**SUPPLEMENTARY MATERIALS**

**Ontological function annotation of long non-coding RNAs through hierarchical multi-label classification**

**Jingpu Zhang, Zuping Zhang, Zixiang Wang, Yuting Liu and Lei Deng**

**Results**

Two lncRNAs are used as instances to further demonstrate the predictive performance and show the application of our method. For each lncRNA, the GO terms predicted, GO names and the paths are listed in the tables. In the column of paths, all the paths from level 1 to the object GO term in the deepest level are shown. For example, in the path: GO0006357@GO0065007/GO0050789/GO0019222/GO0060255/GO0010468/GO0006355/GO0006357, GO0006357 represents the object, and @ is a separator, while the GO terms following the @ separator comprise a specific path from GO0065007 to GO0006357.

**Table S1.** The predicted GO terms for lncRNA HOTAIRM1 by NeuraNetL2GO.

|  |  |  |  |
| --- | --- | --- | --- |
| ID | GO terms | GO names | paths |
| 1 | GO0006357 | regulation of transcription from RNA polymerase II promoter | GO0006357@GO0065007/GO0050789/GO0019222/GO0060255/GO0010468/GO0006355/GO0006357; GO0006357@GO0065007/GO0050789/GO0019222/GO0009889/GO0010556/GO2001141/GO1903506/GO0006355/GO0006357; GO0006357@GO0065007/GO0050789/GO0019222/GO0060255/GO0010556/GO2001141/GO1903506/GO0006355/GO0006357;GO0006357@GO0065007/GO0050789/GO0019222/GO0009889/GO0031326/GO2001141/GO1903506/GO0006355/GO0006357;GO0006357@GO0065007/GO0050789/GO0019222/GO0031323/GO0031326/GO2001141/GO1903506/GO0006355/GO0006357;GO0006357@GO0065007/GO0050789/GO0050794/GO0031323/GO0031326/GO2001141/GO1903506/GO0006355/GO0006357;GO0006357@GO0065007/GO0050789/GO0019222/GO0031323/GO0019219/GO0051252/GO2001141/GO1903506/GO0006355/GO0006357;GO0006357@GO0065007/GO0050789/GO0050794/GO0031323/GO0019219/GO0051252/GO2001141/GO1903506/GO0006355/GO0006357; GO0006357@GO0065007/GO0050789/GO0019222/GO0051171/GO0019219/GO0051252/GO2001141/GO1903506/GO0006355/GO0006357;GO0006357@GO0065007/GO0050789/GO0019222/GO0080090/GO0019219/GO0051252/GO2001141/GO1903506/GO0006355/GO0006357@GO0065007/GO0050789/GO0019222/GO0060255/GO0051252/GO2001141/GO1903506/GO0006355/GO0006357;GO0006357@GO0065007/GO0050789/GO0019222/GO0009889/GO0010556/GO2000112/GO0006355/GO0006357;GO0006357@GO0065007/GO0050789/GO0019222/GO0060255/GO0010556/GO2000112/GO0006355/GO0006357;GO0006357@GO0065007/GO0050789/GO0019222/GO0009889/GO0031326/GO2000112/GO0006355/GO0006357;GO0006357@GO0065007/GO0050789/GO0019222/GO0031323/GO0031326/GO2000112/GO0006355/GO0006357;GO0006357@GO0065007/GO0050789/GO0050794/GO0031323/GO0031326/GO2000112/GO0006355/GO0006357 |
| 2 | GO0006355 | regulation of transcription, DNA-templated |
| 3 | GO1903506 | regulation of nucleic acid-templated transcription |
| 4 | GO2001141 | regulation of RNA biosynthetic process |
| 5 | GO2000112 | regulation of cellular macromolecule biosynthetic process |
| 6 | GO0051252 | regulation of RNA metabolic process |
| 7 | GO0010468 | regulation of gene expression |
| 8 | GO0010556 | regulation of macromolecule biosynthetic process |
| 9 | GO0031326 | regulation of cellular biosynthetic process |
| 10 | GO0019219 | regulation of nucleobase-containing compound metabolic process |
| 11 | GO0060255 | regulation of macromolecule metabolic process |
| 12 | GO0009889 | regulation of biosynthetic process |
| 13 | GO0031323 | regulation of cellular metabolic process |
| 14 | GO0051171 | regulation of nitrogen compound metabolic process |
| 15 | GO0080090 | regulation of primary metabolic process |
| 16 | GO0019222 | regulation of metabolic process |
| 17 | GO0050794 | regulation of cellular process |
| 18 | GO0050789 | regulation of biological process |
| 19 | GO0065007 | biological regulation |
| 20 | GO0007165 | signal transduction | GO0007165@GO0009987/GO0007165;GO0007165@GO0065007/GO0050789/GO0050794/GO0007165 |
| 17 | GO0050794 | regulation of cellular process |
| 18 | GO0050789 | regulation of biological process |
| 21 | GO0009987 | cellular process |
| 19 | GO0065007 | biological regulation |
| 22 | GO0006810 | transport | GO0006810@GO0051179/GO0051234/GO0006810 |
| 23 | GO0051234 | establishment of localization |
| 24 | GO0051179 | localization |

**Table S2.** The predicted GO terms for lncRNA GAS5 by NeuraNetL2GO.

|  |  |  |  |
| --- | --- | --- | --- |
| ID | GO term | GO name | path |
| 1 | GO0015031 | protein transport | GO0015031@GO0051179/GO0051234/GO0006810/GO0071705/GO0042886/GO0015833/GO0015031;GO0015031@GO0051179/GO0051234/GO0006810/GO0071702/GO0015833/GO0015031;GO0015031@GO0051179/GO0033036/GO0008104/GO0045184/GO0015031;GO0015031@GO0051179/GO0051234/GO0045184/GO0015031 |
| 2 | GO0015833 | peptide transport |
| 3 | GO0042886 | amide transport |
| 4 | GO0071702 | organic substance transport |
| 5 | GO0071705 | nitrogen compound transport |
| 6 | GO0045184 | establishment of protein localization |
| 7 | GO0006810 | transport |
| 8 | GO0008104 | protein localization |
| 9 | GO0051234 | establishment of localization |
| 10 | GO0033036 | macromolecule localization |
| 11 | GO0051179 | localization |
| 12 | GO0006915 | apoptotic process | GO0006915@GO0009987/GO0044763/GO0008219/GO0012501/GO0006915;GO0006915@GO0044699/GO0044763/GO0008219/GO0012501/GO0006915 |
| 13 | GO0012501 | programmed cell death |
| 14 | GO0008219 | cell death |
| 15 | GO0044763 | single-organism cellular process |
| 16 | GO0009987 | cellular process |
| 17 | GO0044699 | single-organism process |
| 18 | GO0007165 | signal transduction | GO0007165@GO0009987/GO0007165;GO0007165@GO0065007/GO0050789/GO0050794/GO0007165 |
| 19 | GO0050794 | regulation of cellular process |
| 20 | GO0050789 | regulation of biological process |
| 16 | GO0009987 | cellular process |
| 21 | GO0065007 | biological regulation |
| 22 | GO0007275 | multicellular organism development | GO0007275@GO0032501/GO0044707/GO0007275;GO0007275@GO0044699/GO0044707/GO0007275GO0007275@GO0032502/GO0044767/GO0007275GO0007275@GO0044699/GO0044767/GO0007275;GO0007275@GO0032502/GO0048856/GO0007275 |
| 23 | GO0048856 | anatomical structure development |
| 24 | GO0044707 | single-multicellular organism process |
| 25 | GO0044767 | single-organism developmental process |
| 17 | GO0044699 | single-organism process |
| 26 | GO0032502 | developmental process |
| 27 | GO0032501 | multicellular organismal process |
| 28 | GO0007049 | cell cycle | GO0007049@GO0009987/GO0044763/GO0007049;/GO0007049@GO0044699/GO0044763/GO0007049 |
| 15 | GO0044763 | single-organism cellular process |
| 16 | GO0009987 | cellular process |
| 17 | GO0044699 | single-organism process |

**Dataset**

**Table S3.** The lncRNA2GO-55 dataset

|  |  |  |
| --- | --- | --- |
| 　lncRNA | 　GO term | Pubmed IDs |
| tusc7 | GO:0051726 | 23558749 |
| linc00951 | GO:0051726 | 23872665 |
| linc00951 | GO:0070883 | 23872665 |
| ccat2 | GO:0017147 | 23796952 |
| ccat2 | GO:0051726 | 23796952,24504682 |
| PCAT1 | GO:0051726 | 23640607 |
| PCAT1 | GO:0006281 | 21804560,24473064 |
| LUCAT1 | GO:0051726 | 23672216 |
| LSINCT5 | GO:0051726 | 21532345 |
| Xist | GO:0060817 | 8538762,9009199 |
| Xist | GO:0005515 | 18974356,21123648,20542000,21172659,22659877 |
| Xist | GO:0031491 | 20833368 |
| HOTAIR | GO:0045814 | 17604720,20616235 |
| HOTAIR | GO:0032947 | 20616235 |
| HOTAIR | GO:0051726 | 20393566,21637793 |
| LINC00538 | GO:0051726 | 22258142 |
| uca1 | GO:0051726 | 17416635,18501714 |
| IFNG-AS1 | GO:0010468 | 24056746 |
| Jpx | GO:1900095 | 21029862 |
| SNHG3 | GO:0019827 | 21874018 |
| SNHG3 | GO:0003682 | 21874018 |
| SNHG3 | GO:0032403 | 21874018 |
| GACAT1 | GO:0051726 | 23645148 |
| HULC | GO:0035195 | 20423907 |
| HULC | GO:0010468 | 17241883 |
| HULC | GO:0051726 | 17241883 |
| CDR1-AS | GO:0035198 | 21964070 |
| CDR1-AS | GO:0010468 | 23446346 |
| ADAMTS9-AS2 | GO:0010468 | 24833086 |
| ADAMTS9-AS2 | GO:0051726 | 24833086 |
| MALAT1 | GO:0051726 | 19625619,20797886 |
| MALAT1 | GO:0010468 | 20937273 |
| MALAT1 | GO:1990935 | 19116412,20797886 |
| MALAT1 | GO:0043484 | 20797886 |
| MALAT1 | GO:0032403 | 22659877 |
| MALAT1 | GO:0003682 | 22659877 |
| HOTAIRM1 | GO:0010468 | 19144990 |
| HOTAIRM1 | GO:0003682 | 21874018 |
| HOTAIRM1 | GO:0032403 | 21874018 |
| DHRS4-AS1 | GO:0010468 | 22891334 |
| DHRS4-AS1 | GO:0031056 | 22891334 |
| DHRS4-AS1 | GO:0035064 | 22891334 |
| SPRY4-IT1 | GO:0051726 | 21558391 |
| SPRY4-IT1 | GO:0010941 | 21558391 |
| PRNCR1 | GO:0005102 | 23945587 |
| PRNCR1 | GO:0051726 | 23945587 |
| PRNCR1 | GO:0010468 | 23945587 |
| PRNCR1 | GO:0003723 | 23945587 |
| BLACAT1 | GO:0051726 | 23688781 |
| BLACAT1 | GO:0003682 | 23688781 |
| BLACAT1 | GO:0032403 | 23688781 |
| PTCSC3 | GO:0030308 | 22586128 |
| PTCSC3 | GO:0010468 | 22586128 |
| H19 | GO:0001558 | 9203585 |
| H19 | GO:0010468 | 9203585 |
| H19 | GO:0051726 | 7692308,18719115 |
| H19 | GO:0044030 | 9203585,9294195 |
| H19 | GO:0005520 | 10875929 |
| H19 | GO:0003682 | 21172659 |
| H19 | GO:0032403 | 21172659 |
| PVT1 | GO:0051726 | 25043044 |
| PVT1 | GO:0031323 | 2725491 |
| PVT1 | GO:0010468 | 24204837 |
| HOXA11-AS | GO:0050686 | 12050232 |
| WT1-AS | GO:0003723 | 17940140 |
| WT1-AS | GO:0010468 | 17940140 |
| FOXCUT | GO:0010629 | 24889262 |
| FOXCUT | GO:0042127 | 24889262 |
| HOTTIP | GO:0048097 | 21423168 |
| HOTTIP | GO:0003682 | 21423168 |
| LINC00467 | GO:0010468 | 24586304 |
| LINC00467 | GO:0050794 | 24586304 |
| GAS5 | GO:0051726 | 18354083 |
| GAS5 | GO:0001558 | 18354083 |
| GAS5 | GO:0005102 | 20124551 |
| KCNQ1OT1 | GO:0060968 | 18848501,18951091,15516932,21172659 |
| KCNQ1OT1 | GO:0003682 | 18848501,18951091,15516932,21172659 |
| KCNQ1OT1 | GO:0032403 | 18848501,18951091,15516932,21172659 |
| KCNQ1OT1 | GO:0044030 | 20573698 |
| CCAT1 | GO:0042127 | 23143645,24594601 |
| CCAT1 | GO:0051726 | 24594601 |
| LINC00299 | GO:0007420 | 23217328 |
| TUG1 | GO:0061074 | 15797018 |
| TUG1 | GO:0002039 | 19571010 |
| TUG1 | GO:0007049 | 19571010 |
| TUG1 | GO:0003682 | 19571010 |
| TUG1 | GO:0032403 | 19571010 |
| XACT | GO:0016456 | 23334669 |
| NEAT1 | GO:0003682 | 21874018 |
| NEAT1 | GO:0032403 | 21874018 |
| NEAT1 | GO:0006259 | 21170033 |
| LINC00570 | GO:0010628 | 20887892 |
| HEIH | GO:0042127 | 21769904 |
| HEIH | GO:0051726 | 21769904 |
| HEIH | GO:0003682 | 21769904 |
| HEIH | GO:0032403 | 21769904 |
| CRNDE | GO:0019827 | 21874018 |
| CRNDE | GO:0007049 | 21874018 |
| CRNDE | GO:0019222 | 24184209 |
| CCDC26 | GO:0042127 | 16449964 |
| CCDC26 | GO:0045595 | 16449964 |
| CCDC26 | GO:0010941 | 16449964 |
| LINC00970 | GO:0010468 | 23741487 |
| LINC00970 | GO:0030178 | 23741487 |
| ATXN8OS | GO:0003723 | 14972680 |
| ATXN8OS | GO:0043523 | 14972680 |
| MYCNUT | GO:0010468 | 24906397 |
| MYCNUT | GO:0050794 | 24906397 |
| TINCR | GO:0010468 | 23201690 |
| TINCR | GO:0005515 | 23201690 |
| TINCR | GO:0003729 | 23201690 |
| CYP4A22-AS1 | GO:0010468 | 20887892 |
| NPTN-IT1 | GO:0051726 | 23395002 |
| NPTN-IT1 | GO:0042127 | 23395002 |
| NPTN-IT1 | GO:0005515 | 23395002 |
| NPTN-IT1 | GO:0043488 | 23395002 |
| CTBP1-AS | GO:0005515 | 23644382 |
| CTBP1-AS | GO:0010468 | 23644382 |
| CTBP1-AS | GO:0051726 | 23644382 |
| PCGEM1 | GO:0042981 | 16569192 |
| GHET1 | GO:0042127 | 24397586 |
| GHET1 | GO:0043488 | 24397586 |
| GHET1 | GO:0051098 | 24397586 |
| TRERNA1 | GO:0010468 | 20887892 |
| LINC-ROR | GO:0048863 | 21057500 |
| LINC-ROR | GO:0042981 | 21057500 |
| EMX2OS | GO:0010468 | 20066053 |
| DBET | GO:0006338 | 22541069 |
| DBET | GO:0010468 | 22541069 |