Chapter from the book *Oncogene and Cancer - From Bench to Clinic*
Downloaded from: http://www.intechopen.com/books/oncogene-and-cancer-from-bench-to-clinic

Interested in publishing with IntechOpen?
Contact us at book.department@intechopen.com
Is CCDC26 a Novel Cancer-Associated Long-Chain Non-Coding RNA?

Tetsuo Hirano

1. Introduction

Large-scale analysis of total genome transcripts (transcriptome) in organisms including human and mouse has revealed that many RNAs are transcribed from genomic regions that encode no proteins (referred to as ncRNA) (1-5). Among such ncRNAs, microRNAs (miRNAs), small molecule RNAs 18-28 bases long, have been extensively studied over the past decade, and a gene regulatory system called “RNA silencing” has been revealed. In humans, more than 400 miRNAs are known to regulate at least one-third of protein-encoding genes (6-10). Most miRNAs are generated by processing of long miRNA precursors (pri-miRNAs) (6, 9). Pri-miRNAs are transcribed by RNA polymerase II and 5’ cap structures and poly A tails are added, similarly to protein-encoding mRNAs. Pri-miRNAs are further processed in the nucleus into pre-miRNAs with an approximately 70 base hairpin structure and are then exported to the cytoplasm. pre-miRNAs are finally processed into mature miRNAs by the enzyme, Dicer. It is noteworthy that miRNAs are sometimes encoded in the introns of other genes. A mature miRNA is incorporated into the RNA-induced silencing complex to act on its target mRNA. Broadly speaking, miRNAs can act on mRNAs in two ways. If there is limited homology between an miRNA and a target mRNA, the miRNA suppresses translation of the mRNA. However, if the miRNA has complete or nearly complete homology with a target mRNA, the mRNA is rapidly degraded. In animal cells, the former scenario usually occurs (7, 10-12). Many miRNAs have been reported to be associated with tumors, including AML and glioma; however, it is still unclear how predominant miRNAs are in tumorigenesis.

Relatively large ncRNAs of over several hundred bases, which are longer than pri-miRNAs whose length is usually 200-300 bases, are called long-chain non-coding RNAs (lncRNAs). Despite their somewhat unclear definition and their largely undetermined functions (13), the public databases for lncRNAs, for example, lncRNAdb (http://www.lncrnadb.org/) (14)
or NONCODE (http://www.noncode.org) (15), contain several hundred mammalian lncRNAs, including more than 100 from human (16). The RNAs included are heterologous; some localize in the nucleus to form certain structures, others interact with chromatin modifying enzymes such as p300, while others function in the cytoplasm (Fig. 1).

Both miRNAs and lncRNAs are physiologically important in many biological processes, including development and cell differentiation. Their association with disease, especially cancers, is of great interest (5). Association of miRNAs with various tumors, including different types of leukemia (Table 1) and glioma (Table 2), has been demonstrated. They sometimes act as tumor-promoting factors and sometimes as tumor suppressors. Expression of many lncRNAs, including NDM29 (neuroblastoma) (17, 18) and MALAT-1 (lung cancer) (19) are correlated with tumor progression, while MEG3 (pituitary tumor) (20, 21), HOTAIR (breast carcinoma) (22), H19 (Wilms’ tumor) (23), AK023948 (papillary thyroid tumor) (24) and LOC285194 (osteosarcoma) (25) are putative tumor suppressors (Table 3). These lncRNAs seem to control cancer cell growth by regulating other genes (NDM29, HOTAIR, H19) or by adjusting the mRNA splicing mechanism (MALAT-1) (Fig. 1) (14).

**Figure 1.** Classification of ncRNAs. (a) Pri-miRNAs are synthesized and processed in the nucleus, then exported to the cytoplasm. They are incorporated in the RISC complex to degrade or inhibit transcription of target mRNAs. However, some synthesized lncRNAs associate with chromatin (b) to silence certain genes. Some lncRNAs are incorporated in intranuclear bodies (c) or make complexes with specific proteins (d). Some are exported to the cytoplasm to work in the RNA-protein complex (e).
<table>
<thead>
<tr>
<th>Name</th>
<th>Loci</th>
<th>Name</th>
<th>Loci</th>
</tr>
</thead>
<tbody>
<tr>
<td>oncogenic or increased expression in AML</td>
<td></td>
<td>tumor suppressive or decreased expression in AML</td>
<td></td>
</tr>
<tr>
<td>let-7b</td>
<td>22q13.31</td>
<td>let-7</td>
<td>9q22.32</td>
</tr>
<tr>
<td>let-7e</td>
<td>19q13.41</td>
<td>let-7b</td>
<td>22q13.31</td>
</tr>
<tr>
<td>miR-10a</td>
<td>17q21.32</td>
<td>miR-9*</td>
<td>1q22</td>
</tr>
<tr>
<td>miR-10b</td>
<td>2q31.1</td>
<td>miR-15a</td>
<td>13q14.2</td>
</tr>
<tr>
<td>miR-27a</td>
<td>19p13.13</td>
<td>miR-15b</td>
<td>3p25.33</td>
</tr>
<tr>
<td>miR-30d</td>
<td>8q24.22</td>
<td>miR-16</td>
<td>13q14.2</td>
</tr>
<tr>
<td>miR-126</td>
<td>9q34.3</td>
<td>miR-19a</td>
<td>13q31.3</td>
</tr>
<tr>
<td>miR-129-5p</td>
<td>7q32.1</td>
<td>miR-20a</td>
<td>13q31.3</td>
</tr>
<tr>
<td>miR-130b</td>
<td>22</td>
<td>miR-26a</td>
<td>3p22.2</td>
</tr>
<tr>
<td>miR-142-5p</td>
<td>17q22</td>
<td>miR-29a</td>
<td>7q32.3</td>
</tr>
<tr>
<td>miR-155</td>
<td>21q21.3</td>
<td>miR-29b</td>
<td>7q32.3</td>
</tr>
<tr>
<td>miR-181a</td>
<td>1q31.3/9q33.3</td>
<td>miR-29c</td>
<td>1q32.2</td>
</tr>
<tr>
<td>miR-181b</td>
<td>1q31.3/9q33.3</td>
<td>miR-30a-3p</td>
<td>6q13</td>
</tr>
<tr>
<td>miR-181c</td>
<td>19p13.13</td>
<td>miR-34b</td>
<td>11q23.1</td>
</tr>
<tr>
<td>miR-181d</td>
<td>19p13.13</td>
<td>miR-34c</td>
<td>11q23.1</td>
</tr>
<tr>
<td>miR-195</td>
<td>17p13.1</td>
<td>miR-124</td>
<td>8q23.1</td>
</tr>
<tr>
<td>miR-221</td>
<td>Xp11.3</td>
<td>miR-128-1</td>
<td>2q21.3</td>
</tr>
<tr>
<td>miR-223</td>
<td>Xq12</td>
<td>miR-145</td>
<td>5q32</td>
</tr>
<tr>
<td>miR-221/222</td>
<td>Xp11.3</td>
<td>miR-147</td>
<td>9q33.2</td>
</tr>
<tr>
<td>miR-324-5p</td>
<td>17p13.1</td>
<td>miR-148a</td>
<td>7p15.2</td>
</tr>
<tr>
<td>miR-326</td>
<td>11q13.4</td>
<td>miR-151</td>
<td>8q24.3</td>
</tr>
<tr>
<td>miR-328</td>
<td>16q22.1</td>
<td>miR-181a</td>
<td>1q31.3/9q33.3</td>
</tr>
<tr>
<td>miR-331</td>
<td>12q22</td>
<td>miR-181b</td>
<td>1q31.3/9q33.3</td>
</tr>
<tr>
<td>miR-340</td>
<td>5q35.3</td>
<td>miR-182</td>
<td>7q32.2</td>
</tr>
<tr>
<td>miR-374</td>
<td>Xq13.2</td>
<td>miR-184</td>
<td>15q25.1</td>
</tr>
<tr>
<td>miR-424</td>
<td>Xq26.3</td>
<td>miR-194</td>
<td>1q41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-196a</td>
<td>17q21.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-196a</td>
<td>17q21.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-199a</td>
<td>19p13.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-204</td>
<td>9q21.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-219-5p</td>
<td>6q21.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-220a</td>
<td>Xq25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-302b*</td>
<td>4q25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-302d</td>
<td>4q25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-320</td>
<td>8q21.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-325</td>
<td>Xq21.1</td>
</tr>
</tbody>
</table>

Data are chosen from references 54, 55, 62, 63, 65, 66. Data confined to cytogenetically normal AML where possible. Note some miRNA appeared both oncogenic and tumor-suppressive.

**Table 1.** Examples of miRNAs associated with AML that change expression level
<table>
<thead>
<tr>
<th>Name</th>
<th>Genetic Locus</th>
<th>Name</th>
<th>Genetic Locus</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-9*</td>
<td>1q22</td>
<td>let-7 family</td>
<td>9q22.32</td>
</tr>
<tr>
<td>miR-10a*</td>
<td>17q21.32</td>
<td>miR-7</td>
<td>9q21.32</td>
</tr>
<tr>
<td>miR-10b</td>
<td>2q31.1</td>
<td>miR-15b</td>
<td>3q25.33</td>
</tr>
<tr>
<td>miR-17/92 cluster</td>
<td>13q31.3</td>
<td>miR-17</td>
<td>13q31.3</td>
</tr>
<tr>
<td>miR-21</td>
<td>17q23.1</td>
<td>miR-26b</td>
<td>2q35</td>
</tr>
<tr>
<td>miR-25</td>
<td>7q22.1</td>
<td>miR-29b</td>
<td>7q32.3</td>
</tr>
<tr>
<td>miR-26a</td>
<td>3p22.2</td>
<td>miR-34a</td>
<td>1p36.22</td>
</tr>
<tr>
<td>miR-93</td>
<td>7q22.1</td>
<td>miR-101</td>
<td>1p31.3</td>
</tr>
<tr>
<td>miR-125b</td>
<td>11q24.1</td>
<td>miR-106a</td>
<td>Xq26.2</td>
</tr>
<tr>
<td>miR-182</td>
<td>7q32.2</td>
<td>miR-124</td>
<td>8p23.1</td>
</tr>
<tr>
<td>miR-195</td>
<td>17p13.1</td>
<td>miR-125a</td>
<td>19q13.41</td>
</tr>
<tr>
<td>miR-196a</td>
<td>17q21.32</td>
<td>miR-128</td>
<td>2q21.3</td>
</tr>
<tr>
<td>miR-196b</td>
<td>7p15.2</td>
<td>miR-137</td>
<td>1p21.3</td>
</tr>
<tr>
<td>miR-221/222</td>
<td>Xp11.3</td>
<td>miR-146b/146b-5p</td>
<td>10q24.32</td>
</tr>
<tr>
<td>miR-296</td>
<td>20q13.32</td>
<td>miR-153</td>
<td>2q35</td>
</tr>
<tr>
<td>miR-381</td>
<td>14q32.31</td>
<td>miR-181</td>
<td>1q31.3/9q33.3</td>
</tr>
<tr>
<td>miR-455-3p</td>
<td>9q32</td>
<td>miR-184</td>
<td>15q25.1</td>
</tr>
<tr>
<td>miR-486</td>
<td>8p11.21</td>
<td>miR-195</td>
<td>17p13.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-199b-5p</td>
<td>9q34.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-218</td>
<td>4p15.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-326</td>
<td>11q13.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-451</td>
<td>17q11.2</td>
</tr>
</tbody>
</table>

Data are chosen from references 67 and 68. Data confined to cases of low grade gliomas but exclusion of data from high grade glioblastoma is not necessarily complete. Note some miRNAs appeared both oncogenic and tumor-suppressive.

**Table 2.** miRNAs that show altered expression levels in glioma cells
Is CCDC26 a Novel Cancer-Associated Long-Chain Non-Coding RNA?

<table>
<thead>
<tr>
<th>Name</th>
<th>Alias</th>
<th>Mouse Homolog</th>
<th>Genetic Locus</th>
<th>Product Length (bp)</th>
<th>Tumor</th>
<th>Function</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tumor promoting or Increased Expression</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIRN</td>
<td>Airc</td>
<td>6q26</td>
<td>NA</td>
<td>200</td>
<td>Wilms’ tumor</td>
<td>NA</td>
<td>(59)</td>
</tr>
<tr>
<td>BC200</td>
<td>BCYRN1</td>
<td>2p21</td>
<td>14q23.2</td>
<td>2051</td>
<td>Multiple cancers</td>
<td>Decoy of mRNA</td>
<td>(71)</td>
</tr>
<tr>
<td>HIF1A-AS2</td>
<td>aHIF</td>
<td>6q26</td>
<td>1q12-13.1</td>
<td>2364</td>
<td>Multiple cancers</td>
<td>Epigenetic silencing of HOXD gene through histoneH3K27 methylation</td>
<td>(72)</td>
</tr>
<tr>
<td>HOTAIR</td>
<td>Gm16258</td>
<td>1q21.3</td>
<td>2q13.3</td>
<td>2364</td>
<td>Multiple cancers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HULC</td>
<td>NA</td>
<td>6p24.3</td>
<td>500</td>
<td>Hepatocellular carcinoma</td>
<td>Post-transcriptional regulation</td>
<td>(73)</td>
<td></td>
</tr>
<tr>
<td>IGF2AS</td>
<td>Igh2as</td>
<td>11p15.5</td>
<td>2091</td>
<td>Wilms’ tumor</td>
<td>NA</td>
<td>(74)</td>
<td></td>
</tr>
<tr>
<td>KRASPI1</td>
<td>NA</td>
<td>6p12-p11</td>
<td>5178</td>
<td>Prostate cancer</td>
<td>Decoy of miRNA</td>
<td>(75)</td>
<td></td>
</tr>
<tr>
<td>L1PA16</td>
<td>VL30-1 1060</td>
<td>3q26.3</td>
<td>833</td>
<td>Many tumor cell lines</td>
<td>Activation of proto-oncogene</td>
<td>(76)</td>
<td></td>
</tr>
<tr>
<td>MALAT1</td>
<td>Neat2</td>
<td>11q13.1</td>
<td>8708</td>
<td>Multiple cancer</td>
<td>Control of RNA procession</td>
<td>(19, 77)</td>
<td></td>
</tr>
<tr>
<td>MER11C</td>
<td>HERVK11</td>
<td>11p11.1</td>
<td>1060</td>
<td>Many tumor cell lines</td>
<td>Activation of proto-oncogene</td>
<td>(76)</td>
<td></td>
</tr>
<tr>
<td>PCA3</td>
<td>DD3</td>
<td>9q12-q22</td>
<td>3735</td>
<td>Prostate cancer</td>
<td>NA</td>
<td>(78)</td>
<td></td>
</tr>
<tr>
<td>PCGEM1</td>
<td>NA</td>
<td>2q32</td>
<td>1603</td>
<td>Prostate cancer</td>
<td>NA</td>
<td>(79)</td>
<td></td>
</tr>
<tr>
<td>PRNCR1</td>
<td>NA</td>
<td>8q24</td>
<td>&gt;12756</td>
<td>Prostate cancer</td>
<td>NA</td>
<td>(80)</td>
<td></td>
</tr>
<tr>
<td>SRA1</td>
<td>Sra1</td>
<td>5q31.3</td>
<td>1955</td>
<td>Breast cancer</td>
<td>Activation of nuclear receptors</td>
<td>(81)</td>
<td></td>
</tr>
<tr>
<td>TERC</td>
<td>Terc</td>
<td>3q26</td>
<td>451</td>
<td>Multiple cancer</td>
<td>Telomere template</td>
<td>(82)</td>
<td></td>
</tr>
<tr>
<td>UCA1</td>
<td>CUDR</td>
<td>19p13.12</td>
<td>1591</td>
<td>Bladder cancer</td>
<td>Regulation of cell cycle</td>
<td>(83)</td>
<td></td>
</tr>
<tr>
<td>WT1-AS</td>
<td>WIT1</td>
<td>11p13</td>
<td>1333</td>
<td>Wilms’ tumor</td>
<td>Downregulation of WT1, tumor suppressor</td>
<td>(84)</td>
<td></td>
</tr>
<tr>
<td><strong>Tumor Suppressing or Decreased Expression in Tumor</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AK023948</td>
<td>NA</td>
<td>8q24</td>
<td>2807</td>
<td>Papillary thyroid carcinoma</td>
<td>NA</td>
<td>(24)</td>
<td></td>
</tr>
<tr>
<td>ANRIL</td>
<td>CDK2BAS, p15AS</td>
<td>9q21</td>
<td>944</td>
<td>Prostate cancer, breast cancer, melanoma, and other tumors</td>
<td>Regulation of epigenetic transcriptional repression</td>
<td>(58)</td>
<td></td>
</tr>
<tr>
<td>BC040587</td>
<td>NA</td>
<td>3q13.31</td>
<td>NA</td>
<td>Osteosarcoma</td>
<td>NA</td>
<td>(25)</td>
<td></td>
</tr>
<tr>
<td>DLEU2</td>
<td>Dleu2</td>
<td>13q14.3</td>
<td>2768</td>
<td>Chronic lymphocytic leukemia</td>
<td>pri-miRNA for miR15a and miR16</td>
<td>(86)</td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Alias</td>
<td>Mouse Homolog</td>
<td>Genetic Locus</td>
<td>Product Length (bp)</td>
<td>Tumor</td>
<td>Function</td>
<td>Refs</td>
</tr>
<tr>
<td>----------</td>
<td>----------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------------</td>
<td>------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>GAS5</td>
<td>Gas5</td>
<td>1q25.1</td>
<td>651</td>
<td>Breast cancer</td>
<td>Decoy of glucocorticoid receptor</td>
<td>(87)</td>
<td></td>
</tr>
<tr>
<td>H19</td>
<td>H19</td>
<td>11p15.5</td>
<td>2322</td>
<td>Wilms' tumor</td>
<td>Epigenetic regulation through DNA methylation</td>
<td>(88)</td>
<td></td>
</tr>
<tr>
<td>KCNQ1OT1</td>
<td>LIT1, KvLQT1-AS, KvLQT1OT1</td>
<td>Kcnq1ot1</td>
<td>11p15</td>
<td>91671</td>
<td>Embryonal cancer associated with Beckwith-Wiedemann syndrome</td>
<td>Epigenetic imprinting through H3K27 methylation</td>
<td>(57)</td>
</tr>
<tr>
<td>LOC285194</td>
<td>NA</td>
<td>3q13.31</td>
<td>NA</td>
<td>Osteosarcoma</td>
<td>NA</td>
<td>(25)</td>
<td></td>
</tr>
<tr>
<td>MEG3</td>
<td>Gli2</td>
<td>Mrg3</td>
<td>14q32</td>
<td>Glioma, pituitary adenoma and other tumor</td>
<td>Regulation of p53 target proteins</td>
<td>(89)</td>
<td></td>
</tr>
<tr>
<td>NDM29</td>
<td>29A</td>
<td>NA</td>
<td>11p15.3</td>
<td>Neuroblastoma</td>
<td>Induction the appearance of neuronal-like properties</td>
<td>(18)</td>
<td></td>
</tr>
<tr>
<td>p53 mRNA</td>
<td>Tp53</td>
<td>17p13.1</td>
<td>19144</td>
<td>Multiple cancer</td>
<td>RNA protein binding, MDM3</td>
<td>(90)</td>
<td></td>
</tr>
<tr>
<td>PTENP1</td>
<td>NA</td>
<td>9p21</td>
<td>3932</td>
<td>Prostate cancer</td>
<td>Decoy for PTEN-targeting miRNAs</td>
<td>(75)</td>
<td></td>
</tr>
<tr>
<td>RMRP</td>
<td>Rmrp</td>
<td>9p21-p12</td>
<td>267</td>
<td>Leukemia and lymphoma</td>
<td>Mitochondrial RNA processing endoribonuclease, hTERT-dependent</td>
<td>(91)</td>
<td></td>
</tr>
<tr>
<td>TERRA</td>
<td>TelRNAs</td>
<td>telomere repeats</td>
<td>NA</td>
<td>Many cancer cell lines</td>
<td>Interaction with the TRF1</td>
<td>(92)</td>
<td></td>
</tr>
<tr>
<td>vtRNA2-1</td>
<td>NA</td>
<td>5q31.1</td>
<td>100</td>
<td>AML, papillary thyroid cancer</td>
<td>Regulation of RNA dependent protein kinase (pPKR)</td>
<td>(93)</td>
<td></td>
</tr>
<tr>
<td>ZNFX1-AS1</td>
<td>Zfas1</td>
<td>1500012</td>
<td>1020</td>
<td>Breast cancer</td>
<td>NA</td>
<td>(94)</td>
<td></td>
</tr>
</tbody>
</table>

(a) no homologous RNA but binds to PSF, a transcriptional repressor. NA, not available.

**Table 3.** Human IncRNAs associated with tumors described in public data bases.

### 2. Genetic abnormality observed in acute myeloid leukemia (AML)

AML, which comprises approximately 25% of hematopoietic malignancies, has heterogeneous clinical features and variable responses to contemporary therapy (26). Genetic alterations are often observed in AML cells and the clinical heterogeneity of the disease is considered to reflect the genetic diversity of these cells (27, 28). It is very important to study the genetic mutations in AML cells to fully understand the cause of the disease. However, genetic lesion(s) responsible for AML, such as the loss or gain of a certain gene, have not yet been fully elucidated. Indeed, the complex features of AML suggest that the genetic cause of this disease is multifactorial (29). Several protein-encoding genes have been identified that are useful for indicating the prognosis of the disease (30-32).
include RUNX1 (AML1)-RUNX1T1 (ETO) and CBFB-MYH11, which are associated with specific chromosomal mutations, t(8;21)(q22;q22) and inv(16)(p13;q12)/t(16;16)(p13;q22), respectively. AML with these cytogenetic features (singly or together) represents about 15% of de novo AML. The patients with these diagnostic criteria are classified in the favorable clinical outcome group (standard-risk group). Several other chromosomal abnormalities have been recurrently observed, as described in the WHO classification. AML with balanced or unbalanced translocations involving the MLL gene located on chromosome 11 are also well documented and are mostly classified in the intermediate-risk group. Meanwhile, AML patients with a normal karyotype and no cytological abnormality include cases classified in the unfavorable (adverse-risk) or intermediate-risk group. Moreover, a genetic abnormality of the FL3 gene (internal tandem repeat) is found in many AML subtypes and, in combination with a wild-type NPM gene, contributes to poor prognosis (31). Recently, Paschka and colleagues have revealed that the genes encoding the metabolic enzymes, isocitrate dehydrogenase 1 and 2 (IDH1/2) are important for diagnosis and prognosis prediction of AML patients (33). These mutations of IDH1/2 change the activity of the enzymes to reduce α-ketoglutarate levels and to elevate 2-hydroxyglutarate levels. This results in changes to chromatin structure and destabilization of certain gene-regulatory proteins, including HIF-1 (34). While cytogenetically normal AML patients with an NPM mutation and a normal FL3 gene tend to show favorable outcomes, AML patients with the same genetic profile but also with IDH1/2 mutation showed adverse prognosis with poorer remission. IDH1/2 mutation was also found in several other tumors, including glioma (35). Therefore, a combination of genetic alterations resulting in mutation of specific genes as well as cytogenetically apparent chromosomal changes are important for AML malignancy.

3. AML and CCDC26

In HL-60 cells derived from AML, a small part of chromosome 8 is excised and amplified as an extrachromosomal element, or double minute chromosome (dmin). Dmin is a cytogenetic abnormality infrequently observed in AML. The dmin of HL-60 cells consists of several repeats of an amplification unit (referred as amplicon) of about 2 million base pairs. The amplicon, which is derived from several areas of an approximately 4.6 million base pair region of chromosome 8q24, contains an intact MYC oncogene. Besides MYC, several other genes, including CCDC26 and tribbles homolog 1 (TRIB1), are also encoded on the amplicon (Fig. 2). All are actively transcribed in HL-60 cells. The drug-induced differentiation of HL-60 cells suppressed the expression of all these genes, indicating that they might be related to the cancerous nature of the cells. Some types of cancer cell respond to the anticancer drug hydroxyurea by excluding unstable extrachromosomal elements, which then lose their proliferative nature. In HL-60 cells, the original MYC genetic locus remained intact after dmin was excluded, but was no longer transcribed (36). These observations suggest that the expression of genes from dmin, with its altered DNA structure, and from the intact chromosome are different, and can be interpreted as being due to aberrant gene expression from dmin (including the MYC oncogene). Interestingly, in HL-60 cells, the CCDC26 gene on dmin is rearranged as a result of chromosomal rejoining and is amplified in an incomplete form to produce abnormal transcripts (37).
Figure 2. Depiction of the generation of the discontinuous amplicon unit of dmin in HL-60 cells. Several replication initiation bubbles collapse and “corrupted” bubbles reunite to form an amplicon unit. Excision of an initial large amplicon (possibly as an episome) might precede replication. The arrows numbered 1 through 6 indicate regions reunited in the amplicon. Note that the lengths of the various regions are not to scale. Once the amplicon unit has formed, its multimerization results in a dmin.

A common change occurs at the CCDC26 locus in cytologically dmin-positive AML patients. This chromosomal change occurs at a position consistent with the amplified region observed in HL-60 cells (38, 39). Furthermore, destruction of the internal structure of the CCDC26 gene seems to underlie the common mechanism behind the generation of dmin-positive AML cells.

A comprehensive genome-wide study of a group of childhood AML patients revealed that CCDC26 was one of the genes with the highest increase in copy number in AML cells. Radtke and colleagues investigated chromosome number alteration (CNA) in pediatric AML using a comprehensive single nucleotide polymorphism (SNP) array analysis. They found the most common CNA, in 14% (15 in 111) of pediatric AML patients, to be in
chromosome band 8q24 with a low-burden copy number increase (2.83-3.77 copies) (40). These included cases of trisomy 8, which frequently occurs in AML (41). The minimum altered region common in all 15 of these patients was located in a 20-megabase region of 8q24, which contains CCDC26.

Originally, CCDC26 was reported as a gene associated with differentiation and apoptosis of PLB985 cells (an HL-60 subclone) following induction by treatment with retinoic acid (CCDC26 is also known as RAM, retinoic acid modifying). In cells that have become resistant to differentiation and apoptosis after infection of retrovirus, the viral genome was seen to be inserted in the intron of CCDC26. Retinoic acid promotes differentiation and apoptosis of not only many leukemia cells but also of neuroblastoma and glioblastoma cells through transcriptional regulation of many other genes. CCDC26 may have a role with retinoic acid in differentiation and growth arrest of these cells (42).

4. Glioma and CCDC26

Primary brain tumor (PBT) is a disease with an incidence of 12 in 100,000 per year. Glioma accounts for a major part of PBT, and contains cases with different grades of malignancy, namely (I) benign glioma, (II) diffuse astrocytoma, (III) anaplastic astrocytoma and (IV) glioblastoma (43). Although many genetic abnormalities have been reported in gliomas, a single critical lesion responsible for tumorigenesis has not been found. Among these abnormalities, mutations occur in genes for DNA repair enzymes, including PRKDC, XRCC, PARP1, MGMT, ERCC1, ERCC2, epidermal growth factor and the inflammatory cytokine, IL-13. Furthermore, over-expression or amplification of the epidermal growth factor receptor gene and deletion of p16INK are correlated with poor survival (43). A genome wide association study using SNPs revealed the association of several genes with glioma, including telomerase regulating gene TERT, RTEL1, tumor suppressor gene CDKN2A/2B, pleckstrin homology-like domain family B member 1 (a protein with unknown function) and CCDC26 (44). The CCDC26 gene locus was strongly linked with this glioma by several SNPs, including rs4295627, rs16904140, rs6470745, rs891835, and rs10464870 (see Fig. 3a). A different SNP in the intergenic region bordering CCDC26, rs987525, was linked to cleft palate (45). Notably, cleft palate is also a risk factor of PBT. CCDC26 is, therefore, a potential common factor of both conditions. CCDC26 is just one of the risk factors for glioma and other genetic risk factors increase glioma incidence cumulatively. Therefore, there might be a synergistic effect with other genetic risk factors (46). CCDC26 is not necessarily a risk factor of high grade (III-IV) glioma (47). Interestingly, in concordance with the situation for AML, the CCDC26 genotype is associated with IDH1/2 mutation in low grade glioma. Considering the synergy of CCDC26 with IDH1/2, CCDC26 may have linkage to a subpopulation of gliomas with relatively lower grade (46, 48).

The Gene Expression Omnibus database (GEO; http://www.ncbi.nlm.nih.gov/geo/) (49) contains data showing altered CCDC26 expression between normal and tumorigenic cells. Expression of CCDC26 is higher in myeloid leukemia cell lines, namely KG-1, THP-1 and U937, compared with normal monocytes (GEO dataset accession ID; GDS2251), and is higher in sporadic basal-like cancer compared with normal cells (GD2250). On the other
Figure 3. (A) Summary of genetic locus 8q24.31. The scale indicates the region 8q24.31 (130.35 Mb to 130.71 Mb; numbering is based on human genome assembly 37.1). Marks on the scale indicate the locations of five SNPs that are linked to glioma (open circles; rs4295627, rs16904140, rs6470745, rs891835 and rs10464870, left to right), three miRNAs registered in the public database (filled circles; mir-3669, mir-3673 and mir-3686, left to right) and a retrovirus insertion site (filled square) within the introns of CCDC26. Below the scale, which covers a 350-kb region, major variants of CCDC26 mRNA (long and short forms) are shown, along with the transcriptional activity in mammals. (B) Sequence alignment of CCDC26 mRNA in mammals, showing conserved regions and variable substitutions. The sequence alignment indicates that CCDC26 mRNA is conserved across species, with some variations in specific regions.

Histone H3K27 Acetylated

Transcription

Mammals synteny
short) are shown. The long transcript consists of four (1-2-3-4) exons, and the short transcripts consist of three (1a-3-4) or four (1a-2a-3-4) exons. All variants share exon 3 and 4, in which the hypothetical open reading frame is encoded. Locations of the amplified region in HL-60 cells and the commonly amplified region (MAR) in childhood AML are shown with filled rectangles. The hypothetical open reading frame encoded in exons 3 and 4 is not included in the region amplified in HL-60 cells. The pattern of histone H3K27 acetylation, activity of transcription in leukemia cell lines, and bar plots for conserved synteny between human and each organism (m: mouse, d: dog, e: elephant and o: opossum) obtained from the UCSC Genome Browser (95) are shown. Actively transcribed regions that are not the major exons of CCDC26 mRNA are indicated by grey rectangles. (B) Optimal alignment of the CCDC26 exon 3-4 encoded ORF and the region of conserved synten on mouse chromosome 15. A possible ORF (94 amino acids) in the mouse sequence is totally mismatched with that of human by frameshift changes.

hand, CCDC26 expression is decreased in hyperplastic enlarged lobular units considered as the earliest precursors of breast cancer compared with normal units (GDS2739). Increased CCDC26 expression is associated with malignancy progression in some cancerous cells. CCDC26 expression was increased in CD133 positive neurosphere-like glioma cell lines compared with CD133 negative adherent glioma cell lines (GDS2728), and was increased in alveolar macrophages of cigarette smokers comparison with macrophages of non-smokers (GDS3496). Increased expression of CCDC26 might mean this gene is tumorigenic or oncogenic. However, the relationship of altered CCDC26 expression to malignancy is still ambiguous.

5. Overview of the CCDC26 genetic locus

As described in the previous section, all SNPs associated with glioma, and a retrovirus insertion site where virus insertion makes AML cells resistant to retinoic acid (42) are located in the intron of CCDC26 (Fig. 3a). Exon 4, which encodes the majority of a hypothetical open reading frame (ORF), is not amplified in pediatric AML or in AML-derived HL-60 cells. The exonic sequence of CCDC26 is not well conserved in other species, including mouse, and an ORF has no homology with known proteins. These data strongly suggest that CCDC26 does not function as a protein-encoding RNA; rather it functions as a ncRNA. Highly conserved regions in the intron sequence of CCDC26 suggest the existence of another intronic ncRNA. As mentioned above, the CCDC26 locus is rearranged in the genome of HL-60 cells. It is plausible that the ncRNA encoded by this locus is important for the growth of these cells.

A short putative ORF encoding a protein or with a length of 109 amino acids is present in the CCDC26 exons; there is no other ORF of more than 50 amino acids. This actual protein, however, has not been observed. Moreover, orthologous proteins are not found in any other organism. For example, a loosely homologous sequence of human exon 4, found in the mouse chromosome 15 region of conserved synteny, with an ORF of 94 amino acids is actively transcribed in mouse leukemia cells (T. Hirano unpublished observation). However, this ORF is completely different from the human sequence and even contains frame shift alterations (Fig. 3b). This indicates that the putative protein encoded by CCDC26 has no conserved function among species. Although this ORF may be coincidental due to the
absence of stop codons, an interesting possibility is that this unique protein has newly emerged during human evolution. mRNA stability is influenced by whether an ORF is encoded because nonsense mediated RNA decay, a mechanism associated with quality control of mRNA, rapidly degrades mRNAs that are not useful as templates for protein synthesis. Absence of an ORF in an mRNA promotes degradation by this mechanism, however, the existence of a CCDC26 protein will prolong the lifetime of CCDC26 mRNA and may maintain the function (if any) of the RNA itself.

Because of the considerable length of the CCDC26 intron (330 kbp versus 1200 bp exons), it is very difficult to ignore the possibility that there is another transcript(s) within this intron with important function. Possible encoded ncRNAs within the CCDC26 exon-intron region are summarized in Fig. 4 and include, mRNA (a), intronic encoded ncRNA (b), intronic lariat RNA (c-d) and miRNA independently transcribed or processed from the precursor of the CCDC26 mRNA (e). Actually there are several regions in the CCDC26 intron where nucleosomal histones undergo high levels of methylation and acetylation, meaning that these locations may be actively transcribed (Fig.3a). Furthermore, most of these regions are

Figure 4. A possible function of CCDC26. The 330-kb precursor RNA transcribed from the gene is processed into mature mRNA. It then forms a complex with proteins to perform its biological function, for example, silencing a certain genetic locus (a). Alternatively, the ncRNA independently encoded in the intronic region (b) or the processed intron lariat (c-d) could have biological functions. The intronic microRNA could be transcribed directly from the genome using its own promoter or processed from the precursor of the CCDC26 mRNA(e). Note that lengths of nucleic acid chains are not to scale. Length of the precursor RNA is approximately 330 kb; CCDC26 mRNA is 1.3 kb; the spliced lariat is more than 100 kb; intronic microRNA is 18-23 bases.
CCDC26 as a Novel Cancer-Associated Long-Chain Non-Coding RNA

highly conserved among mammals, suggesting that function is encoded. Also, expressed sequence tags other than known spliced CCDC26 mRNAs have been reported in the intron. There are three miRNAs (miR-3669, 3673 and 3686) in the intron of CCDC26 that are registered in the miRNA database (miRBase; http://www.mirbase.org/) (50). Although their functions are unknown, they may act as oncogenic or tumor suppressive ncRNAs.

6. Hypothetical function of CCDC26 as a non-coding RNA

Although many ncRNAs are registered in databases, only a few have clearly demonstrated functions and detailed mechanisms of action. CCDC26 might be a new ncRNA that is associated with cancer, including AML. Interestingly, expression of an miRNA, miR-21, is observed in many malignant cells, including AML cells (51). Also phorbol ester-induced differentiation of HL-60 cells into macrophage-like cells is accompanied by up-regulation of miR-21 (52). There are several reports suggesting that miRNAs act as oncogenic or tumor suppressive miRNAs in AML, as reviewed in (53, 54). Recently, Marcucci and colleagues used 305 different probes to search for miRNA expression in favorable and adverse-risk groups of normal karyotype AML (monocytic leukemia). They then used these data to link expression profiles with the cohort analysis of the patients. They identified a certain pattern of miRNA expression in the adverse-risk group and linked the expression level of eight types of miRNA to AML prognosis (55). It is possible that an unknown miRNA in the CCDC26 locus affects cancer malignancy through the regulation of other genes. But all miRNAs described so far in the CCDC26 locus (mir-3669, mir-3673 and mir-3686) show no expression in leukemia cells and no conservation among mammals in contrast to other oncogenic miRNAs; for example miR21 and let7, are actively transcribed and strongly conserved.

Within the CCDC26 intronic region, there are some long regions (>10 kb) that are actively transcribed in leukemia cells (Fig. 3a). They seem to be too long for pri-miRNAs but could encode lncRNAs. Indeed, active transcription occurs in the CCDC26 region in cells derived from AML (T. Hirano unpublished observation), meaning that these transcripts might function as a tumor promoting or oncogenic lncRNAs. In contrast, if the original function of CCDC26, or of lncRNAs associated with CCDC26, was lost by chromosomal abnormality (for example in dmin of HL-60 cells), then they might function naturally as tumor suppressors. Some lncRNAs including XIST (56), KCNQ1OT1 (57), ANRIL (58) and AIRN (59) are known to suppress (in cis) the expression of neighboring gene. It is well known that genes located in extrachromosomal elements such as dmin are actively transcribed, but the mechanism behind this phenomenon is not well understood (60, 61). Differences between dmin and an intact chromosome are caused by differences in chromatin structure, which is indicated by differences in DNase I hypersensitivity (36). Similarly to other gene silencing lncRNAs, an ncRNA encoded by the CCDC26 locus might suppress the expression of other nearby genes. The hypothesis that neighboring genes, including the MYC oncogene, are activated when the normal CCDC26 locus structure is destroyed by a chromosomal abnormality could explain the high transcriptional activity of genes in extrachromosomal elements (Fig.5). Further evidence is needed to determine whether CCDC26 mRNA and/or its transcripts encoded in its intron are oncogenic or tumor suppressive.
Figure 5. Hypothetical mechanism by which an oncogene (for example MYC) is inactivated in an intact chromosome (a) by CCDC26 regulation. This suppression does not function on extrachromosomal chromatin (b).

7. Future perspectives

The size of the CCDC26 locus, spanning over 330,000 base pairs, makes it difficult to study. If the ORF of the gene is not functional then it is unclear which part(s) of the locus are functional. Therefore, to study this gene, it is first necessary to determine all transcripts produced by the CCDC26 locus and then to analyze their function. Comprehensive analysis of transcriptome of the relevant region using tiling microarray analysis is needed. Although lncRNA orthologs are frequently not found between species, homology analysis of this region between human and mouse could be helpful to identify functional sequences. Once transcripts are identified, we will be able to perform in situ hybridization to determine subcellular localization. Knock-down of transcripts will be useful to investigate their functions. Proteins interacting with the RNA transcripts will be identifiable by pull-down assays and mass spectrometry analysis. Finally, gene targeting should be used to investigate the effects of disruption of the region encoding the transcript. It will be of special interest if transcription of neighboring genes is activated or inactivated (in particular MYC), suggesting a regulatory function of the ncRNA encoded in the CCDC26 locus. If an ortholog of the gene is found in mice, making a knock-out mouse of the ncRNA or a transgenic mouse with forced expression of the ncRNA will help to demonstrate its relationship to disease.

8. Conclusion

As a conclusion, the CCDC26 locus is considered to encode an lncRNA involved in tumorigenesis. CCDC26 itself might be an lncRNA or its intron might contain a functional miRNA or lncRNA. The study of this gene will bring new knowledge to gene regulation and to cancer treatment strategies targeting lncRNAs. Further in vitro and in vivo study is needed to prove the relationship between transcripts from the locus and disease, such as leukemia and glioma.
Author details

Tetsuo Hirano

Life Science Group, Graduate School of Integrated Arts and Sciences, Hiroshima University, Kagamiyama, Higashihiroshima, Hiroshima, Japan

9. References

[33] Paschka P, Schlenk RF, Gaidzik VI, Habdank M, Krönke J, Bullinger L, et al. IDH1 and IDH2 mutations are frequent genetic alterations in acute myeloid leukemia and confer


[37] Hirano T, Ike F, Murata T, Obata Y, Utiyama H, Yokoyama KK. Genes encoded within 8q24 on the amplicon of a large extrachromosomal element are selectively repressed during the terminal differentiation of HL-60 cells. Mutat Res. 2008;640:97-106.


[64] Langer C, Marcucci G, Holland KB, Radmacher MD, Maharry K, Paschka P, et al. Prognostic importance of MN1 transcript levels, and biologic insights from MN1-


