In silico prediction of long intergenic non-coding RNAs in sheep

Journal:	<i>Genome</i>
Manuscript ID	gen-2015-0141.R1
Manuscript Type:	Article
Date Submitted by the Author:	19-Jan-2016
Complete List of Authors:	Bakhtiarzadeh, Mohammad Reza; University of Tehran, Animal Science Hosseinpour, Batool; Iranian Research Organization for Science and Technology Arefnezhad, Babak; Royan Institute Shamabadi, Narges; University of Qom Salami, Seyed Alireza; University of Tehran
Keyword:	lncRNAs, RNA-seq, comparative genomics, sheep
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
gen-2015-0141.R1Supplement 1.gtf	

SCHOLARONE™
Manuscripts

***In silico* prediction of long intergenic non-coding RNAs in sheep**

Mohammad Reza Bakhtiarizadeh^{a,e*}, Batool Hosseinpour^{b,e}, Babak Arefnezhad^{c,e}, Narges Shamabadi^d, Seyed Alireza Salami^{e,f}

^a Department of Animal and Poultry Science, College of Aburaihan, University of Tehran, Tehran, Iran, Pakdasht, Tehran, Iran ; P.O.Box: 3391653755

^b Department of Agriculture, Iranian Research Organization for Science and Technology (IROST), P.O.Box: 33535111, Tehran, Iran

^c Department of Molecular Systems Biology, Cell Science Research Center, Royan Institute for Stem Cell Biology and Technology, ACECR, Tehran, Iran

^d Center of Environmental Researches, university of Qom, Qom, Iran

^e OMICSTM Research Group, Tehran, Iran

^f University of Tehran, Tehran, Iran

* Corresponding author: Mohammad Reza Bakhtiarizadeh

Email: mrbakhtiari@ut.ac.ir

Tel: 0098 912 852 2523

Abstract

Long non-coding RNAs (lncRNAs) are transcribed RNA molecules >200 nucleotides in length that do not encode proteins and serve as key regulators of diverse biological processes. Recently, thousands of long intergenic non-coding RNAs (lincRNAs), a type of lncRNAs, have been identified in mammals using massive parallel large sequencing technologies. The availability of the genome sequence of sheep (*Ovis aries*) has allowed us genomic prediction of non-coding RNAs. This is the first study to identify lincRNAs using RNA-seq data of eight different tissues of sheep including brain, heart, kidney, liver, lung, ovary, skin and white adipose. A computational pipeline was employed to characterize 325 putative lincRNAs with high confidence from eight important tissues of sheep using different criteria such as GC content, exon number, gene length, co-expression analysis, stability and tissue-specific scores. 64 putative lincRNAs displayed tissues-specific expression. The highest number of tissues-specific lincRNAs was found in skin and brain. All novel lincRNAs that aligned to the human and mouse lincRNAs had conserved synteny. These closest protein-coding genes were enriched in 11 significant GO terms such as limb development, appendage development, striated muscle tissue development and multicellular organismal development. The findings reported here have important implications for the study of sheep genome.

Keywords: lncRNAs; RNA-seq; comparative genomics; sheep.

Introduction

Massive large cDNA sequencing technology or RNA-seq enables comprehensive identification, annotation and quantification of transcriptome. Human genome-wide projects, such as ENCODE, GENCODE and FANTOM, have clearly shown that a considerable fraction of genome are transcribed into non-coding RNAs (ncRNAs), the fraction is much higher than previously thought (Birney et al. 2007). ncRNAs play a variety of important regulatory roles during diverse biological processes (Ulitsky and Bartel 2013). In general, there are two ncRNA classes based on their transcript's length: 1) ncRNAs shorter than 200 nucleotides (nt) are usually recognized as small/short ncRNAs, including microRNAs (miRNAs), endogenous small interfering RNAs (siRNAs), PIWI-interacting RNAs (piRNAs), ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs) (Ma et al. 2012; Skroblin and Mayr 2014); 2) ncRNAs longer than 200 nt, referring to mRNA-like long ncRNAs (lncRNAs) (Ulitsky and Bartel 2013).

Since the first report of lncRNAs in human by Lukiw et al. (Lukiw et al. 1992), lncRNAs have been found as a main class of novel regulating transcripts and have been identified in various organisms, ranging from nematode to human (Derrien et al. 2012; Nam and Bartel 2012). According to NONCODE v4.0 database (a database of literature documented lncRNAs), to date, there are 93,135 and 67,628 lncRNA entries in human and mouse, respectively (Xie et al. 2014).

lncRNAs are transcribed by RNA polymerase II, so they share many features of mRNAs. However, this is not a fast rule and some lncRNAs are transcribed by RNA polymerase III. They can be capped, polyadenylated (or non-polyadenylated) and spliced (or mono-exonic unspliced) (Skroblin and Mayr 2014; Ulitsky and Bartel 2013). Compared to most mRNAs, lncRNAs have

limited coding potential, indicating the lack of significant open reading frames (ORFs) for coding protein products. lncRNAs show a lower sequence conservation and mostly display a tendency to be expressed at low levels and in a specific-tissue manner (often mistakenly considered as transcriptional noises). In addition, they are mostly localized in the nucleus and only a few lncRNAs can be detected in the cytoplasm (Skroblin and Mayr 2014). Despite the fact that there is poor interspecies conservation and low expression of lncRNAs in different tissues, the tissue-specific nature of lncRNAs suggests that they have been involved in various biological processes (Derrien et al. 2012). Recently, it has been found that several human lncRNAs interact with chromatin remodeling complexes; moreover, recent studies indicate important implications of lncRNAs in dosage compensation, gene imprinting, cell fate specification, cell cycle and apoptosis, RNA processing (transcription, splicing and translation), protein localization, stem cell pluripotency and reprogramming, heat shock response and development of different human diseases (Ma et al. 2013; Yan and Wang 2012). For instance, *KCNQ1OT1* and *Air* lncRNAs are required for silencing autosomal imprinted genes by recruiting chromatin modifying machinery (Korostowski et al. 2012). In bovine, ectopic overexpression of a long intergenic ncRNA (lincRNA) suggested that this type of lncRNA may have a regulatory role in horn bud differentiation (Allais-Bonnet et al. 2013).

There are different methods to categorize lncRNAs based on their various characteristics. On the basis of their genomic proximity to the nearest protein-coding genes, lncRNAs fall into four categories: 1) sense and antisense lncRNAs, are locating on the same strand and on the opposite strand (antisense strand) of a nearest protein-coding genes, respectively; 2) bidirectional lncRNAs that share promoters with protein-coding genes, though they are transcribed from the opposite direction; 3) intronic lncRNAs, locating on the introns of protein-coding genes; and 4)

long intergenic non-coding RNAs which are not in the proximity of a protein-coding gene at all (Ma et al. 2013; Skroblin and Mayr 2014). lincRNAs are more appropriate for experimental manipulation and computational analysis with no interference of annotated protein-coding regions than other lncRNAs (Cabili et al. 2011). They also participate in many different biological processes from embryonic stem cell pluripotency to cell proliferation and cancer progression (Wang et al. 2014). For instance, *XIST*, *HOTAIR* and *H19* are some of the best known lincRNAs (Allais-Bonnet et al. 2013).

Recent studies have demonstrated that the number of lincRNAs is at least twice the number of protein-coding genes in mammalian genomes, most of which are still undiscovered. RNA-Seq is a widely used high-throughput technology to transcriptome profiling of rare transcripts and detecting novel RNAs, such as lncRNAs, with no need of gene annotations (Lv et al. 2013). This technology has been applied to identify thousands of lncRNAs in many species including human (Wang et al. 2014), mouse (Luo et al. 2013), cattle (Weikard et al. 2013), chicken (Li et al. 2012), Zebrafish (Kaushik et al. 2013), maize (Li et al. 2014b) and *Caenorhabditis elegans* (Nam and Bartel 2012). Luo *et al.* predicted 3,965 putative lincRNAs genes across multiple mouse tissues (Luo et al. 2013). By using a computational approach, Li *et al.* identified 281 new lincRNAs in chicken muscle (Li et al. 2012). Weikard *et al.* predicted more than 4,000 potential lncRNAs in bovine skin (Weikard et al. 2013). More recently, Billerey *et al.* reported 584 different lincRNAs in bovine muscle (Billerey et al. 2014). Recent studies revealed that there are still many novel lncRNAs to be discovered for the well-studied transcriptomes like human and mouse (Derrien et al. 2012; Luo et al. 2013), then, more efforts are necessarily needed to discover all other lncRNAs.

To the best of our knowledge, there is no available report on lincRNAs and their biological functions in sheep. Deciphering of lincRNAs and their expression profiles in different tissues of sheep would enable us to get a better understanding of the regulatory function and genome annotation of lincRNAs in sheep.

In this study, we used publicly available RNA-Seq data of eight different tissues of sheep for the identification and preliminary characterization of lincRNAs applying a computational pipeline. Later on, gene expression profiling was performed on identified lincRNAs in the tissues, so that the first repertoire of lincRNAs expressed in sheep was generated. The identified lincRNAs provide a basis for an expanded understanding of lincRNAs in farm animals, and a deeper functional annotation in sheep genome.

Material and methods

Datasets

All PolyA⁺ RNA-seq data of sheep tissues were downloaded from the Gene Expression Omnibus (GEO) database (accession number GSE56643). Deposited non-strand specific RNA-seq data, from eight distinct sheep tissues, including brain, heart, kidney, liver, lung ovary, skin and white adipose had been generated by Illumina HiSeq 2000 platform. 16 million reads per sample, on average, were generated by each tissue, and the length of paired reads was 76 base pairs (bp). Skin tissue belonged to a Gansu alpine fine wool female sheep and the other seven tissues were obtained from the reference female Texel.

Pipeline for identifying lincRNAs

A computational pipeline was used to detect the putative lincRNAs of each tissue (Figure 1). The pipeline minimizes false positives and maximizes the prediction of true putative lincRNAs by implementing the following steps:

RNA-Seq read mapping and transcriptome assembly

The Alignment and analysis were performed using the Tuxedo Suite which contains Tophat, Cufflinks, Cuffcompare and Cuffnorm programs. After discarding the low-quality reads (the quality values less than 20 and the reads less than 40 bp in length) and trimming the adaptor sequences, trimmed reads of each tissue were independently aligned with sheep reference genome (Ova ver. 3.1 from Ensembl database) (Jiang et al. 2014) using the spliced read aligner TopHat (ver. 2.0.13) (Kim et al. 2013). By removing the reads mapped to the mitochondrial genome (ChrM), the transcriptome of each tissue was assembled separately using Cufflinks (ver. 2.2.1) in *de novo* mode, with the sheep Ensembl annotations (release 78). Subsequently, Cuffmerge program was used to merge transcriptome data from all the eight tissues to generate a reference transcriptome. Then, the lincRNAs detection pipeline was applied to filter the merged assembly.

The expression levels (FPKM, Fragments Per Kilobase of transcript per Million mapped reads) of all transcripts in the eight tissues were measured and normalized for tissue-specific expression using Cuffnorm program.

Classification of unknown transcripts

To obtain the putative novel non-coding transcripts and filtered out the known transcripts, the unique transcriptome dataset was compared with the Ensembl sheep genome annotation by using

the Cuffcompare program. Consequently, the assemblies that matched annotations can be identified; moreover, the new transcripts can be clustered into different classes based on their locations relative to the known genes. One of which was ‘i’ class contained transcripts locating within an intron of a known gene; the ‘o’ class contained the transcripts having overlap between generic exonic and a known transcript; the ‘x’ class contained transcripts that having overlap between exons and known transcripts; meanwhile, on the opposite strand; the ‘u’ class contained intergenic transcripts. This study mainly focused on lincRNAs, avoiding the complications arising from other types of genes overlap. Therefore, from all classes of transcripts, ‘u’ class code were screened for putative lincRNAs. This subset was defined as candidate lincRNAs and such transcripts can potentially include putative lincRNAs.

Evaluation of size and exon numbers

As previously stated, an arbitrarily selected cut-off of 200 nt is commonly used to distinguish lincRNAs from small ncRNAs. So, only candidate lincRNAs which are longer than or equal to 200 bp were retained. Also, to yield high-quality dataset of putative lincRNAs, single exon candidate lincRNAs were filtered out.

Comparative sequence analysis

To discard the protein coding sequences and ncRNAs (including rRNA, tRNA, snRNA, snoRNA and miRNAs) from candidate lincRNAs sequences, which are not yet annotated in sheep genome, sequence homology searches on candidate lincRNAs sequences were performed using BLAST on several different publicly available databases. First, BLASTX search was performed using candidate lincRNA sequences as the query and the Uniref 90 database was set as target.

Then, Rfam database (ver.12) was searched by BLASTN to remove ncRNAs, including rRNA, tRNA, snRNA and snoRNA. Finally, the sequences with no hits at E-value $1e-5$ were kept and searched against all the known metazoan precursors' miRNAs in miRBase database (ver. 21) using BLASTN. Subsequently, the transcripts candidate lincRNAs which was blasted against only one of the three mentioned databases (E value $1e-5$), were eliminated from candidate lincRNAs list and the remaining candidate lincRNAs were served as input for subsequent analysis.

Evaluation of coding potential

Protein-coding potential of candidate lincRNAs was analysed by integrating the results of three softwares including; CPC (Coding Potential Calculator), CNCI (Coding Noncoding Index) and PLEK (Predictor of Long non-coding RNAs and mEssenger RNAs based on an improved K-mer scheme) (Li et al. 2014a). The overlap between the results, owing to different sensitivity and specificity of softwares, was small (Li et al. 2014a). Therefore, all of the softwares were used to reduce false positive results and thereby more reliable putative lincRNAs were attained. CPC has been widely used to discover lncRNAs. CPC incorporates six biologically meaningful sequence features into a support vector machine to predict the protein-coding potential of transcripts; transcripts with a score of <-0.5 were considered as non-coding transcripts. CNCI, is a robust signature software to find adjoining nucleotide triplets (ANT) to discriminate protein-coding from non-coding sequences with no need of known annotations. PLEK is a powerful alignment-free computational tool to distinguish lncRNAs from protein-coding transcripts (Li et al. 2014a). In order to reduce false-positive results as well as extracting potential non-coding transcripts

with a high reliability, only the candidate lincRNAs sequences which were predicted as non-coding transcripts by three softwares, were considered for further analysis.

Evaluation of ORFs

As ncRNAs do not encode functional proteins (Lv et al. 2013), candidate lincRNAs were screened for transcripts either lacking complete or short ORF with no homology to known proteins. Later on, transcripts with ORF longer than 300 nt were eliminated by using Getorf software in the EMBOSS package.

Evaluation of candidate lincRNAs distances with nearest protein-coding genes

In the last step, candidate lincRNAs, which were located in a distance lower than 1000 bp to a known protein-coding gene were omitted. Finally, the remaining candidate lincRNAs were defined as putative lincRNAs.

Comparisons of GC content, exon number, gene, expression level, transcript and exon length

The putative lincRNAs were compared with 20921 protein-coding genes which were extracted from the Ensembl gene annotation (sheep genome v3.1) in terms of GC content, exon number, gene, transcript and exon length. Moreover, the quantified expression levels (FPKM) of the putative lincRNAs, for all of the tissues, were compared with those of known protein-coding genes; 14331 expressed protein-coding genes from Cufflinks results were extracted.

Tissue specificity score and co-expression analysis

To assess the tissue specificity of transcripts, “rsgcc” package of R software was used. This package calculates the tissue specificity score by using the formula $1 - \min(R(1), R(2), \dots, R(i), \dots, R(n))$, where $R(i) = M(i)/E(i)$, $E(i)$ is the mean expression value of tissue i , and $M(i)$ is the maximal expression value of other tissues. If the tissue specificity score is higher than TS threshold, the gene is considered as tissue specifically expressed. Here, the mean expression value was used and TS Threshold was set to 0.75. In addition, the gene expression values were scaled across tissues. Using the FPKM levels, the Pearson correlation coefficient was calculated for each protein-encoding gene with each putative lincRNAs as the co-expression measurement, using CoExpress software. CoExpress can be applied to build pairwise gene co-expression matrices. The correlation coefficients were further validated by bootstrapping algorithm. For each lincRNA, protein coding genes with validated correlation coefficient ≥ 0.9 were taken as the pair with significant expression correlation.

Stability evaluation

The stability of putative lincRNAs and protein-coding genes was evaluated based on the minimum free energy of secondary structure and calculated by RNAfold software from the Vienna RNA package (Lorenz et al. 2011).

Homology search for putative lincRNAs

To identify the homologs of sheep putative lincRNAs in humans and mouse, putative lincRNAs were aligned with human and mouse lincRNAs, using BLASTN (E value $1e-5$). The sequences of human (release 21) and mouse (release M4) lincRNAs were downloaded from GENCODE database.

Conservation analysis of putative lincRNAs

LiftOver software (a command-line executable software from UCSC database) was applied to compare the conservation properties of putative sheep lincRNAs with protein-coding genes. The liftOver utility translates genomic coordinates between different genomes. The exonic regions of lincRNAs and protein-coding genes were mapped to genomes of five different species (human, mouse, cow, pig and horse), according to genome alignments. The genome assembly build versions used in the study were hg19 for human, Mm10 for mouse, BosTau8 for cow, EquCab2 for horse and SusScr3 for pig. Liftover chains were extracted from UCSC nets generated from blastZ alignments. The minimum ratio of bases that must be remapped was set to 0.80 and 0.40. In other words, in case of genome alignments between any species and sheep sequences (lincRNAs or protein-coding genes), with covering >80% (or >40%) base pairs of sheep sequences, the aligned regions within the specified specie were taken as sheep sequences orthologs. The conservation percentage of exonic regions of lincRNAs and protein-coding genes were compared in each different genome species.

Results

Mapping and transcriptome assembly

RNA-seq data of brain, heart, kidney, liver, lung, ovary, skin and white adipose comprised a total of 126,224,726 raw paired-end reads with a length of 75 bp. Approximately, 105 million trimmed reads (91.09 %) were successfully aligned to sheep genome (Table 1). The assembled transcripts of eight tissues got merged so that a unique dataset of 94,761 nonredundant transcript isoforms were obtained from 56,181 unique genes. Then, the unique dataset of transcripts were

divided into different classes, based on their relative positions. Interestingly, about 28% (26,282 transcripts) of unique dataset were considered as unknown intergenic transcripts (“u” class) and used for further analysis.

Identification of putative lincRNAs

Putative lincRNAs were discriminated from other types of transcripts, like protein coding and small RNA transcripts, in seven steps. First, after discarding candidate lincRNAs based on a minimum length threshold of 200 nt and single exon transcripts, 2546 transcripts were acquired. Second, 1,518 of candidate lincRNAs sharing significant homology with a protein-coding sequence were removed so that 1,028 candidates passed filtering. Third, a set of 971 candidate lincRNAs, with no sequence similarity to known classes of ncRNAs, was retained. Fourth, by comparing miRNA sequences (miRBase) and candidate lincRNAs having significant homology (E value $1e-5$), 959 candidate lincRNAs remained. Fifth, 147 of candidates were predicted to possess coding potential and 812 were assumed as non-coding lincRNAs (related to 715 genes) (Figure 2). Sixth, 436 candidate lincRNAs having uncertain coding potential with maximal ORF >300 nt were removed, as they could potentially encode one hundred or more amino acids producing potential small peptides. As a result, 376 candidate lincRNAs remained. Finally, any candidate lincRNAs having overlap <1000 bp with a protein-coding gene were removed. Approximately, 14% of the candidate lincRNAs (51) were found to be located within a 1000 nt distance to an annotated protein-coding gene (upstream or downstream). At last, a total of 325 candidate lincRNAs representing potential lincRNAs were identified as putative lincRNAs in sheep genome (Supplementary file 1). Also, about 30% of putative lincRNAs (98) overlapped annotated protein-coding gene within a distance of 10000 nt.

Characterization of putative lincRNAs

The chromosomal locations of putative lincRNAs have been shown in Figure 3. The distribution of lincRNAs in sheep chromosomes is consistent with the chromosomal distribution of known lincRNAs in human, mouse (based on GENCODE database) and bovine genome (Billerey et al. 2014). The GC content of protein-coding genes was significantly higher than predicted lincRNA (44.6% vs 48.1%; Mann-Whitney test, $P < 3.167\text{e-}09$) (Figure 4a) and is conserved in the lincRNA of plant (Zhang et al. 2014) and animal (Huang et al. 2012; Nam and Bartel 2012). Such conservation may suggest its impact on their function. Previous reports in mammals have demonstrated that genes encoding lincRNAs are shorter in length, and have shorter transcripts as well as fewer exons than protein-coding genes (Billerey et al. 2014; Li et al. 2012; Weikard et al. 2013). The results showed that the average length of putative lincRNAs genes was 3.3-fold shorter than the average length of protein-coding genes (14,052.5 vs 38,875.8 nt; Mann-Whitney test, $P < 2.2\text{e-}16$) (Figure 4b). In addition, the putative lincRNAs represent much shorter transcript in length, on average, than protein-coding transcripts (755.7 vs 1943.3 nt; Mann-Whitney test, $P < 2.2\text{e-}16$) (Figure 4c). Also, About 19% of putative lincRNAs (62) were longer than 1000 nt. It has been reported that the longer human lincRNAs are associated with chromatin-modifying complexes regulating gene expression (Khalil et al. 2009). Moreover, the putative lincRNAs had fewer exons per transcript than protein-coding genes (2.4 vs 9.7; Mann-Whitney test, $P < 2.2\text{e-}16$) (Figure 4d). In consistent with GENCODE v7 project (Derrien et al. 2012), most of the putative lincRNAs displayed a striking tendency to have only two exons (75% of putative lincRNAs have only two exons compared to 9% of protein-coding genes). Moreover, similar to human lincRNAs (Derrien et al. 2012), the exon length of putative lincRNAs, on average, were longer than those of protein-coding transcripts (310.4 vs 197.1 nt; Mann-Whitney

test, $P < 2.2e-16$). The results are in full agreement with the findings in bovine (Billerey et al. 2014), zebrafish (Pauli et al. 2012) and human (Cabili et al. 2011), providing evidences to confirms that the candidates are indeed lincRNAs.

261 lincRNAs were expressed in more than one tissue (out of 325 putative lincRNAs). The other 64 novel lincRNAs displayed a specific expression in an individual tissue and were marked as tissue-specific lincRNAs (Supplementary file 2, Table 2). The maximum number of expressed lincRNAs was found in skin (209) followed by brain (208), kidney (189), ovary (188), lung (184), white adipose (179), liver (151) and heart (141). The results showed that the majority of the putative lincRNAs were expressed in more than one tissue. Also, the highest number of tissue-specific lincRNAs was found in skin (25) and brain (17), respectively.

The expression levels of putative lincRNAs and protein-coding genes in different tissues, have been demonstrated in Figure 5. The maximal expression levels of protein-coding genes were higher than those of putative lincRNAs across the eight tissues. The mean expression level of lincRNAs was lower than that of mRNAs in all of tissues, except in heart and lung. In other words, there was a lincRNA with a high expression in heart and lung tissues. By removing such lincRNA from heart and lung, the mean expression of lincRNAs in both tissues became similar to other tissues (Figure 5). Notably, putative lincRNAs genes were expressed at a low level, as 77% have a maximum FPKM <5 across all tissues on average (supplementary file 3). The expression level of putative lincRNAs was significantly lower than that of protein-coding genes

(24.9 vs 43.8, Mann-Whitney test, $P < 2.776 \times 10^{-8}$), indicating a common feature in lincRNAs (Billerey et al. 2014; Birney et al. 2007; Li et al. 2012; Lv et al. 2013; Weikard et al. 2013).

It was found that the tissue specificity scores of putative lincRNAs were significantly higher than those of protein-coding genes (0.70 vs 0.53, Mann-Whitney test, $P < 2.2 \times 10^{-16}$) (Figure 6). In agreement with previous reports (Derrien et al. 2012; Guttman et al. 2009; Luo et al. 2013), the findings showed that most of lincRNAs (56%) were tissue-specific, compared to protein-coding genes (only 32%). To understand whether the low expression levels of lincRNAs are not the cause of these differences, the tissue specificity scores were calculated for putative lincRNAs and protein-coding genes with expression level of 5-20 FPKM. The results again confirmed previous findings that putative lincRNAs had significantly higher tissue specificity score than protein-coding genes (0.72 vs 0.42, Mann-Whitney test, $P < 3.661 \times 10^{-13}$). Furthermore, the findings suggest that lincRNAs expressions are under more specific regulation in sheep tissues.

Co-expression analysis showed that 510 pairs of novel lincRNA/protein-coding genes have highly correlated expressions. Two of which were belonged to anti-correlated novel lincRNA/protein-coding gene. The gene expression of 125 different novel lincRNA (out of 510 pairs) and 80 different protein-coding were highly correlated. Some protein-coding genes were correlated with more than one novel lincRNA. About 2 % of putative lincRNAs (8) showed correlations with more than 10 protein-coding genes (Supplementary file 4). In consistent with some previous studies (Billerey et al. 2014; Cabili et al. 2011; Pauli et al. 2012), most of the putative lincRNAs and their overlapping protein-coding genes had a low co-expression correlation (< 0.9), whereas only 124 putative lincRNAs showed a co-expression correlation more than 0.9.

There are different methods for predicting putative function of lincRNAs, but these methods are still in their infancy, and generally, lincRNAs have been annotated based on their proximity to protein-coding genes and their co-expressed protein-coding genes (Guttman et al. 2009; Ilott and Ponting 2013; Weikard et al. 2013). Here, to determine the putative function of lincRNAs, two methods were used. At first, for identifying the putative function of lincRNAs, based on their co-expressed protein-coding genes, it have been focused on the lincRNAs which had more than 10 co-expressed protein-coding genes (8 lincRNAs) and assigned the corresponding GO terms of these protein-coding genes as the annotations of this lincRNA (using DAVID database). This analysis identified several significant GO terms, such as translational elongation, translation, cellular protein metabolic process and gene expression for each lincRNA (Supplementary file 4). It has been revealed that lincRNAs have important roles in transcriptional regulation and translational control (Ma et al. 2013). Second, the putative functions of lincRNAs based on their closest protein-coding genes were predicted. Among the putative lincRNAs, 140 (43%) were located within 25 kb away and 185 (57%) were at least 25 kb away from the nearest protein-coding genes. Then, GO analysis of the closest protein-coding genes (<25 kb) of putative lincRNAs were investigated to consider whether they are enriched in specific GO function terms (biological processes). Earlier studies have revealed that mammalian lincRNAs are preferentially located next to genes with developmental functions (Cabili et al. 2011; Guttman et al. 2009; Pauli et al. 2012). Interestingly, in consistent with these studies, our results showed that these closest protein-coding genes were enriched in 11 significant GO terms such as limb development, striated muscle tissue development and multicellular organismal development. All enriched GO terms in biological processes are provided in Supplementary file 5.

In this study, minimum free energy was considered as a measurement of stability. The results, in agreement with previous studies (Aiso et al. 2005; Wang et al. 2014), showed that putative lincRNAs had significantly higher minimum free energy than protein-coding genes (-199.3 vs -536.1, Mann-Whitney test, $p < 2.2e-16$) (Figure 7). Also, it has been reported that the length of RNA molecule affects the minimum free energy of secondary structure as the longer sequences are more stable, on average (Trotta 2014). As explained above, protein-coding genes were longer than putative lincRNAs and the putative lincRNAs had significantly lower stability than protein-coding genes (-200.3 vs -400.1, Mann-Whitney test, $P < 2.2e-16$), demonstrating that the identified significant difference is not the result of the longer length of protein-coding genes compared to putative lincRNAs.

Although, lincRNAs have lower sequence conservation than protein-coding genes, recent studies showed that some of lincRNAs are evolutionary conserved, indicating their essential functions (Ma et al. 2012; Qu and Adelson 2012). Six (~2%) putative lincRNAs had significant homology with human lincRNAs and the other 6 (~2%) putative lincRNAs showed homology with mouse lincRNAs. Moreover, 759 (~3%) human lincRNAs had significant homology with mouse lincRNAs, using the same method. Interestingly, all novel lincRNAs that showed alignment to the human and mouse lincRNAs had conserved synteny; For example, novel lincRNA CUFF.38336 was aligned to human lincRNA “AC005550.4” and mouse lincRNA “Gm29007”. *MEOX2* and *ISPD* are located in the vicinity of upstream and downstream of protein-coding genes of AC005550.4, respectively. *ISPD* and *MEOX2* are also located in the vicinity of upstream and downstream of protein-coding genes of Gm29007, respectively. Interestingly, *MEOX2* and *ISPD* are located in the vicinity of upstream and downstream of protein-coding genes of novel lincRNA CUFF.38336, respectively. These findings provided strong supports to

confirm the results of this study that assert the putative lincRNAs could represent biologically relevant sequences (supplementary file 6).

Previous reports indicated that, in spite of low sequence conservation of lincRNAs, they share conserved genomic locations or synteny (Tan et al. 2013; Ulitsky and Bartel 2013; Ulitsky et al. 2011). Interestingly, our study identified 39 (12%) putative lincRNAs that shared the same upstream and downstream protein-coding genes with human lincRNAs; 35 of which were located on the same strand as one or both of their neighboring genes (Supplementary file 7).

The results of conservation analysis showed that 0.69%, 0.27%, 0.96%, 0.75% and 0.68% of exonic regions of putative lincRNAs and 0.94%, 0.88%, 0.97%, 0.93% and 0.89% of exonic regions of protein-coding genes in sheep have orthologous regions, with 0.80 identity, in human, mouse, cow, horse and pig genomes, respectively. The same trend was obtained, with 0.40 identity, as 0.80%, 0.36%, 0.98%, 0.82% and 0.76% of exonic regions of putative lincRNAs and 0.96%, 0.90%, 0.98%, 0.95% and 0.92% of exonic regions of protein-coding genes in sheep have orthologous regions in human, mouse, cow, horse and pig genomes, respectively. As expected, putative lincRNAs showed lower conservation percentage than protein-coding genes in comparison with all five species. These results were consistent with previous reports, suggesting that putative lincRNAs are less conserved than protein-coding genes (Li et al. 2012; Luo et al. 2013; Lv et al. 2013; Weikard et al. 2013).

Discussion

To the best of our knowledge, there is no catalog of sheep lincRNAs available, although human, mouse, cattle, zebrafish and some other animal catalogs are already accessible. For the first time, RNA-Seq data was used to identify the polyadenylated transcripts, followed by computational analysis to detect putative lincRNAs across eight tissues of diverse function in sheep. A total of 94,761 transcripts were reconstructed from data. In comparison with 27099 sheep transcripts (based on ESEMBL gtf file for sheep), the number of identified transcripts, by assembled transcriptome, increased three-fold (94,761) due to both novel isoforms of known genes as well as new genes. 26,282 transcripts (out of 94,761) were predicted as unknown intergenic transcripts and analyzed to identify putative lincRNAs using computational methods. The data indicates that only small numbers of new transcripts can be annotated due to incomplete annotation of sheep genome. The present study is mainly focused on lincRNAs, because the lincRNAs have simpler surrounding transcript structure than other types of lincRNAs overlapping genic regions. Furthermore, multi-exon lincRNAs were considered to ensure annotations of high-confidence lincRNA candidates.

The identification and characterization of lincRNAs needs both experimental and computational analyses (Ilott and Ponting 2013; Li et al. 2012; Ulitsky and Bartel 2013). Then, a highly stringent filtering pipeline was employed to minimize false positive and maximize the prediction of lincRNAs, aimed at removing transcripts with evidence for protein-coding potential. Several important criteria were applied to discriminate lincRNAs from other types of transcripts, including transcript length, exon number, homology with known genes (protein-coding or small ncRNAs), coding potential of transcripts, ORF size as well as proximity to known protein-coding genes.

We identified 325 putative lincRNAs with high confidence across eight important tissues of sheep. The RNA-seq library sizes were not that large (16 million reads per tissue, in average); hence the identification of lincRNA with low expression may have been limited. In agreement with similar studies on different organisms, the identified putative lincRNAs represent lower GC content, fewer exon number, shorter gene and transcript length, longer exon length, lower expression, lower stability, more tissue-specific than protein-coding genes and less conserved than protein-coding genes (Billerey et al. 2014; Wang et al. 2014). Also, the numbers of putative lincRNAs in this study were in line with previous studies in cattle (Billerey et al. 2014), chicken (Li et al. 2012) and Zebrafish (Kaushik et al. 2013). The findings indicated that most of identified putative lincRNAs in present study are likely to be genuine candidates.

Our study showed that 64 putative lincRNAs display tissues-specific expression, suggesting specific biological function in each tissue type. The results also showed that the highest number of tissue-specific lincRNAs was in skin (25) and brain (17) was in the second place; an early study has been reported that a large fraction of human tissue-specific lincRNAs are expressed in brain tissue (Derrien et al. 2012; Lv et al. 2013; Ulitsky and Bartel 2013). Most of putative lincRNAs (261) showed expression in more than one tissue, suggesting that they may be important regulators of protein-coding genes required for maintenance of the corresponding tissues. Furthermore, calculating a tissue specificity score for each putative lincRNA showed that the expression pattern of sheep lincRNAs tends to be more tissue-specific than protein-coding genes, which fully confirmed previous studies (Derrien et al. 2012; Guttman et al. 2009; Luo et al. 2013; Marques and Ponting 2009). The findings also showed that the expression of lincRNAs is substantially varied among different tissues which suggests that the expression of lincRNAs is highly regulated (Cabili et al. 2011).

It is well known that the secondary structure of protein-coding genes and lincRNAs is one of the important determinant of their stability (Aiso et al. 2005; Wang et al. 2014). It has been demonstrated that there is more stability of protein-coding genes rather than lincRNAs (Seffens and Digby 1999; Wang et al. 2014). Our findings confirmed previous studies that protein-coding genes have more stable secondary structure than lincRNAs, according to comparing their minimum free energy. However, low stability does not mean lack of function (Clark et al. 2012). The stability of each protein-coding gene is closely related to its physiological function. It has been revealed that RNAs (coding or non-coding genes) with high stability are involved in housekeeping functions, and RNAs with low stability have regulatory function (Clark et al. 2012).

Previous studies have revealed that not only the genomic location of lincRNAs is not random, but also they tend to act in *cis* with neighboring protein-coding genes (Ponjavic et al. 2009; Sun et al. 2013). Interestingly, in line with these studies, 39 putative lincRNAs were identified in present study that shared the same syntenic region with human lincRNAs. Syntenically conserved transcripts across phylogenetically different mammalian species may have functional roles in these species (Khachane and Harrison 2010). Therefore, these putative lincRNAs might have close functional relationships with overlapping protein-coding genes.

This study, for the first time, has generated the catalog of 325 sheep putative lincRNAs. The list of putative lincRNAs is available as a GTF file (Supplementary file 1). In this study, a cutoff of 300 nt was set to discard transcripts with coding potential. However, there are some previously characterized lincRNAs in other species that possess potential for coding peptides >100 amino acids in length (e.g. XIST with 136 amino acids (Duret et al. 2006) and HOTAIR with 106

amino acids (Ranganna et al. 2013)), but they don't function as proteins. Therefore, although the cutoff is more stringent filtering, it eliminates the lincRNAs having long putative ORF (>300 nt). It has been reported that some lincRNAs would be precursors of small ncRNAs (Harrow et al. 2012; Pauli et al. 2012). Generating miRNAs through sequential cleavage of lincRNAs, or producing Piwi-interacting RNAs (piRNAs) by processing a single lincRNA transcript have been previously revealed (Ma et al. 2012). Here, we identified 12 candidate lincRNAs having significant homology with miRNAs, all of which belonged to cattle, except one which belonged to human. However, there is no convincing evidence to classify these genes into precursor's lincRNAs for small ncRNAs or precursor's miRNAs. Therefore, to avoid false positive results, candidate lincRNAs with homology to small ncRNAs (Rfam and miRBase databases) were removed.

It is suggested that the predicted putative lincRNAs should be confirmed by experimental evidences. We also used a PolyA⁺ RNA-Seq data which are selected for polyadenylated transcripts and therefore, some lincRNAs lacking polyadenylation might have been missed.

This study provides an evidence for lincRNA content of eight different tissues in sheep which is a starting point for understanding of their regulatory mechanism. The identification of the novel lincRNAs have greatly improved the genome annotation of sheep. Also we believe that the such putative lincRNA may help to a better understanding of the biological basis of regulatory interactions amongst mRNA, miRNA and lincRNA (Le et al. 2014). To the best of our knowledge, this is the first report on lincRNAs of sheep and would also encourage experimental analysis to elucidate the function and identification of more lincRNAs in sheep.

Conflict of interest: Authors declares that they have no conflict of interest. This article does not contain any studies with animals performed by any of the authors.

References

- Aiso, T., Yoshida, H., Wada, A., and Ohki, R. 2005. Modulation of mRNA stability participates in stationary-phase-specific expression of ribosome modulation factor. *Journal of bacteriology* **187**(6): 1951-1958.
- Allais-Bonnet, A., Grohs, C., Medugorac, I., Krebs, S., Djari, A., Graf, A., Fritz, S., Seichter, D., Baur, A., and Russ, I. 2013. Novel insights into the bovine polled phenotype and horn ontogenesis in Bovidae.
- Billerey, C., Boussaha, M., Esquerré, D., Rebours, E., Djari, A., Meersseman, C., Klopp, C., Gautheret, D., and Rocha, D. 2014. Identification of large intergenic non-coding RNAs in bovine muscle using next-generation transcriptomic sequencing. *BMC genomics* **15**(1): 499.
- Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., and Stamatoyannopoulos, J.A. 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**(7146): 799-816.
- Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes & development* **25**(18): 1915-1927.
- Clark, M.B., Johnston, R.L., Inostroza-Ponta, M., Fox, A.H., Fortini, E., Moscato, P., Dinger, M.E., and Mattick, J.S. 2012. Genome-wide analysis of long noncoding RNA stability. *Genome research* **22**(5): 885-898.

- Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., and Knowles, D.G. 2012. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome research* **22**(9): 1775-1789.
- Duret, L., Chureau, C., Samain, S., Weissenbach, J., and Avner, P. 2006. The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. *Science* **312**(5780): 1653-1655.
- Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., and Cassady, J.P. 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**(7235): 223-227.
- Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., and Searle, S. 2012. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome research* **22**(9): 1760-1774.
- Huang, W., Long, N., and Khatib, H. 2012. Genome-wide identification and initial characterization of bovine long non-coding RNAs from EST data. *Animal genetics* **43**(6): 674-682.
- Ilott, N.E., and Ponting, C.P. 2013. Predicting long non-coding RNAs using RNA sequencing. *Methods* **63**(1): 50-59.
- Jiang, Y., Xie, M., Chen, W., Talbot, R., Maddox, J.F., Faraut, T., Wu, C., Muzny, D.M., Li, Y., and Zhang, W. 2014. The sheep genome illuminates biology of the rumen and lipid metabolism. *Science* **344**(6188): 1168-1173.
- Kaushik, K., Leonard, V.E., Shamsudheen, K., Lalwani, M.K., Jalali, S., Patowary, A., Joshi, A., Scaria, V., and Sivasubbu, S. 2013. Dynamic Expression of Long Non-Coding RNAs (lncRNAs) in Adult Zebrafish. *PloS one* **8**(12): e83616.
- Khachane, A.N., and Harrison, P.M. 2010. Mining mammalian transcript data for functional long non-coding RNAs. *PLoS One* **5**(4): e10316.
- Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Morales, D.R., Thomas, K., Presser, A., Bernstein, B.E., and van Oudenaarden, A. 2009. Many human large intergenic noncoding RNAs associate

with chromatin-modifying complexes and affect gene expression. *Proceedings of the National Academy of Sciences* **106**(28): 11667-11672.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**(4): R36.

Korostowski, L., Sedlak, N., and Engel, N. 2012. The Kcnq1ot1 long non-coding RNA affects chromatin conformation and expression of Kcnq1, but does not regulate its imprinting in the developing heart. *PLoS genetics* **8**(9): e1002956.

Le, T.D., Liu, L., Zhang, J., Liu, B., and Li, J. 2014. From miRNA regulation to miRNA-TF co-regulation: computational approaches and challenges. *Briefings in bioinformatics*: bbu023.

Li, A., Zhang, J., and Zhou, Z. 2014a. PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC bioinformatics* **15**(1): 311.

Li, L., Eichten, S.R., Shimizu, R., Petsch, K., Yeh, C.-T., Wu, W., Chetoor, A.M., Givan, S.A., Cole, R.A., and Fowler, J.E. 2014b. Genome-wide discovery and characterization of maize long non-coding RNAs. *Genome biology* **15**(2): R40.

Li, T., Wang, S., Wu, R., Zhou, X., Zhu, D., and Zhang, Y. 2012. Identification of long non-protein coding RNAs in chicken skeletal muscle using next generation sequencing. *Genomics* **99**(5): 292-298.

Lorenz, R., Bernhart, S.H., Zu Siederdisen, C.H., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. 2011. ViennaRNA Package 2.0. *Algorithms for Molecular Biology* **6**(1): 26.

Lukiw, W., Handley, P., Wong, L., and McLachlan, D.C. 1992. BC200 RNA in normal human neocortex, non-Alzheimer dementia (NAD), and senile dementia of the Alzheimer type (AD). *Neurochemical research* **17**(6): 591-597.

Luo, H., Sun, S., Li, P., Bu, D., Cao, H., and Zhao, Y. 2013. Comprehensive characterization of 10,571 mouse large intergenic noncoding RNAs from whole transcriptome sequencing. *PloS one* **8**(8): e70835.

- Lv, J., Cui, W., Liu, H., He, H., Xiu, Y., Guo, J., Liu, H., Liu, Q., Zeng, T., and Chen, Y. 2013. Identification and characterization of long non-coding RNAs related to mouse embryonic brain development from available transcriptomic data. *PloS one* **8**(8): e71152.
- Ma, H., Hao, Y., Dong, X., Gong, Q., Chen, J., Zhang, J., and Tian, W. 2012. Molecular mechanisms and function prediction of long noncoding RNA. *The Scientific World Journal* **2012**.
- Ma, L., Bajic, V.B., and Zhang, Z. 2013. On the classification of long non-coding RNAs. *RNA biology* **10**(6): 925-934.
- Marques, A.C., and Ponting, C.P. 2009. Catalogues of mammalian long noncoding RNAs: modest conservation and incompleteness. *Genome Biol* **10**(11): R124.
- Nam, J.-W., and Bartel, D.P. 2012. Long noncoding RNAs in *C. elegans*. *Genome research* **22**(12): 2529-2540.
- Pauli, A., Valen, E., Lin, M.F., Garber, M., Vastenhouw, N.L., Levin, J.Z., Fan, L., Sandelin, A., Rinn, J.L., and Regev, A. 2012. Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Research* **22**(3): 577-591.
- Ponjavic, J., Oliver, P.L., Lunter, G., and Ponting, C.P. 2009. Genomic and transcriptional co-localization of protein-coding and long non-coding RNA pairs in the developing brain. *PLoS genetics* **5**(8): e1000617.
- Qu, Z., and Adelson, D.L. 2012. Bovine ncRNAs are abundant, primarily intergenic, conserved and associated with regulatory genes. *PloS one* **7**(8): e42638.
- Ranganna, K., Mathew, O.P., Milton, S.G., and Hayes, B.E. 2013. MicroRNAome of Vascular Smooth Muscle Cells: Potential for MicroRNA-Based Vascular Therapies.
- Seffens, W., and Digby, D. 1999. mRNAs have greater negative folding free energies than shuffled or codon choice randomized sequences. *Nucleic acids research* **27**(7): 1578-1584.
- Skroblin, P., and Mayr, M. 2014. "Going Long": Long Non-Coding RNAs as Biomarkers. *Circulation research* **115**(7): 607-609.
- Sun, J., Lin, Y., and Wu, J. 2013. Long non-coding RNA expression profiling of mouse testis during postnatal development. *PloS one* **8**(10): e75750.

- Tan, M.H., Au, K.F., Yablonovitch, A.L., Wills, A.E., Chuang, J., Baker, J.C., Wong, W.H., and Li, J.B. 2013. RNA sequencing reveals a diverse and dynamic repertoire of the *Xenopus tropicalis* transcriptome over development. *Genome research* **23**(1): 201-216.
- Trotta, E. 2014. On the Normalization of the Minimum Free Energy of RNAs by Sequence Length. *PloS one* **9**(11): e113380.
- Ulitsky, I., and Bartel, D.P. 2013. lincRNAs: genomics, evolution, and mechanisms. *Cell* **154**(1): 26-46.
- Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., and Bartel, D.P. 2011. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**(7): 1537-1550.
- Wang, L., Zhou, D., Tu, J., Wang, Y., and Lu, Z. 2014. Exploring the stability of long intergenic non-coding RNA in K562 cells by comparative studies of RNA-Seq datasets. *Biology direct* **9**(1): 15.
- Weikard, R., Hadlich, F., and Kuehn, C. 2013. Identification of novel transcripts and noncoding RNAs in bovine skin by deep next generation sequencing. *BMC genomics* **14**(1): 789.
- Xie, C., Yuan, J., Li, H., Li, M., Zhao, G., Bu, D., Zhu, W., Wu, W., Chen, R., and Zhao, Y. 2014. NONCODEv4: exploring the world of long non-coding RNA genes. *Nucleic acids research* **42**(D1): D98-D103.
- Yan, B., and Wang, Z. 2012. Long noncoding RNA: its physiological and pathological roles. *DNA and cell biology* **31**(S1): S-34-S-41.
- Zhang, Y.-C., Liao, J.-Y., Li, Z.-Y., Yu, Y., Zhang, J.-P., Li, Q.-F., Qu, L.-H., Shu, W.-S., and Chen, Y.-Q. 2014. Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice. *Genome biology* **15**(12): 512.

Table 1 Summary of reads mapping to the sheep genome.

Tissue	Raw Reads	Trimmed reads	Mapped reads (%)	Concordant alignment* (%)
Brain	15,923,182	14,506,121	90.7	85.4
Heart	14,155,651	12,675,258	87.6	82.3
Kidney	14,329,986	12,899,357	90.9	85.6
Liver	12,988,187	12,422,798	90.3	85.2
Lung	13,885,975	12,538,951	92.4	86.7
Ovary	15,103,630	13,551,486	92.5	87
Skin	23,579,808	21,917,771	90.8	82.8
White adipose	16,258,307	14,455,351	93.5	88.2

* A concordant alignment is defined as a pair on the same chromosome with the proper orientation.

Draft

Table 2 Distribution of expressed novel lincRNAs across the eight tissues of sheep. The lincRNAs are expressed in one or more tissues.

Number of tissues	1	2	3	4	5	6	7	8
Number of expressed lincRNAs	64	48	23	32	26	27	40	65

Draft

Figure captions:

Figure 1 Computational pipeline for identification of long intergenic non-coding RNAs in eight tissues of sheep.

Figure 2 Venn diagram showing the numbers of candidate lincRNAs with coding potential detected by CNCI, PLEK and CPC softwares. Of the total 959 candidate lincRNAs, 6, 101 and 32 were uniquely detected as coding by CNCI, PLEK and CPC, respectively. Only one candidate lincRNA was detected as coding by three softwares.

Figure 3 The chromosome locations of predicted lincRNAs.

Figure 4 Comparison of a) GC contents, b) gene length, c) transcript length and d) exon number of the predicted lincRNAs and protein-coding genes.

Figure 5 Comparison of expression levels (FPKM) of novel lincRNAs and protein-coding genes in different tissues of sheep. The expression of novel lincRNA genes is significantly lower than protein-coding genes.

Figure 6 Tissue-specific expression of lincRNAs. Density plot shows distribution of mean tissue specificity scores calculated for each transcript across eight tissues. The dashed line represents the mean.

Figure 7 The comparison of minimum free energy between the putative lincRNAs and protein-coding genes.

121x121mm (100 x 100 DPI)

29x13mm (300 x 300 DPI)

Draft

173x82mm (300 x 300 DPI)

Draft

173x80mm (300 x 300 DPI)

Draft

173x80mm (300 x 300 DPI)

Draft

173x230mm (300 x 300 DPI)

Supplementary file 2: The expression levels of the identified putative lincRAs across the eight tissue

Transcript id	Brain-FPKM	Heart-FPKM	Kidney-FPKM	Liver-FPKM	Lung-FPKM
CUFF.9390.1	0	13.8265	0	0	0
CUFF.9287.2	0	0	0	0	0
CUFF.8997.1	0	0	0	0	0
CUFF.8164.1	0	0	0	0	0
CUFF.7463.1	0	0	0	0	12817.9
CUFF.74.1	0	0	0	0	0
CUFF.6784.1	0	0	0	0	0
CUFF.5809.1	0	0	0	2.03758	0
CUFF.5096.1	0	0	0	28.264	0
CUFF.50693.1	13.7029	0	0	0	0
CUFF.49664.1	1.54963	0	0	0	0
CUFF.49165.1	0	0	2.60442	0	0
CUFF.49109.1	0	0	0	0	4.7758
CUFF.48797.1	0	0	0	0	0
CUFF.48463.1	0	0	0	0	0
CUFF.48218.2	0	0	0	0	0
CUFF.48218.1	0	0	0	0	0
CUFF.48111.1	7.35353	0	0	0	0
CUFF.4772.1	0	0	0	4.50022	0
CUFF.46517.1	0	0	0	0	0
CUFF.45638.1	1.90578	0	0	0	0
CUFF.43826.1	0	0	0	0	0
CUFF.4158.1	0	0	0	0	0
CUFF.4083.1	0	0	0	0	0
CUFF.40184.1	0	0	0	0	0
CUFF.39343.1	0	0	0	5.88613	0
CUFF.39330.1	0	0	0	4.63403	0
CUFF.3868.1	0	0	0	0	1.86335
CUFF.38508.1	6.00064	0	0	0	0
CUFF.38420.1	0	0	0	0	0
CUFF.38227.1	0	0	0	10.0861	0
CUFF.3811.1	2.75455	0	0	0	0
CUFF.36747.1	0	0	0	0	0
CUFF.36701.1	4.52658	0	0	0	0
CUFF.35948.1	0	0	0	0	0
CUFF.35849.1	0	0	0	0	0
CUFF.35761.1	8.92224	0	0	0	0
CUFF.34739.1	30.7653	0	0	0	0
CUFF.34667.1	0	0	13.9825	0	0
CUFF.34667.2	0	0	10.7575	0	0
CUFF.33831.1	0	0	0	0	0
CUFF.32466.1	8.65773	0	0	0	0
CUFF.30001.1	0	0	0	0	0
CUFF.29175.1	0	0	0	0	0
CUFF.29003.1	0	0	8.19026	0	0

CUFF.27530.1	0	0	0	0	0
CUFF.25809.1	2.18102	0	0	0	0
CUFF.23573.1	0	0	0	0	2.31792
CUFF.2342.2	4.02807	0	0	0	0
CUFF.23411.1	0	0	0	0	0
CUFF.23155.1	4.92199	0	0	0	0
CUFF.22799.1	3.36837	0	0	0	0
CUFF.20442.1	0	0	0	0	0
CUFF.1889.1	0	0	0	0	0
CUFF.17773.1	0	0	0	0	0
CUFF.17065.1	0	0	0	0	0
CUFF.16715.1	0	0	0	0	0
CUFF.16671.1	5.10637	0	0	0	0
CUFF.14163.1	17.5191	0	0	0	0
CUFF.1250.2	0	0	0	0	0
CUFF.12383.1	0	0	0	0	0
CUFF.12219.2	0	0	0	0	0
CUFF.12219.1	0	0	0	0	0
CUFF.10217.1	7.96174	0	0	0	0
CUFF.9684.1	0.364915	0	0	0	0
CUFF.7378.1	0.209389	8.89985	0	0	0
CUFF.6511.2	0	0.713097	0	0	0
CUFF.5512.1	3.61534	0	0	0	0
CUFF.51159.1	18.0313	0	0.930186	0	0
CUFF.47654.1	0	0	4.2374	1.06582E-06	0
CUFF.4762.4	0.266443	0	0	0	0
CUFF.46704.2	0	65.4647	0	0	0.660898
CUFF.46524.1	0.11381	2.62653	0	0	0
CUFF.46290.1	0	0	0	0	0
CUFF.46167.1	0	0	0	0.447785	1.57441
CUFF.44461.2	20.5192	0	0	0	0
CUFF.44461.1	10.5817	0	0	0	0
CUFF.43784.1	5.44962	0	0	0	0
CUFF.43596.1	0	0	0	0.401844	0
CUFF.43360.1	3.60865	0	0.199896	0	0
CUFF.42787.1	5.58722	0	0	0	0
CUFF.42731.1	0	0	0.427062	19.1249	0
CUFF.42195.1	0	0	0	1.09075	0
CUFF.42039.1	8.70941	6.13135	0	0	0
CUFF.41701.1	0	0	0	5.32607	0
CUFF.41470.1	2.88076	0	0	0	0
CUFF.4097.1	0	0	0	0	0.318202
CUFF.39522.1	0	0	1.93643	9.76762	0
CUFF.38420.2	0	0	1.56778	0	0
CUFF.35331.1	7.64438	0	2.04474	0	0
CUFF.35318.1	9.84295	0	3.90113E-05	0	0
CUFF.34977.1	0	0	0	0	5.87318

CUFF.32907.1	0	0	0.486766	0	0
CUFF.31851.1	0	0	2.86119	1.11593	0
CUFF.3120.1	1.7234	0.255183	0	0	0
CUFF.30780.1	0	0	5.79435	0	0.519605
CUFF.30499.1	1.15866	0	256.953	0	0
CUFF.29948.1	0	0	0	0	0
CUFF.26155.1	0.272094	0	0	0	0
CUFF.26021.1	4.99752	0.606445	0	0	0
CUFF.24964.1	0	0	1.15043	12.2398	0
CUFF.23641.1	84.4308	0	0	19.6126	0
CUFF.21553.1	0	0	0	0	0.225031
CUFF.20106.1	301.134	0	0	0	4.94192
CUFF.20106.2	94.7605	0	0	0	0.00113239
CUFF.18719.1	2.30461	0	0	0.691157	0
CUFF.18700.1	0	0	1.44053	0.272014	0
CUFF.17798.1	2.41757	0	0	0	0
CUFF.16281.1	0	0	3.93863	2.67335	0
CUFF.15892.1	0	0	0	0	0
CUFF.1399.1	0	0	2.39325	0.0616616	0
CUFF.139.1	4.62316	0	0	0	0.180348
CUFF.9529.1	0.75945	0	0	0	0.5347
CUFF.890.1	1.15832	0	0	0	0
CUFF.6511.1	0	0.648822	0	0	0
CUFF.48912.1	0	0	103.174	5.62506	0.183182
CUFF.48201.1	0	0	5.3595	0.240113	0
CUFF.47332.1	0	0	0	0	0.0643299
CUFF.46704.1	0	20.4362	0	0	0.0862635
CUFF.42462.1	0	0	0	10.3032	4.63731
CUFF.42025.1	4.5376	0.00011789	5.34106	0	0
CUFF.42025.2	1.88455	0.352112	0.00010969	0	0
CUFF.41792.1	0	0	0	0	0.355392
CUFF.38336.3	0	2.27964	0	0	0
CUFF.38336.2	0	5.46693E-05	0	0	0
CUFF.31681.2	3.61429	0.00113352	0	0	0
CUFF.31644.1	0	0	0.639665	0	0
CUFF.30009.1	0	0	0	0	0
CUFF.2983.1	0.143644	0	0	0	0.100804
CUFF.26354.1	3.54538	0	0	0	0.141237
CUFF.24971.3	0	0	4.88223	0.907827	0.0777149
CUFF.19617.1	0	0	1.35177	0.262989	0
CUFF.18704.1	0	0	0.247637	0	0.0822022
CUFF.16414.1	0	5.64755	1.41694	13.3333	0
CUFF.14944.1	0	2.39532	0.661962	7.19886	0
CUFF.806.2	0	3.62315	0	0	0
CUFF.6627.2	4.77402	0	0	0	0.279729
CUFF.621.1	0.137504	0	6.19968	0.377825	0
CUFF.49979.1	0	0	0	0	1.1158

CUFF.48910.1	0	0.200476	7.26746	0.187999	0
CUFF.48093.1	1.55085	0.0449893	0.0246669	0	0.0164596
CUFF.48092.1	0.499237	0.0442359	0.0242553	0	0.0161867
CUFF.47308.1	0.4701	0	0	0	0.165028
CUFF.4682.1	0.338878	0.658804	0	0	0
CUFF.46388.1	0	5.69015	0.142191	0	0
CUFF.44995.1	0	0	0	0.434817	0
CUFF.44763.2	0	0.000717454	0	0	0.000269937
CUFF.44763.1	0	0.19016	0	0	0.03464
CUFF.43336.1	3.2641	0	0.719776	0	0
CUFF.38336.1	0	0.505941	0	0	0.0520057
CUFF.36041.1	0	0	0	0	0.960487
CUFF.35915.1	0	0	0.319627	0	0.63503
CUFF.35868.1	0	0	0.903334	0	0
CUFF.35050.1	0	2.40999	0.264909	0	0.0878835
CUFF.35006.1	0	0	0	1.08478	1.26814
CUFF.35005.1	0	0	10.4719	0	0.419309
CUFF.34936.1	0	0	0	0	0.245483
CUFF.33008.1	0.311217	0	3.0664	7.51926	0
CUFF.30184.1	0.575446	0	0.630786	18.4936	0
CUFF.24390.1	0	0	0.548057	1.79847	0
CUFF.21157.1	0.114463	0	1.33046	0	0.0803142
CUFF.20864.1	0	0	0.498378	0	0
CUFF.19414.1	0.389118	0	0	0	3.82785
CUFF.19196.1	0.353282	0	0	0	0.24819
CUFF.1838.1	0	0	0	1.47235	6.48733
CUFF.16034.1	0.630177	0	0	0	0.265318
CUFF.15937.1	7.80267	0	0	1.87325	0.406054
CUFF.8015.1	0.16235	5.63318	0.172197	0	0
CUFF.7946.1	24.8763	0	16.5424	0	8.00452
CUFF.6287.1	6.26919	0.273381	0	0	0
CUFF.48400.2	0	3.06912	2.26017	0	0.737687
CUFF.48129.1	0.685089	0	4.36288	0	0.48084
CUFF.44234.1	0.258172	0	0	0.463927	0.725132
CUFF.44117.1	0.840541	0	0	0	0.295357
CUFF.42375.1	2.46878	0.263204	0.0722055	0	0
CUFF.3702.2	0	0.000132553	6.58511	0	0.00128327
CUFF.36137.1	3.72655	0.299211	0.164599	0	0
CUFF.3581.1	9.68047	0	0.989069	0	0.620006
CUFF.34917.1	1.65772	1.65065	0	0	0
CUFF.29848.1	6.97878	3.48204	0.979794	0	0.614559
CUFF.28890.1	0.0910543	0	2.78515	0.164186	0.511051
CUFF.28564.1	0	0	0.286497	4.81277	0.188109
CUFF.28506.6	0.0161622	0	0.000026938	2.56522E-89	0.000248874
CUFF.25254.1	0.311364	0	0.334151	0	0
CUFF.25015.1	2.63055	0	0.268378	0	0.178594
CUFF.23095.1	0.217139	0	0.462602	0	0

CUFF.22959.1	5.41108	0	0.5915	0	0
CUFF.17799.3	1.00355	2.77779E-05	3.87677E-07	0	2.83635E-10
CUFF.17359.1	1.79886	0	0	0	2.10403
CUFF.17080.1	0	0.582108	2.10422	0	3.18393
CUFF.13211.1	0.380438	0.367165	0	0	0
CUFF.12168.2	0	0.967381	0	9.82568	6.01689
CUFF.12168.1	0	1.16793	0	9.13684	11.7538
CUFF.9991.1	4.51982	7.2279	1.01798	0	0.636942
CUFF.8537.1	1.23339	0	0	0.443377	0.519605
CUFF.8298.1	0.343547	0	0.739267	0	3.1374
CUFF.48601.1	14.1547	11.0929	12.8852	12.3421	1.2528
CUFF.47846.1	0.85057	0	0	0.254354	0.00167037
CUFF.46841.1	0	0.589577	0.651637	23.9966	0.42672
CUFF.43860.1	0.0858055	0	0.0904666	0.154738	1.20395
CUFF.42038.1	1739.88	4932.37	665.465	0	0
CUFF.39592.1	0.388414	0.249183	0	0.233273	0.0908519
CUFF.3702.3	0	3.4195	0.447956	1.52209	3.01851
CUFF.32184.2	9.28001	0	2.32807	2.6179	4.05594
CUFF.30641.1	2.4992	0	0.73497	0	0.584602
CUFF.29850.1	0.300126	0	3.86164	0	1.89709
CUFF.28506.4	0.0748668	0	1.55405	5.7706	0.193709
CUFF.25769.1	2.44579	0	0.317886	0	1.18446
CUFF.25564.1	5.01165	0	0	0.529354	1.03522
CUFF.24275.9	0	4.63319	1.30969E-54	7.63885	3.64267
CUFF.23210.1	0	18.2884	1.75171	5.19325	1.04669
CUFF.2309.1	1.08751	0.130432	0	0	0.14304
CUFF.18862.1	0.141203	0.271915	0.149519	0	0.0990902
CUFF.18798.1	1.16125	1.49455	0.411565	0	0
CUFF.17357.1	2.01356	0	0.428446	0.362254	0.424041
CUFF.16639.1	0.887875	0	0.99852	0	1.2511
CUFF.14818.1	0	0	0.612095	0.260227	0.405555
CUFF.13832.1	0.708538	0	1.389	0	0.284068
CUFF.13439.1	0.61862	0	0.108851	0.185866	0.217018
CUFF.12663.2	0	0	3.21404E-05	7.07498	3.8627E-07
CUFF.9539.1	1.86795	1.30984	0.360437	0	0.476743
CUFF.9302.1	2.41117	0	0.506212	0.000435991	0.0332001
CUFF.9114.1	0.0634796	0.510118	0.139904	1.17916	1.09108
CUFF.8234.2	0.955639	2.30156	0.874246	3.36191E-06	0
CUFF.7156.1	0.783741	8.16398	0.106464	0	1.85277E-06
CUFF.6248.2	7.03314	5.08337E-07	2.15807	14.8437	0
CUFF.6248.3	0.000116117	2.37032	5.0444	7.86243	0
CUFF.5933.1	0.0876147	0	0.0923871	0.157995	0.245869
CUFF.50685.1	0.65662	0	1.83337E-05	4.48763E-06	1.09784E-06
CUFF.49338.1	0.443101	0.426023	0	0.333575	1.64427E-05
CUFF.48400.3	0.675804	1.48984E-07	6.56327E-26	0	2.03843E-11
CUFF.47764.1	22.0071	0.130343	0.286053	0.122535	0.0476476
CUFF.43074.1	0.637427	0	0.269697	1.83763	0.178915

CUFF.42143.1	40.0935	2.30898E-14	2.90042E-06	0	0.000670536
CUFF.38537.1	1.14231	0.783392	0.0859976	0	0.0572412
CUFF.37359.1	0.584024	0	0.414012	0.700563	0.409955
CUFF.3704.1	0.998861	0	1.34086	0.654953	0.191121
CUFF.3702.1	3.05737E-06	0.689771	0.0119801	0	2.8881E-22
CUFF.36923.5	1.28926	18.3977	9.31711	3.58081	40.4663
CUFF.36010.1	0	0.288792	0.635366	0.269976	1.47282
CUFF.35609.1	0.953941	0.430486	0.649284	0.101256	0.865947
CUFF.35070.1	0.569756	0.625173	0.428929	0	0.057101
CUFF.34348.1	1.42406	0	0.838856	1.13967	1.22153
CUFF.32424.1	0.127976	0	0.315781	0.319243	0.622109
CUFF.31853.1	0.000807578	0	12.9036	4.03001E-06	0.33235
CUFF.31596.1	3.06696	2.90372E-81	0.26555	0	0.150327
CUFF.30314.1	2.41135	0	1.79274	0.501364	0.566721
CUFF.28297.1	3.4396	0.798546	1.11065	0	3.35711
CUFF.27119.1	5.26819	2.0672	0.575216	0	1.4818
CUFF.23934.4	0.764087	0	0.423189	3.26028	0.561471
CUFF.23294.1	0.675804	0	2.33077	1.21663	1.06721
CUFF.20975.1	1.78509	0	0.726787	0.247365	0.963597
CUFF.18692.1	0.375189	0.165513	5.47633	0.774587	0.172675
CUFF.17379.1	0.226509	0	0.24146	0.814589	0.636095
CUFF.1619.1	5.72906	1.52422	0.477859	0	0.878212
CUFF.13878.1	57.9706	0.584958	1.00097	0.627999	0
CUFF.13878.2	16.9109	0.0060231	0.000013853	0.00575607	0
CUFF.11659.1	2.42102	0.548521	0.452443	0	0.19988
CUFF.11350.2	1.01331E-16	0	1.01879	9.99649E-06	1.86689
CUFF.10429.1	0.468995	1.36079	0.750402	5.05908	0
CUFF.9763.1	0.232983	0.277547	0.309097	1.42934	2.79306
CUFF.966.1	2.82828	3.33575	3.71738	6.06085	1.59138
CUFF.9561.1	37.1675	27.3292	8.32323	13.4822	4.87544
CUFF.9041.1	0.726246	4.20371	1.54291	6.20866	4.58779
CUFF.8234.1	0.119457	0.823584	0.0712679	3.67035E-05	0.222234
CUFF.7188.1	8.18945	34.7827	1.3872	12.3586	0.886444
CUFF.5997.1	14.4833	12.3595	10.8616	7.28257	10.3749
CUFF.544.5	4.40302	0.0653946	0.034921	9.94389	0.0241444
CUFF.544.4	0.0281138	29.7145	32.5678	0.0086512	15.8354
CUFF.50917.1	0.366255	2.13892	1.9739	0.656695	1.54393
CUFF.50586.1	15.0602	9.17193	13.5127	14.8954	6.10364
CUFF.4966.7	2.58285E-93	0.941479	1.72261E-94	5.26251E-45	0.772525
CUFF.49525.1	0.246393	1.57715	0.259703	0.888738	0.172855
CUFF.49118.1	5.81675	1.37303	1.50819	2.57555	2.07636
CUFF.48805.1	3.61581	0.248534	0.136612	0.232671	0.362466
CUFF.47507.1	3.3161	0.144677	0.714558	0.271882	0.951671
CUFF.47119.1	35.7935	14.4671	16.1332	12.136	11.4728
CUFF.46717.6	10.2058	5.34602	6.66313	1.13438	6.10164
CUFF.46299.2	3.55475E-05	7.6574	0.336403	0.880625	0.943011
CUFF.45686.1	0.284507	7.55756E-15	6.21784E-25	4.51667E-20	1.57158E-15

CUFF.44707.1	0.166438	6.45233	0.447135	0.00594386	0.304987
CUFF.44648.3	175.986	163.331	166.409	129.093	108.771
CUFF.44606.5	5.93339	0.241726	0.448784	0.158509	0.126144
CUFF.43625.2	2.14826	0.753676	0.54757	0.412838	0.662811
CUFF.43625.1	1.86334	3.55033E-05	0.539646	2.69408E-05	1.12122
CUFF.43305.1	47.8489	6.6982	3.73649	5.37674E-05	5.13481E-06
CUFF.43.1	337.44	192.811	215.944	112.954	67.5032
CUFF.42237.2	19.9269	0.0164671	1.8193	0.00504417	12.679
CUFF.41662.7	0.223866	0.14871	35.2697	1.11088	10.8753
CUFF.41604.1	0.101323	0.584237	4.17128	0.182662	0.710891
CUFF.41391.1	14.9669	35.7993	22.935	27.3339	3.15582
CUFF.41391.2	0.000253994	0.000926597	0.000135712	0.000374328	0.000126601
CUFF.41054.3	3.32632	0.000121074	0.191025	2.52349E-05	1.03689E-07
CUFF.39851.4	37.4014	9.12125	5.13996	8.08866	41.1389
CUFF.39633.2	5.40727	2.83824E-33	1.60102E-13	1.20659E-06	1.45949E-05
CUFF.39579.1	2.12837	0.684283	0.75337	0.319248	1.12041
CUFF.39574.1	2.49248	0.416718	0.228914	2.53938	0.456214
CUFF.3832.1	6.56694	4.9828	11.0084	3.07669	4.61182
CUFF.37765.3	3.4051	20.0555	7.59329	7.275	5.87148
CUFF.36923.7	1.03407	1.34076	0.727985	6.01153E-05	5.77785
CUFF.35535.1	2.14531	4.61395	1.06776	0.909541	3.27639
CUFF.34229.1	0.243804	4.83393E-26	0.633929	2.56933	0.500732
CUFF.31247.1	8.17216	1.22065	0.674791	0.564162	2.42846
CUFF.3025.2	0.733159	0.617106	0.357195	0.0815859	0.380915
CUFF.27786.5	4.88241	13.6661	18.9285	1.77513	7.35604
CUFF.2570.1	3.64535	1.5507	4.69006E-06	0.000228888	1.33866
CUFF.23934.3	0.000063676	1.40701E-09	1.34453	0.301869	0.00174262
CUFF.23934.7	1.19512E-06	6.29948E-11	0.0588037	1.80216	0.459678
CUFF.23934.1	7.33933E-08	0.0126953	2.35056	3.99136E-39	0.488448
CUFF.23671.3	1.5983	1.06867E-06	5.70264E-75	0.132741	6.59445E-08
CUFF.2354.1	28.1	5.14564	7.43599	4.67218	3.60684
CUFF.23318.2	2.24227	7.5579E-06	1.01561	0.00147456	0.451922
CUFF.22023.1	1.20072	0.464643	0.256271	1.29507	0.505826
CUFF.21962.1	4.79816	5.81672	4.39892	11.5042	5.44157
CUFF.21921.9	0.299902	3.85944	3.54117E-76	6.7606E-146	4.9546E-198
CUFF.20256.3	3.87598	1.16128	1.71316	0.865673	0.367116
CUFF.19876.1	43.994	28.1365	23.3344	8.85559	27.5093
CUFF.19478.2	1.42409	0.321492	0.43507	0.841231	1.55706
CUFF.19476.1	0.372328	0.285885	0.235324	0.268639	0.261194
CUFF.18800.16	73.4269	48.2093	34.1756	20.3939	43.823
CUFF.18800.14	5.19692	3.16238	12.5027	8.26196	4.90562
CUFF.18800.15	5.40199E-17	4.8838E-11	8.21342E-11	7.70189E-09	1.30165E-21
CUFF.18658.2	0.562998	0.755335	0.470125	0.533678	0.205965
CUFF.18479.1	2.18481	5.55273	4.64731	5.86184	8.71964
CUFF.16104.1	16.6749	0.00018877	1.20255E-07	9.875	18.3585

s of sheep.

Ovary-TPKM	Skin-TPKM	White adipose-TPKM	Number tissue
0	0	0	1
0.151701	0	0	1
0	9.49625	0	1
0	11.9593	0	1
0	0	0	1
0	3.61772	0	1
0	1.89038	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	4.86768	0	1
0	10.941	0	1
0	8.23251	0	1
0	6.02371	0	1
0	0	0	1
0	0	0	1
0	0	22.5496	1
0	0	0	1
0	0	7.93024	1
0	4.72187	0	1
0	7.80002	0	1
7.66391	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	31.2799	0	1
0	0	0	1
0	0	0	1
0	0	2.27701	1
0	0	0	1
0	1.45627	0	1
0	3207.78	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	2.73123	0	1
0	0	0	1
0	2.00189	0	1
0	15.3132	0	1
0	0	0	1

0	1.98664	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	1.48668	0	1
0	0	0	1
0	0	0	1
8.38579	0	0	1
0	2.54606	0	1
0	2.21381	0	1
0	1.05493	0	1
0	2.03389	0	1
0	0	0	1
0	0	0	1
3.87976	0	0	1
0	5.1142	0	1
0	5.4543	0	1
0	4.34058	0	1
0	0	0	1
0	3.73872	0	2
0	0	0	2
0	0	12.9918	2
0	1.38191	0	2
0	0	0	2
0	0	0	2
0	18.6359	0	2
0	0	0	2
0	0	0	2
0	0.598599	16.8898	2
0	0	0	2
0	0.259121	0	2
0	0.692781	0	2
0	0.325411	0	2
0	0	2.47379	2
0	0	0	2
0	0.5901	0	2
0	0	0	2
0	16.3424	0	2
0	0	0	2
0.248053	0	0	2
0	0.238043	0	2
0	7.07672	0	2
0	0	0	2
0	20.1061	0	2
0	0	0	2
0	0	0	2
0	0.667055	0	2

0	12.847	0	2
0	0	0	2
0	0	0	2
0	0	0	2
0	0	0	2
182.437	0	0.616068	2
0	11.4965	0	2
0	0	0	2
0	0	0	2
0	0	0	2
0	2.44085	0	2
0	0	0	2
0	0	0	2
0	0	0	2
0	0	0	2
0	1.30985	0	2
0	0	0	2
1.27167	0	0.388812	2
0	0	0	2
0	0	0	2
0	5.26215	0	3
0	7.94404	1.31971	3
0.103142	0	7.98623	3
0	0	0	3
0.201821	0	0	3
0.0691536	2.06924	0	3
0	0.134802	0	3
0	68.295	0	3
0	0	0	3
0	0	0	3
0	0.332667	5.19635	3
0	8.11611E-05	5.67193	3
0	3.4081E-10	2.90146	3
1.06228	0	0	3
0	0.188618	6.10675	3
3.35262	16.3868	3.56485	3
0	3.08796	0	3
0.178317	0	0	3
0	0	0	3
0.0353323	0	0	3
0.885732	0	0	3
0	0	0	3
0	0	0	3
3.51001	0.167919	1.82172	4
0.217718	6.22695	0	4
0	0	0.134411	4
0.120653	0.293783	0.405603	4

0	0.127188	0	4
0	0	0	4
0	0	0	4
0.179605	2.20268	0	4
0	4.75014	0.28928	4
0.101736	0	0.114196	4
0.739787	2.2697	0.206064	4
0.000286417	0	2.65799	4
0.0371167	0	1.13233	4
0.103	14.2644	0	4
0	0.359646	2.89682	4
5.78456	11.8207	1.18685	4
0.228732	4.17387	0	4
0.35731	3.86635	0.266956	4
0.379038	0	0	4
0.456763	0	0.384101	4
0	0.399463	0.511661	4
1.86062	0.216066	1.6364	4
1.69556E-05	0	0	4
0.453698	0	0	4
0.393839	7.60892	0	4
0.173038	0	0	4
0.715891	5.68669	0.392786	4
0.301535	0	0.332524	4
0.545732	4.96575	0	4
1.01475	0	1.11444E-26	4
0.095365	4.93834	0	4
0	0	0.00231346	4
0.246487	0.300187	0	5
0	40.4363	4.86944	5
4.29128E-06	0.259008	0.316904	5
34.8376	1.15385	0	5
0.260241	0	0.874123	5
2.37223	0	0.220009	5
0.326611	4.08022	0.359386	5
0	0.747822	0.0582067	5
7.93148E-23	0	11.4001	5
0	0.477876	0.264067	5
0.714185	2.50164	0	5
0.669352	41.9542	1.43015	5
0	0	1.50661	5
0	0	0.0773214	5
0.410601	0	0.22833	5
1.30505E-64	0	0	5
0.718721	28.0117	0.265639	5
0.0639709	1.23746	0	5
1.6565	0.135201	0.184881	5

1.27578	1.07597	0.463924	5
7.56305E-09	0	0	5
0.113669	0.460898	0.127454	5
0	9.24307E-06	1.81725E-06	5
0.289524	2.12124	0.161868	5
1.45617E-06	0	0.825256	5
0.505496	0	0.354637	5
5.14704	7.10262	0	6
0.943441	4.0156	0.420326	6
0.530221	0.876345	0.58659	6
12.5645	0	0	6
1.15709	12.6277	0.135419	6
0.700729	0.192264	0	6
0.194076	0	0.509994	6
1575.96	860.229	678.685	6
0.097995	1.82425	0	6
1.38851	0	10.6493	6
0	2.35375	2.8847	6
0.42071	5.87908	0.236045	6
0.230688	1.70836	8.70363	6
0.335156	0	0.0615105	6
1.8386	3.82498	0.14926	6
0.67946	1.86263	1.50852	6
3.74455	0	6.023E-135	6
6.37745	0	19.4397	6
0.0511633	0.53525	0.115381	6
0.106987	1.04027	0	6
0.294669	1.19979	0.494121	6
0.306783	0	0.514159	6
0.721084	0.632028	8.43671	6
0.547495	2.92882	0.368473	6
0.305673	1.78894	1.97736	6
0.155707	3.01714	0	6
2.23854E-05	0.424806	2.46846E-05	6
0.773976	0.314422	0.722212	7
0.283272	0.0377911	0.445398	7
0	0.683707	0.242699	7
0.97387	1.10133E-08	0.71605	7
0.230403	1.17884E-07	0.111826	7
8.10735E-06	0.943944	1.54483	7
0.638167	11.4974	0.000087354	7
0.132133	0.0532635	0.892742	7
0.250708	0.135429	8.67079E-06	7
0.167409	0.681111	1.16441	7
0.868556	2.24182	0.143706	7
0	0.123434	0.230604	7
1.73654	0.390351	1.51678	7

0.181429	115.254	0.360063	7
2.21376	0.297292	0.346369	7
1.77855	0.603535	0.497027	7
0.273969	0.828533	0.462659	7
0.000481336	7.93895E-07	5.26472	7
17.8371	0	26.8774	7
1.02302	0.184359	0.509817	7
0.928284	0	0.761855	7
0.79744	0.889659	0.0691038	7
1.92113	0.292344	0.538222	7
0.976749	6.94969E-32	0.435037	7
0.000596123	2.13988E-08	0.0546181	7
0.196815	0.720957	2.00162E-85	7
0.642413	2.95124	0.849637	7
0.757925	2.7235	2.51283	7
1.24044	3.48194	3.61227	7
0.0598017	0.252652	0.260325	7
0.898314	0.417032	1.43706	7
0.416019	0.421209	1.40063	7
5.46807	0	1.68327	7
0.518834	3.10822	0.192897	7
4.65865	0.85912	3.75938	7
1.71627	20.8989	1.33429	7
0.000652814	9.57173	4.47197	7
0.323746	0.174903	0.363196	7
7.17769E-05	1.68888E-18	0.000022514	7
0.179174	2.05079	0.399469	7
4.87174	0.00446327	0.309219	8
0.445572	1.50707	2.42614	8
11.9813	1.70025	14.0696	8
3.58971	0.898417	5.25199	8
1.62979E-05	1.61616E-05	1.03502	8
0.998442	2.12531	4.32911	8
16.0154	37.2991	6.91634	8
0.258927	2.40307	0.0289447	8
8.19534	2.56518	19.7944	8
1.13289	0.234589	3.44117	8
11.9323	22.0457	22.9291	8
1.62802	3.38184E-67	1.6353E-197	8
0.433309	0.748199	1.32489	8
4.23773	0.995231	3.03371	8
0.879645	0.15821	0.658487	8
0.283841	0.365719	0.31989	8
16.848	23.4519	10.9545	8
9.12854	6.40868	7.88417	8
0.0283507	0.373721	0.000317513	8
1.83393E-08	1.06074	0.000255301	8

0.280322	0.349416	0.777307	8
278.273	246.004	85.3658	8
0.139075	0.848222	1.48913	8
1.41867	1.05818E-05	2.27487E-05	8
0.444377	4.23867	1.15771	8
4.05164	2.37593E-06	7.47763	8
133.13	538.264	124.118	8
18.5108	1.82234	21.3126	8
4.18757	10.2073	0.959862	8
0.535477	0.123503	2.15149	8
12.3346	14.5118	3.16855	8
0.000160233	7.57716	0.000215183	8
2.99744	5.19178	1.15348	8
9.65388	11.7495	3.15372	8
7.7799E-15	1.59125	0.161845	8
1.48305	0.438481	0.452665	8
3.02912	0.793572	0.460291	8
13.3693	16.5674	21.6675	8
4.08045	2.43538	6.31065	8
5.25373	0.219422	2.94316	8
21.962	13.7526	1.93035	8
0.384814	0.322657	1.06735	8
4.59621	3.5879	1.8772	8
0.305817	0.185235	0.889291	8
24.8019	10.1277	56.9797	8
0.171604	1.02724	0.407327	8
5.34346E-09	1.33904E-10	2.91453E-10	8
1.26493E-09	1.9006E-11	4.32258E-11	8
1.57598E-25	3.3514E-06	5.48669E-06	8
1.1435E-07	1.02982	4.99556E-07	8
8.79264	19.6475	6.21435	8
4.10827E-05	2.73284	0.157107	8
1.65224	0.450513	2.25025	8
12.3467	4.80243	13.7854	8
0.28286	0.195551	0.707122	8
7.35016E-92	0.837842	0.768642	8
26.4625	31.0577	13.1359	8
0.00623119	1.5017	0.206734	8
0.168257	0.0451614	0.75855	8
29.6741	41.0889	31.2266	8
26.9953	3.52668	15.9174	8
4.45663E-17	2.66715	1.49085E-08	8
3.25792	1.72038	0.827501	8
5.89229	1.53562	8.20383	8
0.000195119	19.2195	15.7111	8

Supplementary file 3: The means and standard deviation of putative lincRNAs and protein-coding genes across

Tissue	Type	Mean	SD
Brain	lincRNAs	18.38377084	125.0179382
Brain	mRNA	40.73382838	2979.883683
Heart	lincRNAs	41.90503174	415.4121547
Heart	mRNA	34.77728049	859.7725087
Kidney	lincRNAs	10.36884709	55.60573484
Kidney	mRNA	63.68922993	3954.829662
Liver	lincRNAs	5.313326367	14.62833152
Liver	mRNA	59.10818056	2022.483828
Lung	lincRNAs	73.121692	944.7561287
Lung	mRNA	18.25984857	385.4917293
Ovary	lincRNAs	14.318287	117.4881268
Ovary	mRNA	21.22693468	485.6480883
Skin	lincRNAs	28.62954155	232.4962963
Skin	mRNA	90.50841863	3440.083026
White adipose	lincRNAs	8.387134181	51.99866603
White adipose	mRNA	23.63093919	486.3847365

Draft

in different tissues.

Tissue	Type	Mean (After disc	SD (After discarding the lincRNA with very high expression
Brain	lincRNAs	18.38377084	125.0179382
Brain	mRNA	40.73382838	2979.883683
Heart	lincRNAs	6.97313911	22.7490783
Heart	mRNA	34.77728049	859.7725087
Kidney	lincRNAs	10.36884709	55.60573484
Kidney	mRNA	63.68922993	3954.829662
Liver	lincRNAs	5.313326367	14.62833152
Liver	mRNA	59.10818056	2022.483828
Lung	lincRNAs	3.478094693	11.06221154
Lung	mRNA	18.25984857	385.4917293
Ovary	lincRNAs	14.318287	117.4881268
Ovary	mRNA	21.22693468	485.6480883
Skin	lincRNAs	28.62954155	232.4962963
Skin	mRNA	90.50841863	3440.083026
White adipose	lincRNAs	8.387134181	51.99866603
White adipose	mRNA	23.63093919	486.3847365

Draft

1 from heart and lung tissues)

Draft

Supplementary file 4: The list of co-expressed genes with lincRNAs and GO analysis of co-expressed pro

lincRNA	Coexpressed gene	Validated correlation	lincRNA
CUFF.46704.1	ENSOART000000019	0.9425	CUFF.30009.1
CUFF.46704.1	ENSOART000000165	0.9511	CUFF.20864.1
CUFF.47119.1	ENSOART000000203	0.9176	CUFF.18800.15
CUFF.5997.1	ENSOART000000164	0.94283	CUFF.14818.1
CUFF.5997.1	ENSOART000000011	0.90793	CUFF.25254.1
CUFF.48601.1	ENSOART000000217	-0.96893	CUFF.47332.1
CUFF.7378.1	ENSOART000000193	0.9895	CUFF.44995.1
CUFF.7378.1	ENSOART000000046	0.9475	CUFF.890.1
CUFF.7378.1	ENSOART000000219	0.9265	CUFF.33008.1
CUFF.7156.1	ENSOART000000193	0.97463	CUFF.28564.1
CUFF.7156.1	ENSOART000000046	0.93827	CUFF.28506.4
CUFF.7156.1	ENSOART000000219	0.92647	CUFF.39522.1
CUFF.44707.1	ENSOART000000046	0.91883	CUFF.24964.1
CUFF.46388.1	ENSOART000000204	0.93197	CUFF.42731.1
CUFF.46388.1	ENSOART000000173	0.92647	CUFF.36041.1
CUFF.46388.1	ENSOART000000202	0.92127	CUFF.42195.1
CUFF.46388.1	ENSOART000000091	0.9441	CUFF.34229.1
CUFF.46388.1	ENSOART000000059	0.9517	CUFF.46388.1
CUFF.46388.1	ENSOART000000031	0.94787	CUFF.35050.1
CUFF.46388.1	ENSOART000000030	0.95673	CUFF.8015.1
CUFF.8015.1	ENSOART000000019	0.98513	CUFF.47846.1
CUFF.8015.1	ENSOART000000173	0.91673	CUFF.17379.1
CUFF.8015.1	ENSOART000000202	0.93023	CUFF.42462.1
CUFF.8015.1	ENSOART000000165	0.9358	CUFF.19196.1
CUFF.8015.1	ENSOART000000091	0.91063	CUFF.30184.1
CUFF.8015.1	ENSOART000000059	0.96017	CUFF.35868.1
CUFF.8015.1	ENSOART000000030	0.91777	CUFF.34917.1
CUFF.46524.1	ENSOART000000193	0.98893	CUFF.35915.1
CUFF.46524.1	ENSOART000000046	0.94697	CUFF.24390.1
CUFF.46524.1	ENSOART000000219	0.92767	CUFF.46841.1
CUFF.35050.1	ENSOART000000204	0.90487	CUFF.41391.2
CUFF.35050.1	ENSOART000000173	0.97793	CUFF.38420.2
CUFF.35050.1	ENSOART000000202	0.90113	CUFF.32907.1
CUFF.35050.1	ENSOART000000091	0.9501	CUFF.4682.1
CUFF.35050.1	ENSOART000000059	0.92373	CUFF.21553.1
CUFF.35050.1	ENSOART000000031	0.95277	CUFF.4097.1
CUFF.35050.1	ENSOART000000030	0.92557	CUFF.46524.1
CUFF.34917.1	ENSOART000000100	0.92817	CUFF.7378.1
CUFF.34917.1	ENSOART000000007	0.93107	CUFF.7156.1
CUFF.34917.1	ENSOART000000164	0.92317	CUFF.23934.7
CUFF.34917.1	ENSOART000000141	0.90457	CUFF.50693.1
CUFF.34917.1	ENSOART000000109	0.9004	CUFF.49664.1
CUFF.34917.1	ENSOART000000012	0.90873	CUFF.48111.1
CUFF.10429.1	ENSOART000000171	0.91157	CUFF.45638.1
CUFF.12168.1	ENSOART000000195	0.90407	CUFF.38508.1

CUFF.12168.1	ENSOART0000000068	0.93	CUFF.3811.1
CUFF.12168.2	ENSOART0000000195	0.99157	CUFF.36701.1
CUFF.12168.2	ENSOART0000000068	0.97373	CUFF.35761.1
CUFF.4682.1	ENSOART0000000100	0.90017	CUFF.34739.1
CUFF.4682.1	ENSOART0000000007	0.90113	CUFF.32466.1
CUFF.4682.1	ENSOART0000000184	0.90957	CUFF.25809.1
CUFF.4682.1	ENSOART0000000144	0.90407	CUFF.2342.2
CUFF.46841.1	ENSOART0000000226	0.9135	CUFF.23155.1
CUFF.46841.1	ENSOART0000000172	0.91853	CUFF.22799.1
CUFF.46841.1	ENSOART0000000068	0.91837	CUFF.16671.1
CUFF.46841.1	ENSOART0000000048	0.9266	CUFF.14163.1
CUFF.46841.1	CUFF.2467.1	0.92397	CUFF.10217.1
CUFF.13878.1	ENSOART0000000203	0.9727	CUFF.44461.2
CUFF.13878.1	ENSOART0000000007	0.93783	CUFF.43784.1
CUFF.38336.1	ENSOART0000000223	0.95033	CUFF.44461.1
CUFF.35609.1	ENSOART0000000171	-0.91507	CUFF.20106.2
CUFF.13211.1	ENSOART0000000164	0.93553	CUFF.20106.1
CUFF.36137.1	ENSOART0000000203	0.96203	CUFF.41470.1
CUFF.36137.1	ENSOART0000000007	0.93167	CUFF.42787.1
CUFF.6287.1	ENSOART0000000203	0.9332	CUFF.139.1
CUFF.6287.1	ENSOART0000000007	0.92003	CUFF.35318.1
CUFF.42375.1	ENSOART0000000203	0.97237	CUFF.5512.1
CUFF.42375.1	ENSOART0000000007	0.93793	CUFF.46290.1
CUFF.39592.1	ENSOART0000000164	0.90457	CUFF.26354.1
CUFF.39592.1	ENSOART0000000014	0.9145	CUFF.9529.1
CUFF.44606.5	ENSOART0000000203	0.91517	CUFF.2983.1
CUFF.47507.1	ENSOART0000000203	0.92297	CUFF.46704.1
CUFF.47507.1	ENSOART0000000007	0.94123	CUFF.41792.1
CUFF.2309.1	ENSOART0000000203	0.94693	CUFF.48093.1
CUFF.2309.1	ENSOART0000000007	0.91603	CUFF.16034.1
CUFF.47764.1	ENSOART0000000203	0.9143	CUFF.28506.6
CUFF.48093.1	ENSOART0000000203	0.90347	CUFF.17799.3
CUFF.48093.1	ENSOART0000000007	0.9051	CUFF.6287.1
CUFF.48092.1	ENSOART0000000007	0.90033	CUFF.36137.1
CUFF.41391.2	ENSOART0000000203	0.9006	CUFF.3581.1
CUFF.41391.2	ENSOART0000000184	0.94787	CUFF.42375.1
CUFF.41391.2	ENSOART0000000144	0.93057	CUFF.25015.1
CUFF.41391.2	ENSOART0000000014	0.91067	CUFF.12168.2
CUFF.41391.2	ENSOART0000000012	0.90447	CUFF.12168.1
CUFF.42025.1	ENSOART0000000195	0.93417	CUFF.2309.1
CUFF.17799.3	ENSOART0000000203	0.90053	CUFF.39592.1
CUFF.17799.3	ENSOART0000000007	0.90193	CUFF.25564.1
CUFF.48400.3	ENSOART0000000164	0.91857	CUFF.31596.1
CUFF.23934.7	ENSOART0000000195	0.97213	CUFF.13878.1
CUFF.23934.7	ENSOART0000000068	0.90547	CUFF.47507.1
CUFF.23934.7	ENSOART0000000058	0.92913	CUFF.39633.2
CUFF.18800.15	ENSOART0000000100	0.98357	CUFF.5997.1

CUFF.18800.15	ENSOART0000000085	0.90933	CUFF.51159.1
CUFF.18800.15	ENSOART0000000007	0.98383	CUFF.43360.1
CUFF.18800.15	ENSOART0000000168	0.97423	CUFF.23641.1
CUFF.18800.15	ENSOART0000000162	0.9405	CUFF.17798.1
CUFF.18800.15	ENSOART0000000155	0.95333	CUFF.9684.1
CUFF.18800.15	ENSOART0000000147	0.91687	CUFF.26155.1
CUFF.18800.15	ENSOART0000000144	0.96217	CUFF.4762.4
CUFF.18800.15	ENSOART0000000143	0.92707	CUFF.16281.1
CUFF.18800.15	ENSOART0000000141	0.91913	CUFF.42025.1
CUFF.18800.15	ENSOART0000000131	0.9397	CUFF.48092.1
CUFF.18800.15	ENSOART0000000131	0.9109	CUFF.15937.1
CUFF.18800.15	ENSOART0000000127	0.9059	CUFF.43336.1
CUFF.18800.15	ENSOART0000000125	0.93607	CUFF.47308.1
CUFF.18800.15	ENSOART0000000119	0.94113	CUFF.38336.1
CUFF.18800.15	ENSOART0000000115	0.93987	CUFF.49979.1
CUFF.18800.15	ENSOART0000000115	0.93537	CUFF.35006.1
CUFF.18800.15	ENSOART0000000109	0.9837	CUFF.22959.1
CUFF.18800.15	ENSOART0000000091	0.95303	CUFF.44117.1
CUFF.18800.15	ENSOART0000000085	0.97437	CUFF.13211.1
CUFF.18800.15	ENSOART0000000079	0.9754	CUFF.13439.1
CUFF.18800.15	ENSOART0000000061	0.9175	CUFF.8537.1
CUFF.18800.15	ENSOART0000000053	0.96293	CUFF.12663.2
CUFF.18800.15	ENSOART0000000053	0.96757	CUFF.48601.1
CUFF.18800.15	ENSOART0000000037	0.9279	CUFF.47764.1
CUFF.18800.15	ENSOART0000000028	0.96833	CUFF.42143.1
CUFF.18800.15	ENSOART0000000014	0.91233	CUFF.48400.3
CUFF.18800.15	ENSOART0000000012	0.91187	CUFF.10429.1
CUFF.42143.1	ENSOART0000000164	0.9088	CUFF.35609.1
CUFF.45686.1	ENSOART0000000164	0.92823	CUFF.44606.5
CUFF.34229.1	CUFF.25010.1	0.91423	CUFF.45686.1
CUFF.34229.1	ENSOART0000000148	0.91383	CUFF.47119.1
CUFF.34229.1	ENSOART0000000068	0.9213	CUFF.44707.1
CUFF.34229.1	ENSOART0000000181	0.91343	CUFF.9390.1
CUFF.34229.1	ENSOART0000000091	0.9129	CUFF.9287.2
CUFF.34229.1	ENSOART0000000058	0.924	CUFF.8997.1
CUFF.34229.1	ENSOART0000000035	0.93833	CUFF.8164.1
CUFF.34229.1	CUFF.29360.1	0.90763	CUFF.7463.1
CUFF.39633.2	ENSOART0000000203	0.98183	CUFF.74.1
CUFF.39633.2	ENSOART0000000007	0.9517	CUFF.6784.1
CUFF.31596.1	ENSOART0000000203	0.99013	CUFF.5809.1
CUFF.31596.1	ENSOART0000000007	0.9598	CUFF.5096.1
CUFF.9684.1	ENSOART0000000164	0.95407	CUFF.49165.1
CUFF.9529.1	ENSOART0000000187	0.9318	CUFF.49109.1
CUFF.9529.1	ENSOART0000000164	0.9299	CUFF.48797.1
CUFF.890.1	ENSOART0000000100	0.95397	CUFF.48463.1
CUFF.890.1	ENSOART0000000007	0.95577	CUFF.48218.2
CUFF.890.1	ENSOART0000000220	0.90013	CUFF.48218.1

CUFF.890.1	ENSOART000000168	0.94327	CUFF.4772.1
CUFF.890.1	ENSOART000000164	0.90177	CUFF.46517.1
CUFF.890.1	ENSOART000000141	0.9693	CUFF.43826.1
CUFF.890.1	ENSOART000000109	0.92763	CUFF.4158.1
CUFF.890.1	ENSOART000000091	0.90853	CUFF.4083.1
CUFF.890.1	ENSOART000000085	0.90953	CUFF.40184.1
CUFF.890.1	ENSOART000000079	0.92047	CUFF.39343.1
CUFF.890.1	ENSOART000000061	0.90273	CUFF.39330.1
CUFF.890.1	ENSOART000000053	0.90283	CUFF.3868.1
CUFF.890.1	ENSOART000000037	0.9181	CUFF.38420.1
CUFF.890.1	ENSOART000000028	0.95063	CUFF.38227.1
CUFF.8537.1	ENSOART000000164	0.93987	CUFF.36747.1
CUFF.5512.1	ENSOART000000203	0.96977	CUFF.35948.1
CUFF.5512.1	ENSOART000000007	0.93567	CUFF.35849.1
CUFF.51159.1	ENSOART000000203	0.91537	CUFF.34667.1
CUFF.50693.1	ENSOART000000203	0.9194	CUFF.34667.2
CUFF.50693.1	ENSOART000000007	0.92377	CUFF.33831.1
CUFF.49979.1	ENSOART000000041	0.90407	CUFF.30001.1
CUFF.49664.1	ENSOART000000203	0.9194	CUFF.29175.1
CUFF.49664.1	ENSOART000000007	0.92377	CUFF.29003.1
CUFF.48111.1	ENSOART000000203	0.9194	CUFF.27530.1
CUFF.48111.1	ENSOART000000007	0.92377	CUFF.23573.1
CUFF.47846.1	ENSOART000000164	0.98617	CUFF.23411.1
CUFF.47846.1	ENSOART000000143	0.9008	CUFF.20442.1
CUFF.47846.1	ENSOART000000131	0.91793	CUFF.1889.1
CUFF.47846.1	ENSOART000000129	0.95653	CUFF.17773.1
CUFF.47846.1	ENSOART000000127	0.92757	CUFF.17065.1
CUFF.47846.1	ENSOART000000115	0.92897	CUFF.16715.1
CUFF.47846.1	ENSOART000000053	0.9031	CUFF.1250.2
CUFF.4762.4	ENSOART000000164	0.95777	CUFF.12383.1
CUFF.47332.1	ENSOART000000187	0.9567	CUFF.12219.2
CUFF.47332.1	ENSOART000000085	0.9155	CUFF.12219.1
CUFF.47332.1	ENSOART000000164	0.93443	CUFF.26021.1
CUFF.47332.1	ENSOART000000162	0.90727	CUFF.3120.1
CUFF.47332.1	ENSOART000000155	0.9104	CUFF.35331.1
CUFF.47332.1	ENSOART000000143	0.90463	CUFF.18719.1
CUFF.47332.1	ENSOART000000143	0.915	CUFF.41701.1
CUFF.47332.1	ENSOART000000131	0.93043	CUFF.31851.1
CUFF.47332.1	ENSOART000000131	0.9056	CUFF.42039.1
CUFF.47332.1	ENSOART000000129	0.96693	CUFF.18700.1
CUFF.47332.1	ENSOART000000127	0.95953	CUFF.46704.2
CUFF.47332.1	ENSOART000000125	0.9379	CUFF.46167.1
CUFF.47332.1	ENSOART000000115	0.92747	CUFF.1399.1
CUFF.47332.1	ENSOART000000053	0.90623	CUFF.47654.1
CUFF.47332.1	ENSOART000000053	0.9101	CUFF.34977.1
CUFF.47332.1	ENSOART000000037	0.90617	CUFF.30780.1
CUFF.47332.1	ENSOART000000012	0.9327	CUFF.15892.1

CUFF.47308.1	ENSOART000000164	0.9614	CUFF.6511.2
CUFF.46290.1	ENSOART000000101	0.99223	CUFF.43596.1
CUFF.46290.1	ENSOART000000223	0.9892	CUFF.29948.1
CUFF.45638.1	ENSOART000000203	0.9194	CUFF.30499.1
CUFF.45638.1	ENSOART000000007	0.92377	CUFF.42025.2
CUFF.44995.1	ENSOART000000162	0.93637	CUFF.31681.2
CUFF.44995.1	ENSOART000000155	0.90867	CUFF.14944.1
CUFF.44995.1	ENSOART000000143	0.90173	CUFF.16414.1
CUFF.44995.1	ENSOART000000143	0.91857	CUFF.19617.1
CUFF.44995.1	ENSOART000000131	0.95653	CUFF.48201.1
CUFF.44995.1	ENSOART000000131	0.92947	CUFF.24971.3
CUFF.44995.1	ENSOART000000129	0.9596	CUFF.48912.1
CUFF.44995.1	ENSOART000000128	0.92603	CUFF.38336.3
CUFF.44995.1	ENSOART000000127	0.95133	CUFF.6511.1
CUFF.44995.1	ENSOART000000125	0.93717	CUFF.38336.2
CUFF.44995.1	ENSOART000000115	0.96473	CUFF.18704.1
CUFF.44995.1	ENSOART000000079	0.90083	CUFF.31644.1
CUFF.44995.1	ENSOART000000053	0.91373	CUFF.6627.2
CUFF.44995.1	ENSOART000000053	0.9283	CUFF.19414.1
CUFF.44995.1	ENSOART000000012	0.90583	CUFF.1838.1
CUFF.44461.2	ENSOART000000203	0.98493	CUFF.48910.1
CUFF.44461.2	ENSOART000000007	0.97063	CUFF.35005.1
CUFF.44461.1	ENSOART000000203	0.98963	CUFF.621.1
CUFF.44461.1	ENSOART000000007	0.97247	CUFF.21157.1
CUFF.44117.1	ENSOART000000164	0.9556	CUFF.44763.1
CUFF.43784.1	ENSOART000000203	0.98923	CUFF.34936.1
CUFF.43784.1	ENSOART000000007	0.9724	CUFF.44763.2
CUFF.43360.1	ENSOART000000203	0.91517	CUFF.806.2
CUFF.43336.1	ENSOART000000164	0.9362	CUFF.29848.1
CUFF.42787.1	ENSOART000000203	0.9916	CUFF.28890.1
CUFF.42787.1	ENSOART000000007	0.9723	CUFF.7946.1
CUFF.42731.1	ENSOART000000226	0.99827	CUFF.48129.1
CUFF.42731.1	ENSOART000000170	0.98803	CUFF.17359.1
CUFF.42731.1	ENSOART000000172	0.9959	CUFF.17080.1
CUFF.42731.1	ENSOART000000136	0.91703	CUFF.44234.1
CUFF.42731.1	ENSOART000000068	0.99933	CUFF.48400.2
CUFF.42731.1	ENSOART000000048	0.91003	CUFF.23095.1
CUFF.42731.1	CUFF.2467.1	0.9999	CUFF.3702.2
CUFF.42731.1	ENSOART000000092	0.98743	CUFF.17357.1
CUFF.42731.1	CUFF.29360.1	0.94513	CUFF.30641.1
CUFF.42462.1	ENSOART000000187	0.91113	CUFF.18862.1
CUFF.42462.1	ENSOART000000203	0.9043	CUFF.43860.1
CUFF.42462.1	ENSOART000000144	0.95603	CUFF.8298.1
CUFF.42462.1	ENSOART000000115	0.92217	CUFF.16639.1
CUFF.42462.1	ENSOART000000079	0.92497	CUFF.23210.1
CUFF.42462.1	ENSOART000000014	0.9716	CUFF.42038.1
CUFF.42195.1	ENSOART000000184	0.91373	CUFF.25769.1

CUFF.42195.1	ENSOART000000144	0.96033	CUFF.32184.2
CUFF.42195.1	ENSOART000000143	0.90077	CUFF.3702.3
CUFF.42195.1	ENSOART000000131	0.9047	CUFF.9991.1
CUFF.42195.1	ENSOART000000115	0.9351	CUFF.18798.1
CUFF.42195.1	ENSOART000000115	0.90307	CUFF.29850.1
CUFF.42195.1	ENSOART000000079	0.92947	CUFF.24275.9
CUFF.42195.1	ENSOART000000014	0.9954	CUFF.13832.1
CUFF.41792.1	ENSOART000000101	0.94747	CUFF.11659.1
CUFF.41792.1	ENSOART000000223	0.93303	CUFF.9302.1
CUFF.41470.1	ENSOART000000203	0.99053	CUFF.50685.1
CUFF.41470.1	ENSOART000000007	0.97257	CUFF.13878.2
CUFF.4097.1	ENSOART000000187	0.9994	CUFF.23934.4
CUFF.4097.1	ENSOART000000085	0.90397	CUFF.6248.2
CUFF.4097.1	ENSOART000000037	0.9212	CUFF.5933.1
CUFF.39522.1	ENSOART000000226	0.9847	CUFF.31853.1
CUFF.39522.1	ENSOART000000170	0.97347	CUFF.11350.2
CUFF.39522.1	ENSOART000000172	0.99853	CUFF.3702.1
CUFF.39522.1	ENSOART000000136	0.96703	CUFF.9539.1
CUFF.39522.1	ENSOART000000068	0.9864	CUFF.30314.1
CUFF.39522.1	ENSOART000000048	0.92603	CUFF.6248.3
CUFF.39522.1	CUFF.2467.1	0.98603	CUFF.36923.5
CUFF.39522.1	ENSOART000000092	0.97777	CUFF.36010.1
CUFF.39522.1	CUFF.29360.1	0.973	CUFF.18692.1
CUFF.38508.1	ENSOART000000203	0.9194	CUFF.38537.1
CUFF.38508.1	ENSOART000000007	0.92377	CUFF.32424.1
CUFF.38420.2	ENSOART000000063	0.9981	CUFF.8234.2
CUFF.38420.2	ENSOART000000085	0.94623	CUFF.20975.1
CUFF.38420.2	ENSOART000000147	0.9176	CUFF.28297.1
CUFF.38420.2	ENSOART000000143	0.92937	CUFF.27119.1
CUFF.3811.1	ENSOART000000203	0.9194	CUFF.1619.1
CUFF.3811.1	ENSOART000000007	0.92377	CUFF.35070.1
CUFF.36701.1	ENSOART000000203	0.9194	CUFF.43074.1
CUFF.36701.1	ENSOART000000007	0.92377	CUFF.23294.1
CUFF.36041.1	ENSOART000000162	0.90027	CUFF.3704.1
CUFF.36041.1	ENSOART000000131	0.91173	CUFF.37359.1
CUFF.36041.1	ENSOART000000131	0.93097	CUFF.9114.1
CUFF.36041.1	ENSOART000000129	0.93013	CUFF.49338.1
CUFF.36041.1	ENSOART000000128	0.91187	CUFF.34348.1
CUFF.36041.1	ENSOART000000127	0.91907	CUFF.43305.1
CUFF.36041.1	ENSOART000000125	0.91553	CUFF.48805.1
CUFF.36041.1	ENSOART000000105	0.93443	CUFF.23671.3
CUFF.36041.1	ENSOART000000012	0.90367	CUFF.2354.1
CUFF.35915.1	ENSOART000000187	0.97297	CUFF.46299.2
CUFF.35915.1	ENSOART000000063	0.91033	CUFF.23318.2
CUFF.35915.1	ENSOART000000085	0.9374	CUFF.21921.9
CUFF.35915.1	ENSOART000000129	0.9004	CUFF.43625.1
CUFF.35915.1	ENSOART000000037	0.92153	CUFF.43.1

CUFF.35868.1	ENSOART000000063	0.967	CUFF.23934.3
CUFF.35868.1	ENSOART000000085	0.9621	CUFF.7188.1
CUFF.35868.1	ENSOART000000147	0.9549	CUFF.37765.3
CUFF.35868.1	ENSOART000000143	0.95437	CUFF.20256.3
CUFF.35868.1	ENSOART000000119	0.90107	CUFF.23934.1
CUFF.35868.1	ENSOART000000061	0.91227	CUFF.35535.1
CUFF.3581.1	ENSOART000000203	0.98843	CUFF.36923.7
CUFF.3581.1	ENSOART000000007	0.95763	CUFF.41662.7
CUFF.35761.1	ENSOART000000203	0.9194	CUFF.31247.1
CUFF.35761.1	ENSOART000000007	0.92377	CUFF.18658.2
CUFF.35318.1	ENSOART000000203	0.9166	CUFF.2570.1
CUFF.35318.1	ENSOART000000007	0.9022	CUFF.9763.1
CUFF.35006.1	ENSOART000000068	0.9103	CUFF.39851.4
CUFF.34739.1	ENSOART000000203	0.9194	CUFF.41604.1
CUFF.34739.1	ENSOART000000007	0.92377	CUFF.544.5
CUFF.33008.1	ENSOART000000226	0.93877	CUFF.8234.1
CUFF.33008.1	ENSOART000000170	0.92633	CUFF.9561.1
CUFF.33008.1	ENSOART000000172	0.97077	CUFF.39579.1
CUFF.33008.1	ENSOART000000136	0.99573	CUFF.18800.16
CUFF.33008.1	ENSOART000000068	0.94227	CUFF.41054.3
CUFF.33008.1	ENSOART000000048	0.9077	CUFF.43625.2
CUFF.33008.1	CUFF.2467.1	0.9405	CUFF.18479.1
CUFF.33008.1	ENSOART000000092	0.9388	CUFF.49118.1
CUFF.33008.1	ENSOART000000081	0.95017	CUFF.27786.5
CUFF.33008.1	CUFF.29360.1	0.97147	CUFF.18800.14
CUFF.32907.1	ENSOART000000063	0.9996	CUFF.9041.1
CUFF.32907.1	ENSOART000000085	0.9462	CUFF.50917.1
CUFF.32907.1	ENSOART000000147	0.9173	CUFF.4966.7
CUFF.32907.1	ENSOART000000143	0.93003	CUFF.41391.1
CUFF.32466.1	ENSOART000000203	0.9194	CUFF.544.4
CUFF.32466.1	ENSOART000000007	0.92377	CUFF.46717.6
CUFF.30184.1	ENSOART000000226	0.9383	CUFF.42237.2
CUFF.30184.1	ENSOART000000170	0.9237	CUFF.966.1
CUFF.30184.1	ENSOART000000172	0.94117	CUFF.16104.1
CUFF.30184.1	ENSOART000000068	0.9493	CUFF.21962.1
CUFF.30184.1	CUFF.2467.1	0.94457	CUFF.49525.1
CUFF.30184.1	ENSOART000000092	0.95287	CUFF.44648.3
CUFF.30009.1	ENSOART000000100	0.93237	CUFF.3832.1
CUFF.30009.1	ENSOART000000085	0.9229	CUFF.3025.2
CUFF.30009.1	ENSOART000000007	0.93567	CUFF.19876.1
CUFF.30009.1	ENSOART000000168	0.95303	CUFF.19478.2
CUFF.30009.1	ENSOART000000162	0.9734	CUFF.19476.1
CUFF.30009.1	ENSOART000000155	0.98167	CUFF.39574.1
CUFF.30009.1	ENSOART000000147	0.95983	CUFF.22023.1
CUFF.30009.1	ENSOART000000143	0.91797	CUFF.50586.1
CUFF.30009.1	ENSOART000000143	0.93793	
CUFF.30009.1	ENSOART000000141	0.91337	

CUFF.30009.1	ENSOART000000131	0.9664
CUFF.30009.1	ENSOART000000131	0.98787
CUFF.30009.1	ENSOART000000129	0.93063
CUFF.30009.1	ENSOART000000128	0.94933
CUFF.30009.1	ENSOART000000127	0.9474
CUFF.30009.1	ENSOART000000125	0.97317
CUFF.30009.1	ENSOART000000119	0.9684
CUFF.30009.1	ENSOART000000119	0.9466
CUFF.30009.1	ENSOART000000119	0.90203
CUFF.30009.1	ENSOART000000109	0.9496
CUFF.30009.1	ENSOART000000105	0.97563
CUFF.30009.1	ENSOART000000091	0.96813
CUFF.30009.1	ENSOART000000085	0.96557
CUFF.30009.1	ENSOART000000079	0.9299
CUFF.30009.1	ENSOART000000061	0.98517
CUFF.30009.1	ENSOART000000053	0.97563
CUFF.30009.1	ENSOART000000053	0.97047
CUFF.30009.1	ENSOART000000037	0.9309
CUFF.30009.1	ENSOART000000028	0.9682
CUFF.30009.1	ENSOART000000012	0.95777
CUFF.2983.1	ENSOART000000187	0.9414
CUFF.2983.1	ENSOART000000164	0.93963
CUFF.28564.1	ENSOART000000226	0.9161
CUFF.28564.1	ENSOART000000170	0.92087
CUFF.28564.1	ENSOART000000172	0.9085
CUFF.28564.1	ENSOART000000068	0.90903
CUFF.28564.1	ENSOART000000068	0.9009
CUFF.28564.1	ENSOART000000048	0.92
CUFF.28564.1	CUFF.2467.1	0.91177
CUFF.28564.1	ENSOART000000092	0.9254
CUFF.28564.1	ENSOART000000058	0.9054
CUFF.28564.1	ENSOART000000035	0.90983
CUFF.28506.4	ENSOART000000226	0.97023
CUFF.28506.4	ENSOART000000170	0.95787
CUFF.28506.4	ENSOART000000172	0.98953
CUFF.28506.4	ENSOART000000136	0.9744
CUFF.28506.4	ENSOART000000068	0.9726
CUFF.28506.4	ENSOART000000048	0.9375
CUFF.28506.4	CUFF.2467.1	0.97237
CUFF.28506.4	ENSOART000000092	0.9666
CUFF.28506.4	ENSOART000000081	0.90153
CUFF.28506.4	CUFF.29360.1	0.9718
CUFF.28506.6	ENSOART000000203	0.9335
CUFF.28506.6	ENSOART000000007	0.96433
CUFF.26354.1	ENSOART000000203	0.9107
CUFF.26354.1	ENSOART000000007	0.93323
CUFF.26155.1	ENSOART000000164	0.95763

CUFF.25809.1	ENSOART000000203	0.9194
CUFF.25809.1	ENSOART000000007	0.92377
CUFF.25564.1	ENSOART000000203	0.92763
CUFF.25564.1	ENSOART000000007	0.92233
CUFF.25254.1	ENSOART000000063	0.90837
CUFF.25254.1	ENSOART000000085	0.93197
CUFF.25254.1	ENSOART000000164	0.97457
CUFF.25254.1	ENSOART000000155	0.9158
CUFF.25254.1	ENSOART000000147	0.94003
CUFF.25254.1	ENSOART000000143	0.95753
CUFF.25254.1	ENSOART000000131	0.9445
CUFF.25254.1	ENSOART000000131	0.92467
CUFF.25254.1	ENSOART000000129	0.96163
CUFF.25254.1	ENSOART000000127	0.94753
CUFF.25254.1	ENSOART000000125	0.91363
CUFF.25254.1	ENSOART000000119	0.90323
CUFF.25254.1	ENSOART000000115	0.93267
CUFF.25254.1	ENSOART000000105	0.90543
CUFF.25254.1	ENSOART000000085	0.909
CUFF.25254.1	ENSOART000000061	0.93213
CUFF.25254.1	ENSOART000000053	0.90547
CUFF.25254.1	ENSOART000000053	0.93303
CUFF.25254.1	ENSOART000000028	0.9281
CUFF.25254.1	ENSOART000000012	0.9136
CUFF.25015.1	ENSOART000000203	0.95867
CUFF.25015.1	ENSOART000000007	0.91807
CUFF.24964.1	ENSOART000000226	0.99533
CUFF.24964.1	ENSOART000000170	0.98493
CUFF.24964.1	ENSOART000000172	0.99967
CUFF.24964.1	ENSOART000000136	0.93993
CUFF.24964.1	ENSOART000000068	0.99737
CUFF.24964.1	ENSOART000000048	0.91917
CUFF.24964.1	CUFF.2467.1	0.9975
CUFF.24964.1	ENSOART000000092	0.98637
CUFF.24964.1	CUFF.29360.1	0.95907
CUFF.24390.1	ENSOART000000144	0.91603
CUFF.24390.1	ENSOART000000131	0.90207
CUFF.24390.1	ENSOART000000115	0.93077
CUFF.24390.1	ENSOART000000079	0.90967
CUFF.24390.1	ENSOART000000014	0.9706
CUFF.23641.1	ENSOART000000007	0.9004
CUFF.2342.2	ENSOART000000203	0.9194
CUFF.2342.2	ENSOART000000007	0.92377
CUFF.23155.1	ENSOART000000203	0.9194
CUFF.23155.1	ENSOART000000007	0.92377
CUFF.22959.1	ENSOART000000203	0.92767
CUFF.22799.1	ENSOART000000203	0.9194

CUFF.22799.1	ENSOART000000007	0.92377
CUFF.21553.1	ENSOART000000187	0.997
CUFF.21553.1	ENSOART000000085	0.90227
CUFF.21553.1	ENSOART000000037	0.92123
CUFF.20864.1	ENSOART000000063	0.9255
CUFF.20864.1	ENSOART000000100	0.9119
CUFF.20864.1	ENSOART000000085	0.96597
CUFF.20864.1	ENSOART000000007	0.91003
CUFF.20864.1	ENSOART000000168	0.91067
CUFF.20864.1	ENSOART000000164	0.9217
CUFF.20864.1	ENSOART000000162	0.93023
CUFF.20864.1	ENSOART000000155	0.94457
CUFF.20864.1	ENSOART000000147	0.9841
CUFF.20864.1	ENSOART000000143	0.92493
CUFF.20864.1	ENSOART000000143	0.98783
CUFF.20864.1	ENSOART000000131	0.96527
CUFF.20864.1	ENSOART000000131	0.96347
CUFF.20864.1	ENSOART000000129	0.9576
CUFF.20864.1	ENSOART000000128	0.9342
CUFF.20864.1	ENSOART000000127	0.9603
CUFF.20864.1	ENSOART000000125	0.93587
CUFF.20864.1	ENSOART000000119	0.94687
CUFF.20864.1	ENSOART000000115	0.94083
CUFF.20864.1	ENSOART000000109	0.90517
CUFF.20864.1	ENSOART000000105	0.94007
CUFF.20864.1	ENSOART000000085	0.9391
CUFF.20864.1	ENSOART000000079	0.9041
CUFF.20864.1	ENSOART000000061	0.96367
CUFF.20864.1	ENSOART000000053	0.93207
CUFF.20864.1	ENSOART000000053	0.95347
CUFF.20864.1	ENSOART000000037	0.91903
CUFF.20864.1	ENSOART000000028	0.9423
CUFF.20864.1	ENSOART000000012	0.93263
CUFF.20106.1	ENSOART000000203	0.93367
CUFF.20106.1	ENSOART000000007	0.96627
CUFF.20106.2	ENSOART000000203	0.93337
CUFF.20106.2	ENSOART000000007	0.9655
CUFF.19196.1	ENSOART000000187	0.90303
CUFF.19196.1	ENSOART000000164	0.98453
CUFF.19196.1	ENSOART000000131	0.90347
CUFF.19196.1	ENSOART000000129	0.95557
CUFF.19196.1	ENSOART000000127	0.9268
CUFF.19196.1	ENSOART000000115	0.91017
CUFF.17798.1	ENSOART000000203	0.93293
CUFF.17379.1	ENSOART000000187	0.9066
CUFF.17379.1	ENSOART000000131	0.9082
CUFF.17379.1	ENSOART000000129	0.9203

CUFF.17379.1	ENSOART000000127	0.90973
CUFF.17379.1	ENSOART000000115	0.9428
CUFF.17379.1	ENSOART000000079	0.92567
CUFF.17379.1	ENSOART000000014	0.9303
CUFF.16671.1	ENSOART000000203	0.9194
CUFF.16671.1	ENSOART000000007	0.92377
CUFF.16281.1	ENSOART000000081	0.9602
CUFF.16034.1	ENSOART000000187	0.90873
CUFF.16034.1	ENSOART000000164	0.9557
CUFF.15937.1	ENSOART000000007	0.91123
CUFF.14818.1	ENSOART000000063	0.9277
CUFF.14818.1	ENSOART000000085	0.97467
CUFF.14818.1	ENSOART000000168	0.90273
CUFF.14818.1	ENSOART000000162	0.92767
CUFF.14818.1	ENSOART000000155	0.92043
CUFF.14818.1	ENSOART000000147	0.9589
CUFF.14818.1	ENSOART000000143	0.9319
CUFF.14818.1	ENSOART000000143	0.95683
CUFF.14818.1	ENSOART000000131	0.94547
CUFF.14818.1	ENSOART000000131	0.9357
CUFF.14818.1	ENSOART000000129	0.9322
CUFF.14818.1	ENSOART000000128	0.94183
CUFF.14818.1	ENSOART000000127	0.9389
CUFF.14818.1	ENSOART000000125	0.92213
CUFF.14818.1	ENSOART000000119	0.93577
CUFF.14818.1	ENSOART000000115	0.9319
CUFF.14818.1	ENSOART000000105	0.90987
CUFF.14818.1	ENSOART000000085	0.92927
CUFF.14818.1	ENSOART000000079	0.91163
CUFF.14818.1	ENSOART000000061	0.9344
CUFF.14818.1	ENSOART000000053	0.91347
CUFF.14818.1	ENSOART000000053	0.93707
CUFF.14818.1	ENSOART000000037	0.92887
CUFF.14818.1	ENSOART000000028	0.91567
CUFF.14163.1	ENSOART000000203	0.9194
CUFF.14163.1	ENSOART000000007	0.92377
CUFF.139.1	ENSOART000000203	0.9337
CUFF.139.1	ENSOART000000007	0.9671
CUFF.13439.1	ENSOART000000164	0.94503
CUFF.12663.2	ENSOART000000223	0.90477
CUFF.10217.1	ENSOART000000203	0.9194
CUFF.10217.1	ENSOART000000007	0.92377

protein-coding genes of 8 lincRNAs had more than 10 co-expressed protein-coding genes.
Number of coexpressed genes

- 30
- 29
- 28
- 24
- 20
- 17
- 15
- 14
- 10
- 10
- 10
- 9
- 9
- 9
- 9
- 8
- 8
- 7
- 7
- 7
- 7
- 7
- 6
- 6
- 6
- 6
- 6
- 5
- 5
- 5
- 5
- 4
- 4
- 4
- 3
- 3
- 3
- 3
- 3
- 3
- 3
- 2
- 2
- 2
- 2
- 2

Draft

[illegible]

1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
0
0
0
0
0
0
0
0
0
0
0
0

[illegible]

[illegible]

[illegible]

[illegible]

Category	Term	Count	%	PValue
GOTERM_	GO:0006414~translational elongation	16	94.11764706	3.54E-32
GOTERM_	GO:0006412~translation	16	94.11764706	4.07E-24
GOTERM_	GO:0044267~cellular protein metabolic proces	16	94.11764706	2.81E-11
GOTERM_	GO:0034645~cellular macromolecule biosynth	16	94.11764706	3.90E-10
GOTERM_	GO:0019538~protein metabolic process	16	94.11764706	3.90E-10
GOTERM_	GO:0009059~macromolecule biosynthetic pro	16	94.11764706	4.33E-10
GOTERM_	GO:0010467~gene expression	16	94.11764706	1.01E-09
GOTERM_	GO:0044249~cellular biosynthetic process	16	94.11764706	7.73E-09
GOTERM_	GO:0009058~biosynthetic process	16	94.11764706	1.18E-08
GOTERM_	GO:0042274~ribosomal small subunit biogene	4	23.52941176	1.96E-07
GOTERM_	GO:0044260~cellular macromolecule metaboli	16	94.11764706	3.36E-06
GOTERM_	GO:0042254~ribosome biogenesis	5	29.41176471	8.92E-06
GOTERM_	GO:0043170~macromolecule metabolic proce	16	94.11764706	1.25E-05
GOTERM_	GO:0022613~ribonucleoprotein complex bioge	5	29.41176471	4.13E-05
GOTERM_	GO:0044237~cellular metabolic process	16	94.11764706	1.08E-04
GOTERM_	GO:0006364~rRNA processing	4	23.52941176	1.41E-04
GOTERM_	GO:0016072~rRNA metabolic process	4	23.52941176	1.60E-04
GOTERM_	GO:0044238~primary metabolic process	16	94.11764706	1.96E-04
GOTERM_	GO:0008152~metabolic process	16	94.11764706	7.95E-04
GOTERM_	GO:0034470~ncRNA processing	4	23.52941176	1.13E-03
GOTERM_	GO:0034101~erythrocyte homeostasis	3	17.64705882	1.37E-03
GOTERM_	GO:0034660~ncRNA metabolic process	4	23.52941176	2.04E-03
GOTERM_	GO:0044085~cellular component biogenesis	6	35.29411765	4.01E-03
GOTERM_	GO:0048872~homeostasis of number of cells	3	17.64705882	5.59E-03
GOTERM_	GO:0033119~negative regulation of RNA splici	2	11.76470588	5.66E-03
GOTERM_	GO:0009987~cellular process	17	100	9.32E-03
GOTERM_	GO:0045934~negative regulation of nucleobas	4	23.52941176	1.87E-02
GOTERM_	GO:0051172~negative regulation of nitrogen c	4	23.52941176	1.94E-02
GOTERM_	GO:0006396~RNA processing	4	23.52941176	2.22E-02
GOTERM_	GO:0043484~regulation of RNA splicing	2	11.76470588	2.24E-02
GOTERM_	GO:0030097~hemopoiesis	3	17.64705882	0.028628
GOTERM_	GO:0048534~hemopoietic or lymphoid organ i	3	17.64705882	0.034217
GOTERM_	GO:0002520~immune system development	3	17.64705882	0.038164
GOTERM_	GO:0031324~negative regulation of cellular m	4	23.52941176	0.045039
GOTERM_	GO:0010605~negative regulation of macromol	4	23.52941176	0.04726
GOTERM_	GO:0030218~erythrocyte differentiation	2	11.76470588	0.047666
GOTERM_	GO:0006413~translational initiation	2	11.76470588	0.04983
GOTERM_	GO:0042592~homeostatic process	4	23.52941176	0.050031
GOTERM_	GO:0009892~negative regulation of metabolic	4	23.52941176	0.054946
GOTERM_	GO:0051253~negative regulation of RNA meta	3	17.64705882	0.062114
GOTERM_	GO:0016070~RNA metabolic process	4	23.52941176	0.085695

Genes	List Total	Pop Hits	Pop Total	Fold Enrich	Bonferroni	Benjamini	FDR
RPL19, RPL	17	101	14116	131.5411	6.55E-30	6.55E-30	4.38E-29
RPL19, RPL	17	331	14116	40.13791	7.53E-22	3.76E-22	5.03E-21
RPL19, RPL	17	2355	14116	5.641464	5.20E-09	1.73E-09	3.48E-08
RPL19, RPL	17	2812	14116	4.724626	7.22E-08	1.80E-08	4.83E-07
RPL19, RPL	17	2812	14116	4.724626	7.22E-08	1.80E-08	4.83E-07
RPL19, RPL	17	2832	14116	4.69126	8.02E-08	1.60E-08	5.36E-07
RPL19, RPL	17	2999	14116	4.430026	1.87E-07	3.12E-08	1.25E-06
RPL19, RPL	17	3442	14116	3.859863	1.43E-06	2.04E-07	9.57E-06
RPL19, RPL	17	3542	14116	3.750888	2.18E-06	2.73E-07	1.46E-05
RPS19, RPS	17	11	14116	301.9465	3.63E-05	4.03E-06	2.43E-04
RPL19, RPL	17	5214	14116	2.548072	6.21E-04	6.22E-05	0.004157
RPS19, RPS	17	122	14116	34.03086	0.001649	1.50E-04	0.011036
RPL19, RPL	17	5710	14116	2.326733	0.002308	1.93E-04	0.015454
RPS19, RPS	17	180	14116	23.06536	0.007608	5.87E-04	0.051071
RPL19, RPL	17	6636	14116	2.002057	0.019693	0.00142	0.132945
RPS19, RPS	17	92	14116	36.1023	0.025766	0.001739	0.174441
RPS19, RPS	17	96	14116	34.59804	0.029183	0.001849	0.197902
RPL19, RPL	17	6923	14116	1.919059	0.035632	0.002132	0.242379
RPL19, RPL	17	7647	14116	1.737367	0.136878	0.008144	0.979719
RPS19, RPS	17	187	14116	17.76156	0.188475	0.010931	1.387128
RPS19, RPS	17	49	14116	50.83794	0.224476	0.01263	1.685964
RPS19, RPS	17	230	14116	14.44092	0.315105	0.017862	2.499778
RPS19, FAB	17	1001	14116	4.977141	0.524309	0.033208	4.848051
RPS19, RPS	17	100	14116	24.91059	0.645416	0.044078	6.699775
RPS26, RPS	17	5	14116	332.1412	0.649782	0.042775	6.77705
FABP9, RPL	17	10541	14116	1.339152	0.823158	0.066953	10.94172
RPS26, RPS	17	512	14116	6.487132	0.969439	0.125545	20.80823
RPS26, RPS	17	519	14116	6.399637	0.973137	0.125377	21.48845
RPS19, RPS	17	547	14116	6.072051	0.984399	0.138075	24.29096
RPS26, RPS	17	20	14116	83.03529	0.98499	0.1348	24.48617
RPS19, RPL	17	236	14116	10.55533	0.995362	0.163991	30.19103
RPS19, RPL	17	260	14116	9.580995	0.998405	0.18761	35.00189
RPS19, RPL	17	276	14116	9.025575	0.999252	0.20145	38.21357
RPS26, RPS	17	720	14116	4.613072	0.999802	0.227676	43.46095
RPS26, RPS	17	734	14116	4.525084	0.999871	0.231584	45.06689
RPS19, RPS	17	43	14116	38.62107	0.999881	0.227522	45.35597
RPS17, RPS	17	45	14116	36.90458	0.999922	0.231006	46.87293
RPS19, RPS	17	751	14116	4.422652	0.999925	0.226347	47.01188
RPS26, RPS	17	780	14116	4.25822	0.999971	0.240526	50.30579
RPS26, RPS	17	362	14116	6.881378	0.999993	0.262281	54.77387
RPS19, RPS	17	938	14116	3.540951	1	0.339235	66.9973

Category	Term	Count	%	PValue	Genes	List Total
GOTERM_	GO:0006414~translational elongation	13	86.66667	8.24E-25	RPL23A, RP	15
GOTERM_	GO:0006412~translation	13	86.66667	1.97E-18	RPL23A, RP	15
GOTERM_	GO:0044267~cellular protein metabolic pro	13	86.66667	2.96E-08	RPL23A, RP	15
GOTERM_	GO:0019538~protein metabolic process	13	86.66667	2.33E-07	RPL23A, RP	15
GOTERM_	GO:0034645~cellular macromolecule biosyr	13	86.66667	2.33E-07	RPL23A, RP	15
GOTERM_	GO:0009059~macromolecule biosynthetic p	13	86.66667	2.53E-07	RPL23A, RP	15
GOTERM_	GO:0010467~gene expression	13	86.66667	4.90E-07	RPL23A, RP	15
GOTERM_	GO:0044249~cellular biosynthetic process	13	86.66667	2.38E-06	RPL23A, RP	15
GOTERM_	GO:0009058~biosynthetic process	13	86.66667	3.31E-06	RPL23A, RP	15
GOTERM_	GO:0042274~ribosomal small subunit bioge	3	20	5.00E-05	RPS17, RPS	15
GOTERM_	GO:0042254~ribosome biogenesis	4	26.66667	2.14E-04	RPS17, RPS	15
GOTERM_	GO:0044260~cellular macromolecule metat	13	86.66667	2.54E-04	RPL23A, RP	15
GOTERM_	GO:0022613~ribonucleoprotein complex bic	4	26.66667	6.69E-04	RPS17, RPS	15
GOTERM_	GO:0043170~macromolecule metabolic pro	13	86.66667	6.83E-04	RPL23A, RP	15
GOTERM_	GO:0044237~cellular metabolic process	13	86.66667	0.003396	RPL23A, RP	15
GOTERM_	GO:0006364~rRNA processing	3	20	0.003634	RPS17, RPS	15
GOTERM_	GO:0016072~rRNA metabolic process	3	20	0.003949	RPS17, RPS	15
GOTERM_	GO:0033119~negative regulation of RNA spl	2	13.33333	0.00495	RPS26, RPS	15
GOTERM_	GO:0044238~primary metabolic process	13	86.66667	0.005281	RPL23A, RP	15
GOTERM_	GO:0045934~negative regulation of nucleot	4	26.66667	0.01281	RPS26, RPS	15
GOTERM_	GO:0051172~negative regulation of nitroge	4	26.66667	0.013288	RPS26, RPS	15
GOTERM_	GO:0044085~cellular component biogenesis	5	33.33333	0.014164	FABP9, RPS	15
GOTERM_	GO:0034470~ncRNA processing	3	20	0.014306	RPS17, RPS	15
GOTERM_	GO:0008152~metabolic process	13	86.66667	0.014578	RPL23A, RP	15
GOTERM_	GO:0009987~cellular process	15	100	0.016728	FABP9, RPL	15
GOTERM_	GO:0043484~regulation of RNA splicing	2	13.33333	0.019663	RPS26, RPS	15
GOTERM_	GO:0034660~ncRNA metabolic process	3	20	0.021144	RPS17, RPS	15
GOTERM_	GO:0031324~negative regulation of cellular	4	26.66667	0.031562	RPS26, RPS	15
GOTERM_	GO:0010605~negative regulation of macron	4	26.66667	0.033167	RPS26, RPS	15
GOTERM_	GO:0009892~negative regulation of metabo	4	26.66667	0.038744	RPS26, RPS	15
GOTERM_	GO:0006413~translational initiation	2	13.33333	0.043737	RPS17, RPS	15
GOTERM_	GO:0034101~erythrocyte homeostasis	2	13.33333	0.047537	RPS17, RPS	15
GOTERM_	GO:0051253~negative regulation of RNA me	3	20	0.048701	RPS26, RPS	15
GOTERM_	GO:0048872~homeostasis of number of cell	2	13.33333	0.09478	RPS17, RPS	15

Pop Hits	Pop Total	Fold Enrich	Bonferroni	Benjamini	FDR
101	14116	121.1274	1.53E-22	1.53E-22	1.02E-21
331	14116	36.96032	3.67E-16	1.83E-16	2.44E-15
2355	14116	5.194848	5.51E-06	1.84E-06	3.67E-05
2812	14116	4.350593	4.33E-05	1.08E-05	2.88E-04
2812	14116	4.350593	4.33E-05	1.08E-05	2.88E-04
2832	14116	4.319868	4.70E-05	9.40E-06	3.13E-04
2999	14116	4.079315	9.11E-05	1.52E-05	6.06E-04
3442	14116	3.55429	4.43E-04	6.33E-05	0.002953
3542	14116	3.453943	6.15E-04	7.69E-05	0.004096
11	14116	256.6545	0.009254	0.001032	0.061886
122	14116	30.85464	0.039002	0.00397	0.264544
5214	14116	2.34635	0.04609	0.00428	0.313694
180	14116	20.91259	0.117081	0.010323	0.825708
5710	14116	2.142534	0.119406	0.009734	0.843115
6636	14116	1.84356	0.468837	0.044186	4.125272
92	14116	30.68696	0.4919	0.044135	4.408239
96	14116	29.40833	0.520983	0.044959	4.782674
5	14116	376.4267	0.602653	0.052844	5.96046
6923	14116	1.767134	0.626523	0.053247	6.347581
512	14116	7.352083	0.909098	0.118569	14.75752
519	14116	7.252922	0.91694	0.116982	15.26808
1001	14116	4.700633	0.929585	0.118693	16.19473
187	14116	15.09733	0.931451	0.114699	16.34449
7647	14116	1.599826	0.934873	0.111977	16.6293
10541	14116	1.339152	0.956615	0.12255	18.85409
20	14116	94.10667	0.975121	0.137353	21.80388
230	14116	12.27478	0.981221	0.141772	23.25476
720	14116	5.228148	0.997434	0.198231	32.77995
734	14116	5.128429	0.998115	0.200732	34.14628
780	14116	4.825983	0.999357	0.223871	38.69942
45	14116	41.82519	0.999756	0.242156	42.52841
49	14116	38.41088	0.999884	0.2534	45.29411
362	14116	7.798895	0.999907	0.251888	46.1165
100	14116	18.82133	1	0.429506	70.86572

Category	Term	Count	%	PValue	Genes
GOTERM_	GO:0006414~translational elongation	14	87.5	6.00E-27	RPL19, RPL
GOTERM_	GO:0006412~translation	14	87.5	5.15E-20	RPL19, RPL
GOTERM_	GO:0044267~cellular protein metabolic process	14	87.5	5.67E-09	RPL19, RPL
GOTERM_	GO:0034645~cellular macromolecule biosynthetic	14	87.5	5.32E-08	RPL19, RPL
GOTERM_	GO:0019538~protein metabolic process	14	87.5	5.32E-08	RPL19, RPL
GOTERM_	GO:0009059~macromolecule biosynthetic proces	14	87.5	5.81E-08	RPL19, RPL
GOTERM_	GO:0010467~gene expression	14	87.5	1.19E-07	RPL19, RPL
GOTERM_	GO:0042274~ribosomal small subunit biogenesis	4	25	1.59E-07	RPS19, RPS
GOTERM_	GO:0044249~cellular biosynthetic process	14	87.5	6.67E-07	RPL19, RPL
GOTERM_	GO:0009058~biosynthetic process	14	87.5	9.52E-07	RPL19, RPL
GOTERM_	GO:0044260~cellular macromolecule metabolic p	14	87.5	1.07E-04	RPL19, RPL
GOTERM_	GO:0006364~rRNA processing	4	25	1.15E-04	RPS19, RPS
GOTERM_	GO:0016072~rRNA metabolic process	4	25	1.31E-04	RPS19, RPS
GOTERM_	GO:0042254~ribosome biogenesis	4	25	2.66E-04	RPS19, RPS
GOTERM_	GO:0043170~macromolecule metabolic process	14	87.5	3.16E-04	RPL19, RPL
GOTERM_	GO:0022613~ribonucleoprotein complex biogene	4	25	8.29E-04	RPS19, RPS
GOTERM_	GO:0034470~ncRNA processing	4	25	9.26E-04	RPS19, RPS
GOTERM_	GO:0034101~erythrocyte homeostasis	3	18.75	0.001204	RPS19, RPS
GOTERM_	GO:0034660~ncRNA metabolic process	4	25	0.001681	RPS19, RPS
GOTERM_	GO:0044237~cellular metabolic process	14	87.5	0.001823	RPL19, RPL
GOTERM_	GO:0044238~primary metabolic process	14	87.5	0.002956	RPL19, RPL
GOTERM_	GO:0048872~homeostasis of number of cells	3	18.75	0.004913	RPS19, RPS
GOTERM_	GO:0033119~negative regulation of RNA splicing	2	12.5	0.005303	RPS26, RPS
GOTERM_	GO:0008152~metabolic process	14	87.5	0.008992	RPL19, RPL
GOTERM_	GO:0045934~negative regulation of nucleobase, r	4	25	0.015588	RPS26, RPS
GOTERM_	GO:0051172~negative regulation of nitrogen com	4	25	0.016164	RPS26, RPS
GOTERM_	GO:0044085~cellular component biogenesis	5	31.25	0.018251	RPS19, FAB
GOTERM_	GO:0006396~RNA processing	4	25	0.018593	RPS19, RPS
GOTERM_	GO:0043484~regulation of RNA splicing	2	12.5	0.021053	RPS26, RPS
GOTERM_	GO:0030097~hemopoiesis	3	18.75	0.025324	RPS19, RPL
GOTERM_	GO:0048534~hemopoietic or lymphoid organ dev	3	18.75	0.030302	RPS19, RPL
GOTERM_	GO:0002520~immune system development	3	18.75	0.033822	RPS19, RPL
GOTERM_	GO:0031324~negative regulation of cellular meta	4	25	0.037994	RPS26, RPS
GOTERM_	GO:0010605~negative regulation of macromolecu	4	25	0.039897	RPS26, RPS
GOTERM_	GO:0042592~homeostatic process	4	25	0.042273	RPS19, RPS
GOTERM_	GO:0030218~erythrocyte differentiation	2	12.5	0.044753	RPS19, RPS
GOTERM_	GO:0009892~negative regulation of metabolic pr	4	25	0.046495	RPS26, RPS
GOTERM_	GO:0006413~translational initiation	2	12.5	0.046788	RPS17, RPS
GOTERM_	GO:0051253~negative regulation of RNA metabol	3	18.75	0.055263	RPS26, RPS
GOTERM_	GO:0016070~RNA metabolic process	4	25	0.0731	RPS19, RPS
GOTERM_	GO:0009987~cellular process	15	93.75	0.076096	FABP9, RPL
GOTERM_	GO:0030099~myeloid cell differentiation	2	12.5	0.094439	RPS19, RPS

List Total	Pop Hits	Pop Total	Fold Enrich	Bonferroni	Benjamini	FDR
16	101	14116	122.2921	1.12E-24	1.12E-24	7.43E-24
16	331	14116	37.31571	9.58E-18	4.79E-18	6.38E-17
16	2355	14116	5.244798	1.05E-06	3.51E-07	7.02E-06
16	2812	14116	4.392425	9.89E-06	2.47E-06	6.58E-05
16	2812	14116	4.392425	9.89E-06	2.47E-06	6.58E-05
16	2832	14116	4.361405	1.08E-05	2.16E-06	7.20E-05
16	2999	14116	4.11854	2.22E-05	3.70E-06	1.48E-04
16	11	14116	320.8182	2.96E-05	4.23E-06	1.97E-04
16	3442	14116	3.588466	1.24E-04	1.55E-05	8.26E-04
16	3542	14116	3.487154	1.77E-04	1.97E-05	0.001179
16	5214	14116	2.368911	0.019757	0.001993	0.132782
16	92	14116	38.3587	0.021196	0.001946	0.142553
16	96	14116	36.76042	0.024021	0.002024	0.161763
16	122	14116	28.92623	0.048215	0.003794	0.328501
16	5710	14116	2.163135	0.05714	0.004194	0.391007
16	180	14116	19.60556	0.142908	0.010228	1.021562
16	187	14116	18.87166	0.158235	0.010708	1.140413
16	49	14116	54.01531	0.20078	0.013097	1.481226
16	230	14116	15.34348	0.268654	0.017231	2.061703
16	6636	14116	1.861287	0.287784	0.017703	2.234396
16	6923	14116	1.784125	0.423365	0.027152	3.599458
16	100	14116	26.4675	0.59989	0.042682	5.917047
16	5	14116	352.9	0.628017	0.043955	6.372573
16	7647	14116	1.615209	0.813655	0.070446	10.58431
16	512	14116	6.892578	0.946182	0.114636	17.68136
16	519	14116	6.799615	0.951739	0.114183	18.27659
16	1001	14116	4.406843	0.967486	0.123459	20.39768
16	547	14116	6.451554	0.969524	0.12128	20.74011
16	20	14116	88.225	0.980894	0.131812	23.16635
16	236	14116	11.21504	0.991528	0.151697	27.2165
16	260	14116	10.17981	0.996731	0.173684	31.68853
16	276	14116	9.589674	0.998338	0.186529	34.69715
16	720	14116	4.901389	0.999257	0.201601	38.10488
16	734	14116	4.807902	0.999486	0.205055	39.60394
16	751	14116	4.699068	0.999676	0.210449	41.42966
16	43	14116	41.03488	0.9998	0.215976	43.2802
16	780	14116	4.524359	0.999857	0.218068	44.548
16	45	14116	39.21111	0.999865	0.214069	44.75892
16	362	14116	7.311464	0.999974	0.2429	50.54307
16	938	14116	3.76226	0.999999	0.303739	60.94246
16	10541	14116	1.255455	1	0.307906	62.47748
16	93	14116	18.97312	1	0.362392	70.72943

Category	Term	Count	%	PValue	Genes	List Total
GOTERM_	GO:0006414~translational elongation	12	100	1.44E-24	RPS26, RPS	12
GOTERM_	GO:0006412~translation	12	100	1.00E-18	RPS26, RPS	12
GOTERM_	GO:0044267~cellular protein metabolic pr	12	100	2.73E-09	RPS26, RPS	12
GOTERM_	GO:0019538~protein metabolic process	12	100	1.93E-08	RPS26, RPS	12
GOTERM_	GO:0034645~cellular macromolecule biosy	12	100	1.93E-08	RPS26, RPS	12
GOTERM_	GO:0009059~macromolecule biosynthetic	12	100	2.09E-08	RPS26, RPS	12
GOTERM_	GO:0010467~gene expression	12	100	3.92E-08	RPS26, RPS	12
GOTERM_	GO:0044249~cellular biosynthetic process	12	100	1.79E-07	RPS26, RPS	12
GOTERM_	GO:0009058~biosynthetic process	12	100	2.45E-07	RPS26, RPS	12
GOTERM_	GO:0044260~cellular macromolecule met	12	100	1.73E-05	RPS26, RPS	12
GOTERM_	GO:0042274~ribosomal small subunit biog	3	25	3.02E-05	RPS17, RPS	12
GOTERM_	GO:0043170~macromolecule metabolic pr	12	100	4.72E-05	RPS26, RPS	12
GOTERM_	GO:0042254~ribosome biogenesis	4	33.33333	9.88E-05	RPS17, RPS	12
GOTERM_	GO:0044237~cellular metabolic process	12	100	2.47E-04	RPS26, RPS	12
GOTERM_	GO:0022613~ribonucleoprotein complex k	4	33.33333	3.12E-04	RPS17, RPS	12
GOTERM_	GO:0044238~primary metabolic process	12	100	3.93E-04	RPS26, RPS	12
GOTERM_	GO:0008152~metabolic process	12	100	0.001175	RPS26, RPS	12
GOTERM_	GO:0006364~rRNA processing	3	25	0.002224	RPS17, RPS	12
GOTERM_	GO:0016072~rRNA metabolic process	3	25	0.002419	RPS17, RPS	12
GOTERM_	GO:0033119~negative regulation of RNA s	2	16.66667	0.003891	RPS26, RPS	12
GOTERM_	GO:0045934~negative regulation of nucle	4	33.33333	0.006296	RPS26, RPS	12
GOTERM_	GO:0051172~negative regulation of nitrog	4	33.33333	0.006538	RPS26, RPS	12
GOTERM_	GO:0034470~ncRNA processing	3	25	0.008875	RPS17, RPS	12
GOTERM_	GO:0034660~ncRNA metabolic process	3	25	0.013196	RPS17, RPS	12
GOTERM_	GO:0043484~regulation of RNA splicing	2	16.66667	0.015481	RPS26, RPS	12
GOTERM_	GO:0031324~negative regulation of cellula	4	33.33333	0.01603	RPS26, RPS	12
GOTERM_	GO:0010605~negative regulation of macrc	4	33.33333	0.016882	RPS26, RPS	12
GOTERM_	GO:0009892~negative regulation of metak	4	33.33333	0.019863	RPS26, RPS	12
GOTERM_	GO:0051253~negative regulation of RNA n	3	25	0.030955	RPS26, RPS	12
GOTERM_	GO:0006413~translational initiation	2	16.66667	0.034525	RPS17, RPS	12
GOTERM_	GO:0034101~erythrocyte homeostasis	2	16.66667	0.037541	RPS17, RPS	12
GOTERM_	GO:0044085~cellular component biogene	4	33.33333	0.038166	RPS17, RPS	12
GOTERM_	GO:0009987~cellular process	12	100	0.040207	RPS26, RPS	12
GOTERM_	GO:0006396~RNA processing	3	25	0.065399	RPS17, RPS	12
GOTERM_	GO:0048872~homeostasis of number of ce	2	16.66667	0.075249	RPS17, RPS	12

Pop Hits	Pop Total	Fold Enrich	Bonferroni	Benjamini	FDR
101	14116	139.7624	2.13E-22	2.13E-22	1.71E-21
331	14116	42.64653	1.48E-16	7.40E-17	1.19E-15
2355	14116	5.994055	4.04E-07	1.35E-07	3.26E-06
2812	14116	5.019915	2.86E-06	7.14E-07	2.30E-05
2812	14116	5.019915	2.86E-06	7.14E-07	2.30E-05
2832	14116	4.984463	3.09E-06	6.18E-07	2.49E-05
2999	14116	4.706902	5.81E-06	9.68E-07	4.67E-05
3442	14116	4.101104	2.65E-05	3.78E-06	2.13E-04
3542	14116	3.985319	3.63E-05	4.54E-06	2.92E-04
5214	14116	2.707326	0.002564	2.85E-04	0.020662
11	14116	320.8182	0.004467	4.48E-04	0.036032
5710	14116	2.472154	0.006957	6.34E-04	0.056185
122	14116	38.56831	0.014517	0.001218	0.117646
6636	14116	2.127185	0.035863	0.002805	0.293564
180	14116	26.14074	0.045139	0.003294	0.371137
6923	14116	2.039	0.056548	0.003873	0.467491
7647	14116	1.845953	0.15974	0.010819	1.391274
92	14116	38.3587	0.280755	0.019199	2.618005
96	14116	36.76042	0.301222	0.019715	2.844052
5	14116	470.5333	0.438395	0.02991	4.538271
512	14116	9.190104	0.6073	0.04566	7.248244
519	14116	9.066153	0.621233	0.045178	7.517583
187	14116	18.87166	0.732692	0.058208	10.07617
230	14116	15.34348	0.859981	0.081926	14.63722
20	14116	117.6333	0.900645	0.091727	16.9625
720	14116	6.535185	0.908517	0.091231	17.51243
734	14116	6.410536	0.919524	0.092367	18.35933
780	14116	6.032479	0.948662	0.104144	21.26086
362	14116	9.748619	0.990474	0.153126	31.24492
45	14116	52.28148	0.994483	0.164153	34.20304
49	14116	48.01361	0.996528	0.172021	36.61037
1001	14116	4.700633	0.996846	0.169543	37.09903
10541	14116	1.339152	0.997697	0.172873	38.67084
547	14116	6.451554	0.999955	0.261649	55.32728
100	14116	23.52667	0.999991	0.28861	60.62492

Category	Term	Count	%	PValue	Genes	List Total
GOTERM_	GO:0006414~translational elonga	8	88.88889	6.18E-15	RPS26, RPS	9
GOTERM_	GO:0006412~translation	8	88.88889	2.87E-11	RPS26, RPS	9
GOTERM_	GO:0044267~cellular protein me	8	88.88889	2.44E-05	RPS26, RPS	9
GOTERM_	GO:0034645~cellular macromole	8	88.88889	8.18E-05	RPS26, RPS	9
GOTERM_	GO:0019538~protein metabolic p	8	88.88889	8.18E-05	RPS26, RPS	9
GOTERM_	GO:0009059~macromolecule bio	8	88.88889	8.58E-05	RPS26, RPS	9
GOTERM_	GO:0010467~gene expression	8	88.88889	1.27E-04	RPS26, RPS	9
GOTERM_	GO:0044249~cellular biosynthesi	8	88.88889	3.21E-04	RPS26, RPS	9
GOTERM_	GO:0009058~biosynthetic proces	8	88.88889	3.89E-04	RPS26, RPS	9
GOTERM_	GO:0033119~negative regulation	2	22.22222	0.002831	RPS26, RPS	9
GOTERM_	GO:0044260~cellular macromole	8	88.88889	0.005068	RPS26, RPS	9
GOTERM_	GO:0042274~ribosomal small sub	2	22.22222	0.006219	RPS17, RPS	9
GOTERM_	GO:0043170~macromolecule me	8	88.88889	0.009142	RPS26, RPS	9
GOTERM_	GO:0043484~regulation of RNA s	2	22.22222	0.011281	RPS26, RPS	9
GOTERM_	GO:0044237~cellular metabolic p	8	88.88889	0.023865	RPS26, RPS	9
GOTERM_	GO:0006413~translational initiati	2	22.22222	0.025226	RPS17, RPS	9
GOTERM_	GO:0044238~primary metabolic p	8	88.88889	0.031132	RPS26, RPS	9
GOTERM_	GO:0045934~negative regulation	3	33.33333	0.031799	RPS26, RPS	9
GOTERM_	GO:0051172~negative regulation	3	33.33333	0.03261	RPS26, RPS	9
GOTERM_	GO:0006364~rRNA processing	2	22.22222	0.050978	RPS17, RPS	9
GOTERM_	GO:0016072~rRNA metabolic pro	2	22.22222	0.053142	RPS17, RPS	9
GOTERM_	GO:0008152~metabolic process	8	88.88889	0.057563	RPS26, RPS	9
GOTERM_	GO:0031324~negative regulation	3	33.33333	0.059282	RPS26, RPS	9
GOTERM_	GO:0010605~negative regulation	3	33.33333	0.061364	RPS26, RPS	9
GOTERM_	GO:0042254~ribosome biogenes	2	22.22222	0.067102	RPS17, RPS	9
GOTERM_	GO:0009892~negative regulation	3	33.33333	0.068394	RPS26, RPS	9
GOTERM_	GO:0009987~cellular process	9	100	0.09662	KRT25, RPS	9
GOTERM_	GO:0022613~ribonucleoprotein c	2	22.22222	0.097596	RPS17, RPS	9

Pop Hits	Pop Total	Fold Enrich	Bonferroni	Benjamini	FDR
101	14116	124.2332	8.46E-13	8.46E-13	7.29E-12
331	14116	37.90802	3.90E-09	1.95E-09	3.37E-08
2355	14116	5.328049	0.003314	0.001106	0.028645
2812	14116	4.462146	0.01106	0.002777	0.095953
2812	14116	4.462146	0.01106	0.002777	0.095953
2832	14116	4.430634	0.011603	0.002331	0.100685
2999	14116	4.183913	0.01707	0.002865	0.148506
3442	14116	3.645426	0.042746	0.006221	0.376374
3542	14116	3.542506	0.051596	0.0066	0.456209
5	14116	627.3778	0.319918	0.041933	3.273087
5214	14116	2.406512	0.498928	0.066767	5.790101
11	14116	285.1717	0.571889	0.074226	7.061137
5710	14116	2.19747	0.713228	0.098855	10.22061
20	14116	156.8444	0.786259	0.111918	12.46967
6636	14116	1.890831	0.962558	0.209145	24.68911
45	14116	69.70864	0.969032	0.206779	25.91314
6923	14116	1.812445	0.986448	0.235723	31.01375
512	14116	9.190104	0.987659	0.227807	31.56916
519	14116	9.066153	0.988989	0.221581	32.23902
92	14116	34.09662	0.999188	0.312383	45.89374
96	14116	32.67593	0.999405	0.310176	47.32435
7647	14116	1.640847	0.999685	0.318832	50.14026
720	14116	6.535185	0.999754	0.314619	51.19752
734	14116	6.410536	0.999818	0.31234	52.45089
122	14116	25.7122	0.999921	0.325378	55.75294
780	14116	6.032479	0.999935	0.319821	56.46721
10541	14116	1.339152	0.999999	0.412284	69.6638
180	14116	17.42716	0.999999	0.403855	70.04634

Category	Term	Count	%	PValue	Genes	List Total
GOTERM_	GO:0006414~translational elongation	10	83.33333	1.85E-18	RPS26, RPS	12
GOTERM_	GO:0006412~translation	10	83.33333	1.02E-13	RPS26, RPS	12
GOTERM_	GO:0044267~cellular protein metabolism	10	83.33333	3.93E-06	RPS26, RPS	12
GOTERM_	GO:0034645~cellular macromolecule biosynthesis	10	83.33333	1.81E-05	RPS26, RPS	12
GOTERM_	GO:0019538~protein metabolic process	10	83.33333	1.81E-05	RPS26, RPS	12
GOTERM_	GO:0009059~macromolecule biosynthesis	10	83.33333	1.93E-05	RPS26, RPS	12
GOTERM_	GO:0010467~gene expression	10	83.33333	3.15E-05	RPS26, RPS	12
GOTERM_	GO:0044249~cellular biosynthetic process	10	83.33333	1.01E-04	RPS26, RPS	12
GOTERM_	GO:0009058~biosynthetic process	10	83.33333	1.29E-04	RPS26, RPS	12
GOTERM_	GO:0044260~cellular macromolecule biosynthesis	10	83.33333	0.003134	RPS26, RPS	12
GOTERM_	GO:0033119~negative regulation of transcription	2	16.66667	0.003891	RPS26, RPS	12
GOTERM_	GO:0043170~macromolecule metabolism	10	83.33333	0.006453	RPS26, RPS	12
GOTERM_	GO:0045109~intermediate filament cytoskeleton	2	16.66667	0.007768	KRT25, KRT	12
GOTERM_	GO:0043484~regulation of RNA splicing	2	16.66667	0.015481	RPS26, RPS	12
GOTERM_	GO:0045104~intermediate filament cytoskeleton	2	16.66667	0.015481	KRT25, KRT	12
GOTERM_	GO:0045103~intermediate filament cytoskeleton	2	16.66667	0.017017	KRT25, KRT	12
GOTERM_	GO:0044237~cellular metabolic process	10	83.33333	0.0206	RPS26, RPS	12
GOTERM_	GO:0044238~primary metabolic process	10	83.33333	0.028303	RPS26, RPS	12
GOTERM_	GO:0009987~cellular process	12	100	0.040207	KRT25, RPS	12
GOTERM_	GO:0045934~negative regulation of transcription	3	25	0.058148	RPS26, RPS	12
GOTERM_	GO:0008152~metabolic process	10	83.33333	0.058493	RPS26, RPS	12
GOTERM_	GO:0051172~negative regulation of transcription	3	25	0.059573	RPS26, RPS	12

Pop Hits	Pop Total	Fold Enrich	Bonferroni	Benjamini	FDR
101	14116	116.4686	2.72E-16	2.72E-16	2.21E-15
331	14116	35.53877	1.49E-11	7.47E-12	1.21E-10
2355	14116	4.995046	5.78E-04	1.93E-04	0.004679
2812	14116	4.183262	0.002663	6.67E-04	0.021588
2812	14116	4.183262	0.002663	6.67E-04	0.021588
2832	14116	4.153719	0.00283	5.67E-04	0.02294
2999	14116	3.922419	0.004617	7.71E-04	0.037454
3442	14116	3.417587	0.014809	0.002129	0.120709
3542	14116	3.321099	0.01882	0.002372	0.153689
5214	14116	2.256105	0.369639	0.049981	3.666707
5	14116	470.5333	0.436202	0.055695	4.533056
5710	14116	2.060128	0.613912	0.08288	7.414814
10	14116	235.2667	0.682196	0.091106	8.862126
20	14116	117.6333	0.899083	0.161734	16.94435
20	14116	117.6333	0.899083	0.161734	16.94435
22	14116	106.9394	0.91978	0.164908	18.47331
6636	14116	1.772654	0.953102	0.184526	21.94023
6923	14116	1.699167	0.985309	0.231859	28.94082
10541	14116	1.339152	0.9976	0.298722	38.63556
512	14116	6.892578	0.99985	0.386909	50.97744
7647	14116	1.538294	0.999858	0.372698	51.19078
519	14116	6.799615	0.99988	0.363294	51.85291

Category	Term	Count	%	PValue	Genes
GOTERM_	GO:0006414~translational elongation	11	100	2.23E-22	RPS26, RPS
GOTERM_	GO:0006412~translation	11	100	4.39E-17	RPS26, RPS
GOTERM_	GO:0044267~cellular protein metabolic proc	11	100	1.64E-08	RPS26, RPS
GOTERM_	GO:0019538~protein metabolic process	11	100	9.72E-08	RPS26, RPS
GOTERM_	GO:0034645~cellular macromolecule biosynt	11	100	9.72E-08	RPS26, RPS
GOTERM_	GO:0009059~macromolecule biosynthetic pr	11	100	1.04E-07	RPS26, RPS
GOTERM_	GO:0010467~gene expression	11	100	1.85E-07	RPS26, RPS
GOTERM_	GO:0044249~cellular biosynthetic process	11	100	7.36E-07	RPS26, RPS
GOTERM_	GO:0009058~biosynthetic process	11	100	9.80E-07	RPS26, RPS
GOTERM_	GO:0044260~cellular macromolecule metabi	11	100	4.70E-05	RPS26, RPS
GOTERM_	GO:0043170~macromolecule metabolic proc	11	100	1.17E-04	RPS26, RPS
GOTERM_	GO:0044237~cellular metabolic process	11	100	5.25E-04	RPS26, RPS
GOTERM_	GO:0044238~primary metabolic process	11	100	8.02E-04	RPS26, RPS
GOTERM_	GO:0008152~metabolic process	11	100	0.002171	RPS26, RPS
GOTERM_	GO:0042254~ribosome biogenesis	3	27.27273	0.003186	RPS14, RPL
GOTERM_	GO:0033119~negative regulation of RNA spli	2	18.18182	0.003538	RPS26, RPS
GOTERM_	GO:0045934~negative regulation of nucleob	4	36.36364	0.004704	RPS26, RPS
GOTERM_	GO:0051172~negative regulation of nitrogen	4	36.36364	0.004887	RPS26, RPS
GOTERM_	GO:0022613~ribonucleoprotein complex bio	3	27.27273	0.006803	RPS14, RPL
GOTERM_	GO:0042274~ribosomal small subunit bioger	2	18.18182	0.007768	RPS14, RPS
GOTERM_	GO:0031324~negative regulation of cellular i	4	36.36364	0.012111	RPS26, RPS
GOTERM_	GO:0010605~negative regulation of macrom	4	36.36364	0.012764	RPS26, RPS
GOTERM_	GO:0043484~regulation of RNA splicing	2	18.18182	0.014083	RPS26, RPS
GOTERM_	GO:0009892~negative regulation of metabol	4	36.36364	0.015055	RPS26, RPS
GOTERM_	GO:0051253~negative regulation of RNA me	3	27.27273	0.025757	RPS26, RPS
GOTERM_	GO:0009987~cellular process	11	100	0.053856	RPS26, RPS
GOTERM_	GO:0006364~rRNA processing	2	18.18182	0.063315	RPS14, RPS
GOTERM_	GO:0016072~rRNA metabolic process	2	18.18182	0.065984	RPS14, RPS

List Total	Pop Hits	Pop Total	Fold Enrich	Bonferroni	Benjamini	FDR
11	101	14116	139.7624	3.30E-20	3.30E-20	2.65E-19
11	331	14116	42.64653	6.50E-15	3.25E-15	5.24E-14
11	2355	14116	5.994055	2.43E-06	8.11E-07	1.96E-05
11	2812	14116	5.019915	1.44E-05	3.59E-06	1.16E-04
11	2812	14116	5.019915	1.44E-05	3.59E-06	1.16E-04
11	2832	14116	4.984463	1.54E-05	3.09E-06	1.24E-04
11	2999	14116	4.706902	2.74E-05	4.57E-06	2.21E-04
11	3442	14116	4.101104	1.09E-04	1.56E-05	8.76E-04
11	3542	14116	3.985319	1.45E-04	1.81E-05	0.001168
11	5214	14116	2.707326	0.006934	7.73E-04	0.055998
11	5710	14116	2.472154	0.01713	0.001726	0.13899
11	6636	14116	2.127185	0.074814	0.007044	0.624016
11	6923	14116	2.039	0.112016	0.009851	0.951791
11	7647	14116	1.845953	0.275033	0.024437	2.555872
11	122	14116	31.55589	0.376433	0.033173	3.73061
11	5	14116	513.3091	0.408142	0.034362	4.134212
11	512	14116	10.02557	0.502361	0.04268	5.463008
11	519	14116	9.890349	0.515724	0.041756	5.669933
11	180	14116	21.38788	0.635897	0.054583	7.811066
11	11	14116	233.3223	0.684665	0.058935	8.872083
11	720	14116	7.129293	0.835252	0.086221	13.51227
11	734	14116	6.993312	0.850614	0.086557	14.19108
11	20	14116	128.3273	0.877429	0.091002	15.54691
11	780	14116	6.580886	0.89408	0.092999	16.53372
11	362	14116	10.63486	0.978974	0.148637	26.72042
11	10541	14116	1.339152	0.999724	0.279444	48.29198
11	92	14116	27.89723	0.999938	0.31087	54.12602
11	96	14116	26.73485	0.999959	0.312145	55.65939

Category	Term	Count	%	PValue	Genes
GOTERM_	GO:0006414~translational elongation	4	66.66667	3.52E-06	RPS19, RPL
GOTERM_	GO:0006412~translation	4	66.66667	1.23E-04	RPS19, RPL
GOTERM_	GO:0030097~hemopoiesis	3	50	0.002692	RPS19, RPL
GOTERM_	GO:0048534~hemopoietic or lymphoid organ	3	50	0.003258	RPS19, RPL
GOTERM_	GO:0002520~immune system development	3	50	0.003664	RPS19, RPL
GOTERM_	GO:0042274~ribosomal small subunit biogen	2	33.33333	0.003891	RPS19, RPS
GOTERM_	GO:0048856~anatomical structure developm	5	83.33333	0.004392	KRT25, RPS
GOTERM_	GO:0032502~developmental process	5	83.33333	0.010148	KRT25, RPS
GOTERM_	GO:0030154~cell differentiation	4	66.66667	0.012992	RPS19, FAB
GOTERM_	GO:0048869~cellular developmental process	4	66.66667	0.014589	RPS19, FAB
GOTERM_	GO:0030218~erythrocyte differentiation	2	33.33333	0.015141	RPS19, RPS
GOTERM_	GO:0048513~organ development	4	66.66667	0.015368	KRT25, RPS
GOTERM_	GO:0034101~erythrocyte homeostasis	2	33.33333	0.017239	RPS19, RPS
GOTERM_	GO:0032501~multicellular organismal proces	5	83.33333	0.031982	KRT25, RPS
GOTERM_	GO:0006364~rRNA processing	2	33.33333	0.03217	RPS19, RPS
GOTERM_	GO:0030099~myeloid cell differentiation	2	33.33333	0.032515	RPS19, RPS
GOTERM_	GO:0016072~rRNA metabolic process	2	33.33333	0.033549	RPS19, RPS
GOTERM_	GO:0048731~system development	4	66.66667	0.034545	KRT25, RPS
GOTERM_	GO:0048872~homeostasis of number of cells	2	33.33333	0.034927	RPS19, RPS
GOTERM_	GO:0044267~cellular protein metabolic proce	4	66.66667	0.035562	RPS19, RPL
GOTERM_	GO:0042254~ribosome biogenesis	2	33.33333	0.042479	RPS19, RPS
GOTERM_	GO:0002376~immune system process	3	50	0.043256	RPS19, RPL
GOTERM_	GO:0044085~cellular component biogenesis	3	50	0.043497	RPS19, FAB
GOTERM_	GO:0019538~protein metabolic process	4	66.66667	0.05728	RPS19, RPL
GOTERM_	GO:0034645~cellular macromolecule biosynt	4	66.66667	0.05728	RPS19, RPL
GOTERM_	GO:0009059~macromolecule biosynthetic pr	4	66.66667	0.058367	RPS19, RPL
GOTERM_	GO:0007275~multicellular organismal develo	4	66.66667	0.060186	KRT25, RPS
GOTERM_	GO:0022613~ribonucleoprotein complex biog	2	33.33333	0.062161	RPS19, RPS
GOTERM_	GO:0034470~ncRNA processing	2	33.33333	0.064514	RPS19, RPS
GOTERM_	GO:0010467~gene expression	4	66.66667	0.067898	RPS19, RPL
GOTERM_	GO:0034660~ncRNA metabolic process	2	33.33333	0.078867	RPS19, RPS
GOTERM_	GO:0044249~cellular biosynthetic process	4	66.66667	0.097085	RPS19, RPL

List Total	Pop Hits	Pop Total	Fold Enrich	Bonferroni	Benjamini	FDR
6	101	14116	93.17492	5.14E-04	5.14E-04	0.004182
6	331	14116	28.43102	0.017853	0.008967	0.146544
6	236	14116	29.90678	0.325382	0.12296	3.153555
6	260	14116	27.14615	0.378998	0.112286	3.804271
6	276	14116	25.57246	0.414836	0.101629	4.268653
6	11	14116	427.7576	0.434	0.0905	4.527805
6	2527	14116	4.655059	0.474102	0.087719	5.097266
6	3148	14116	3.736764	0.774432	0.169844	11.41698
6	1637	14116	5.748727	0.851803	0.191144	14.39516
6	1706	14116	5.516217	0.883008	0.193108	16.02713
6	43	14116	109.4264	0.892195	0.183308	16.58433
6	1738	14116	5.414653	0.895777	0.171746	16.81349
6	49	14116	96.02721	0.921037	0.177404	18.67205
6	4280	14116	2.748442	0.991311	0.287504	32.04667
6	92	14116	51.14493	0.991553	0.27259	32.20266
6	93	14116	50.59498	0.991982	0.260386	32.48941
6	96	14116	49.01389	0.993141	0.254033	33.34251
6	2330	14116	4.038913	0.994099	0.248098	34.15396
6	100	14116	47.05333	0.994431	0.239051	34.46351
6	2355	14116	3.996037	0.994941	0.232282	34.97408
6	122	14116	38.56831	0.998231	0.260503	40.30557
6	998	14116	7.072144	0.998429	0.254317	40.87882
6	1001	14116	7.050949	0.998486	0.245951	41.0559
6	2812	14116	3.34661	0.999818	0.301507	50.3958
6	2812	14116	3.34661	0.999818	0.301507	50.3958
6	2832	14116	3.322976	0.999846	0.296169	51.07151
6	2865	14116	3.2847	0.999884	0.294302	52.18334
6	180	14116	26.14074	0.999915	0.293215	53.36367
6	187	14116	25.16221	0.999941	0.293715	54.73577
6	2999	14116	3.137935	0.999965	0.298117	56.64389
6	230	14116	20.45797	0.999994	0.329546	62.3345
6	3442	14116	2.734069	1	0.381827	70.29532

Supplementary file 5: All enriched

Category	Term	Count	%	PValue
GOTERM_	GO:0003002~regionalization	4	3.921569	0.007609
GOTERM_	GO:0032501~multicellular organismal process	16	15.68627	0.008067
GOTERM_	GO:0007389~pattern specification process	4	3.921569	0.016591
GOTERM_	GO:0060173~limb development	3	2.941176	0.023121
GOTERM_	GO:0048736~appendage development	3	2.941176	0.023121
GOTERM_	GO:0009952~anterior/posterior pattern formation	3	2.941176	0.036825
GOTERM_	GO:0006812~cation transport	6	5.882353	0.049247
GOTERM_	GO:0009987~cellular process	38	37.2549	0.050164
GOTERM_	GO:0014706~striated muscle tissue development	3	2.941176	0.051486
GOTERM_	GO:0007275~multicellular organismal development	11	10.78431	0.052237
GOTERM_	GO:0048856~anatomical structure development	10	9.803922	0.05346
GOTERM_	GO:0060537~muscle tissue development	3	2.941176	0.055795
GOTERM_	GO:0051270~regulation of cell motion	3	2.941176	0.058735
GOTERM_	GO:0006811~ion transport	7	6.862745	0.065548
GOTERM_	GO:0001756~somitogenesis	2	1.960784	0.070694
GOTERM_	GO:0030001~metal ion transport	5	4.901961	0.075532
GOTERM_	GO:0048513~organ development	8	7.843137	0.076607
GOTERM_	GO:0048731~system development	9	8.823529	0.07744
GOTERM_	GO:0035282~segmentation	2	1.960784	0.081126
GOTERM_	GO:0007517~muscle organ development	3	2.941176	0.084044
GOTERM_	GO:0002366~leukocyte activation during immune response	2	1.960784	0.091443
GOTERM_	GO:0002263~cell activation during immune response	2	1.960784	0.091443
GOTERM_	GO:0032502~developmental process	11	10.78431	0.091497

GO terms of the closest protein-coding genes (<25 kb) of putative lincRNAs.

Genes	List Total	Pop Hits	Pop Total	Fold Enrich	FDR
MEOX2, PS	54	72	9430	9.701646	10.69718
ADORA3, N	54	1393	9430	2.005796	11.30575
MEOX2, PS	54	96	9430	7.276235	21.94915
MEOX2, PS	54	42	9430	12.47354	29.28445
MEOX2, PS	54	42	9430	12.47354	29.28445
MEOX2, PS	54	54	9430	9.701646	42.63725
TF, FXYD2,	54	357	9430	2.934952	52.67103
PDP1, ALDI	54	5457	9430	1.216039	53.3431
MEOX2, CB	54	65	9430	8.059829	54.29575
LAMA4, CD	54	1010	9430	1.901907	54.82893
LAMA4, CD	54	876	9430	1.993489	55.68434
MEOX2, CB	54	68	9430	7.704248	57.27617
LAMA4, AD	54	70	9430	7.484127	59.20529
TF, FXYD2,	54	511	9430	2.392187	63.36695
MEOX2, PS	54	13	9430	26.8661	66.24432
TF, FXYD2,	54	285	9430	3.063678	68.75587
LAMA4, CD	54	664	9430	2.103971	69.28956
LAMA4, CD	54	802	9430	1.959684	69.69744
MEOX2, PS	54	15	9430	23.28395	71.44206
MEOX2, CB	54	86	9430	6.091731	72.75642
ADORA3, P	54	17	9430	20.54466	75.8403
ADORA3, P	54	17	9430	20.54466	75.8403
LAMA4, CD	54	1120	9430	1.715112	75.86189

Supplementary file 6: The list of neighboring protein-coding genes

Novel lincRNAs	Human lincRNA	Upstream-prot	Chr	start	end	Novel lincRNA
CUFF.890.1	ENST00000423403.1 ENSG00000234421.1	C1orf87	1	34768542	34828225	CUFF.890.1
CUFF.35535.1	ENST00000409569.2 ENSG00000234421.1	BIM	3	105205578	105251184	CUFF.35535.1
CUFF.38336.1	ENST00000442176.1 ENSG00000234421.1	MEOX2	4	24100798	24175729	CUFF.38336.1
CUFF.38336.2	ENST00000442176.1 ENSG00000234421.1	MEOX2	4	24100798	24175729	CUFF.38336.2
CUFF.38336.3	ENST00000442176.1 ENSG00000234421.1	MEOX2	4	24100798	24175729	CUFF.38336.3
CUFF.41054.3	ENST00000511422.1 ENSG00000234421.1	FTMT	5	29459887	29460615	CUFF.41054.3

Novel lincRNAs	Mouse lincRNA	Upstream-prot	Chr	start	end	Novel lincRNA
CUFF.18704.1	ENSMUST00000186618.1 ENSMUST00000186618.1	NR2F2	18	9705041	9713831	CUFF.18704.1
CUFF.38336.1	ENSMUST00000190633.1 ENSMUST00000190633.1	MEOX2	4	24100798	24175729	CUFF.38336.1
CUFF.38336.2	ENSMUST00000190633.1 ENSMUST00000190633.1	MEOX2	4	24100798	24175729	CUFF.38336.2
CUFF.38336.3	ENSMUST00000190633.1 ENSMUST00000190633.1	MEOX2	4	24100798	24175729	CUFF.38336.3
CUFF.42025.1	ENSMUST00000182477.1 ENSMUST00000182477.1	TMEM161B	5	85105482	85176705	CUFF.42025.1
CUFF.42025.2	ENSMUST00000182477.1 ENSMUST00000182477.1	TMEM161B	5	85105482	85176705	CUFF.42025.2

Draft

ies of putative aligned lincRNAs to human and mouse lincRNAs.

st	end	Downstream	st	end	Upstream Chr	st	end
35660168	35677729	NFIA	35994682	36385631	C1orf87	1	60476517 60456505
105540222	1.06E+08	ANAPC1	105840829	105932962	BCL2L11	2	111881329 111903861
24177920	24186703	ISPD	24627906	25005090	MEOX2	7	15725511 15652014
24177920	24211143	ISPD	24627906	25005090	MEOX2	7	15725511 15652014
24177920	24230429	ISPD	24627906	25005090	MEOX2	7	15725511 15652014
30572212	30661429	PRR16	30674570	30852936	PRR16	5	119799973 120022404
st	end	Downstream	st	end	Upstream Chr	st	end
10755913	10775257	MCTP2	11749418	11953144	Nr2f2	7	70354439 70354656
24177920	24186703	ISPD	24627906	25005090	lspd	12	36381519 36689444
24177920	24211143	ISPD	24627906	25005090	lspd	12	36381519 36689444
24177920	24230429	ISPD	24627906	25005090	lspd	12	36381519 36689444
85290086	85330388	MEF2C	85606652	85706150	Mef2c	13	83573607 83662631
85293233	85330388	MEF2C	85606652	85706150	Mef2c	13	83573607 83662631

Draft

lincRNA-Hu	st	end	Downstream	st	end
RP11-436K8	61291008	61127142	NFIA	61330931	61920992
MIR4435-1	112252462	112187197	ANAPC1	112614373	112524572
AC005550.4	15728003	15735116	ISPD	16460691	16131322
AC005550.4	15728003	15735116	ISPD	16460691	16131322
AC005550.4	15728003	15735116	ISPD	16460691	16131322
CTD-2334D	120116913	120126473	FTMT	121187650	121188387

lincRNA-Mu	st	end	Downstream	st	end
6030442E23	71217049	71224409	Gm10295	71348961	71351485
Gm29007	37098550	37106789	Meox2	37108540	37179534
Gm29007	37098550	37106789	Meox2	37108540	37179534
Gm29007	37098550	37106789	Meox2	37108540	37179534
Gm17750	84025297	84064772	Tmem161b	84222314	84295870
Gm17750	84025297	84064772	Tmem161b	84222314	84295870

Draft

Supplementary file 7: The list of upstream and downstream protein-coding genes of predicted and human linc

chr	transcriptid-linc	st	end	Upstream protein coding gene	Downstream protein coding gene
5	ENSG0000027211	1.41E+08	141120765	ARAP3	PCDH1
4	ENSG0000024761	1.4909961	15002045	BOD1L1	CPEB2
1	ENSG0000022495	1.77E+08	177366272	BRINP2	SEC16B
1	ENSG0000022646	60540249	60640491	C1orf87	NFIA
8	ENSG0000025313	34228439	34346731	DUSP26	UNC5D
16	ENSG0000026181	14018880	14021077	ERCC4	MKL2
8	ENSG0000025411	1246789	1248760	ERICH1	DLGAP2
10	ENSG0000022541	1.29E+08	129291722	FAM196A	NPS
5	ENSG0000025111	1.54E+08	153898987	HAND1	LARP1
12	ENSG0000025771	1.03E+08	103178675	IGF1	PAH
8	ENSG0000020552	57978358	57984126	IMPAD1	FAM110B
1	ENSG0000023751	2.35E+08	234959989	IRF2BP2	TOMM20
4	ENSG0000024951	16973275	17073903	LDB2	QDPR
16	ENSG0000026021	80040662	80045745	MAF	DYNLRB2
5	ENSG0000025018	8408982	88439090	MEF2C	CETN3
7	ENSG0000022951	15841297	15842360	MEOX2	ISPD
1	ENSG0000022421	1.7E+08	170284208	METTTL11B	GORAB
3	ENSG0000022653	34159334	34562877	PDCD6IP	ARPP21
15	ENSG0000025921	70195638	70198509	RPLP1	TLE3
15	ENSG0000023571	25484831	25578791	SNRPN	UBE3A
20	ENSG0000022451	47391929	47412327	SULF2	PREX1
14	ENSG0000025851	63298126	63299547	SYT16	KCNH5
18	ENSG0000026041	53579623	53581059	TCF4	TXNL1
4	ENSG0000024841	65702202	65705553	TECRL	EPHA5
5	ENSG0000024871	87713135	87733269	TMEM161F	MEF2C
2	ENSG0000022771	13723048	13758152	TRIB2	FAM84A
1	ENSG0000022351	56414963	56415966	USP24	PPAP2B
3	ENSG0000023951	1.15E+08	114876514	ZBTB20	GAP43

lncRNAs.

chr	transcriptid-linc	gene id lincRN	st	end	strand	Upstream strand	Downstream strand
chr24	CUFF.30641.1	CUFF.30641	13040405	13043643	+	ERCC4	MKL2
chr6	CUFF.43860.1	CUFF.43860	1.11E+08	1.11E+08	-	LDB2	QDPR
chr12	CUFF.9684.1	CUFF.9684	36022512	36023183	+	METTL11B	GORAB
chr19	CUFF.20106.1	CUFF.20106	9566928	9567227	+	PDCD6IP	ARPP21
chr19	CUFF.20106.2	CUFF.20106	9566928	9567227	+	PDCD6IP	ARPP21
chr13	CUFF.12383.1	CUFF.12383	75851917	75853076	-	SULF2	PREX1
chr23	CUFF.30009.1	CUFF.30009	56019369	56037967	-	TCF4	TXNL1
chr23	CUFF.30001.1	CUFF.30001	55642792	55681188	-	TCF4	TXNL1
chr1	CUFF.806.2	CUFF.806	30549105	30576752	-	USP24	PPAP2B
chr5	CUFF.41470.1	CUFF.41470	50226788	50233681	+	ARAP3	PCDH1
chr6	CUFF.43826.1	CUFF.43826	1.1E+08	1.1E+08	+	BOD1L1	CPEB2
chr12	CUFF.10217.1	CUFF.10217	56631815	56678310	+	BRINP2	SEC16B
chr1	CUFF.890.1	CUFF.890	35660168	35677729	+	C1orf87	NFIA
chr26	CUFF.33008.1	CUFF.33008	28055710	28057365	+	DUSP26	UNC5D
chr26	CUFF.32466.1	CUFF.32466	485489	486039	-	ERICH1	DLGAP2
chr22	CUFF.29175.1	CUFF.29175	45951599	45956962	-	FAM196A	NPS
chr5	CUFF.41701.1	CUFF.41701	63477464	63479690	+	HAND1	LARP1
chr3	CUFF.36747.1	CUFF.36747	1.72E+08	1.72E+08	+	IGF1	PAH
chr9	CUFF.48400.2	CUFF.48400	37151944	37214145	+	IMPAD1	FAM110B
chr9	CUFF.48400.3	CUFF.48400	37160006	37203843	-	IMPAD1	FAM110B
chr25	CUFF.31644.1	CUFF.31644	7613190	7615165	+	IRF2BP2	TOMM20
chr14	CUFF.12663.2	CUFF.12663	6052589	6067837	+	MAF	DYNLRB2
chr5	CUFF.42038.1	CUFF.42038	85766610	85778514	+	MEF2C	CETN3
chr5	CUFF.42039.1	CUFF.42039	85782820	85847140	+	MEF2C	CETN3
chr4	CUFF.38336.1	CUFF.38336	24177920	24186703	+	MEOX2	ISPD
chr4	CUFF.38336.3	CUFF.38336	24177920	24230429	+	MEOX2	ISPD
chr4	CUFF.38336.2	CUFF.38336	24177920	24211143	+	MEOX2	ISPD
chr7	CUFF.44461.2	CUFF.44461	16157260	16160851	+	RPLP1	TLE3
chr7	CUFF.44461.1	CUFF.44461	16156471	16160851	+	RPLP1	TLE3
chr18	CUFF.18479.1	CUFF.18479	996771	1002862	+	SNRPN	UBE3A
chr7	CUFF.45638.1	CUFF.45638	71203649	71204718	-	SYT16	KCNH5
chr6	CUFF.43360.1	CUFF.43360	80706321	80708123	+	TECRL	EPHA5
chr5	CUFF.42025.2	CUFF.42025	85293233	85330388	+	TMEM161E	MEF2C
chr5	CUFF.42025.1	CUFF.42025	85290086	85330388	+	TMEM161E	MEF2C
chr3	CUFF.33831.1	CUFF.33831	22979007	22981342	-	TRIB2	FAM84A
chr1	CUFF.3702.3	CUFF.3702	1.78E+08	1.78E+08	-	ZBTB20	GAP43
chr1	CUFF.3704.1	CUFF.3704	1.78E+08	1.78E+08	-	ZBTB20	GAP43
chr1	CUFF.3702.2	CUFF.3702	1.78E+08	1.78E+08	-	ZBTB20	GAP43
chr1	CUFF.3702.1	CUFF.3702	1.78E+08	1.78E+08	-	ZBTB20	GAP43