wilms’ tumor</td>
<td>[27, 280]</td>
</tr>
<tr>
<td>α-globulin &amp; LUC7L</td>
<td>α- Thalassemia</td>
<td>[56]</td>
</tr>
<tr>
<td>KvLQT1</td>
<td>Beckwith-Wiedemann syndrome</td>
<td>[59]</td>
</tr>
<tr>
<td>SNURF-SNRPN &amp; UBE3A</td>
<td>Prader-Willi and Angelman syndrom</td>
<td>[60]</td>
</tr>
<tr>
<td>Bcl-2 &amp; IgH</td>
<td>Follicular B-cell lymphoma</td>
<td>[64]</td>
</tr>
<tr>
<td>Zeb2 &amp; Zeb2 NAT</td>
<td>Epithelial- mesenchymal transition</td>
<td>[37, 38]</td>
</tr>
<tr>
<td>RMRP</td>
<td>Cartilage-hair hypoplasia</td>
<td>[281]</td>
</tr>
</tbody>
</table>

### 1.9.1.2 NATs reported in various cancers:

**aHIF transcript:**
The antisense transcript for hypoxia inducible factor, aHIF, has been reported as a marker of poor prognosis in human breast cancer and shown to be inversely related to disease free survival of the patient as well as proliferation of cancerous cells [253].

**FGF-2/GFG transcripts:**
FGF-2/GFG is another well-characterized S-AS pair, suggested to be linked to tumor progression [282]. The sense protein fibroblast growth factor2, FGF-2, induces growth and proliferation. The FGF-2 antisense transcript (GFG) regulates the levels of the FGF-2 sense mRNA discordantly, *i.e.* sense and antisense transcripts are inversely expressed in many tissues and cell types. Reduced levels of the NAT increase FGF-2 mRNA and promotes carcinogenic progression [283, 284] or implantation of ectopic tissues such as in endometriosis [277]. Reduction of FGF-AS is a negative prognostic factor for esophageal adenocarcinoma [278].
FGF-AS (GFG) also inhibits cell cycle-dependent nuclear accumulation of FGF-2, and this is associated with a marked delay in S-phase progression leading to tumor suppression. This may play a significant functional role in the regulation of FGF-2 dependent cell proliferation in C6 glioma cells [279] as well as lymphoid and myeloid tumor cells [285].

**EPR-1/Survivin transcripts and other examples**

The effector cell protease receptor-1 (EPR-1) cDNA, an antisense transcript for Surviving gene, has been shown *in vitro* to reduce tumor growth potential and enhance the response to anti-cancer drugs [55]. The NAT for the Bcl-2/IgH hybrid gene has been proposed to underlie follicular B-cell leukemia [64]. Antisense transcript for zinc finger homeobox 1b, Zeb2-AS, has been linked to colon cancer and Hirschsprung’s disease. Specifically, antisense RNA, Zeb2-AS, induces Zeb2 protein up-regulation and subsequently induces epithelial-mesenchymal transition [37].

1.9.2 NAT involvement in certain forms of anemia

In one inherited forms of anemia, α-thalassemia, a NAT has been reported to cause silencing of the α-globulin gene via methylation [56]. Tufarelli *et al.* found that a deletion in the globin gene locus of thalassemic patients relocates the constitutively active LUC7L gene 300 nucleotides downstream of alpha-2 globin (HBA2) gene. HBA2 encode hemoglobin alpha chain and antisense RNA causing promoter methylation and transcriptional silencing of HBA2 gene. Reduction in hemoglobin alpha chain, which is major constitute of adult hemoglobin, causes anemia.

A mouse model for genomic rearrangement (relocation of LUC7L) recapitulated the α-thalassemia disease phenotype and confirmed the role of *cis*-NAT in early developmental CpG island methylation [56].

1.9.3 NAT-related disorders linked to imprinting

Alterations in NAT expression patterns are commonly linked to imprinted gene disorders. For example, in both Prader-Willi and Angelman syndromes the SNURF-SNRPN sense and UBE3A antisense RNA are implicated in the disease pathophysiology [208]. Additionally, antisense RNA, Kcnq1ot1, is shown to be involved in sporadic cases of Beckwith-Wiedemann syndrome (BWS) patients [59, 216].

1.9.4 Involvement of NATs in neurological disorders

1.9.4.1 FMR4 transcript:

Fragile X syndrome, the most common cause of mental retardation, is caused by the expansion of CGG trinucleotide repeats (>200 repeats) in the 5’ UTR of the fragile X mental retardation 1 (FMR1) gene [286]. On the other hand, individuals with 55–200 repeats are premutation carriers and generally express higher levels of FMR1 mRNA than normal individuals with less than 55 repeats and may develop a clinical condition termed fragile X tremor and ataxia syndrome (FXTAS) [287].

We reported recently the discovery of FMR4, a long ncRNA (2.4 kb), resides upstream of the FMR1 gene, which may shares a bidirectional promoter with the FMR1 gene [32]. The antisense transcript for FMR1 (ASFMR1) has also been reported recently, which overlaps the CGG repeat region of FMR1 and is transcribed in the opposite direction [31]. The CGG expansion appears to affect transcription in
both directions as FMR4 and ASFMR1 similar to FMR1, are silenced in fragile X patients and up-regulated in permutation carriers [31, 32].

The similar expression pattern of FMR4, ASFMR1 and FMR1, suggests that FMR4 and ASFMR1 may also contribute to aspects of the clinical presentation of fragile X syndrome and FXTAS arguing that these disorders should not continue to be considered single gene disorders.

1.9.4.2 naPINK1 transcript:
Mutations in the PTEN induced putative kinase 1 (PINK1) are implicated in early-onset Parkinson's disease. PINK1 is expressed abundantly in mitochondria-dense tissues, such as skeletal muscle. We characterized a novel NAT at the PINK1 locus (naPINK1) [30]. The naPINK1 transcript induces alteration in PINK1 splice variant in neuronal cell lines.

Moreover, the PINK1 and naPINK1 transcripts display discordant regulation during in vivo exercise-induced mitochondrial biogenesis, suggesting that the noncoding naPINK1 might play a role in Parkinson's disease.

1.9.4.3 BACE1-AS transcript:
We recently identified BACE1-AS as a noncoding NAT for /secretase-1 (BACE1), a crucial enzyme in β-amyloid biosynthesis related to Alzheimer’s disease pathophysiology [16]. We showed that BACE1-AS up-regulates BACE1 mRNA and protein in vitro and in vivo. Upon exposure to various cell stressors including amyloid-β 1–42 (Aβ 1–42), expression of BACE1-AS becomes elevated, increasing BACE1 mRNA stability and generating additional Aβ 1–42 through a post-transcriptional feed-forward mechanism. We showed that BACE1-AS concentrations are significantly elevated in Alzheimer’s disease patients [16].

1.10 ALZHEIMER’S DISEASE

1.10.1 Alzheimer’s disease with an enormous unmet medical need
Alzheimer’s disease is a devastating age-related neurodegenerative disorder characterized by progressive impairment of cognition and short-term memory. It is the most common form of dementia affecting 5% of adults over 65 years. More than 4.5 million people in the United States alone currently suffer from Alzheimer’s disease [288]. Worldwide, approximately 18 million people suffer from Alzheimer’s disease, with this number projected to increase to approximately 34 million people by the year 2025 [288]. Direct and indirect annual costs of caring for individuals with Alzheimer’s disease are at least $100 billion in the US alone [289]. In view of our ageing society the number of patients, as well as the economical and social impact, is expected to grow dramatically in the future.

Currently available medications appear to be able to produce moderate symptomatic benefits but not to stop disease progression. Indeed, most therapeutic entities for the treatment of Alzheimer’s disease are designed to alleviate Alzheimer’s disease-related symptomatology, and not block the mechanisms that underlie Alzheimer’s disease pathology. Until very recently the majority of compounds considered candidate drugs for treatment of Alzheimer’s disease were designed to modulate the synthesis, release, or degradation of various brain neurotransmitters, for review see [290]. Thus, an improved strategy for developing novel agents for treatment of Alzheimer’s disease
should include therapeutic targets located directly within cellular pathways involved in the neuronal dysfunction and death that drives Alzheimer’s disease pathology, potentially slowing or halting the progression of Alzheimer’s disease.

1.10.2 Amyloid hypothesis related to Alzheimer’s disease

The sequential cleavage of the Amyloid precursor protein (APP) by the β-site cleaving enzyme (BACE1) followed by γ-secretase initiates the “Amyloid Cascade” which is central to Alzheimer’s disease pathophysiology [291-293].

Controversy still exists in the link between amyloid pathway and Alzheimer’s disease and in the precedence of events leading to Alzheimer’s disease; however, deposition of myloid-β 1-42 (Aβ 1-42) into senile plaques is a proven feature of Alzheimer’s disease neuropathology. Moreover, a recent study has reported that Amyloid-beta protein dimers isolated directly from human Alzheimer’s brains impair synaptic plasticity and memory in the rat brain [294]. BACE1 is essential for Aβ 1-42 biosynthesis [295, 296], which is the main component of senile plaques found in Alzheimer’s disease brain.

Oligomers of Aβ 1-42 produced by BACE1 influence virtually all the other observed components of Alzheimer’s disease pathophysiology, such as mitochondrial function [297, 298], tau hyperphosphorylation [299], NMDAR endocytosis, excessive calcium influx [300, 301], synaptic dysfunction [302], neuronal stress [303], and apoptosis [304]. Numerous studies have documented up-regulation of BACE1 concentrations in the brain of Alzheimer’s disease patients compared with normal controls [305-310]. Furthermore, the amyloid pathology usually observed in transgenic mice overexpressing mutant human APP which is completely ameliorated when these mice are crossed with BACE1 knockout mice [311].

1.10.2.1 Alteration in APP cleavage products in late-onset Alzheimer’s disease

Alterations in the processing and clearance of proteolytic products of the APP, such as C-terminal fragments and Aβ peptides, likely play a key role in the pathogenesis of late-onset Alzheimer’s disease [312-314]. Indeed, recent studies suggest that Aβ causes neuritic dystrophy and interferes with mechanisms of synaptic plasticity such as long-term potentiation (LTP) [315].

Transgenic mouse models of Alzheimer’s disease have been generated that express familial Alzheimer’s disease mutations in the APP gene, such as the Tg19959 mice that overexpress a doubly mutated human APP [316]. Such Alzheimer’s disease mouse models recapitulate many aspects of human Alzheimer’s disease pathology, including Aβ plaques in hippocampal and cortical regions [317-319].

The precise relationship between Aβ peptides and the emergence of Alzheimer’s disease-related cognitive deficits remains unclear. Nevertheless, Aβ-associated memory deficits in mice have been observed in a range of hippocampus-dependent place learning paradigms, including the Morris water maze, Y-maze, radial arm maze and Barnes maze tasks [311, 320-327]. Based on these and similar observations, it is likely that Aβ production is closely associated with the physiological and cognitive impairment so characteristic of Alzheimer’s disease [314, 328, 329].
1.10.3 APP cleavage and generation of Aβ peptides

APP cleavage and its enzymatic cleavage byproducts are depicted in figure-6. In this model APP C-terminal fragments and Aβ peptides are generated by sequential cleavage of the transmembrane APP at the β and γ locations. The initial cleavage, which generates C-terminal APP fragments, is accomplished by BACE1 [330-332]. Subsequently, the C-terminal APP is further cleaved by a second protease known as γ-secretase [333]. Thus, β- and γ-secretase activities are required for the production of the Aβ peptides.

Figure 6: (A) One structural model of BACE1; (B) schematic model for APP cleavage and its byproduct of enzymatic cleavage; (C) amyloid plaques in Alzheimer’s disease brain.

1.10.4 BACE1 expression is tightly regulated

In addition to its role in Alzheimer’s disease pathology, BACE1 performs several important functions in mammalian brain since the ablation of BACE1, results in a range of deficits, including memory loss [334], emotional disturbances [290], myelination defects in peripheral nerves [335, 336], and loss of synaptic plasticity [290]. The subtle but critical boundaries between BACE1 physiology and pathology indicate that BACE1 expression and activity levels must be tightly regulated, both temporally and spatially. Such a well articulated regulation machinery would allow the enzyme to perform its important physiological functions while avoiding serious consequences of deregulation such as Aβ 1-42 accumulation.

1.10.5 Unexpected complexity in the BACE1 gene locus

We have recently identified a noncoding cis-antisense transcript to BACE1. We have termed this transcript, BACE1-AS, and shown that it is highly conserved. We have also conducted a bioinformatics search for miRNA binding sites in BACE1 mRNA and predicted the presence of a binding site for miR-485-5p in the sixth exon of
BACE1 mRNA, i.e. exactly on the overlapping regions of the BACE1-AS and BACE1 transcripts. We postulated that at least part of regulatory function of BACE1-AS may be “masking” of the miR-485-5p binding site and thereby blocking the inhibitory effects of this miRNA on BACE-1 translation.

1.10.6 **BACE1-AS has potentials as an Alzheimer’s disease biomarkers**

The increasing prevalence of Alzheimer’s disease and the devastating consequences of late-life dementia motivate the drive to develop diagnostic biomarkers to reliably identify the pathology associated with this disorder prior to the start of manifestations.

Therapeutic approaches and novel medications targeting the presumed underlying pathogenic mechanisms need to be tested on clearly diagnosed Alzheimer’s disease patients as early as possible. Accessible, sensitive, and specific biomarkers need to be also employed to monitor the response to treatment.

Of great value as diagnostic tools, ideal biomarkers for the disease should be directed toward basic neuropathological pathways and detect a fundamental characteristic neuropathology, instead of secondary markers prone to inform non-specific changes.

Some methods show promise as diagnostic tools for the disease, including plasma signaling proteins [337] and CSF protein panel [338]. Although, neuroimaging, CT and MRI, plays an important part in the diagnosis of Alzheimer’s disease to exclude alternative causes of dementia, such as brain tumor and subdural haematoma, but the overlap with normal ageing and other dementias is too large to have any diagnostic value. However, at present, none of these are recommended as routine diagnosis methods for Alzheimer’s disease. Blood levels of Aβ 1-42 were found to be significantly elevated in the family of late onset Alzheimer’s disease patients, compared to non-blood relatives, such as spouses [339].

Peripheral blood cells have already shown promising data as reporters of neural pathology. Gene expression profiles from lymphoblastoid cell lines of autistic twins have yielded a series of informative biomarker candidates [340]. Gene expression studies of patient samples with bipolar disorder have yielded similar results [341, 342].

Considering marked up-regulation of BACE1-AS in Alzheimer’s disease patients, which is significantly higher than BACE1 changes, we postulated that BACE1-AS might also be useful as a readily accessible peripheral biomarker candidate, presumably ultimately as part of a signature of biomarkers. The stress response features of BACE1-AS further enhance its candidacy as a peripheral biomarker of the early stages of CNS pathology. We have tested BACE1-AS ratio to β- Actin as a molecular diagnostic marker of Alzheimer’s disease.

1.10.6.1 **BACE1-AS level is high in subjects with Alzheimer’s disease**

In fact, BACE1-AS is markedly elevated in the brain of Alzheimer’s disease patients in each of two independent sets of Alzheimer’s disease brain samples tested so far.

A peripheral blood mononuclear cell (PBMC) is a blood cell having a round nucleus, such as a lymphocyte or a monocyte. BACE1-AS is also readily detectable in PBMC which make it suitable for diagnostic purposes. BACE-AS alteration in PBMC of subjects with Alzheimer’s disease and elderly patients with mild cognitive impairment (MCI) might be proven beneficial as a diagnostic biomarker.
1.11 CONCLUSION

In the light of this present thesis, we know that natural antisense transcripts are functional RNA molecules that control sense mRNA expression at different levels.

Most NAT are noncoding RNA, NATs help to mediate sense gene expression in response to a variety of environmental stimuli and to keep tight regulation of sense protein expression (allowing proteins to perform their physiological functions while avoiding the serious consequences of over or under expression).

We present several examples of functional NATs to show multilayer involvement of these molecules in regulation of gene expression.

Although protein synthesis was the first assigned function for RNA molecules, it is become increasingly evident that the more pervasive function of RNA molecules is a regulatory one. This hypothesis is considerably supported with the content of this current thesis in which we summarize proposed regulatory functions of naturally occurring antisense transcripts. Considering other reported functional long ncRNA (macroRNA), and small regulatory RNA (such as miRNA, piRNA, rasiRNA) and their enormous physiological impact it is feasible to claim that the more frequent function of RNA molecules is a regulatory role, which is far greater than their assigned functions in protein synthesis as messenger, transfer and ribosomal RNA
2 PRESENT INVESTIGATION

2.1 SCIENTIFIC BACKGROUND

I received my medical doctorate (M.D.) degree in 1994 from Shiraz University of Medical Sciences, Iran. Initially, I entered the public health sector and worked on practical public health issues like preventing HIV transmission among prisoners. I was successful in improving the living conditions of my target groups, establishing higher health standards. Those dedicated efforts did not satisfy my eagerness to study Neuroscience, so I moved to the Karolinska Institutet in Sweden to join the laboratory of Professor Claes Wahlestedt, where I have been deeply involved in projects related to natural antisense transcripts, with a particular focus on their involvement in neurological disorders. I have been able to show a critical role for these long non-protein-coding RNA transcripts in regulation of major neurological disorders, like Parkinson’s disease, Fragile-X mental retardation and Alzheimer’s disease.

2.2 FANTOM PROJECT REVEALED HIGH ABUNDANCE OF NATS

I made contributions to the FANTOM-3 project (Functional annotation of mammalian transcriptomes), which was the largest transcriptome profiling effort in the world and resulted in the discovery of NATs for more than 70% of transcription units (Paper V).

This work was published in “Science Magazine” by this title: “Antisense transcription in the mammalian transcriptome” Science 309:1564-6, 2005.

We showed that antisense transcription is a common feature for many transcriptional units in mammalian transcriptome. We also presented experimental evidence that perturbation of an antisense RNA can alter the expression of sense messenger RNAs, suggesting that antisense transcription contributes to control of transcriptional outputs in mammals.

Antisense transcripts were previously suggested for up to 20% of transcripts. The FANTOM-3 consortium conducted a large-scale cDNA sequencing approach, which revealed antisense transcription for up to 72% of transcriptional units. Interestingly, imprinted loci showed much higher rate of antisense transcription than previous reports.

Table-5  Total number of sense-antisense (S-AS) overlapping transcriptional units (TU)

<table>
<thead>
<tr>
<th></th>
<th>Total TU</th>
<th>S-AS overlapping</th>
<th>Percent of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coding TU</td>
<td>20,714</td>
<td>18,021</td>
<td>87%</td>
</tr>
<tr>
<td>Noncoding TU</td>
<td>22,839</td>
<td>13,401</td>
<td>59%</td>
</tr>
<tr>
<td>Total</td>
<td>43,553</td>
<td>31,422</td>
<td>72%</td>
</tr>
</tbody>
</table>
In this study we clustered transcripts into the transcriptional units (TUs), which we defined as all the EST or mRNA in one direction, with at least one nucleotide exonic overlap. We found more than 50,000 overlapping pairs, grouped into about 30,000 non-redundant different overlapping regions in close to 10,000 TU pairs. We showed that 4,520 TU pairs contain full-length transcripts, which forms exon-overlapping sense-antisense pair. There were additional 4129 TU pairs, in the sense-antisense direction, without any apparent exon overlapping regions. Although conservative, the combined NAT prediction were 1.5- to 2-fold greater than that from previous studies of mouse [12] and human [6]. Our finding is to this time considered the largest reported collection of NATs (Table-5).

Head-to-head NATs (see NAT classification), were slightly more frequent than 3'UTR overlapping partners, which suggest an important role for NATs in control of transcriptional output (Table-6).

Table-6: Percent of sense-antisense overlapping transcripts based on coding properties and transcript orientation

<table>
<thead>
<tr>
<th>Transcription Unit</th>
<th>Head-to-head (5')</th>
<th>Tail-to-tail (3')</th>
<th>Full</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coding-coding</td>
<td>37%</td>
<td>45%</td>
<td>19%</td>
</tr>
<tr>
<td>Coding-noncoding</td>
<td>36%</td>
<td>27%</td>
<td>37%</td>
</tr>
<tr>
<td>Noncoding-noncoding</td>
<td>29%</td>
<td>34%</td>
<td>36%</td>
</tr>
<tr>
<td>Total</td>
<td>36%</td>
<td>34%</td>
<td>30%</td>
</tr>
</tbody>
</table>

Interestingly, overlapping sense-antisense pairs, in our study, were not evenly distributed. Chromosome X showed the fewest bidirectional pairs, which could be related to monallelic inactivation. Another finding was the evidence for antisense transcription in more than 80% of imprinted loci, emphasizing the proposed role for NATs in genomic imprinting.

Expression profiling reveals frequent concordant regulation of sense/antisense pairs. Among the functionally validated NATs, we found both types of discordant and concordant regulation, consistent with the other published works. For instance, two protein coding S-AS transcripts, Ddx39 and CD97, showed a reciprocal regulation pattern. We observed that siRNA-mediated knockdown of Ddx39 transcript cause up-regulation of CD97, which is a G protein-coupled receptor.

Global transcriptome analysis of the mouse by FANTOM-3 revealed that antisense transcription is widespread in the mammalian genome. This study provides evidence that a large proportion of the genome can produce transcripts from both strands of a gene locus, and those NATs commonly link neighboring genes in complex loci into chains of linked transcriptional units.
2.3 MECHANISTIC ASPECTS OF NATS

Next, I showed that NAT-mediated regulation of gene expression predominantly occurs through a pathway independent of Dicer associated RNAi, (Paper IV). This work was published in “Genome Biology” by this title: “RNA interference is not involved in natural antisense mediated regulation of gene expression in mammals” Genome Biology 7(5):R38, 2006 [251].

We investigated functional properties of natural antisense transcripts (NATs) to explore the mechanism of reciprocal S-AS regulation, and to investigate the possible involvement of endogenous RNA interference (RNAi) in S-AS interactions. We selected two examples from each coding and noncoding NATs to study effects of antisense transcript modulation on corresponding gene expression. Particularly, we investigated the possible involvement of endogenous RNA interference (RNAi) in S-AS interactions.

The simplistic assumption, at the time, was that all NATs should produce endo-siRNA and therefore should down regulate the corresponding sense mRNA levels. We selected two functional NATs and investigated the presence of endogenous siRNA in cells originating from the overlapping region of S-AS transcripts. We found that at least in our cellular model and for our selected candidates, RNA interference is not involved in antisense-mediated regulation of the sense mRNA. This finding argued against the simplistic assumption of a negative regulatory role of antisense transcription. In the light of this study, researchers in antisense field start to realize that NATs can be functional even if they do not cause down-regulation of corresponding sense mRNA partner.

We examined the mechanism of S-AS RNA base pairing, using thymidylate synthase and hypoxia inducible factor-1α as primary examples of endogenous genes with coding and noncoding NAT partners, respectively. We provided direct evidence against activation of RNA interference (RNAi) and generation of endogenous siRNA. Our data demonstrated that NAT regulation of gene expression occurs through a pathway independent of Dicer associated RNAi.

Endogenous siRNA originating from overlapping region of NATs have recently reported for a few NATs. Endo-siRNA or piRNA found for few genes (17 NATs [134]), which is a very small fraction of NATs in any given cell. Endo-siRNA, in many cases, originated from intra-molecular hairpin formation instead of intermolecular S-AS duplex formation. However, Watanabe et al. shown that Pdzd11/Kif4 loci can generate endo-siRNA from intermolecular S-AS duplex RNA.

Majority of reported cases of endo-siRNA are originated from pseudogenes, transposable elements or repeat regions [133, 135-137]. These regions are shown to have different characteristics than the rest of genome and some reports state that endo-siRNAs, in these cases, used as a defense mechanism to silence selfish genetic elements [135, 343].

It is not clear that how the majority of co-expressed NATs escape the endo-siRNA formation pathway. Most co-expressed NATs in Drosophila did not generate endo-
siRNAs [269, 270]. In our study, we tested endogenous expression of S-AS transcripts and found that both transcripts co-exist in a single cell, but do not generate endo-siRNA. We have also tested overexpression of both S-AS transcripts and found that, even when we have a very high concentrations of both S-AS transcripts, there was not any detectable endogenous siRNA’s corresponding to the overlapping region. Our finding showed that production of endo-siRNA is not a prime route of NAT-mediated regulatory function.

Furthermore, frequently observed positive regulation of sense and antisense RNAs in many tissues and cell lines are against endogenous siRNA as a prime mechanism of NATs mediated regulation of gene expression. Examples of positive regulations, which are explained in the introduction section, include but not limited to aHIF1 [174], asEPO-R [33] BACE1-AS [275] and Zeb2-AS [37]. Therefore, endo-siRNA formation is a valid regulatory mechanism for NAT-mediated regulation of the sense transcripts, but it is not definitely the predominant mechanism.

2.4 NATURAL ANTISENSE TRANSCRIPT FOR PINK1

I selected several interesting NATs identified by the FANTOM-3 consortium for detailed studies of their role in neurological disorders. In these studies I have characterized a NAT for the PINK1 gene, implicated in early-onset Parkinson's disease, (Paper III). This work was published in “BMC Genomics” by this title: “The human PINK1 locus is regulated in vivo by a non-coding natural antisense RNA during modulation of mitochondrial function.” BMC Genomics 15:8:74, 2007.

Mutations in the PTEN induced putative kinase 1 (PINK1) are implicated in early-onset Parkinson's disease. PINK1 is expressed abundantly in mitochondria rich tissues, such as skeletal muscle, where it plays a critical role determining mitochondrial structural integrity in Drosophila. We characterize a novel splice variant of PINK1 (svPINK1) that is homologous to the C-terminus regulatory domain of the protein kinase. Moreover, we describe a human specific noncoding antisense expressed at the PINK1 locus (naPINK1). The observation of regulation of svPINK1 and naPINK1 during in vivo mitochondrial biogenesis was confirmed using RNAi, where selective targeting of naPINK1 results in loss of the PINK1 splice variant in neuronal cell lines. Our report provided description of a novel ncRNA, which is the in vivo regulator of PINK1 gene and could potentially be involved in Parkinson's disease pathophysiology.

2.5 FMR4 NONCODING RNA INVOLVED IN FRAGILE X MENTAL RETARDATION

Furthermore, I discovered FMR4, a ncRNA transcript that shares a bidirectional promoter with FMR1, a gene implicated in the fragile-X syndrome, (Paper II). This work was published in “PLoS ONE” by this title: “A Novel RNA Transcript with Antiapoptotic Function Is Silenced in Fragile X Syndrome” PLoS ONE 3(1):e1486, 2008.
In this publication we described FMR4, a ncRNA transcript (2.4 kb) that resides upstream and likely shares a bidirectional promoter with FMR1, a gene heavily implicated in the fragile-X syndrome.

We showed that FMR4 is a product of RNA polymerase II and has a similar half-life to FMR1. The CGG expansion in the 5'UTR of FMR1 appears to affect transcription in both directions as we found FMR4, similar to FMR1, to be silenced in fragile X patients and up-regulated in premutation carriers.

Knockdown of FMR4 by several siRNAs did not affect FMR1 expression, nor vice versa, suggesting that FMR4 is not a direct regulatory transcript for FMR1. However, FMR4 markedly affected human cell proliferation in vitro; siRNAs knockdown of FMR4 resulted in alterations in the cell cycle and increased apoptosis, while the overexpression of FMR4 caused an increase in cell proliferation.

2.6 BACE1-AS INVOLVED IN ALZHEIMER’S DISEASE

I recently identified a conserved noncoding antisense transcript for β-secretase-1 (BACE1), a critical enzyme in Alzheimer’s disease pathophysiology [291]. I generated compelling data indicating that the BACE1-antisense transcript (BACE1-AS) concordantly regulates BACE1 expression (Paper I). This work was published in “Nature Medicine” by this title: “A noncoding RNA is elevated in Alzheimer’s disease and drives rapid feed-forward regulation of β-secretase expression” Nature Medicine 14(7):723-30, 2008.

We showed that BACE1-AS levels are dramatically up-regulated in several brain regions of individuals with Alzheimer’s disease. We proposed a model in which exposure to various cell stressors results in BACE1-AS levels becoming elevated, increasing BACE1 mRNA stability and generating additional Amyloid-beta 1–42 (Aβ 1–42) through a post-transcriptional feed-forward mechanism.

We demonstrated a putative role for a noncoding RNA transcript in Alzheimer’s disease pathophysiology. We showed that BACE1-AS transcript is indeed elevated in human Alzheimer’s disease brain samples and that this regulatory RNA contributes to a feed-forward mechanism that underlies β-amyloid formation in the Alzheimer’s disease senile plaques.

The BACE1-antisense transcript (BACE1-AS) regulates BACE1 mRNA and subsequently BACE1 protein expression in vitro and in vivo. Upon exposure to various cell stressors including amyloid-b 1–42 (Aβ 1–42), expression of BACE1-AS becomes elevated, increasing BACE1 mRNA stability and generating additional Aβ 1–42 through a post-transcriptional feed-forward mechanism. BACE1-AS concentrations were elevated in subjects with Alzheimer’s disease and in amyloid precursor protein transgenic mice.

Our finding showed that BACE1 mRNA expression is under the control of a regulatory noncoding RNA that may drive Alzheimer’s disease-associated pathophysiology. In summary, we reported that a long noncoding RNA is directly implicated in the increased abundance of Ab 1–42 in Alzheimer’s disease.
2.7 FUNCTIONAL RELEVANCE OF NATs

Next, we planned to find out if these scattered reports on functional NATs are exceptions or they are representing general rules. Then we took all the reported conserved NATs between human and mouse [188] and designed siRNA for the entire collection.

We performed a cell viability screening, using this aforementioned library of siRNA. Large-scale screening with siRNA targeting 794 conserved natural antisense transcripts (NATs) revealed a potential role for a considerable number of these NATs in regulating cell viability and proliferation. We prepared this work for publication in a manuscript with this title: “A High-Throughput RNAi Screen Reveals Widespread Biological Function For Mammalian Noncoding Antisense Transcripts”

In this study we focused only on well-conserved S-AS transcript pairs and showed functional role for these RNA molecules. Particularly, we showed a potential role for a significant number of these NATs in cell viability and proliferation.

It is worth noting that only a fraction of all NATs were expressed in our cellular model and we have only tested cell viability phenotype. Therefore, considering the whole phenotypic space in various cell line, we argue that perhaps all of these RNA transcripts are functional elements. Collectively, our screening revealed a prominent role for NATs in basic cellular pathways, much more prevalent than previously appreciated.
3 PROSPECTIVE

3.1 FUTURE RESEARCH GOALS

My current projects focus on the expanding role of ncRNAs in the regulation of nervous system gene expression, especially in the delicate balance between physiological and pathological information flow that could define the onset of chronic disease.

I am working to define new BACE1-AS related therapeutic targets and biomarkers for Alzheimer’s disease.

I am participating in a project aiming to characterize novel NAT-mediated regulation of brain derived neurotrophic factor (BDNF) in human cortical neurons.

Furthermore, I have preliminary data of what may be the first evidence of direct communication and thermodynamic interaction between members of two distinct families of ncRNAs in the computation of nervous systems gene expression.

Based on what I learned about various families of ncRNA, their involvement in neurophysiology and neuropathologies, I am planning to build my career on characterizing ncRNAs in neurological functions and their dysregulation in various neuropathologies. I am very determined to drive this area of research and to become a leader in this new field.

Keywords:
Natural antisense transcripts, NATs, NAT, Sense-Antisense, S-AS RNA, antisense transcript, noncoding RNA, ncRNA, non-protein-coding RNA, Regulation of gene expression
4 ACKNOWLEDGEMENTS

First and foremost I would like to thank Professor Claes Wahlestedt for taking me as a Ph.D. student and providing brilliant ideas and challenging projects.

I had ample opportunity to get to know and work with Professor Claes Wahlestedt, who is not only an excellent teacher, but also is an outstanding and exceptional human being. He has a great personality, an endless patience and devotion to his science. I am particularly grateful for his excellent guidance and for providing me with a broad range of research topics all these years. Perhaps, the great experience for me was that he has never told me the answers; instead he always showed the way to find them and the way to create more questions. I am so grateful to him for all the experiences and the science that he shared with me.

Special thanks to Professor Gunnar Norstedt for taking responsibility as my co-supervisor and always being helpful.

Many thanks to Omid Faridani for his loyal friendship, for many shared ideas and scientific projects and for his help in arranging my half-time seminar and public defense dissertation.

The first four years of my studies were supported by a scholarship from Ahwaz University, Iranian Ministry of Health and Medical Education. The total of 100K dollars, which I received, had covered my salaries. I am grateful to Ahwaz University and Iranian Ministry of Health and Medical Education for this scholarship, without that my Ph.D. studies and this thesis were not possible.

I would like to thank Drs. Liam Good and Zicai Liang for serving as my co-supervisors for a part of my Ph.D. studies and Salim Mottagui-Tabar for his mentorship during the first year of my Ph.D. studies.

I am grateful to all the past and present colleagues at the Scripps Florida and Karolinska Institute. Specially, I would like to thank Drs. Janett Kocerha, Paul J. Kenny, Georges St. Laurent III, Patricia McDonald, Shun P. Brothers, Jose P. Silva, Marcel P. Van der Brug, Miguel A. Lopez, Maggie Przydzial, Brooke Miller and Jia Huang, for all their help. I also would like to mention and thank Drs. Par G. Engstrom, Camilla Scheele, James A Timmons, Yosuke Mizuno, Alistair Chalk, Sayed Issac Hashemi, Abbas Nickravesh, Mohsen Karimi, and Reza Mohammadi for their company and help.

My dear wife Farzaneh Modarresi is standing beside me with encouragement, and patience. She was supporting me, all these years, from the day that we both decided to move to Sweden for our studies and to these days that I am preparing my thesis. Without her I was not able to finish this job and I am so grateful to her.

Last but not least, I would like to thank my lovely girls, Shamim and Zahra. They turned every single minute of my life to happiness, enjoyment, satisfaction and gratification. I am so proud of them, trying my best to compensate for the lack of time to play with them during my studies.
REFERENCES


endothelial nitric-oxide synthase. 

necrosis factor receptor homologue.


43


